Classical descriptions of quantum computations:

Foundations of quantum computation via hidden variable models, quasiprobability representations, and classical simulation algorithms

by

Michael Zurel

B.Sc., University of British Columbia, 2019

M.Sc., University of British Columbia, 2020

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Physics)

The University of British Columbia
(Vancouver)

August 2024

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the thesis entitled:

Classical descriptions of quantum computations: Foundations of quantum computation via hidden variable models, quasiprobability representations, and classical simulation algorithms

submitted by **Michael Zurel** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy** in **Physics**.

Examining Committee:

Dr. Robert Raussendorf, Alexander von Humboldt Professor, Institute of Theoretical Physics, Leibniz University Hannover
Supervisor

Dr. William G. Unruh, Professor, Department of Physics and Astronomy, University of British Columbia
Co-supervisor

Dr. Joseph Salfi, Assistant Professor, Department of Electrical and Computer Engineering, University of British Columbia
Supervisory Committee Member

Dr. Mark Van Raamsdonk, Professor, Department of Physics and Astronomy, University of British Columbia
Supervisory Committee Member

Dr. Roman Krems, Professor, Department of Chemistry, University of British Columbia
University Examiner

Dr. Fei Zhou, Professor, Department of Physics and Astronomy, University of British Columbia
University Examiner

Dr. Mark Howard, Assistant Professor, School of Mathematical and Statistical Sciences, University of Galway
External Examiner

Additional Supervisory Committee Members:

Dr. Gordon W. Semenoff, Professor, Department of Physics and Astronomy, University of British Columbia
Supervisory Committee Member

Abstract

Quasiprobability representations serve as a bridge between classical and quantum descriptions of physical systems. In these representations, nonnegativity allows for a probabilistic interpretation, aligning the description with classical physics. However, the capacity to model quantum systems hinges on the use of negative quasiprobabilities. Accordingly, negativity is considered a hallmark of genuinely quantum behaviour. This principle has been applied to quantum information processing where negativity in the Wigner function is necessary for a quantum computational advantage. This is demonstrated by an efficient classical simulation algorithm for quantum computations in which all components of the computation remain nonnegative.

However, when constructing quasiprobability representations for quantum computation to which this statement applies, a marked difference arises between the cases of even and odd Hilbert space dimensions. We find that Wigner functions with the properties required to describe quantum computation do not exist in any even dimension. We establish that the obstructions to the existence of such Wigner functions are cohomological.

In order to recover the properties required for classical simulation of quantum computation in any dimension, some constraints that traditionally define a Wigner function must be relaxed. We consider several examples of these generalized quasiprobability representations, and we find that when sufficiently general representations are admitted, no negativity is required to represent universal quantum computation. The result is a hidden variable model that represents all elements of universal quantum computation probabilistically. Since this model can simulate any quantum computation, the simulation must be inefficient in general. However, in certain restricted settings, the simulation is efficient, allowing for a broader class of quantum circuits to be efficiently classically simulated than those covered by the stabilizer formalism and Wigner function methods.

With this hidden variable model, we present a formulation of quantum mechanics that replaces the central notion of state, as a complex vector or density operator, with a bit string. This formulation applies to universal quantum computation, and hence all finite-dimensional quantum mechanics. Thus, we present a surprising response to Wheeler's "It from Bit" challenge. Alongside coherence, entanglement, and contextuality, this provides a new approach to characterizing quantum computational advantage.

Lay Summary

Quantum phenomena like entanglement and contextuality seem to put quantum theory well beyond the grasp of our classical intuition. To get a grip on this problem we often ask: to what extent can we assign a classical explanation to the predictions of quantum theory? While this is already an interesting question, the introduction of quantum computation provides it with a useful application. That is, although there are clear tasks on which quantum computers are known to outperform classical computers, the source of this quantum advantage is unclear. Often the failure of classical explanations heralds a quantum computational advantage over classical computing. Thus, by characterizing features of quantum theory that evade classical explanations, we can get a better grasp on the source of quantum advantage in useful computational tasks.

Preface

This dissertation contains material that was previously published in Refs. [1–11]. The bulk of the content has previously appeared in coauthored papers and preprints. For the most part, these papers were genuinely coauthored, with all authors playing a significant role in their development and writing. A detailed breakdown of each coauthor's contributions is included below. For all previously published content, copyrighted material is included with permission from the copyright owners.

The published papers and preprints, Refs. [1–8], are generally structured with introduction, background/preliminaries, results, and discussion/conclusion sections. This structure is largely preserved where these papers are presented in Chapters 3–10, with one exception: there is some background material that is common to many chapters. To reduce redundancy, this has been removed from the those chapters and placed in a unified background chapter, Chapter 2, along with some additional background information. The rest of the content from these papers is presented more or less unmodified, except for some changes in notation for consistency across chapters, reordering of sections, and corrections of errors.

Chapter 1. This chapter serves to introduce the thesis. It is largely original and written by me, with some parts taken from the introductions of the published papers and preprints [1–8] that are presented in Chapters 3–10.

Chapter 2. This chapter provides general background information required for understanding the main results of this thesis, and it introduces notation and nomenclature used in the rest of the thesis. Most of this chapter was written by me, with some content taken from Chapter 2 of Ref. [9] (also written by me). Some material was also taken from the background sections of Refs. [1–8]. In particular, parts of Section 2.3 were taken from Appendix A of Ref. [7], which was written by me, and Section 2.5.2 is from Ref. [4] and was written by Cihan Okay.

Chapter 3. This chapter was previously published in Ref. [1]. This chapter was coauthored by Robert Raussendorf, Juani Bermejo-Vega, Emily Tyhurst, Cihan Okay, and me. The initial idea as well as most of the the main results of this work are due to Robert Raussendorf. This includes the definition of the generalized phase space and quasiprobability function, and the state update rules for Clifford gates and Pauli measurements. Efficiency of the classical simulation algorithm was proven by Juani Bermejo-Vega. My main contributions to this work were the characterization of the multiqubit phase space in Section 3.3.3, the observation made in Section 3.6.5, and the computational methods

and numerical results regarding volumes of positively representable states and robustness of specific states in Sections 3.3.2, 3.5.1, and 3.6. Section A.2.1 is from Ref. [9] and was written by me.

After publication of Ref. [1], Lemma 10 therein was found to be incorrect. This was pointed out to us by Arne Heimendahl. An erratum was published which amends the statement of the lemma by restricting its scope, and gives a corrected proof [10]. Lemma 10 of Ref. [1] is presented as Lemma 8 in Chapter 3, this is the corrected lemma and proof from the erratum.

Chapter 4. This chapter was previously published in Ref. [2]. This chapter was coauthored by me, Cihan Okay, and Robert Raussendorf. All coauthors contributed significantly to this work. I wrote the first proof of Theorem 6, which was later refined by Robert Raussendorf. I also wrote the proof of Theorem 7. The introduction was written by Robert Raussendorf, and the discussion includes contributions from all coauthors.

Section A.3.1 contains additional material about the Λ polytopes, the main objects of study in Ref. [2], that was left out of the published paper. It is a summary of the content of Section 4.3 from Ref. [9].

Chapter 5. This chapter was previously published in Ref. [4]. It was coauthored by Robert Raussendorf, Cihan Okay, me, and Polina Feldmann. Most of the main cohomological arguments and calculations were by Robert Raussendorf and Cihan Okay, with significant proof checking, writing, and revisions contributed by Polina Feldmann. My contributions include the writing of Section 5.3, and the generalization of the Mermin square state-independent proof of contextuality shown in Figure 5.1 and used in the proof of Lemma 24.

Chapter 6. This chapter was previously published in Ref. [3]. This chapter was coauthored by Cihan Okay, me, and Robert Raussendorf. Cihan Okay wrote the first version of the proof of Theorem 15, which was then refined with input from me and Robert Raussendorf. Robert Raussendorf wrote the proof of Theorem 16 with input from me and Cihan Okay. Cihan Okay wrote Section 6.5.

In the original publication of Ref. [3], there was a gap in the proof of Theorem 2 (this is Theorem 15 in Chapter 6). This was corrected in an erratum [11]. The proof included in Chapter 6 is the corrected version from the erratum.

Chapter 7. This chapter was previously published in Ref. [7]. This chapter was coauthored by me, Cihan Okay, Robert Raussendorf, and Arne Heimendahl. Most of the chapter was written by me, including the introduction, the discussion, and the proofs of the main results, Theorems 18, 19, and 20. Cihan Okay contributed Section 7.5 including the proof of Theorem 21. The proof of Lemma 27 was written by Arne Heimendahl with input from Cihan Okay.

Chapter 8. This chapter was previously published in Ref. [8]. This chapter was coauthored by me and Arne Heimendahl. The original idea for the paper was due to Arne

Heimendahl through pointing out that Lemma 10 of Ref. [1] was incorrect. I wrote most of the paper. Arne Heimendahl contributed the original proofs of Theorems 22 and 24, with significant revisions by me to ensure consistency of notation and nomenclature. Arne Heimendahl wrote Section 8.5.

Chapter 9. This chapter was previously published in Ref. [6]. This chapter was coauthored by me, Larry Cohen, and Robert Raussendorf. The original idea was due to Robert Raussendorf, the main results and their proofs as well as most of the writing were contributed by me and Larry Cohen. I wrote the proof of Theorem 27. Larry Cohen wrote the proof of efficiency of the simulation algorithm given in Section 9.4.

Chapter 10. This chapter was previously published in Ref. [5]. This chapter was coauthored by me, Cihan Okay, and Robert Raussendorf. The original idea was due to Robert Raussendorf. Robert Raussendorf wrote the introduction and Section 10.2. I wrote most of the rest of the paper. The proofs of the main results, Theorem 29 and Corollary 13, were written by me based on discussions with Robert Raussendorf.

Chapters 11. This chapter is original and was written by me.

Table of Contents

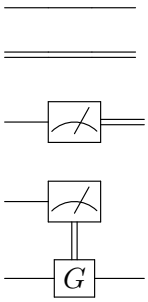
List of Tables

List of Figures

List of Symbols

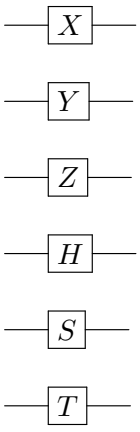
Quantum circuit components

- Quantum wire
- Classical wire
- Measurement of a qubit in the computational basis
- Classical conditioning of gate G on the outcome of a measurement



Single qubit gates

- Pauli-X gate: $X = |0\rangle\langle 1| + |1\rangle\langle 0|$
- Pauli-Y gate: $Y = -i\,|0\rangle\langle 1| + i\,|1\rangle\langle 0|$
- Pauli-Z gate: $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$
- Hadamard gate: $H = \frac{1}{\sqrt{2}}\left(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 1|\right)$
- Phase gate: $S = |0\rangle\langle 0| + i\,|1\rangle\langle 1|$
- T gate: $T = |0\rangle\langle 0| + e^{i\pi/4}\,|1\rangle\langle 1|$

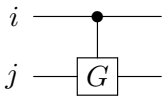


A subscript, e.g. Z_i with $i \in \mathbb{N}$, indicates the qubit on which the gate acts. For example, the n-qubit gate Z_i is defined as

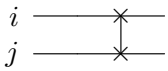
$$Z_i := \underbrace{\mathbb{1} \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}}_{i-1 \text{ times}} \otimes Z \otimes \underbrace{\mathbb{1} \otimes \mathbb{1} \otimes \cdots \otimes \mathbb{1}}_{n-i \text{ times}}.$$

Multiqubit gates

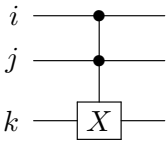
- controlled gates: $CG_{i,j} = |0\rangle\langle 0|_i \mathbb{1}_j + |1\rangle\langle 1|_i G_j$



- Swap gate: $SWAP_{i,j} = |00\rangle\langle 00|_{i,j} + |01\rangle\langle 10|_{i,j} + |10\rangle\langle 01|_{i,j} + |11\rangle\langle 11|_{i,j}$

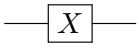


- Toffoli gate: $CCX_{i,j,k} = (\mathbb{1} - |11\rangle\langle 11|)_{i,j}\mathbb{1}_k + |11\rangle\langle 11|_{i,j} X_k$

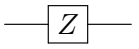


Qu*d*it gates

- Generalized Pauli-X gate: $X = \sum_{k\in\mathbb{Z}_d} |k+1 \mod d\rangle\langle k|$

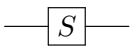


- Generalized Pauli-Z gate: $Z = \sum_{k\in\mathbb{Z}_d} \omega^k |k\rangle\langle k|$, with $\omega = \exp(2\pi i/d)$

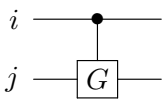


- Hadamard gate: $H = \frac{1}{\sqrt{d}} \sum_{k,\ell\in\mathbb{Z}_d} \omega^{k\ell} |k\rangle\langle \ell|$

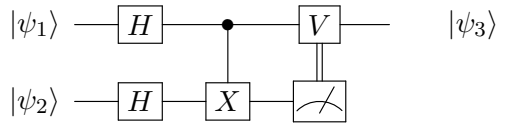
- Phase gate: $S = \sum_{k\in\mathbb{Z}_d} \omega^{k(k+1)2^{-1}} |k\rangle\langle k|$



- controlled gates: $CG_{i,j} = \sum_{k\in\mathbb{Z}_d} |k\rangle\langle k|_i G_j^k$



Circuit diagrams are read from left to right. For example, the circuit



acts on each qubit with a Hadamard gate, followed by a controlled-X gate with the first qubit the control qubit, followed by a measurement of the second qubit in the computational basis and the gate V on the first qubit conditioned on the outcome of the measurement.

Common magic states

- $|H\rangle\langle H| = \frac{1}{2}\mathbb{1} + \frac{1}{2\sqrt{2}}(X+Y)$
- $|T\rangle\langle T| = \frac{1}{2}\mathbb{1} + \frac{1}{2\sqrt{3}}(X+Y+Z)$

Notation for Pauli and Clifford operators

- n is the number of qudits, d is the qudit Hilbert space dimension
- $\omega = \exp(2\pi i/d), \quad \mu = \begin{cases} \omega & \text{if } d \text{ is odd} \\ \sqrt{\omega} & \text{if } d \text{ is even} \end{cases}$
- $T_a = \mu^{\phi(a)} \bigotimes_{k=1}^{n} Z^{[a_z]_k} X^{[a_x]_k}, \quad \forall a = (a_z, a_x) \in \mathbb{Z}_d^n \times \mathbb{Z}_d^n =: E$
- Unless otherwise specified, we use the phase convention
$$\phi(a) = \begin{cases} -\langle a_z | a_x \rangle \cdot 2^{-1} \in \mathbb{Z}_d & \text{if } d \text{ is odd,} \\ -\langle a_z | a_x \rangle \in \mathbb{Z}_{2d} & \text{if } d \text{ is even} \end{cases}$$
- $T_a T_b = \omega^{-\beta(a,b)} T_{a+b} = \omega^{[a,b]} T_b T_a$
- The n-qudit Clifford group is denoted $\mathcal{C}\ell_n$ or $\mathcal{C}\ell$.
- For any $g \in \mathcal{C}\ell$, $g(T_a) := g T_a g^\dagger = \omega^{\Phi_g(a)} T_{S_g a}$ where $S_g \in \mathrm{Sp}(\mathbb{Z}_d^{2n})$
- $\Pi_I^r := \frac{1}{|I|} \sum_{a \in I} \omega^{-r(a)} T_a$ is the projector onto the eigenspace of the Pauli operators $\{T_a \mid a \in I\}$ corresponding to eigenvalues $\{\omega^{r(a)} \mid a \in I\}$. For a single Pauli operator T_a, the projector corresponding to eigenvalue ω^s is $\Pi_a^s = \frac{1}{d} \sum_{j \in \mathbb{Z}_d} \omega^{-js} T_a^j$, and for qubits, $\Pi_a^s = \frac{1}{2}(\mathbb{1} + (-1)^s T_a)$.
- We usually denote the identity operator by $\mathbb{1}$, though we sometimes also denote it by 1 when it is clear from context that this should be interpreted as an operator, or suppress it entirely, for example in $T_a T_b T_a^{-1} T_b^{-1} = \omega^{[a,b]}$.

Miscellanea

- The symbols V and $\mathcal{V}$ refer to different phase spaces in different chapters. V generally refers to the phase space of the standard Wigner function, which is $V = \mathbb{R}^{2D}$ for continuous-variable quantum theory in D dimensions (see Section 2.4), or $V = \mathbb{Z}_d^{2n}$ or $\mathbb{F}_d^{2n}$ for finite-dimensional quantum theory of n qudits each with Hilbert space dimension d. $\mathcal{V}$ refers to various generalized phase spaces. In Chapters 3 and 8, $\mathcal{V}$ (or $\mathcal{V}_n$) is the CNC phase space, in Chapters 4, 7, and 10, it is an index set for the vertices of the Λ polytopes, and in Chapter 9, it is the augmented line graph phase space constructed in Section 9.3.
- Subscripts and superscripts (e.g. $\mathcal{V}_n$, Λ_n) are used to indicate parameters. Generally, Λ_n refers to a specific polytope for a fixed number of qubits (or qudits) n. We refer to these polytopes collectively as the Λ polytopes, or Λ (no subscript). The set of vertices of Λ_n is $\{A_\alpha \mid \alpha \in \mathcal{V}_n\}$. That is, A_α is a vertex of Λ_n and it is labeled by $\alpha \in \mathcal{V}_n$. $\mathcal{V}_n$ is the set of all labels of vertices of Λ_n. In some places, subscripts and superscripts are omitted where the relevant parameters can be inferred from context.
- Following the convention of Ref. [12], in this thesis "z" is pronounced "zed".

List of Abbreviations

CNC Closed under inference and NonContextual

FBQC Fusion-Based Quantum Computation

HVM Hidden Variable Model

LOCC Local Operations and Classical Communication

LOSR Local Operations and Shared Randomness

MBQC Measurement-Based Quantum Computation

NCHVM NonContextual Hidden Variable Model

NCVA NonContextual Value Assignment

NISQ Noisy Intermediate-Scale Quantum computation

OB Operator Basis

PBC Pauli-Based quantum Computation

QCM Quantum Computation with Magic states

QR Quasiprobability representation

QRT Quantum Resource Theory

SW Stratonovich-Weyl

Acknowledgments

First and foremost, I would like to express my deepest gratitude to my supervisors, Robert Raussendorf and William Unruh, for their guidance and mentorship throughout the course of my PhD. Robert was instrumental in my decision to pursue graduate studies, having encouraged me to continue after supervising my undergraduate honours project. When Robert relocated to Hannover, Bill stepped in and took on the responsibility as my co-supervisor. I would also like to extend my thanks to the other members of my supervisory committee, Joe Salfi, Gordon Semenoff, and Mark Van Raamsdonk for their feedback and continuous support. My sincere appreciation also goes out to the other members of my examining committee for contributing their time.

Much of this thesis was written in collaboration with coauthors. I appreciate the contributions of all my coauthors, including Robert Raussendorf, Cihan Okay, Juani Bermejo-Vega, Emily Tyhurst, Polina Feldmann, Arne Heimendahl, and Larry Cohen. A detailed breakdown of their contributions to the work presented here is given in the preface. I am also grateful for discussions with collaborators on work that did not make it into this thesis, including Filipa Peres, Ernesto Galvão, Rui Soares Barbosa, Angela Karanjai, Hakop Pashayan, Stephen Bartlett, David Gross, and Nadish de Silva.

I am immensely grateful to my friends and family for their unconditional support and understanding throughout this journey. I also want to acknowledge the inspiration provided by my heroes of physics, including Albert Einstein, Emmy Noether, Grete Hermann, Eugene Wigner, John Archibald Wheeler, Edwin Thompson Jaynes, John Stewart Bell, and Hugh Everett III.

Most of this work was undertaken at the University of British Columbia Vancouver campus, situated on the traditional, ancestral, and unceded territory of the xʷməθkʷəy̓əm (Musqueam) People. Work was also undertaken in other parts of Vancouver on the unceded territory of the Coast Salish Peoples, including the territories of the xʷməθkʷəy̓əm (Musqueam), Sḵwx̱wú7mesh (Squamish), Stó:lō, and Səlilwətaɬ (Tsleil-Waututh) Nations.

Lastly, I would like to acknowledge the financial support from the Natural Sciences and Engineering Research Council of Canada (NSERC). I was supported by NSERC through the Alexander Graham Bell Canada Graduate Scholarships—Doctoral program (CGS-D) and the Canada Graduate Scholarships—Michael Smith Foreign Study Supplements (CGS-MSFSS). Additional support was provided by the University of British Columbia, Department of Physics & Astronomy, and the Stewart Blusson Quantum Matter Institute. This work was also supported in part through the Canada First Research Excellence Fund, Quantum Materials and Future Technologies Program, and by the Horizon Europe project FoQaCiA, GA no. 101070558.

Chapter 1

Introduction

The field of quantum computation has seen a remarkable surge in interest in recent years. Supplementing academic research, big tech companies and startups have recently joined the effort to realize useful quantum technologies. This surge may not be so surprising given the promise of quantum computers to solve some computational problems that are intractable for any classical computer, even the largest modern supercomputers. Although it was originally conceived in the early 1980s as a means of simulating quantum systems [13, 14], the fire has been fueled by the more recent development of quantum algorithms for useful tasks with broader appeal [15–17].

But in spite of the age of the field and the recent surge in interest, a crucial question at the foundation of quantum computation remains unresolved, namely, *what is the essential quantum resource that provides the quantum computational advantage over classical computation?* This is clearly an important question as its resolution could inform the design of quantum computer architectures and the development of new quantum algorithms. This question is the main focus of this thesis.

Classical simulation of quantum computations

A common way to approach this question is to try to determine families of quantum circuits which can be efficiently simulated by a classical computer [18–21]. Once an efficient classical simulation algorithm is established for a certain family of circuits, it can be concluded that some quantum resource beyond the scope of that family is required for the possibility of a quantum computational advantage.

As a simple motivating example of this approach, consider the family of quantum circuits on n qubits consisting of only local operations and local measurements performed on an unentangled state. Although you may have access to many qubits resulting in a large Hilbert space (dimension exponential in n), if the initial state is unentangled and the operations do not allow for the creation of entanglement, you are confined to a very small subspace of the full Hilbert space (dimension linear in n). We can simulate these circuits classically by keeping track of each qubit separately. Therefore, we can conclude that entanglement is necessary for a quantum computational advantage. The source of this entanglement may be different depending on the model of quantum computation. In measurement-based quantum computation (MBQC) [22], all operations are local measurements performed on

an entangled state, i.e., all the entanglement comes from the initial preparation of the of the input state. On the other hand, in the circuit model the initial state is typically unentangled and the entanglement is generated by the unitary gates of the circuit. But the fact remains entanglement in some form is necessary for a quantum computational advantage [18].

As another example, consider quantum circuits consisting of only Toffoli gates (see the List of Symbols on page xviii). These gates are universal for classical computation, but they cannot perform arbitrary quantum computations because they cannot generate quantum superpositions of computational basis states. This limitation also results in an efficient classical simulation of these circuits, in which quantum Toffoli gates are mapped directly to classical Toffoli gates, and quantum computational basis states are mapped to classical bit strings. Supplementing the Toffoli gates with Hadamard gates, which do generate coherent superpositions of computational basis states, we obtain a universal gate set for quantum computation [23]. Thus, we conclude that superpositions of states, or the capacity to generate superpositions, is also necessary.

In addition to entanglement and superposition, many other candidates for essential quantum properties have been proposed, such as large Hilbert space dimension [24], quantum contextuality [20, 25, 26], and Wigner function negativity [19, 27–32]. These properties give similar necessary conditions for quantum computational advantage.

Resource theories in quantum information processing

One way to systematize this approach is through quantum resource theories [33]. In a quantum resource theory for some quantum information processing task, we partition the quantum states and operations into two classes, (1) the "free" states and operations, which are considered easy to prepare and implement for some system, and (2) the resource states, which are considered essential for accomplishing the task at hand, or for achieving a quantum advantage. Crucially, the set of free states is preserved by all the free operations. The value of resources are then characterized by *resource monotones*—real-valued functions of quantum states which are nonincreasing under the free operations of the resource theory.

This approach has been fruitful, for example, in characterizing entanglement as a resource for quantum communication tasks. There, we typically assume a spatial separation of two parties each with access to a quantum system. In the LOCC resource theory [34], the free operations consist of Local Operations and Classical Communication.[1] No quantum advantage is attainable with just these free operations, but when supplemented with resource states (entangled states), quantum advantage is possible.

Magic and the resource theory of stabilizer quantum computation

One model of quantum computation that lends itself to a resource theoretic perspective is quantum computation with magic states (QCM) [37, 38]. QCM is a universal model of quantum computation closely related to the circuit model and is one of the leading candidates for scalable fault-tolerant quantum computation [39]. In QCM, the allowed operations are restricted to those of the stabilizer subtheory, consisting of so-called Clifford

[1] For a more direct comparison of classical and quantum correlations, the resource theory of Local Operations and Shared Entanglement (LOSR) has also been proposed [35, 36].

gates and Pauli measurements, as well as preparation of stabilizer states and tracing out subsystems. These operations will be introduced in depth in section 2.3, for now we only note that these operations alone are not sufficient for a quantum computational advantage. In fact, any quantum circuit consisting of only these operations can be simulated efficiently on a classical computer [40, 41]. Quantum computational universality is restored in QCM through the inclusion of additional "magic" quantum states from outside the scope of the stabilizer subtheory.

In the resource theory of stabilizer quantum computation [42], the free operations are those of the stabilizer subtheory. The resources are the magic states which allow for the implementation of nonstabilizer operations, and thus allow for universal quantum computation. These are quantified by nonstabilizer monotones (or magic monotones).

Whereas above we were somewhat vague in talking about resources for quantum computation ("entanglement in some form is necessary for a quantum computational advantage"), this model allows us to import the resource theory perspective and focus on the quantum states. Rather than asking broadly "which nonclassical resources are required for a quantum computational advantage?", we ask "which states could be useful for QCM?"

Quasiprobabilities and the Wigner function

A partial answer to this question is provided by quasiprobability representations like the Wigner function [43]. The Wigner function is the closest quantum mechanical counterpart to the classical notion of a probability distribution over a phase space, but unlike a regular probability distribution it can take negative values (making it a *quasi*probability function). Negative quasiprobabilities are essential for the representation to reproduce quantum theory, and so negativity in the Wigner function has traditionally been considered an indicator distinguishing quantum states that look classical from those that exhibit genuinely quantum behaviour [44, 45]. When adapted to finite-dimensional quantum theory, the setting relevant for quantum computation, quantum states are represented by a discrete Wigner function [27–31, 46, 47]—a quasiprobability function over a finite set (a generalized phase space) usually satisfying certain constraints [4, 48].

Veitch et al. [19] applied this idea to QCM and showed that a necessary condition for a quantum computational speedup on odd-prime-dimensional qudits (quantum systems with odd prime Hilbert space dimension) is that the Wigner function of the input state of the quantum circuit must take negative values (this result is easily extended to QCM on qudits with any odd dimension [9]). In particular, the amount of negativity in the Wigner function (a magic monotone) quantifies the cost of classical simulation of a quantum computations [49] with the simulation being efficient if the Wigner function is nonnegative everywhere.

Contextuality as a resource for quantum computation

It turns out that nonnegativity of the discrete Wigner function also implies the existence of a noncontextual (read classical) hidden variable model (HVM) describing the computation [4, 20, 50–52]. Thus, two traditional notions of nonclassicality for quantum systems—Wigner function negativity and failure of a classical HVM description—herald a quantum computational advantage over classical computation.

This aligns with other work that shows contextuality is required for quantum advantage in other settings [26, 53–56]. For example, contextuality is required to compute nonlinear Boolean functions in measurement-based quantum computation [26, 54], and to boost the probability of winning certain nonlocal games [54, 56].

No-go theorems for discrete Wigner functions in even-dimensions

This three-way alignment between Wigner function negativity, contextuality, and quantum computational advantage is a satisfying result, but as mentioned above it applies only to qudits with odd Hilbert space dimension. The case of even dimensions (including systems of multiple qubits) is much trickier [1, 2, 21, 28, 57–61]. Do these results on Wigner function negativity and contextuality carry over to even dimensions? State-independent proofs of quantum contextuality like the Peres-Mermin square argument [62] essentially preclude the latter. Regarding the former, until recently an answer was known only in certain special cases [58, 63].

In Chapter 5, we resolve this question for the general case. At a technical level, establishing this connection between Wigner function negativity and the possibility of a quantum computational advantage rests on two facts about the Wigner function [29–31], namely, (a) that it is covariant with respect to all Clifford transformations, and (b) that its positivity is preserved under all Pauli measurements. We show that, in all even dimensions, Wigner functions constructed from operator bases cannot be Clifford covariant and cannot represent Pauli measurement positively. We found that the source of the difference between even and odd dimensions is a question of cohomology.

This appearance of cohomology may seem surprising, but in fact, homology and cohomology have a long history in physics [64]. Cohomological phenomena have also recently found widespread use in the field of quantum computation, for example: (i) quantum error correction with the Kitaev surface code [65, 66], and its measurement-based counterpart with 3D cluster states [67], (ii) proofs of contextuality of quantum mechanics [53, 68, 69], and (iii) the contextuality of measurement-based quantum computation (MBQC) [26, 70]. Cohomology emerges as a language to navigate this web of fundamental facts about quantum computation, and to explain those facts in a unified fashion.

Alternative quasiprobability representations

The idea of negativity as a precondition for quantum computational advantage can be salvaged for the case of even dimensions *if* one admits more general quasiprobability functions, specifically, quasiprobability functions that do not stem from operator bases, and hence are not uniquely defined for each state. This means representations with enlarged phase spaces, bigger than you would expect for a typical Wigner function for finite-dimensional quantum theory. This enlargement of the phase space is required for consistency with a proven memory lower bound for simulating the kind of quantum contextuality that comes up in quantum computation on even-dimensional systems [71]. One example of this is obtained by defining a quasiprobability function over stabilizer states [21], bypassing Wigner functions entirely. This representation has the feature of efficiently simulating all QCM circuits on all probabilistic mixtures of stabilizer states.

As mentioned above, in odd dimensions, a nonnegative Wigner function implies the existence of a classical HVM description. This is a result of the fact that phase space points of the Wigner function can be identified with noncontextual value assignments (NCVAs) on the set of observables involved in the computation—Pauli measurements. In even dimensions, these noncontextual value assignements don't exists as a result of state-independent proofs of quantum contextuality like the Mermin square [62].

In Chapter 3, we provided the previously missing phase space picture for QCM on multiqubit systems—the so-called CNC construction. Therein, we define a new quasiprobability function for systems of any finite Hilbert space dimension. In this construction, phase space points are identified with certain noncontextual subsets of Pauli observables, and NCVAs on these subsets. In all dimensions, the function is Clifford covariant and the positivity of the representation is preserved under all Pauli measurements, and for multiple qubits, this simulation contains the efficient classical simulation of stabilizer mixtures [21] as a special case. We thus reproduce the essential features of the odd-dimensional scenario in all dimensions.

Efficient classical simulation beyond Wigner positivity

When we apply the CNC construction to odd-dimensions, it reproduces and extends the existing Wigner function arguments [19]. The phase space contains the original phase space of the Wigner function, but it also contains new phase space points that are not Wigner function phase space points or convex mixtures thereof. This results in a new quasiprobability represenation for odd-dimensions that strictly outperforms the previous stabilizer and Wigner function methods in terms of the circuits that can be efficiently classically simulated, and in terms of the negativity based magic monotones. This is shown in Chapter 8.

A hidden variable model for universal quantum computation

As noted above, obtaining a phase space picture of QCM in all dimensions requires relaxing some of the assumptions that traditionally go into the definition of the Wigner function. This is required in order for the representation to effectively describe quantum computation in all dimensions [4, 71], but it leads to some interesting consequences.

In Chapters 4 and 7, we show that if sufficiently general quasiprobability representations are admitted, we can define a fully *probabilistic* representation of QCM.[2] That is, a hidden variable model can be defined which represents all quantum states, operations, and measurements relevant for QCM using only (nonnegative) probabilities. This model is structurally similar to previously defined quasiprobability representations (modulo absence of negativity) and it leads to a classical simulation method for universal quantum computation based on sampling from the defining probability distributions. The model is based on the so-called Λ polytopes—polytopes in the space of Hermitian operators on the multiqudit Hilbert space which contain the space of physical quantum states. We identify the generalized phase space points of the model with the vertices of these polytopes.

[2]The model is defined first for qubits only in Chapter 4, and then for qudits of any dimension in Chapter 7.

In many ways, this model appears classical, and this gives us an interesting puzzle. Namely, if any quantum computation can be simulated by sampling from a finite family of probability distributions, then where is the quantumness that provides the advantage of quantum computation hiding?

Trading negativity for complexity

The absence of negativity in these models comes at a cost. We can no longer guarantee that the generalized phase space points over which the models are defined (i.e. the vertices of the Λ polytopes) have efficiently computable state update rules under Clifford gates and Pauli measurements. Therefore, although the models can classically simulate any universal quantum computation (magic state quantum computations on any input state), the simulation is not guaranteed to be efficient in general.

There are, however, some special cases in which the simulation is efficient. For example, for odd-dimensional qudits, the phase space point operators which define the odd-dimensional Wigner function are vertices of the Λ polytopes [7]. Similarly, in the multiqubit case, the phase space point operators described in the unified phase space picture above are also vertices of Λ [2, 72], these are the CNC vertices. The update of these vertices under Clifford gates and Pauli measurements is efficiently computable classically, and so the simulation algorithm is efficient whenever the support of the probability distribution representing the input state of the circuit is restricted to vertices of this type, and samples can be efficiently obtained from this distribution. In other words, these subsets of Λ vertices define subpolytopes inside the full Λ-polytopes, wherein the simulation remains efficient.

Connections with qubit-fermion mappings

In Chapter 9, we enlarge the known region of efficient classical simulability inside the Λ-polytopes. Specifically, we define a new quasiprobability representation intermediate between those of the CNC construction and the Λ-polytopes. This model has efficiently computable state update rules under the dynamics of quantum computation, and it yields an efficient classical simulation algorithm for quantum comptutations on the subset of input states that are positively represented. It subsumes the CNC construction, and previous methods such as those based on quasimixtures of stabilizer states [21]. Thus, we effectively push out the boundary of quantum computations which can be efficiently simulated classically.

The model is based on generalized Jordan-Wigner transformations [73], and its conception was influenced by the surprising connection to Majorana fermions first realized in the CNC construction [1]. It was also partially inspired by the techniques of mapping to free fermions in the simulation of other computational models such as matchgate circuits [74], which have recently received renewed interest [75–78].

The path forward: expanding the scope of efficient classical simulation

The above observations leave us with a clear path forward. Some subpolytope models inside Λ provide efficient classical simulation of quantum computations, but if a quantum computational advantage is possible, as is commonly believed, then this efficiency must

breakdown somewhere in between the CNC subpolytope model and the full Λ polytope. Now we can probe this phenomenon by characterizing more vertices of Λ. The more vertices we can efficiently characterize, the more quantum computations we can efficiently simulate classically, thus effectively pushing back the boundary between classical and quantum.

Some progress has been made in this regard with the CNC model [1, 8] and the model based on generalized Jordan-Wigner transformations [6] mentioned above. That said, the vertices characterized so far remain a small fraction of the total number of vertices for $n \geq 3$ qubits, and so there is still much work to be done here and much to be gained from it. The Λ polytopes are still relatively new objects of study with much left to be discovered about them.

Perspectives on quantum advantage

Overall, this thesis attempts to approach the problem of characterizing quantum computational advantage through classical simulations of quantum computations. To clarify the discussion so far, we should note that there are a few different things one can mean by the expression *quantum advantage*. This could mean, for example, an advantage in the number of queries to an oracle needed to solve a certain problem, as in the case of Simon's algorithm [79] and the Deutsch-Jozsa algorithm [80], or in complexity theory, it could refer to the capacity to solve decision problems in BQP as opposed to merely BPP. In complexity theoretic terms, it could refer to polynomial or superpolynomial quantum speedups in time-complexity as in the case of Shor's algorithm [15], or an advantage for constant-depth circuits as demonstrated by Bravyi, Gosset, and König [81]. It could also refer to more practical considerations, such as the ability to perform tasks on quantum hardware that cannot practically be simulated on a large supercomputer, such as recent hardware demonstrations of random circuit sampling [82, 83] and Boson sampling [84].

For our purposes, in this thesis we are approaching quantum advantage from a particular model of quantum computation—the magic state model (explained in detail in Chapter 2). We define quantum advantage as the failure of efficient (polynomial-time) classical simulation of computations in this model. The magic state model allows for polynomial-time classical side processing, and so it doesn't make sense to talk about polynomial time quantum advantages, only superpolynomial speedups are considered.

Perspectives on resources for quantum computation

It is also possible to analyze quantum advantage through many different lenses, even within the same computational model. For example, in the magic state model, we usually consider the classical backbone of the computational model to be the efficiently-classically-simulable stabilizer subtheory, and then the resources which provide the quantum computational advantage are the nonstabilizer magic states. On the other hand, local quantum theory is also efficiently simulable. In the magic state model it suffices to use unentangled products of single-qubit magic states as inputs [38], so from this perspective the resource of entanglement is provided by entanglement-generating operations.

To take another example, in the measurement-based model, we often talk about the entangled initial state as being a resource state on which local measurements are performed to enable universal quantum computation (e.g. Refs. [85–92]). From this perspective,

entanglement is a resource and the operations are free. On the other hand, this model can also be analyzed through the lens of magic. For universal MBQC, it suffices to use a cluster state as input [93] (a stabilizer state), and the nonstabilizerness comes from nonPauli measurements.

Any quantum computational model, or more general quantum information processing task, can similarly be analyzed from different perspectives. In general, the appropriate notion of resource may be task-dependent or model-dependent. Taken together, these different tasks, perspectives, and approaches can be used to characterize the broad and widespread phenomenon of quantum advantage.

Necessary versus sufficient conditions for quantum computational advantage

We have so far mostly discussed the problem of characterizing the utility of quantum computation from one direction, namely, finding necessary conditions for a quantum computational advantage by finding families of quantum circuits that can be efficiently simulated classically. Progress can also be made in the opposite direction in a number of ways. First, we can try to establish sufficient conditions for quantum computation. This direction seems to be harder to link to physical phenomena, but there are some concrete, well-known examples, e.g., in MBQC it is known that local measurements on cluster input states, or any quantum state in the same symmetry-protected topological phase, are sufficient [89, 93].

Equally important is progress on the front of quantum algorithms. So far, a handful of basic quantum algorithms are known, Shor's algorithm [15] and its core the quantum Fourier transform, the HHL algorithm for solving linear systems [17], and Grover's algorithm [16] are well-known examples. Many applications and algorithms are being created, but most are based on a relatively small number of essentially quantum subroutines. One hope for the present work is that progress on the abstract problem of characterizing quantum advantage can also have concrete impact by influencing the development of new quantum algorithms and protocols.

Foundations of quantum computation

With the development of quantum theory in the early 20th century, it quickly became clear that quantum physics was drastically different from the kind of physical theories we were used to. Early quantum physicists such as Einstein, Bohr, and de Broglie would debate the interpretation of unintuitive quantum phenomena. There were puzzles suggested by some in the form of thought experiments, such as the Schödinger's cat [94], Wigner's friend [95], and the Einstein-Podolsky-Rosen paradox [96]. Possible resolutions to these unintuitive ideas were suggested by others, such as pilot waves [97–99].

In the early days, the approach to these interpretational questions seemed to mix mathematical rigour and wild speculation. Over time, the approach became more formal and systematic, and eventually spawned a field of research that became known as quantum foundations. An active area of research in this field has been finding properties of quantum systems that distinguish them from classical systems. With the introduction of quantum computation and other quantum information processing tasks, the differences between quantum and classical become a problem of practical significance, as alluded to above.

Essentially, what we're proposing here is a programme analogous to the quantum foundations programme of characterizing properties of quantum systems that distinguish them from classical systems, in the realm quantum computation and quantum information. This is the programme to which this thesis contributes.

Foundations from quantum computation

While the ability to use techniques and concepts developed in quantum foundations over the last number of decades has been fruitful for studying quantum information, we believe it may sometimes be possible to export ideas in the opposite direction as well. This is something we attempt in the last research chapter of this thesis, Chapter 10.

Based on the hidden variable model mentioned above, we present a reponse to the challenge offered by John Wheeler, in which he claims that modern physics requires a new interpretation based on information theoretic principles [100]. He wrote: "No element in the description of physics shows itself as closer to primordial than the elementary quantum phenomenon, that is, the elementary device-intermediated act of posing a yes-no physical question and eliciting an answer or, in brief, the elementary act of observer-participancy. Otherwise stated, every physical quantity, every it, derives its ultimate significance from bits, binary yes-or-no indications, a conclusion we epitomize in the phrase, *it from bit*."

A prototypical realization of this view on physics has been provided in the description of quantum computation with magic states (QCM) through the Λ polytopes [2], in which the quantum phenomena at hand are reproduced, without any deviation or approximation, by repeated sampling processes dependent on and producing bit strings. The "*It*" in this case is universal quantum computation, and hence all non-relativistic quantum mechanics in finite-dimensional Hilbert spaces. The "*Bits*" represent the binary outcomes of Pauli measurements and the labels of the vertices of the Λ polytopes in the repeated sampling.

To conclude this introduction, we note that the Λ-polytopes—a newly introduced mathematical structure and a common theme throughout this thesis—are only beginning to be explored, and they may hold many more surprises. From the perspective of quantum computation, an important question is the following. Two known efficient classical simulation methods, the stabilizer formalism [101] and sampling from nonnegative Wigner functions [19] are reproduced and extended by the Λ-polytope method, at the level of the simplest vertices of the polytopes. Where does efficient classical simulation end?

Part I

Background and Previous Work

Chapter 2

Background

In this chapter, we review the necessary background material required to understand the results presented in the remaining chapters. Since quantum computing is an interdisciplinary field, hosting researchers with backgrounds in physics, mathematics, computer science, electrical and computer engineering, and other areas, it's hard to include sufficient breadth and depth to satisfy everyone. The goal of this chapter is to make the thesis digestible by graduate students with a background in physics. With some extra work it should also be approachable by students and researchers in other areas but, for example, some prior knowledge of basic quantum theory is assumed. For more background on topics not covered in this chapter, we point to standard references in quantum physics [102] and quantum computation [12, 103]. Throughout the text we point to more specific references where applicable. In this chapter we also introduce some notation and nomenclature that will be used throughout the thesis, a summary of which is provided for reference in the List of Symbols on page xviii.

We start by reviewing the basics of quantum computation in Section 2.1. Then we introduce the models of quantum computation that will be the focus of this thesis in Section 2.2, and we review the stabilizer formalism in Section 2.3. Wigner functions and their role in describing quantum computation are described in Section 2.4. In Section 2.5 we discuss a connection to hidden variable models and quantum contextuality, and in Sections 2.6 and 2.7 we review some more general quasiprobability representations that are useful for describing quantum computations.

2.1 Introduction to quantum computation

2.1.1 Basics of quantum theory

States

In quantum theory, every physical system is identified with a complex Hilbert space $\mathcal{H}$. For our purposes, we are only concerned with finite-dimensional quantum theory, where $\mathcal{H} = \mathbb{C}^N$ for some $N \in \mathbb{N}$. The pure states of the system are identified with elements of the projective Hilbert space $P(\mathcal{H})$. That is, for the case $\mathcal{H} = \mathbb{C}^N$, states are equivalence classes of vectors in the Hilbert space up to the equivalence relation $v \sim w$ if and only if $v = aw$

for some $a \in \mathbb{C}$.

We get mixed quantum states by taking probabilistic mixtures of pure quantum states, or by tracing out subsystems of an entangled composite system. These are represented by density operators—Hermitian, positive semi-definite operators on the Hilbert space normalized to have unit trace. In general, we use the following definition of states.

Definition 1. *A bounded linear operator ρ on $\mathcal{H}$ is called a state if*

1. *ρ is Hermitian, $\rho^\dagger = \rho$,*
2. *ρ is positive semi-definite, $\langle\psi|\rho|\psi\rangle \geq 0 \; \forall \, |\psi\rangle \in \mathcal{H}$, and*
3. *ρ is normalized, $\mathrm{Tr}(\rho) = 1$.*

Observables

Every observable of a physical system with Hilbert space $\mathcal{H}$ is represented by a Hermitian operator A on $\mathcal{H}$. By the spectral theorem, these operators have spectral decompositions

$$A = \sum_\lambda \lambda \Pi_A^\lambda,$$

where the sum is over the eigenvalues of A, and Π_A^λ is the projector onto the eigenspace of A corresponding to eigenvalue λ. The eigenvalues of A label the possible measurement outcomes.

The Born rule

Let Π_A^λ denote the projector onto the eigenspace of A corresponding to eigenvalue λ. The probability of getting measurement outcome λ given a measurement of A is performed on a state ρ is given by the Born rule,

$$P(\lambda \mid A, \rho) = \mathrm{Tr}\Big(\Pi_A^\lambda \rho\Big).$$

The expectation value of the measurement is given by

$$\langle A \rangle_\rho = \mathrm{Tr}(A\rho).$$

This can be shown by using the spectral decomposition of A.

$$\langle A \rangle_\rho := \sum_\lambda \lambda P(\lambda \mid A, \rho) = \sum_\lambda \lambda \, \mathrm{Tr}\Big(\Pi_A^\lambda \rho\Big) = \mathrm{Tr}\left[\left(\sum_\lambda \lambda \Pi_A^\lambda\right)\rho\right] = \mathrm{Tr}(A\rho).$$

Here the first equality is the definition of the expectation value, the second equality uses the Born rule for probabilities of obtaining certain measurement outcomes, and the third equality follows from linearity of the trace.

Dynamics

There are two fundamental types of quantum dynamics. The first is unitary evolution. The state ρ_1 of a system at time t_1 is related to the state ρ_2 at some later time $t_2 > t_1$ by

$$\rho_2 = U\rho_1 U^\dagger$$

where U is a unitary operator on $\mathcal{H}$.

The second type of dynamics is the result of measurement. Suppose a measurement A is made on state ρ. As mentioned above, the probabilities for obtaining certain measurement outcomes are given by the Born rule, $P(\lambda|A,\rho) = \mathrm{Tr}\big(\Pi_A^\lambda \rho\big)$. If the measurement is performed and the result λ is returned for some λ satisfying $\mathrm{Tr}\big(\Pi_A^\lambda \rho\big) > 0$, then the post-measurement state is the projected and renormalized state

$$\rho' = \frac{\Pi_A^\lambda \rho \Pi_A^\lambda}{\mathrm{Tr}\big(\Pi_A^\lambda \rho\big)}.$$

Multiple subsystems

Given two quantum systems with Hilbert spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, the Hilbert space of the composite system is $\mathcal{H}_A \otimes \mathcal{H}_B$. In quantum computation, we typically assume access to many subsystems (called qudits) each with the same Hilbert space dimension d. For n qudits each with Hilbert space dimension d, the Hilbert space of the composite system is $\big(\mathbb{C}^d\big)^{\otimes n} \simeq \mathbb{C}^{d^n}$. Qudits with $d = 2$ are called qubits, and with $d = 3$ are called qutrits. In some chapters we will specialize to qubits ($d = 2$), or qudits of odd prime Hilbert space dimension, but unless otherwise specified d can be any integer greater than or equal to 2.

The summary of quantum theory provided here is horribly incomplete. For example, although we introduce mixed quantum states as they will come up in future chapters, we neglect corresponding generalizations of dynamics and measurements, such as the Lindbladian equation and positive operator-valued measures (POVMs). That said, it should be sufficient for our purposes. For a more complete introduction to quantum theory and quantum information, we point to standard textbooks on the subjects, for example, Ref. [102].

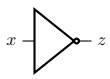
2.1.2 Basics of quantum computation

Qauntum computation can most easily be approached by comparison with the circuit model of classical computation, the quantum analogue of which is called the circuit model of quantum computation. A summary of this comparison is given in Table 2.2.

The circuit model of classical computation

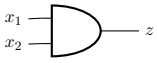
A Boolean circuit is a representation of a Boolean function $f : \{0,1\}^m \to \{0,1\}^n$ as a composition of elementary logic gates. For example, Figure 2.1 shows a circuit that computes the sum of two binary numbers using a circuit consisting of `NAND` and `XOR` gates. The definitions of some common elementary logic gates are given in Table 2.1. It is well known that, if the set of elementary gates is sufficiently expressive, then any Boolean function can be represented as a circuit.

x	z
0	1
1	0



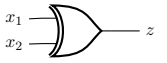
(a) $z = \texttt{NOT}(x)$

x_1	x_2	z
0	0	0
1	0	0
0	1	0
1	1	1



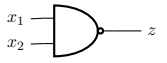
(b) $z = \texttt{AND}(x_1, x_2)$

x_1	x_2	z
0	0	0
1	0	1
0	1	1
1	1	0



(c) $z = \texttt{XOR}(x_1, x_2)$

x_1	x_2	z
0	0	1
1	0	1
0	1	1
1	1	0



(d) $z = \texttt{NAND}(x_1, x_2)$

Table 2.1: Definitions of common elementary logic gates via their truth tables.

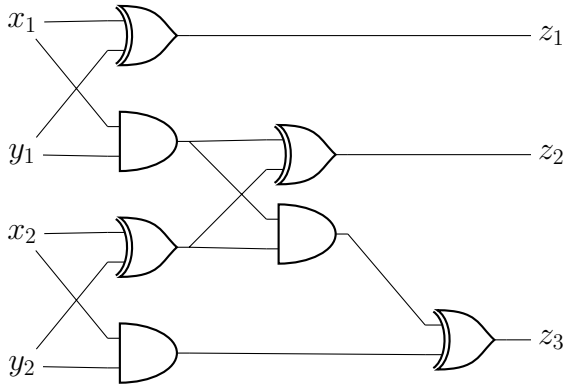


Figure 2.1: Circuit constructed from `XOR` gates and `AND` gates which adds to numbers in binary, $0bx_2x_1 + 0by_2y_1 = 0bz_3z_2z_1$.

Definition 2. *A set of logic gates $\mathcal{G}$ is called* universal *(or* functionally complete*) if for any Boolean function $f : \{0,1\}^m \to \{0,1\}^n$ there is a logic circuit composed of gates in $\mathcal{G}$ that computes f.*

In some cases, even a single gate suffices for universality. For example, the following classic result is well known.

Theorem. *The set $\{\texttt{NAND}\}$ is universal.*

A sketch of a proof of this statement is given in Section A.1.1.

The circuit model of quantum computation

A quantum circuit consists of

1. a set of quantum systems (called qudits), each with Hilbert space $\mathbb{C}^d$ prepared in some fixed state $|0\rangle$,
2. application of a sequence of unitary operations (quantum gates),
3. a measurement of each qudit in the computational basis—an orthonormal basis $\{|s\rangle \mid s \in \mathbb{Z}_d\}$.

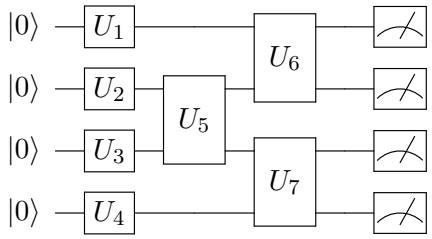


Figure 2.2: A quantum circuit consisting of a layer of single-qubit unitary gates, U_1–U_4, followed by multiqubit gates U_5–U_7, followed by a measurement of each qubit in the computational basis.

	Boolean circuits	Quantum circuits
Input:	A string of input bits $\vec{x} \in \{0,1\}^m$	A quantum state $\rho \in L(\mathcal{H})$
Dynamics:	Logic gates	Unitary gates
Output:	A string of output bits $\vec{z} \in \{0,1\}^n$	Outcomes of quantum measurements

Table 2.2: Comparison of inputs, dynamics, and outputs in classical logic circuits and the quantum circuit model.

The output of the computation is derived from the outcomes of the measurements. An example of such a circuit is shown in Figure 2.2.

In quantum computation there is also a notion of universality. Since the unitary group is a Lie group, it can't be generated by circuits with gates from a discrete set, instead we only require that an arbitrary unitary group element can be approximate by a quantum circuit.

Definition 3. *A set of quantum gates $\mathcal{G}$ is* universal *for quantum computation if any element of $U(d^n)$ can be approximated arbitrarily well by a circuit composed of gates in $\mathcal{G}$ with respect to the metric induced by the operator norm. In other words, the group generated by $\mathcal{G}$ is dense in $U(d^n)$.*

We often consider models of quantum computation in which the allowed operations are restricted to a subset of $U(d^n)$. For example, a quantum computer may only be able to implement certain local gates and two-qudit gates. If the allowed operations are sufficiently expressive, this is not a problem. For example, for a qubit based quantum computer the gate set $\{H, T, CX\}$ is universal for quantum computation [12, §4.5.3] (these gates are defined in the List of Symbols on page xviii).

2.2 Quantum computation with magic states

Quantum computation with magic states (QCM) [38] is a universal model of quantum computation closely related to the circuit model. In QCM, the allowed operations are restricted to those of the stabilizer subtheory—Clifford gates and Pauli measurements.

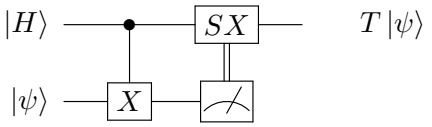


Figure 2.3: An implementation of the T gate via the injection of a magic state. Only Clifford gates, Pauli measurements, and classical conditioning are used but with the inclusion of a nonstabilizer input state, $|H\rangle = T|+\rangle$, the effect of the circuit is a T gate applied to the second qubit. For implementations of other nonClifford gates using different ancillas see [12, §10.6.2]. This circuit reproduced from [12, Fig. 10.25].

These operations will be defined in more detail in the next section, for now we only note that these operations are not by themselves universal for quantum computation. The Clifford group—the group generated by $\{H, S, CX\}$—is a finite subgroup of $U(d^n)$ for each number n of qudits, and the number of states reachable by applying a circuit of these gates to computational basis state $|0\rangle^{\otimes n}$ is finite.[1] These states are called stabilizer states.

The stabilizer subtheory, consisting of stabilizer states, Clifford gates, and Pauli measurements, forms a closed subtheory of quantum theory. According to the Gottesman-Knill theorem [40, 41], any process in this subtheory can be efficiently simulated on a classical computer, and thus by itself it leaves no hope for a quantum computational advantage. Something outside the scope of the stabilizer subtheory needs to be added in order to restore the possibility of quantum computational universality. As mentioned above, the capacity to perform T gates would suffice.

Quantum computation with magic states solves this problem another way. Instead of adding gates to the model, QCM restores universality by instead supplementing the stabilizer subtheory with so-called magic states—nonstabilizer ancillary states. The term "magic state" refers to any nonstabilizer state which allow for universality in QCM. Often tensor products of single-qubit states such as $|H\rangle\langle H| := \frac{1}{2} + \frac{1}{2\sqrt{2}}(X + Y)$ or $|T\rangle\langle T| := \frac{1}{2} + \frac{1}{2\sqrt{3}}(X + Y + Z)$ are used, these are sufficient for computational universality [105].

Computational universality

By allowing nonstabilizer input states, as well as adaptivity—conditioning gates and measurements on the outcomes of prior measurements, we can effectively implement nonstabilizer operations using only Clifford gates and Pauli measurements. An example of a procedure for implementing a T gate in the QCM framework is shown in Figure 2.3. For the case of qubits, $d = 2$, it is well known that the gates $\{CNOT_{ij}, H_i, T_i;\ 1 \leq i, j \leq n, i \neq j\}$ form a universal set [106]. Therefore, this shows that QCM (with suitable magic states) is also universal.

For qudits with Hilbert space dimension $d > 2$, there are correspondingly higher dimensional magic states suitable for promoting the stabilizer subtheory to computational universality [107–110]

[1] The gates H, S, and CX are defined in the List of Symbols on page xviii for the case of qubits. There are suitable generalizations of these gates for higher dimensional qudits [101, 104].

Magic state distillation

One motivation for QCM comes from quantum error correction. Fault-tolerant versions of the Clifford gates and Pauli measurements are relatively easy to produce, but implementation of nonClifford gates is more challenging. In QCM, implementation of nonstabilizer operations is possible if suitable magic states can be prepared. Thus, in QCM the main difficulty of universal fault-tolerant quantum computation is shifted from implementing gates to preparing magic input states.

Without the ability to perform nonstabilizer operations, it's not clear that these states can be prepared. Fortunately, there exist protocols based on the stabilizer formalism which fault-tolerantly produce certain magic states. These work by taking as input a number of noisy magic states, performing some stabilizer operations, and returning smaller number of purified magic states. This is known as magic state distillation [38] (also see Ref. [12, §10.6.3]). This is one of the leading candidates for a model of scalable, fault-tolerant quantum computation [39].

The resource theory of stabilizer quantum computation

A central question of this thesis is: *which quantum states are useful for univeral quantum computation in QCM?* We know that stabilizer states are not useful, and we have certain examples, such as $|H\rangle$ and $|T\rangle$ which are, but we are missing a general characterization. This motivates the resource theory of stabilizer quantum computation [42], also known as the resource theory of magic.

In this resource theory, we split the components of QCM into two types, the "free" operations and the resources. The free operations are (i) preparation of all stabilizer states, (ii) all Clifford unitaries, (iii) measurement of all Pauli observables, and (iv) tracing out subsystems. The resources are copies of the magic state

$$|H\rangle = \frac{|0\rangle + e^{i\pi/4}|1\rangle}{\sqrt{2}}, \tag{2.1}$$

or some other suitable magic state.

Pauli-based quantum computation

Another closely related model of computation is Pauli-based quantum computation (PBC) [111–113]. This model is similar the magic state model, except that the operations are restricted even further without sacrificing computational universality. A Pauli-based quantum computation on n qudits consists of the following steps:

1. preparation of an initial magic state (a n-fold tensor product of a fixed single-qudit magic state is sufficient),

2. a sequence of pair-wise commuting Pauli measurements of length at most n.

Note that the measurement sequence is not necessarily determined in advance. Measurements can depend on the outcomes of previous measurements and classical randomness, and they are determined on the fly via classical side-processing. For a complete description

of the above circuit simplifications, along with a method for compiling a given QCM circuit into an adaptive Pauli circuit, see Ref. [112, 113]. Since this model can be obtained from the magic state model by imposing further restrictions on the operations and states that we allow, all of our results applying to the magic state model also apply to Pauli-based quantum computation.

2.3 The stabilizer formalism

The stabilizer formalism [40, 101, 114] describes a large family of quantum error correcting codes, as well as providing a framework for describing quantum error correction, fault-tolerant quantum computation, and other aspects of quantum information processing more broadly. Here we review some features of the stabilizer formalism which will prove useful in subsequent chapters.

2.3.1 The stabilizer subtheory

The stabilizer formalism describes a closed subtheory of quantum theory, often called the stabilizer subtheory, consisting of a restricted set of states, unitaries, and measurements. The measurements are Pauli measurements, the corresponding observables are identified with elements of the Pauli group. For a single qubit, the Pauli group, $P = \langle X, Y, Z \rangle$, is generated by the familiar Pauli operators,

$$X = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad Y = \begin{bmatrix} 0 & -i \\ i & 0 \end{bmatrix}, \text{ and } \quad Z = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}.$$

To describe the stabilizer formalism for higher dimensional qudits, we need a suitable generalization of these Pauli operators. For d-dimensional qudits, the Pauli X and Z operators are

$$X = \sum_{j \in \mathbb{Z}_d} |j+1\rangle \langle j|, \quad Z = \sum_{j \in \mathbb{Z}_d} \omega^j |j\rangle \langle j|$$

where $\mathbb{Z}_d$ is the ring of integers modulo d and $\omega = \exp(2\pi i/d)$ is a primitive d^{th} root of unity.[2] The generalized X and Z operators are also called Heisenberg-Weyl operators and the group generated by them the Weyl-Heisenberg group for d-dimension Hilbert space. Then the single-qudit generalized Pauli group can be defined as the group $\langle \mu, X, Z \rangle$, where $\mu = \omega$ ($\mu = \sqrt{\omega}$) when d is odd (even).[3]

The n-qudit Pauli group is the group generated by the local Pauli operators acting on n qudits (i.e. the generators of the n-qudit Pauli group are formed from tensor products of the local Pauli operators). Formally, for a system of n d-dimensional qudits, the generalized Pauli group is $P = \langle \mu, X_k, Z_k \mid 1 \leq k \leq n \rangle$.

The unitary operations of the stabilizer subtheory form the *Clifford group*. The Clifford group is defined as the group of unitary operators that map Pauli operators to Pauli oper-

[2] When $d > 2$, these operators are not Hermitian so they do not correspond directly to measurements as they do in the case $d = 2$. They can be measured indirectly through the measurement of Hermitian operators with the same eigenspaces, e.g. $\tilde{Z} = \sum_{j=1}^{d} j |j\rangle \langle j|$.

[3] An explanation of this distinction between even and odd d will be given in the proceeding sections.

ators under conjugation. As global phases are inconsequential in quantum computation we can take the n-qudit Clifford group to be the quotient

$$\mathcal{C}\ell = \mathcal{N}(P)/U(1). \tag{2.2}$$

That is, the normalizer of the Pauli group in the unitary group, $U(d^n)$, modulo phases.

The Clifford group is a finite subgroup of $U(d^n)/U(1)$. For qubits, it is the group generated by the gates H, S, and CX [12, §10.5.2]. For qudits, the group is generated by suitable qudit analogues of these gates [101, 104]. These gates are defined in the List of Symbols on page xviii.

A stabilizer group is an Abelian subgroup of the n-qudit Pauli group which stabilizes a nontrivial subspace of the n-qudit Hilbert space. For example, the 2-qubit stabilizer group $S = \langle X_1X_2, Z_1Z_2\rangle$ stabilizes the Bell state $|\beta_{00}\rangle = (|00\rangle + |11\rangle)/\sqrt{2}$, that is, $A\,|\beta_{00}\rangle = |\beta_{00}\rangle$ for each $A \in S$.

If a stabilizer group S is a maximal Abelian subgroup of the Pauli group then it will stabilize a unique state. The unique state stabilized by a maximal Abelian subgroup of the Pauli group is called a stabilizer state [40, 114]. These are the basic building blocks of stabilizer codes as the codespace of any stabilizer code is the span of some set of stabilizer states. In particular, the codespace of a code with stabilizer group S is the subspace of the Hilbert space stabilized by S.

2.3.2 The $\mathbb{Z}_d$ representation of the stabilizer subtheory

For the Pauli observables, overall phases are inconsequential (or at least they can be tracked separately), leaving the essential part of the Pauli group to be the quotient $\mathcal{P} := P/\mathcal{Z}(P)$, where $\mathcal{Z}(P)$ is the centre of the Pauli group P. With overall phases modded out, the Pauli group is isomorphic to the additive group of the module $\mathbb{Z}_d^{2n}$, $\mathcal{P} \cong \mathbb{Z}_d^{2n}$, and we can index the Pauli operators by elements of the module as follows, for any $a = (a_Z, a_X) \in \mathbb{Z}_d^n \times \mathbb{Z}_d^n =: E$,

$$T_a = \mu^{\phi(a)} \bigotimes_{k=1}^{n} Z^{[a_z]_k} X^{[a_x]_k} \tag{2.3}$$

Here $\mu = \omega$ when d is odd, and $\mu = \sqrt{\omega}$ when d is even. The phase function ϕ can be chosen freely subject to the constraint $(T_a)^d = \mathbb{1}$ for all $a \in E$. This constraint forces the eigenvalues of the operators to be in $\{\omega^j \mid j \in \mathbb{Z}_d\}$.

Note that when d is prime, $\mathbb{Z}_d$ is a field so the index set $\mathbb{Z}_d^{2n}$ forms a vector space. When d is not prime, $\mathbb{Z}_d$ is a commutative ring, not a field, and so $\mathbb{Z}_d^{2n}$ is not a vector space. This distinction can introduce some subtelty (see, for example, Ref. [29]), but for our purposes it is largely inconsequential and we will generally treat $\mathbb{Z}_d^{2n}$ as a linear space with notions of the standard inner product, $\langle a|b\rangle = \sum_k a_k b_k$, and of subspaces as subsets which are closed under addition and multiplication by elements of $\mathbb{Z}_d$.

The labels $E \simeq \mathbb{Z}_d^{2n}$ of the Pauli operators can be equipped with a symplectic form $[\cdot,\cdot] : E \times E \to \mathbb{Z}_d$. For any $a = (a_z, a_x), b = (b_z, b_x) \in \mathbb{Z}_d^n \times \mathbb{Z}_d^n =: E$, the symplectic form is defined by

$$[a, b] := \langle a_z|b_x\rangle - \langle a_x|b_z\rangle \tag{2.4}$$

where these inner products are computed in $\mathbb{Z}_d$. This symplectic inner product tracks the commutator of the Pauli operators labelled by a and b. This is the result of the following lemma.

Lemma 1. *For any $a, b \in E$,*

$$[T_a, T_b] := T_a T_b T_a^{-1} T_b^{-1} = \omega^{[a,b]}, \tag{2.5}$$

and in particular, operators T_a and T_b commute if and only if $[a, b] = 0$.

The proof of this lemma is given in Section B.1.1. Because of this correspondence we will say that elements $a, b \in E$ commute when $[a, b] = 0$.

We define a function $\beta : E \times E \to \mathbb{Z}_d$ which tracks how Pauli operators compose through the relation

$$T_a T_b = \omega^{-\beta(a,b)} T_{a+b}, \quad \forall a, b \in E \text{ with } [a, b] = 0. \tag{2.6}$$

The function β satisfies the relation

$$\beta(a, b) + \beta(a + b, c) - \beta(b, c) - \beta(a, b + c) = 0, \tag{2.7}$$

for $a, b, c \in E$. We state this relation for later reference. It is a consequence of the associativity of operator multiplication. Consider the operator product $T_a T_b T_c = T_a(T_b T_c) = (T_a T_b)T_c$, and expand $T_a(T_b T_c) = \omega^{-\beta(a,b+c)-\beta(b,c)} T_{a+b+c}$, $(T_a T_b)T_c = \omega^{-\beta(a,b)-\beta(a+b,c)} T_{a+b+c}$. Comparing the two equivalent expressions yields Eq. (2.7). It follows straightforwardly from the definition Eq. (2.6) of β that

$$\beta(a, b) = \beta(b, a), \quad \forall a, b \text{ with } [a, b] = 0. \tag{2.8}$$

The Pauli group with phases modded out forms a normal subgroup of the Clifford group such that $\mathcal{C}\ell/\mathcal{P} \cong \mathrm{Sp}(E)$ is the group of symplectic transformations on E. The Clifford group acts on the Pauli group by conjugation as

$$g T_a g^\dagger = \omega^{\tilde{\Phi}_g(a)} T_{S_g(a)} \quad \forall g \in \mathcal{C}\ell \; \forall a \in E \tag{2.9}$$

where $S_g \in \mathrm{Sp}(E)$ is a symplectic map, and the function $\tilde{\Phi}_g : E \to \mathbb{Z}_d$ tracks the extra phases that get picked up in this action. The functions $\tilde{\Phi}$ and β have a cohomological interpretation that we will elucidate further in Section 2.5.2.

For a subspace $W \subset E$ let $W^\perp$ denote the subspace $\{v \in E |\ [v, w] = 0,\ \forall w \in W\}$. A subspace $J \subset E$ is called isotropic if $J \subset J^\perp$. The collection of maximal isotropic subspaces of E is denoted by $\mathcal{I}(E)$.

On the level of the index set E, a stabilizer group is specified by an isotropic subspace—a subspace $I \subset E$ on which the symplectic form vanishes: $[a, b] = 0,\ \forall a, b \in I$. $I \subset E$, together with a function $\gamma : I \to \mathbb{Z}_d$ which specifies the phases of the stabilizer group elements. The stabilizer group corresponding to the pair (I, γ) is $S = \left\{\omega^{-\gamma(a)} T_a \mid a \in I\right\}$.

The function γ is subject to some additional constraints required in order for S to form a group. First, S must contain the identity operator and so we must have $\omega^{-\gamma(0)} T_0 = \mathbb{1}$. From the closure of the group S we have $\omega^{-\gamma(a)} T_a \cdot \omega^{-\gamma(b)} T_b = \omega^{-\gamma(a+b)} T_{a+b}$. This relation

together with Eq. (2.6) gives the constraint that γ must satisfy

$$\gamma(a) + \gamma(b) - \gamma(a+b) = -\beta(a,b), \quad \forall a, b \in I. \tag{2.10}$$

2.3.3 A concrete phase convention for the Pauli operators

In Chapter 5 we will consider the effect that different choices of phases ϕ in Eq. (2.3) can have, but for most of this thesis, we can fix a phase convention. For concreteness, one choice that works is the following. For any $a \in E$, we choose

$$\phi(a) = \begin{cases} -\langle a_z | a_x \rangle \cdot 2^{-1} & \text{if } d \text{ is odd} \\ -\langle a_z | a_x \rangle_{2d} & \text{if } d \text{ is even} \end{cases} \tag{2.11}$$

where in the first case the inner product $\langle a_z | a_x \rangle := \sum_{k=1}^{n} [a_z]_k [a_x]_k$ is computed in $\mathbb{Z}_d$ and 2^{-1} is the multiplicative inverse of 2 in $\mathbb{Z}_d$, and in the second case the inner product is computed mod $2d$.

With this choice of phases, a direct calculation shows that when d is odd, for any $a, b \in E$, $T_a T_b = \omega^{[a,b]2^{-1}} T_{a+b}$. Therefore, $\beta(a,b) = 0$ for all commuting $a, b \in E$. For even d, no choice of phases with this property exists [4, 69]. One explanation for the difference between even and odd dimensions d is that in $\mathbb{Z}_d$ for even d, there is no multiplicative inverse of 2, and so the corresponding phase convention doesn't exist.

2.3.4 Stabilizer states and stabilizer codes

A stabilizer code is specified by a pair (I, r) where I is an isotropic subgroup of E, and $r : I \to \mathbb{Z}_d$ is a noncontextual value assignment for I. The codespace of the code is the simultaneous $+1$-eigenspace of the Pauli operators $S_I^r = \{\omega^{-r(a)} T_a \mid a \in I\}$. S_I^r is called the stabilizer group of the code. The projector onto this eigenspace, or equivalently, the projector onto the eigenspace of the Pauli observables labeled by I corresponding to eigenvalues $\{\omega^{r(a)} \mid a \in I\}$ is

$$\Pi_I^r = \frac{1}{|I|} \sum_{b \in I} \omega^{-r(b)} T_b. \tag{2.12}$$

The dimension of this eigenspace is $d^n/|I|$ [115]. In particular, we have the following lemma.

Lemma 2. *The order of a maximal isotropic subgroup of E is d^n.*

For the proof of Lemma 2 see Theorem 1 in Ref. [115]. If I is a maximal isotropic subgroup of E, $|I| = d^n$, and there is a unique quantum state fixed by a stabilizer group S_I^r. Such states are called stabilizer states.

Stabilizer code projectors can be constructed from products of stabilizer code projectors of higher rank. If $\{I_k \mid I_k \subset I\}$ are such that $I = \bigcup_k I_k$, and the value assignments $r_k : I_k \to \mathbb{Z}_d$ satisfy $r_k = r|_{I_k}$, then

$$\Pi_I^r = \prod_k \Pi_{I_k}^{r_k}. \tag{2.13}$$

Stabilizer measurements can also be coarse-grained to give stabilizer projectors of higher rank. This is the result of the following lemma.

Lemma 3. *If I is a nonmaximal isotropic subgroup of E and Π_I^r is a stabilizer code projector for some noncontextual value assignment $r : I \to \mathbb{Z}_d$, then for any isotropic subgroup I' containing I*

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \Pi_{I'}^{r'} = \Pi_I^r, \tag{2.14}$$

where $\Gamma_{I,r}^{I'}$ is the set of all noncontextual value assignments on I' such that $r'|_I = r$.

Corollary 1. *The stabilizer polytope (the convex hull of projectors onto stabilizer states) is*

$$SP = \operatorname{conv}\left\{ \frac{|I|}{d^n} \Pi_I^r \;\middle|\; \Pi_I^r \text{ stabilizer code projector} \right\}. \tag{2.15}$$

The above lemmas describe structural properties of stabilizer code projectors. It will also be useful for us to describe how stabilizer projectors behave under the dynamical operations of quantum computation with magic states—Clifford gates and Pauli measurements. First, we have the following result regarding Clifford gates.

Lemma 4. *For any Clifford group element $g \in \mathcal{C}\ell$, and any stabilizer state $|\sigma\rangle \in \mathcal{S}$,*

$$g |\sigma\rangle \langle\sigma| g^\dagger = |\sigma'\rangle \langle\sigma'| \tag{2.16}$$

where $|\sigma'\rangle \in \mathcal{S}$ is a stabilizer state. I.e. Clifford group operations map stabilizer states to stabilizer states.

The update of stabilizer states under Pauli measurements is probabilistic in general. It is described in the following lemma.

Lemma 5. *For any isotropic subgroups $I, J \subset E$ and any noncontextual value assignments $r : I \to \mathbb{Z}_d$ and $s : J \to \mathbb{Z}_d$,*

1. *if $r|_{I\cap J} = s|_{I\cap J}$ then*

$$\operatorname{Tr}(\Pi_J^s \Pi_I^r) = \frac{|I \cap J|}{|I||J|} d^n > 0 \tag{2.17}$$

and

$$\frac{\Pi_I^r \Pi_J^s \Pi_I^r}{\operatorname{Tr}\left(\Pi_J^s \Pi_I^r\right)} = \frac{|J \cap I^\perp|}{|J|} \Pi_{I+J\cap I^\perp}^{r\star s} \tag{2.18}$$

where $r \star s$ is the unique noncontextual value assignment on the set $I + J \cap I^\perp$ such that $r \star s|_I = r$ and $r \star s|_{J\cap I^\perp} = s|_{J\cap I^\perp}$.

2. *If $r|_{I\cap J} \neq s|_{I\cap J}$ then*

$$\operatorname{Tr}(\Pi_J^s \Pi_I^r) = 0 \quad \textit{and} \quad \Pi_I^r \Pi_J^s \Pi_I^r = 0. \tag{2.19}$$

The proofs of Lemmas 4 and 5 are given in Sections B.1.3 and B.1.4.

2.4 Wigner functions

Since the operations used in QCM are not by themselves universal, but they become universal when paired with nonstabilizer input states, we can say that the computational power of quantum computation is effectively transferred from the operations to the states. An important question is then which quantum states promote QCM to universality? One approach to studying this question is through *quasiprobability representations* like the discrete Wigner function.

2.4.1 The original Wigner function

The Wigner function [43] forms the basis of an alternative formulation of quantum mechanics. In this formulation, instead of representing a state by a wavefunction, a state is represented by a function over the position-momentum phase space. For the simplest case of one-dimensional continuous-variable quantum mechanics, if the state of the system is represented by the density operator ρ, the Wigner function is defined as

$$W_\rho(x,p) = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty} dy\, \langle x-y|\rho|x+y\rangle\, e^{2ipy/\hbar}$$

This formulation can be generalized to higher dimensions. The map that takes Hilbert space operators ρ to phase space functions W_ρ is called the Wigner transform, and it's inverse is called the Weyl transform.

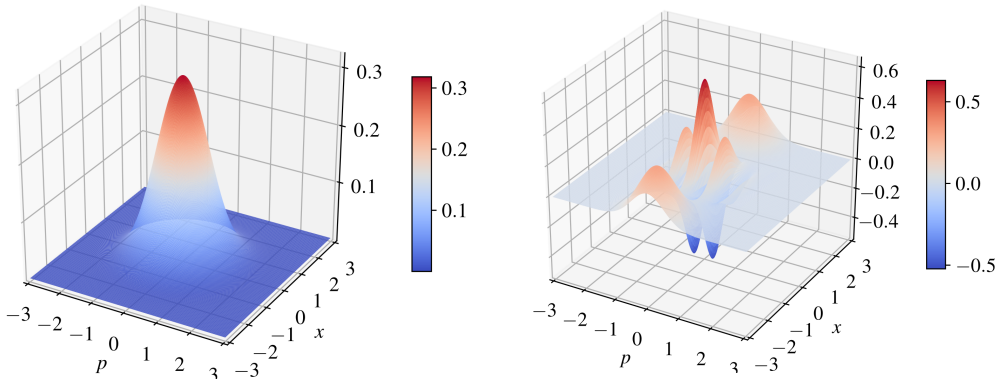
The Wigner function of a state is a quasiprobability function, which behaves much like a probability distribution over phase space, the basis of classical statistical mechanics. For example, it is real-valued and normalized: $\int dxdp\; W(x,p) = 1$. Further, integrating over one variable gives the marginal probability distribution for the other variable.

The essential difference between the Wigner function and a regular phase space distribution is that, unlike a probability distribution, it can take negative values. Such functions are called *quasiprobability functions*. This property allows the Wigner function to represent quantum mechanics, and as a result, negativity in the Wigner functions of states has been proposed as a measure that distinguishes classically behaving subsystems of quantum mechanics from those which are genuinely quantum [44, 45].

Examples of the Wigner function for two particular states are shown in Fig. 2.4. Figure 2.4a shows a Gaussian state. All Gaussian states have nonnegative Wigner functions [44], making them proper phase space distributions. Figure 2.4b show a state whose Wigner function is not nonnegative.

2.4.2 The Stratonovich-Weyl criteria

In quantum optics and quantum statistical mechanics, several quasiprobability representations of quantum states have been studied, for example, the Wigner function [43], the Sudarshan-Glauber P representation [116, 117], and the Husimi Q representation [118]. They are related by the Stratonovich-Weyl (SW) criteria, a set of criteria that reasonable quasiprobability representations over generalized phase spaces should satisfy [48] (also



(a) Gaussian state **(b)** Cat state

Figure 2.4: The Wigner functions for (a) a Gaussian state and (b) a cat state. The Gaussian state has a Wigner function which is everywhere nonnegative making it a probability distribution over the position-momentum phase space. The Wigner function of the cat state takes negative values at some points making it a quasiprobability distribution.

see [119]).

The Wigner transform can be directly applied to other linear operators, in general a linear operator Y is represented by a Wigner function W_Y over the phase space. Quasiprobabililty representations in the SW class have the form $W_Y : \mathcal{V} \to \mathbb{C}$, where W_Y represents the linear operator Y over phase space $\mathcal{V}$, and W_ρ of a state ρ is a quasiprobability distribution. In general, observables are represented by a dual object, the effect function Θ, which maps linear operators Y onto $\Theta_Y : \mathcal{V} \longrightarrow \mathbb{C}$. For example, the P and Q functions are dual to each other, and up to normalization the Wigner function is self dual.

The SW criteria are as follows:

(sw0) (Linearity): the map $Y \rightarrowtail W_Y$ is one-to-one and linear,

(sw1) (Reality):
$$W_{Y^\dagger}(x) = (W_Y(x))^* \quad \forall x \in \mathcal{V},$$

(sw2) (Standardization):
$$\int_{\mathcal{V}} d\mu(x) W_Y(x) = \mathrm{Tr}(Y),$$

(sw3) (Covariance):
$$W_{g\cdot Y}(g \cdot x) = W_Y(x) \quad \forall x \in \mathcal{V}\, \forall g \in G$$
where G is a symmetry group of the phase space.

(sw4) (Traciality):
$$\int_{\mathcal{V}} d\mu(x) W_{Y_1}(x) \Theta_{Y_2}(x) = \mathrm{Tr}(Y_1 Y_2).$$

The purpose of the traciality condition (sw4) is to represent the Born rule in a phase-space formulation of quantum mechanics. Quasiprobability representations satisfying these criteria are structurally similar and thus closely related to the original Wigner function.

2.4.3 Discrete Wigner functions

Attempts to formulate analogues of the Wigner function that apply to quantum mechanics in finite-dimensional Hilbert spaces go back to the 1980's [46]. There is even a paper by Richard Feynman from 1987 discussing the interpretation of negative probabilities in physics which features a discrete Wigner function for one-qubit states [120]. In 2004, a large family of discrete phase spaces and Wigner functions was defined by Gibbons, Hoffman, and Wootters based on finite fields [47]. A particular Wigner function in this family that applies to odd-dimensional Hilbert spaces was found by Gross and shown to have properties which align it with the original Wigner function. This particular discrete Wigner function has found applications in the study of quantum computation on odd-dimensional qudits.

The interpretation of negativity in quasiprobability functions as an indicator of genuine quantumness, with many applications in quantum information processing, has also been proposed for discrete Wigner functions—quasiprobability functions used for describing finite-dimensional quantum mechanics. For a system of n qudits, each with local Hilbert space dimension d, the discrete Wigner function is usually defined over the finite phase space $\mathbb{F}_d^{2n}$ [47] or $\mathbb{Z}_d^{2n}$ [29, 31], where $\mathbb{F}_d$ is the field with d elements and $\mathbb{Z}_d$ is the integers mod d. When $d = p$ is prime, $\mathbb{F}_d \cong \mathbb{Z}_d$, and so the two choices for the phase space are equivalent. The scope of the former choice is limited to the case where the local Hilbert space dimension $d = p^N$ is a power of a prime since a finite field with d elements exists if and only if d is a power of a prime. In this case, there is a map $\iota : \mathbb{F}_{p^N}^2 \to \mathbb{F}_p^{2N}$ which preserves the structure of the phase space, and so Wigner functions defined over $\mathbb{F}_{p^N}^{2n}$ coincide with Wigner functions defined over $\mathbb{Z}_p^{2nN}$ up to relabeling of phase space points [29]. Therefore, choosing the phase space defined over $\mathbb{F}_{p^N}$ is equivalent to representing each $d = p^N$-dimensional qudit as a system of N independent p-dimensional qudits with the overall phase space $\mathbb{Z}_p^{2nN}$. Many other phase spaces for finite-dimensional systems have also been proposed, see [121] for a review. For the purposes of the present section we are interested in Wigner functions defined over the phase space $V := \mathbb{Z}_d^{2n}$. In later sections we consider more general phase spaces.

2.4.4 Gross' Wigner function

Gross' Wigner function [29–31] is defined for systems of d-level qudits where d is odd. To start, we fix the phase convention of the Pauli operators to be that of Section 2.3.3.

The phase space over which the Wigner function is defined is $V := \mathbb{Z}_d^{2n}$ and for each point $u \in V$ there is a corresponding phase space point operator A_u.[4] The phase space

[4]The phase space is isomorphic to the index set for the Pauli operators. To distinguish them we denote the former by V and the latter by E.

point operator corresponding to the zero vector, $0 \in V$, is

$$A_0 := \frac{1}{d^n} \sum_{v \in E} T_v. \tag{2.20}$$

Then all other phase space point operators can be defined as translations from the base point

$$A_u := T_u A_0 T_u^\dagger \quad \forall u \in V. \tag{2.21}$$

Using the commutation relation Eq. (2.5), we can rewrite the phase space point operators in a more convenient form: $A_u = d^{-n} \sum_{v \in E} \omega^{[u,v]} T_v$.

These phase space point operators form an orthogonal operator basis for the space of Hilbert-Schmidt operators on n-qudit Hilbert space: $\mathrm{Tr}\left(A_u^\dagger A_v\right) = d^n \delta_{u,v}$. Therefore, for operators A and B expanded in the phase space point operator basis with coefficients $W_A(u), W_B(u) \in \mathbb{C}$, $\forall u \in V$, the Hilbert-Schmidt inner product is given by $\mathrm{Tr}\big(A^\dagger B\big) = d^n \sum_{u \in V} W_A(u)^* W_B(u)$. Since the phase space point operators are Hermitian, the coefficients $W_A(u)$ are real whenever A is Hermitian.

Any density matrix, ρ, representing a n-qudit quantum state can be expanded in this basis as

$$\rho = \sum_{u \in V} W_\rho(u) A_u. \tag{2.22}$$

The coefficients in this expansion, given by $W_\rho(u) = d^{-n} \mathrm{Tr}(\rho A_u) \in \mathbb{R}$, define the discrete Wigner function of the state ρ. Since the phase space point operators have unit trace, taking a trace of Eq. (2.22) we obtain $1 = \sum_{u \in V} W_\rho(u)$. Therefore, the discrete Wigner function is a quasiprobability representation for n-qudit quantum states.

Measurements also have a representation in this framework. For a POVM $\{E_k\}$, we could define the Wigner function of the element E_k to be the coefficients in the expansion of E_k in the phase space point operator basis as for density matrices, but it will be more convenient to rescale the Wigner function and define the Wigner function representation of E_k as $W_{E_k}(u) = \mathrm{Tr}(E_k A_u)$. Then the probability of obtaining outcome k given the measurement is performed on state ρ is given by

$$\mathrm{Tr}(E_k \rho) = \sum_{u \in V} W_\rho(u) \, \mathrm{Tr}(E_k A_u) = \sum_{u \in V} W_\rho(u) W_{E_k}(u). \tag{2.23}$$

The SW correspondence

To justify calling this quasiprobability representation a discrete Wigner function, we can check that it satisfies a suitable finite-dimensional generalization of the SW criteria. Consider a Hilbert-Schmidt operator B expanded in the phase space point operator basis as

$$B = \sum_{u \in V} W_B(u) A_u. \tag{2.24}$$

The coefficients in this expansion, given by $W_B(u) = d^{-n} \mathrm{Tr}(B A_u)$, define the quasiprobability representation of the operator B.

We will consider the criteria in order: (sw0) Linearity of W_B follows from the linearity of the trace. (sw1) Since the phase space point operators are Hermitian, taking the Hermitian conjugate of Eq. (2.24) we get $B^\dagger = \sum_{u \in V} W_B(u)^* A_u$. Therefore, $W_{B^\dagger}(u) = W_B(u)^*$ for all $u \in V$ and so reality is satisfied. (sw2) Since the phase space point operators have unit trace, taking a trace of Eq. (2.24) we obtain standardization: $\mathrm{Tr}(B) = \sum_{u \in V} W_B(u)$. (sw3) The relevant symmetry group in this case is the Pauli group. Covariance of the discrete Wigner function with respect to the Pauli group is easy to check by considering the result of conjugating a phase space point operator by a Pauli operator.[5] (sw4) Finally, traciality follows from the representation of the Hilbert-Schmidt inner product where we take the dual Wigner function to be equal to the Wigner function except rescaled by a factor d^n like the Wigner function representation of a POVM.

2.4.5 The odd-dimensional Wigner function applied to quantum computation with magic states

The discrete Wigner function representation has several properties that make it useful for describing quantum computation with magic states on odd-dimensional qudits [19, 29–31]. First, consider the representation of states. It can be shown that a pure n-qudit state has a nonnegative Wigner function if and only if it is a stabilizer state. This is the discrete Hudson's theorem[6] [29, 30] (also see [27]).

Nonnegativity of the Wigner function for stabilizer states is the first indication that the discrete Wigner function might be relevant for quantum computation with magic states. In the case of infinite-dimensional quantum mechanics, as mentioned above, negativity in the Wigner function is an indicator of true quantumness whereas states with nonnegative Wigner function are considered more classical. From the perspective of quantum computation with magic states, stabilizer states can be considered classical in the sense that they do not promote QCM to universality.

Probabilistic mixtures of pure stabilizer states clearly also have nonnegative Wigner functions but they are not the only mixed states that are positively represented. Mixed states which are positively represented but are not convex combinations of stabilizer states are called *bound magic states*. Figure 2.5 shows two examples of the discrete Wigner function for one-qutrit ($d = 3$) states.

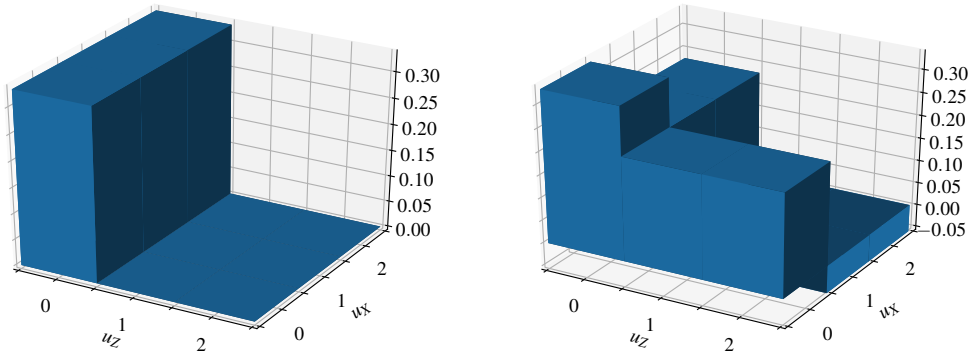
Now we consider the effect of the dynamics of quantum computation with magic states—Clifford unitaries and Pauli measurements—on the phase space and its corresponding Wigner function.

Clifford unitaries

Clifford unitaries map phase space point operators to phase space point operators under conjugation [29]. In particular, for systems of odd-dimensional qudits, the Clifford group

[5] In fact, the Wigner function is covariant with respect to the Clifford group, of which the Pauli group is a subgroup.

[6] This is the analogue for finite-dimensional quantum mechanics of Hudson's theorem, which applies to the original Wigner function defined in Section 2.4 and states that a pure state has nonnegative Wigner function if and only if it is a Gaussian state, i.e. it's wavefunction has the form $\psi(x) = \exp\left(-ax^2 + bx + c\right)$ [44].



(a) Stabilizer state **(b)** Magic state

Figure 2.5: The discrete Wigner functions for two single-Qutrit states: (a) the stabilizer state with stabilizer group $\langle Z \rangle$, and (b) the distillable magic state $|H_+\rangle$ described in Ref. [107]. The Wigner function of the stabilizer state is everywhere nonnegative making it a probability distribution over the phase space $V = \mathbb{Z}_3^2$. The Wigner function of the magic state takes negative values at four points making it a quasiprobability distribution.

splits as $\mathcal{C}\ell \simeq \mathrm{Sp}(\mathbb{Z}_d^{2n}) \ltimes \mathbb{Z}_d^{2n}$ where $\mathrm{Sp}(\mathbb{Z}_d^{2n})$ is the symplectic group on $\mathbb{Z}_d^{2n}$. Then Clifford unitaries $g \in \mathcal{C}\ell$ can be identified with $S_g \in \mathrm{Sp}(\mathbb{Z}_d^{2n})$ and $a_g \in \mathbb{Z}_d^{2n}$, and the action of the Clifford group on phase space point operators is

$$g A_u g^\dagger = A_{S_g u + a + g}. \tag{2.25}$$

As a result of this fact, it can be easily shown that the Wigner function is Clifford covariant:

$$W_{g\rho g^\dagger}(u) = W_\rho\left(g^{-1} \cdot u\right) = W_\rho\left(S_g^{-1}(u - a_g)\right). \tag{2.26}$$

An example of a mapping of a qutrit phase space point operator under a Clifford unitary is given in Figure 2.6.

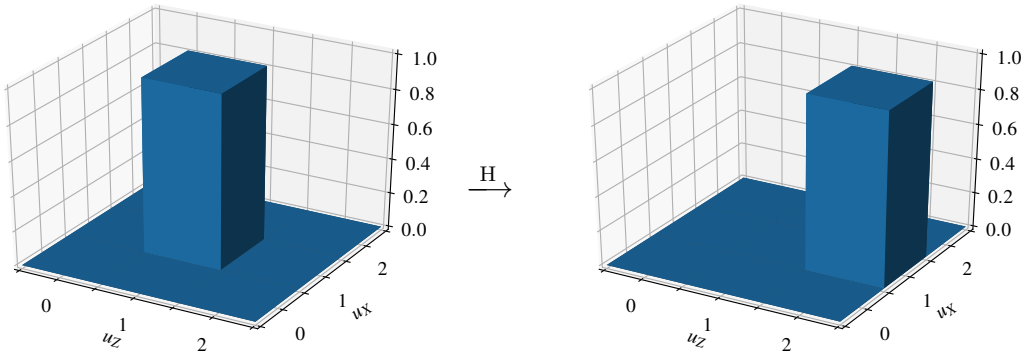
For a more complete discussion of the interaction between the Clifford group and the discrete Wigner function see Ref. [29].

Pauli measurements

As a special case of the general POVM measurements described above we can consider the case of projective Pauli measurements. For a Pauli measurement T_a, $a \in E$, the eigenprojector corresponding to the measurement outcome ω^s, $s \in \mathbb{Z}_d$ is

$$\Pi_a^s = \frac{1}{d} \sum_{k \in \mathbb{Z}_d} \omega^{-ks} T_a^k.$$

For any $a \in E$, the set $\{\Pi_a^s \mid s \in \mathbb{Z}_d\}$ is a projector-valued measure [51]. That is, $\Pi_a^{s\dagger} = \Pi_a^s = \Pi_a^{s\,2}$ for all $a \in E$ and all $s \in \mathbb{Z}_d$, and for any $a \in E$, $\sum_{s \in \mathbb{Z}_d} \Pi_a^s = I$. The discrete



(a) Wigner function of $A_{(1,1)}$ **(b)** Wigner function of $HA_{(1,1)}H^\dagger$

Figure 2.6: Under conjugation by the qutrit Clifford unitary H, the phase space point operator $A_{(1,1)}$ maps to the phase space point operator $A_{(2,1)}$. In general, phase space point operators map deterministically to phase space point operators under conjugation by Clifford unitaries.

Wigner function representation of a measure $\{\Pi_a^s\}$ is given by $W_{\Pi_a^s}(u) = \mathrm{Tr}(\Pi_a^s A_u)$.

In order to describe quantum computation, we must allow for the possibility of sequential measurements. That is, we need a way of extracting probabilities for measurement outcomes as well as the postmeasurement state from the discrete Wigner function. Both of these are obtained from the following lemma.

Lemma 6. *Let $A_u,\ u \in V$ be a phase space point operator and Π_a^s be an eigenprojector corresponding to a Pauli measurement $T_a,\ a \in E$ giving outcome $\omega^s,\ s \in \mathbb{Z}_d$. Then $\mathrm{Tr}(\Pi_a^s A_u) = 1$ if $s = [a,u]$ and $\mathrm{Tr}(\Pi_a^s A_u) = 0$ otherwise. Further, let $a^\perp := \{v \in E \mid [a,v] = 0\}$ and $\Gamma_{a,u} := \{w \in E \mid [w,v] = [u,v]\ \forall v \in a^\perp\}$. Then*

$$\Pi_a^s A_u \Pi_a^s = \frac{\delta_{s,[a,u]}}{|\Gamma_{a,u}|} \sum_{w \in \Gamma_{a,u}} A_w. \tag{2.27}$$

Proof of Lemma 6. By a direct calculation we obtain

$$\begin{aligned}
\Pi_a^s A_u \Pi_a^s &= \frac{1}{d^{n+2}} \sum_{v\in E} \sum_{k,\ell\in\mathbb{Z}_d} \omega^{-(k+\ell)s+[u,v]} T_a^k T_v T_a^\ell \\
&= \frac{1}{d^{n+2}} \sum_{v\in E} \sum_{k,\ell\in\mathbb{Z}_d} \omega^{-(k+\ell)s+[u,v]+(k-\ell)[a,v]2^{-1}} T_{v+(k+\ell)a} \\
&= \frac{1}{d^{n+2}} \sum_{v\in E} \left(\sum_{k,\ell\in\mathbb{Z}_d} \omega^{-(k+\ell)s+[u,v]-(k+\ell)[u,a]+(k-\ell)[a,v]2^{-1}} \right) T_v \\
&= \frac{1}{d^{n+2}} \sum_{v\in E} \delta_{[a,v],0} \left(\sum_{k,\ell\in\mathbb{Z}_d} \omega^{-(k+\ell)s+[u,v]-(k+\ell)[u,a]} \right) T_v \\
&= \frac{1}{d^{n+1}} \sum_{v\in E} \delta_{[a,v],0} \left(\sum_{k\in\mathbb{Z}_d} \omega^{[u,v]-k(s+[u,a])} \right) T_v \\
&= \frac{1}{d^n} \sum_{v\in E} \delta_{[a,v],0} \delta_{s,[a,u]} \omega^{[u,v]} T_v
\end{aligned} \tag{2.28}$$

Here in the first line we use the definitions of Π_a^s and A_u, in the second line we use the relation $T_a T_b = \omega^{[a,b]2^{-1}} T_{a+b}$, in the third line we relabel $v \to v-(k+\ell)a$. The fourth line is obtained by observing that if $[a,v] \neq 0$ then regardless of the values of s and $[u,a]$, the coefficient in the parentheses is a uniformly weighted sum over all d^{th} roots of unity. In the fifth line we sum over ℓ, and in the last line observe that if $s+[u,v] \neq 0$ then the coefficient in parentheses is a sum over all d^{th} roots of unity.

Since all the generalized Pauli operators except $T_0 = I$ are traceless, taking a trace of this final expression we obtain the first statement of the lemma: $\mathrm{Tr}(\Pi_a^s A_u) = \mathrm{Tr}(\Pi_a^s A_u \Pi_a^s) = \delta_{s,[a,u]}$. For the second statement of the lemma, from Eq. (2.28), we have

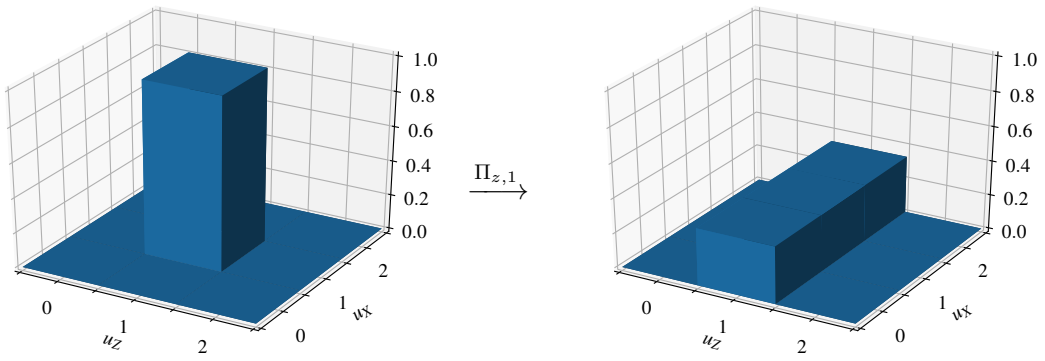
$$\Pi_a^s A_u \Pi_a^s = \frac{\delta_{s,[a,u]}}{d^n} \sum_{v\in E} \delta_{[a,v],0} \omega^{[u,v]} T_v.$$

Also,

$$\begin{aligned}
\frac{\delta_{s,[a,u]}}{|\Gamma_{a,u}|} \sum_{w\in\Gamma_{a,u}} A_w &= \frac{\delta_{s,[a,u]}}{|\Gamma_{a,u}| \cdot d^n} \sum_{v\in E} \left(\sum_{w\in\Gamma_{a,u}} \omega^{[w,v]} \right) T_v \\
&= \frac{\delta_{s,[a,u]}}{d^n} \sum_{v\in E} \delta_{[a,v],0} \omega^{[u,v]} T_v.
\end{aligned}$$

Comparing this with the expression above we obtain the second statement of the lemma. □

The implication of this lemma is that with respect to Pauli measurements, phase space point operators are associated with deterministic assignments of measurement outcomes, and under Pauli measurements phase space point operators map to probabilistic combina-



(a) Wigner function of $A_{(1,1)}$

(b) Wigner function of $\Pi_{z,1}A_{(1,1)}\Pi_{z,1}$

Figure 2.7: A Pauli Z measurement on the qutrit phase space point operator $A_{(1,1)}$ gives outcome ω with probability $\mathrm{Tr}\left(\Pi_{z,1}A_{(1,1)}\right) = 1$. Under the corresponding projection, the phase space point operator $A_{(1,1)}$ maps to the probabilistic combination of the phase space point operators $A_{(1,0)}$, $A_{(1,1)}$, and $A_{(1,2)}$ with equal weight. In general, phase space point operators map to probabilistic combinations of phase space point operators under Pauli measurements.

tions of phase space point operators. An example of a Pauli measurement on a qutrit phase space point operator is given in Figure 2.7.

This lemma together with covariance of the Wigner function with respect to Clifford unitaries leads to the following lemma.

Lemma 7. *The set of positively represented states is closed under Clifford unitaries and Pauli measurements.*

Proof of Lemma 7. Suppose a state ρ is positively represented. That is, ρ can be expanded in the phase space point operator basis as in Eq. (2.22) with $W_\rho(u) \geq 0$ for all $u \in V$. Preservation of nonnegativity of the discrete Wigner function under Clifford unitaries is clear from Eq. (2.26)—Clifford covariance of the discrete Wigner function. Since the Wigner function $W_\rho(u)$ is nonnegative, under a Clifford unitary $g \in \mathcal{C}\ell$, the Wigner function $W_{g\rho g^\dagger}(u) = W_\rho(S_g^{-1}(u - a_g))$ is also nonnegative.

Now consider a Pauli measurement performed on state ρ giving outcome ω^s with probability $\mathrm{Tr}(\Pi_a^s \rho) > 0$. The Wigner function of the postmeasurement state corresponding to

this measurement outcome is

$$\begin{aligned} W_{\frac{\Pi_a^s \rho \Pi_a^s}{\mathrm{Tr}(\Pi_a^s \rho)}}(u) &= d^{-n} \,\mathrm{Tr}\left(\frac{\Pi_a^s \rho \Pi_a^s}{\mathrm{Tr}(\Pi_a^s \rho)} A_u\right) \\ &= \frac{d^{-n}}{\mathrm{Tr}(\Pi_a^s \rho)} \,\mathrm{Tr}\left(\rho \Pi_a^s A_u \Pi_a^s\right) \\ &= \frac{d^{-n}}{\mathrm{Tr}(\Pi_a^s \rho)} \,\mathrm{Tr}\left(\rho \frac{\mathrm{Tr}(\Pi_a^s A_u)}{|\Gamma_{a,u}|} \sum_{w \in \Gamma_{a,u}} A_w\right) \\ &= \frac{d^{-n} \,\mathrm{Tr}(\Pi_a^s A_u)}{\mathrm{Tr}(\Pi_a^s \rho) \cdot |\Gamma_{a,u}|} \sum_{w \in \Gamma_{a,u}} \mathrm{Tr}(\rho A_w) \\ &= \frac{\delta_{s,[a,u]}}{\mathrm{Tr}(\Pi_a^s \rho) \cdot |\Gamma_{a,u}|} \sum_{w \in \Gamma_{a,u}} W_\rho(w). \end{aligned} \tag{2.29}$$

Here, in the second line we use linearity and the cyclic property of the trace, in third line Eq. (2.27), in the fourth line linearity of the trace again, and in the last line the definition of the Wigner function. Since $W_\rho(v) \geq 0$ for all $v \in V$, each term in this sum is nonnegative. Therefore,

$$W_{\frac{\Pi_a^s \rho \Pi_a^s}{\mathrm{Tr}(\Pi_a^s \rho)}}(u) \geq 0$$

for all $u \in V$. That is, nonnegativity of the Wigner function is also preserved under Pauli measurements. □

Phase-space-simulation for quantum computation with magic states on odd-dimensional qudits

Lemma 6 gives a way of extracting the probabilities associated with Pauli measurements from the Wigner function of the state. In particular, for a Pauli measurement T_a performed on state ρ, the probability of obtaining outcome ω^s is given by

$$\mathrm{Tr}(\Pi_a^s \rho) = \sum_{u \in V} W_\rho(u) \,\mathrm{Tr}(\Pi_a^s A_u) = \sum_{u \in V} W_\rho(u) \delta_{s,[a,u]}.$$

Further, Eq. (2.26) and Eq. (2.29) give ways of extracting the Wigner function of the updated state after Clifford unitaries and Pauli measurements from the Wigner function of the initial state alone. In the case of Clifford unitaries, the Wigner function of the updated state at a point $u \in V$ is obtained deterministically from the Wigner function of the initial state at a single phase space point. In the case of Pauli measurements, the Wigner function at a point is a probabilistic combination of the Wigner function of the initial state at multiple points.

These ideas lead to an efficient classical simulation algorithm for quantum computation with magic states based on the discrete Wigner function that applies whenever the input state to the quantum circuit, ρ_{in}, is positively represented.

When the Wigner function of the input quantum state is nonnegative, it is simply a probability distribution over the discrete phase space, meaning it can be sampled from.

The simulation algorithm proceeds by sampling a phase space point $u \in V$ according to the Wigner function distribution $W_{\rho_{in}}(u)$. The phase space point is then propagated through the circuit. When a Clifford unitary $g \in \mathcal{C}\ell$ is encountered, the phase space point is updated as $u \to S_g u + a_g$. When a Pauli measurement $T_a,\ a \in E$ is encountered, we output $s = [a, u]$ as the measurement outcome and update u by sampling a new phase space point according to the distribution implied by Eq. (2.27). This provides samples from the joint probability distribution of the measurement outcomes for the Pauli measurement in the circuit. A complete description of the classical simulation algorithm is given in Algorithm 1. A similar classical simulation algorithm that applies to a slightly different family of circuits was described first in Ref. [19].

Input: $W_{\rho_{in}} \geq 0$
1: sample a point $u \in V$ according to the probability distribution $W_{\rho_{in}}$
2: **while** end of circuit has not been reached **do**
3: **if** a Clifford unitary $g \in \mathcal{C}\ell$ is encountered **then**
4: update $u \leftarrow S_g u + a_g$
5: **end if**
6: **if** a Pauli measurements $T_a,\ a \in E$ is encountered **then**
7: **Output:** $\omega^{[a,u]}$
8: sample w uniformly from $\Gamma_{a,u}$
9: update $u \leftarrow w$
10: **end if**
11: **end while**

Algorithm 1: One run of the classical simulation algorithm for quantum computation with magic states on odd-dimensional qudits based on the discrete Wigner function. The algorithm provides samples from the joint probability distribution of the Pauli measurements in a quantum circuit consisting of Clifford unitaries and Pauli measurements applied to an input state ρ_{in} with $W_{\rho_{in}}(u) \geq 0$ for all $u \in V$.

We have the following result.

Theorem 1. *For any number $n \in \mathbb{N}$ of qudits with odd local Hilbert space dimension d, and any n-qudit quantum state ρ_{in} with $W_{\rho_{in}}(u) \geq 0\ \forall u \in V$, the classical simulation algorithm of Algorithm 1 for sampling the outcomes of the Pauli measurements in a circuit consisting of Clifford gates and Pauli measurements applied to input state ρ_{in} agrees with the predictions of quantum mechanics. When the Wigner function of the input state ρ_{in} is nonnegative and can be efficiently sampled from, the simulation algorithm is efficient.*

The proof of correctness for the classical simulation algorithm, Algorithm 1, is similar in structure to the proof of Theorem 7 given in Chapter 4 so we leave it out.[7] To prove

[7] In fact, the proof of Theomem 7 can be applied almost without modification to prove correctness of Algorithm 1 with a simple reinterpretation of the symbols. In particular, $A_\alpha \to A_u,\ u \in V$ are the phase space point operators, the action of the Clifford group is as $A_{g\cdot\alpha} \to A_{g\cdot u} = A_{S_g u + a_g}$ with $u \in V$, $S_g \in \mathrm{Sp}(\mathbb{Z}_d^{2n})$, and $a_g \in \mathbb{Z}_d^{2n}$, and the functions Q and q have particular forms based on Lemma 6. Namely, $Q_a(s\,|\,\alpha) \to Q_a(s\,|\,u) = \delta_{[a,u],s}$ and $q_{\alpha,a}(\beta, s) \to q_{u,a}(w, s) = \frac{1}{|\Gamma_{a,u}|}\delta_{[a,u],s}\delta_{w\in\Gamma_{a,u}}$ with $u, w \in V$, $a \in E$, and $s \in \mathbb{Z}_d$.

efficiency of the classical simulation algorithm, we must prove four properties:

(E0) An efficient description for points in phase space,

(E1) The capacity to efficiently obtain samples from the probability distribution describing the input state of the quantum circuit,

(E2) Efficient update of phase space points under Clifford unitaries,

(E3) Efficient extraction of Pauli measurement outcomes as well as efficient updating of phase space points under Pauli measurements.

These are the efficiency criteria for any classical simulation algorithm for QCM based on sampling from quasiprobability functions like Algorithm 1. First, for (E0), specifying a point in phase space is equivalent to specifying an element of $\mathbb{Z}_d^{2n}$. (E2) follows immediately from Eq. (2.26): update under Clifford unitaries relies only on being able perform matrix multiplication and vector addition in $\mathbb{Z}_d^{2n}$. Finally, for (E3), returning measurement outcomes relies on calculation of the symplectic inner product in $\mathbb{Z}_d^{2n}$, and update of phase space points after Pauli measurements relies on being able to sample an element uniformly from a subset of $\mathbb{Z}_d^{2n}$. Each of these tasks can be performed efficiently in the number n of qudits.

Property (E1) is included as an assumption of Theorem 1. We could prove that this property holds for certain states of interest. For example, often n-qudit magic states are taken to be tensor products of single-qudit magic states. In the case of product states, the Wigner function is a product of single-qudit Wigner functions [29]. Then, if the Wigner function of the input state is nonnegative, each single-qudit Wigner function could be sampled independently giving efficient samples from the product distribution.

Note that the classical simulation algorithm given in Algorithm 1 is slightly different from the one described in Ref. [19] but the implication is the same: any quantum circuit consisting of only Clifford unitaries and Pauli measurements applied to an input state with a nonnegative Wigner function can be classically simulated efficiently.

When the Wigner function of the input state of a QCM circuit takes negative values, Algorithm 1 does not apply but simulation is possible via probability estimation [49, 122] which is generally inefficient. This algorithm is based on sampling from a probability distribution derived from the discrete Wigner function with the number of samples required to estimate a probability to a given error ϵ proportional to $\mathfrak{R}(\rho_{in})^2/\epsilon^2$ where

$$\mathfrak{R}(\rho_{in}) = \sum_{u \in V} |W_{\rho_{in}}(u)|$$

is a measure of the amount of negativity in the Wigner function of the input state. In this sense, negativity in the Wigner function of the input state to the circuit determines the complexity of classical simulation of a quantum computation.

To summarize, QCM circuits can be efficiently classically simulated when the Wigner function of the input state is positively represented, but generally not when the Wigner function takes negative values. This means Wigner function negativity is *necessary* for quantum computation with magic states on odd-dimensional qudits to exhibit a speedup over classical computation.

This result provides a partial characterization of the states which are useful for QCM. However, as noted above, it applies only when the local Hilbert space dimension is odd. This restriction seems strange at first sight. An explanation for this distinction between even and odd dimensions comes from a reinterpretation of the discrete Wigner function and associated classical simulation algorithm for quantum computation as a *hidden variable model*.

2.5 Hidden variable models

Attempts to formulate hidden variable models for quantum theory are almost as old as quantum theory itself [96, 97]. Many of the early twentieth century physicists who led the development of quantum mechanics were dissatisfied with the indeterminism and nonlocality that seemed inherent to the theory. This led many to suggest that quantum mechanics could not be the final and complete theory that they were looking for, and that it must be augmented through the inclusion of "*hidden variables*" which aligned more closely with their intuitions.

The most famous example of a successful hidden variable model is de Broglie-Bohm theory [97–99], also known as the pilot-wave interpretation of quantum mechanics. This model skirts indeterminism but at the cost of requiring the hidden variables to be explicitly nonlocal.

The answer to the question of whether we can formulate a hidden variable model of quantum mechanics that fully aligns with intuitions of locality and determinism turns out to be a resounding no. This was first proven by John Bell in 1964 [123]. Bell proved that there are correlations between quantum observables on entangled and spatially separated subsystems that cannot be reproduced by any local theory. This was the first valid no-go result restricting hidden variable models for quantum mechanics.[8]

A similar result, proven by Bell [126] and independently by Kochen and Specker [127] states that there can be no noncontextual hidden variable model for quantum theory. This result and *contextuality* are the focus of the next section.[9]

2.5.1 Contextuality

The motivation behind deterministic hidden variable models for quantum theory is the desire to think of the probabilistic predictions of quantum theory as a result of epistemic uncertainty rather than of a more fundamental indeterminism. That is, we would like to

[8]There was an earlier no-go result for hidden variable models by von Neumann in 1932 [124] but this result relied on a faulty assumption as pointed out by Hermann [125]. See Ref. [62] for a review.

[9]People often refer to contextuality as a generalization of nonlocality, but, in fact, Bell's result on contextuality [126] was proven before his result on nonlocality [123], and was only published afterwards. These results both rule out "classical" models for quantum theory, which make different assumptions about what "classicality" means. Bell thought the assumption of locality for classical models were more justifiable than those of noncontextuality (the constraints that locality puts on a model are a strict subset of those imposed by noncontextuality), and so Ref. [123] provides a stronger argument for the nonclassicality of quantum theory. After these results were proven, the failure of these classical models was taken to be an indicator of quantum behaviour of physical systems (we talk about nonlocality and contextuality of quantum systems as equivalent to the failure of locality and noncontextuality of the classical models). For more on the history of the subject, see Ref. [62].

think that all measurements that can be performed on a system have outcomes which are the uniquely determined by the physical state of the system prior to the measurement. There can be uncertainty about the physical state of the system, but the act of measurement should play no role in producing the measurement outcome.

Mathematically, this idea is represented by deterministic value assignments—functions that map quantum observables to the possible outcomes of those observables. These deterministic value assignments are associated with the "hidden variables" of a hidden variable model. Quantum states are then represented by probability distributions over the hidden variables.

The impossibility of representing quantum theory by a hidden variable model of this form was proven by Kochen and Specker [127], and independently by John Bell [126]. In particular, what they showed is that a hidden variable model representing quantum theory must be contextual. A contextual hidden variable model is one in which the value assignments are not simply specified by functions from observables to their possible measurement outcomes, instead the value assigned to an observable can depend on the measurement context. Formally, a measurement context is a collection of simultaneously measurable observables. An observable may belong to many mutually incompatible contexts with a contextual value assignment assigning a potentially different value to the observable depending on the context in which it is measured.

In order for a hidden variable model to accurately represent quantum mechanics, it must reproduce the measurement statistics predicted by quantum mechanics. The original proofs of contextuality of quantum mechanics involved complicated geometric arguments which ultimately proved the inability of any noncontextual hidden variable models to accurately reproduce the predictions of quantum mechanics. A much simpler proof of contextuality that applies when the dimension of the Hilbert space is at least 4 is known as the Mermin square.

The Mermin square

A particularly simple proof of the contextuality of quantum mechanics with Hilbert space dimension at least four is given by the so called Mermin square [62]. The Mermin square consists of the nine two-qubit Pauli observables shown in Figure 2.8.

The observables in each row and in each column commute and so they can be measured simultaneously—these are the six maximal measurement contexts. Multiplying the operators in each row and in each column we obtain the relations

$$\begin{cases} X_1 \cdot X_2 \cdot X_1 X_2 = \mathbb{1}, \\ Z_2 \cdot Z_1 \cdot Z_1 Z_2 = \mathbb{1}, \\ X_1 Z_2 \cdot Z_1 X_2 \cdot Y_1 Y_2 = \mathbb{1}, \\ X_1 \cdot Z_2 \cdot X_1 Z_2 = \mathbb{1}, \\ X_2 \cdot Z_1 \cdot Z_1 X_2 = \mathbb{1}, \\ X_1 X_2 \cdot Z_1 Z_2 \cdot Y_1 Y_2 = -\mathbb{1}. \end{cases} \tag{2.30}$$

In order to define a noncontextual hidden variable, we would would first need to define a deterministic noncontextual value assignment for these observables. That is, a function λ

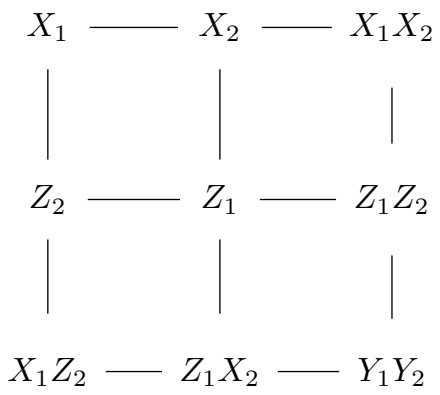


Figure 2.8: The two-qubit Pauli observables constituting the Mermin square, a set of observables which gives a simple state-independent proof of contextuality for quantum mechanics in four-dimensional Hilbert space [62]. The observables, A, B, C, in each row/column give a constraint that a deterministic noncontextual value assignment must satisfy of the form $\lambda(A)\lambda(B)\lambda(C) = \lambda(ABC)$. The six constraints corresponding to the rows and columns are inconsistent, proving there can be no deterministic noncontextual value assignment for these observables.

from these observables to their eigenvalues that is consistent with the predictions of quantum mechanics. First, since the only eigenvalue of the identity operator is 1,[10] we must have $\lambda(\mathbb{1}) = 1$. Then, for consistency with the relations of Eq. (2.30), λ must map the Mermin square observables to $\{+1, -1\}$ subject to the constraints

$$\begin{cases} \lambda(X_1)\lambda(X_2)\lambda(X_1X_2) = 1, \\ \lambda(Z_2)\lambda(Z_1)\lambda(Z_1Z_2) = 1, \\ \lambda(X_1Z_2)\lambda(Z_1X_2)\lambda(Y_1Y_2) = 1, \\ \lambda(X_1)\lambda(Z_2)\lambda(X_1Z_2) = 1, \\ \lambda(X_2)\lambda(Z_1)\lambda(Z_1X_2) = 1, \\ \lambda(X_1X_2)\lambda(Z_1Z_2)\lambda(Y_1Y_2) = -1. \end{cases}$$

But these six equations are inconsistent. To see this, note that if we multiply the six equations together, on the left hand side there are exactly two factors of $\lambda(O)$ for each observable O, and since $\lambda(O) = \pm 1$ for each observable, the left hand side is $+1$. The right hand side is -1. This is a contradiction. Therefore, there can be no deterministic noncontextual value assignment for the Mermin square observables. This constitutes a *state-independent proof of contextuality* as it does not rely on a hidden variable model reproducing

[10] The identity operator is Hermitian, and thus it corresponds to a valid quantum mechanical observables. But since it has only one eigenvalue, namely 1, a measurement of the identity operator on any state gives 1 with probability 1.

the measurement statistics associated with any state, it relies only on the properties of the observables.

Although a noncontextual value assignment may exist for a subtheory of quantum mechanics consisting of a restricted set of observables (no state-independent contextuality), it could still be impossible for a noncontextual hidden variable model to reproduce the measurement statistics associated with those observables on particular states. This is *state-dependent contextuality* [69].

For a more complete review of contextuality, see Ref. [62]. For other examples and types of contextuality, see Refs. [53, 68, 69, 128–130].

2.5.2 A cohomological description of contextuality

Contextuality among generalized Pauli observables can be described using group cohomology [69]. In this description, a chain complex $\mathcal{C}_* = \mathcal{C}_*(E)$ is formed from the index set E of the Pauli observables. Edges of the complex are associated with the labels of the observables $a \in E$, faces are associated with commuting pairs $[a|b] \in E \times E$, and volumes with pairwise commuting triples $[a|b|c] \in E \times E \times E$. Dual to $\mathcal{C}_*$ there is a cochain complex $\mathcal{C}^*$ formed from maps from edges, faces, and volumes of $\mathcal{C}_*$ to $\mathbb{Z}_d$.

Then the function β defined in Eq. (2.6) is a 2-cocycle with the cocycle condition enforced by the associativity of multiplication for the operators corresponding to edges of the complex. I.e., the relation Eq. (2.7) is equivalent to the statement that for any volume $[a|b|c] \in \mathcal{C}_*$, the coboundary $d\beta$ evaluated on $[a|b|c]$ satisfies

$$d\beta(a,b,c) := \beta(a,b) + \beta(a+b,c) - \beta(b,c) - \beta(a,b+c) = 0. \tag{2.31}$$

Noncontextual value assignments can be associated with 1-cochains. The relation between a noncontextual value assignment for the Pauli observables $\lambda : \mathcal{P} \to \{\omega^k \mid k \in \mathbb{Z}_d\}$ and a 1-cochain $\gamma : E \to \mathbb{Z}_d$ of $\mathcal{C}^*$ is given by $\lambda(T_a) = \omega^{\gamma(a)}$. For any commuting pair of observables, T_a and T_b, there is a consistency condition on the noncontextual value assignment λ: $\lambda(T_a)\lambda(T_b) = \lambda(T_aT_b)$. Constraints of this form are enforced on the level of γ by the statement that for any face $(a,b) \in \mathcal{C}_*$, the coboundary $d\gamma$ evaluated on $[a|b]$ satisfies

$$d\gamma(a,b) := \gamma(a) + \gamma(b) - \gamma(a+b) = -\beta(a,b). \tag{2.32}$$

The existence of a function γ on the labels of a set of Pauli observables satisfying this constraint is equivalent to the statement that there is no parity based state-independent proof of contextuality like the Mermin square argument among the corresponding observables.

Theorem 1 of Ref. [69] gives a relation between contextuality and the second cohomology group $H^2(\mathcal{C}, \mathbb{Z}_d)$. It states that if $H^2(\mathcal{C}, \mathbb{Z}_d) \ni [\beta] \neq 0$, then the corresponding observables exhibit state-independent contextuality. With this we can see a key difference between the generalized Pauli observables in even and odd-dimensional Hilbert spaces. For odd-dimensional qudits, $[\beta] = 0$ but for even-dimensional qudits (including qubits), $[\beta] \neq 0$ [69].

This is a special case of a more general framework for describing contextuality using group cohomology. See Ref. [69] for details. This framework will be expanded in much more detail in Chapter 5.

2.5.3 The odd-dimensional Wigner function as a noncontextual hidden variable model

Lemma 6 shows that phase space point operators are associated with deterministic assignments of measurement outcomes for the generalized Pauli observables. In particular, the phase space point operator $A_u = d^{-n} \sum_{v \in V} \omega^{[u,v]} T_v$ is associated with the value assignment $T_a \mapsto \omega^{-[u,a]}$. It can be shown that for systems of multiple qudits, these coincide precisely with the noncontextual value assignments of a hidden variable model. That is, the discrete Wigner function defines a noncontextual hidden variable model for the subtheory of quantum mechanics on systems of multiple odd-dimensional qudits consisting of states with nonnegative Wigner function, Clifford unitaries, and generalized Pauli measurements [20, 50, 51].

In this hidden variable model there is a hidden variable (or ontic state) corresponding to each point in the phase space V. The admissible states, those with nonnegative Wigner function, are represented by probability distributions (the Wigner function) over the hidden variables. The Wigner function representation of a POVM element E_k, $W_{E_k}(u) = \mathrm{Tr}(E_k A_u)$ can be interpreted as the probability of obtaining outcome k given the measurement is performed on the "state" A_u.[11] Then Eq. (2.23) is simply the law of total probability:

$$P_\rho(k) = \sum_{u \in V} P_\rho(u) P(k \,|\, u) = \sum_{u \in V} W_\rho(u) W_{E_k}(u).$$

Here $P_\rho(u)$ is the probability of obtaining the ontic state A_u given a preparation of the quantum state ρ, and $P(k \,|\, u)$ is the probability of obtaining measument outcome k for a POVM measurement $\{E_k\}$ on state A_u.

With this reinterpretation of the discrete Wigner function, the classical simulation algorithm, Algorithm 1, amounts to treating the hidden variable model as physical [19]. The update rules for Clifford unitaries and Pauli measurements stated in Eqs. (2.25) and (2.27) and used in the classical simulation algorithm represent transition probabilities for the operations performed on the ontic states themselves.

Since not all physical states have a nonnegative Wigner function, not all states can be represented by this noncontextual hidden variable model. States with Wigner functions which take negative values are classified as contextual,[12] and thus the two traditional indicators of quantumness, negativity in the Wigner function and contextuality, agree [20, 50, 51]. With Algorithm 1, both are necessary conditions for a quantum computational speedup over classical computation in QCM on odd-dimensional qudits [19].

This interpretation of the discrete Wigner function as a hidden variable model partially explains the difficulty in extending the results of the odd-dimensional Wigner function to systems of qubits. For odd-dimensional qudits, the points in phase space are associated with deterministic noncontextual value assignments for the Pauli observables. For systems of multiple qubits, noncontextual value assignments for the Pauli observables do not exist as evidenced by the Mermin square proof of contextuality.

[11] The phase space point operators are Hermitian and have unit trace like density matrices but they are not positive semidefinite so they do not correspond to physical states.

[12] To say a state is contextual here is to say that it admits a state-dependent proof of contextuality with the generalized Pauli observables.

2.6 Quasiprobability representations

The term quasiprobability representation means different things to different people. For example, some define quasiprobability representations as representations satisfying the Stratonowich-Weyl criteria [48, 131], or finite-dimensional generalizations thereof [4]. Others admit only frame representations [132, 133]. See, for example, Refs. [4, 121, 134, 135] for other types of quasiprobability representations in physics, and in particular, quasiprobability representations for finite-dimensional quantum theory.

As shown above, there is some difficulty in defining Wigner functions that are useful for describing quantum computation on even-dimensional qudits (this will be formalized in Chapter 5). Therefore, in order to obtain useful descriptions of quantum computations in this setting, we need to allow more general quasiprobability representations.

Here we impose very few restrictions on the types of quasiprobability representations we allow. We impose only the constraints required for the representation to be useful for describing quantum computations. Denote by $\mathrm{Herm}(\mathcal{H})$ the real vector space of Hermitian operators on Hilbert space $\mathcal{H}$. We start by defining a finite set of operators $\{A_\alpha \mid \alpha \in \mathcal{V}\}$ on the n-qudit Hilbert space $(\mathbb{C}^d)^{\otimes n} \simeq \mathbb{C}^{d^n}$, with the following properties:

(QR1) $A_\alpha^\dagger = A_\alpha, \quad \forall \alpha \in \mathcal{V}$,

(QR2) $\mathrm{Tr}(A_\alpha) = 1, \quad \forall \alpha \in \mathcal{V}$,

(QR3) $\mathrm{Span}(\{A_\alpha \mid \alpha \in \mathcal{V}\}) = \mathrm{Herm}(\mathcal{H})$.

As a result of property (QR3), any n-qudit quantum state represented by a density matrix ρ can be expanded in these operators as

$$\rho = \sum_{\alpha \in \mathcal{V}} Q_\rho(\alpha) A_\alpha. \tag{2.33}$$

From (QR1), the expansion coefficients are real for Hermitian operators, and with (QR2), taking a trace of this equation we obtain $1 = \sum_{\alpha \in \mathcal{V}} Q_\rho(\alpha)$. This is the motivation for the term quasiprobability, the function $Q_\rho : \mathcal{V} \to \mathbb{R}$ looks like a probability distribution over the generalized phase space $\mathcal{V}$, except for the fact that it can take negative values. This function is how we represent states.

In order to describe quantum computations, we will be interested only in quasiprobability representation which are closed under the dynamics of quantum computation with magic states—Clifford gates and Pauli measurements. Therefore, we consider only quasiprobability representations satisfying the following additional constraints.

(QR4) For any Clifford group element $g \in \mathcal{C}\ell$, $gA_\alpha g^\dagger = A_{g\cdot\alpha}$ with $g \cdot \alpha \in \mathcal{V}$

(QR5) For any Pauli measurement $a \in E$ and any measurement outcome $s \in \mathbb{Z}_d$, we have

$$\Pi_a^s A_\alpha \Pi_a^s = \sum_{\beta \in \mathcal{V}} q_{\alpha,a}(\beta, s) A_\beta \tag{2.34}$$

with $q_{\alpha,a}(\beta, s) \geq 0, \ \forall \alpha, a, \beta, s$ and $\sum_{\beta,s} q_{\alpha,a}(\beta, s) = 1$.

That is, according to (QR4) Clifford operations are represented by a deterministic update map $\alpha \mapsto g \cdot \alpha$ on the phase space, and according to (QR5) Pauli measurements are represented by a stochastic map $q_{\alpha,a}$. Once the operators $\{A_\alpha\}_{\alpha\in\mathcal{V}}$ are defined, the quasiprobability representation is determined, with Q_ρ in Eq. (2.33) giving the representation of states, and (QR4) and (QR5) giving the representation of computational dynamics.

2.7 Quasiprobability representations for qubits

The discrete Wigner function defined in Section 2.4.4 applies only to systems of odd-dimensional qudits. Attempting to formulate a phase space and discrete Wigner function for qubits, or for systems of even local Hilbert space dimension, one immediately runs into several issues. In this section we discuss the issues with a naïve generalization of the odd-dimensional Wigner function to qubits, and we review some previously defined quasiprobability representations for quantum computation with magic states on qubits.

2.7.1 The trouble with qubits

An ostensible issue with generalizing the odd-dimensional Wigner function to systems of even-dimensional qudits arises from the choice of phase used in the definition of the generalized Pauli operators that were then used in the definition of the phase space point operators in Section 2.4.4. In Eq. (2.11), the phase function ϕ was chosen so that for any $a \in E$, $e^{i\phi(a)} = \omega^{-\langle a_z|a_x\rangle 2^{-1}}$. This is a convenient choice because for any Pauli operators T_a and T_b, we have $T_aT_b = \omega^{[a,b]/2}T_{a+b}$. In particular, the function β is zero whenever a and b commute. Defining the phase space point operators as in Eq. (2.20) and Eq. (2.21) with this choice of phase for the Pauli operators leads to all the properties of the discrete Wigner function that make it useful for describing of quantum computation with magic states, e.g. Clifford covariance and positivity preservation under Pauli measurements.

A similar choice of phase is not possible for the multiqubit Pauli observables, on the surface because there is no multiplicative inverse of 2 in $\mathbb{Z}_2$, but more fundamentally, because for the multiqubit Pauli operators $H^2(\mathcal{C}, \mathbb{Z}_d) \ni [\beta] \neq 0$ [69].

The closest analogue of the phase convention chosen above that works for qubits is to set $e^{i\phi(a)} = i^{\langle a_z|a_x\rangle}$ with the inner product taken mod 4. This gives a choice of phase for multiqubit Pauli operators that aligns with simply taking tensor products of the standard Pauli matrices, $\mathbb{1}, X, Y, Z$. Unfortunately, the Wigner function resulting from this choice of phase lacks the crucial properties that make it useful for describing quantum computation with magic states.

First, it is not Clifford covariant. This is true already for one qubit. To see this, note that the one-qubit phase space point operators defined like in Eqs. (2.20) and (2.21) are

$$A_{(0,0)} = \frac{1}{2}(\mathbb{1} + X + Y + Z) \qquad A_{(0,1)} = \frac{1}{2}(\mathbb{1} + X - Y - Z)$$
$$A_{(1,0)} = \frac{1}{2}(\mathbb{1} - X - Y + Z) \qquad A_{(1,1)} = \frac{1}{2}(\mathbb{1} - X + Y - Z).$$

Under conjugation by the Clifford unitary H, $HA_{(0,0)}H^\dagger = \frac{1}{2}(\mathbb{1} + X - Y + Z)$ is not a phase space point operator.

This is not a problem that can be remedied simply by choosing a better phase convention. In fact, it has been proven that no qubit Wigner function in which the phase space point operators form an operator basis is Clifford covariant [58, 61]. It turns out that Clifford unitaries are not necessary for the universality of quantum computation with magic states. Pauli measurements are sufficient since any QCM circuit containing Clifford gates and Pauli measurements is equivalent to another QCM circuit containing only Pauli measurements [61]. Therefore, Clifford covariance is not strictly needed for simulation of universal quantum computation with magic states.

A more significant obstruction to obtaining a useful discrete Wigner function for qubits from a computational perspective comes from the fact that in qudits with even-dimensional Hilbert spaces, the Pauli operators exhibit state-independent contextuality. Karanjai et al. proved a memory lower bound of $O(n^2)$ bits required to simulate contextuality [71]. The implication of this result is again, if a discrete Wigner function is going to be useful for simulating quantum computation with magic states on qubits then the phase space point operators over which it is defined cannot form an operator basis. The cardinality of the phase space must be strictly larger than 2^{2n}.

Several quasiprobability representations for systems of qubits have been defined [21, 57, 60, 136, 137]. Some of these require restrictions on the allowed states or operations, others apply only in limited cases. We review some of these in more detail below.

2.7.2 The eight state model

In the definition of the one-qubit phase space point operators above we made a choice of phase for the Pauli operators: $\mu^{\phi(a)} = i^{-\langle a_z|a_x\rangle}$ in Eq. (2.3) with the inner product taken mod 4. This choice is not unique. An equally valid choice would be $\mu^{\phi(a)} = i^{\langle a_z|a_x\rangle}$, this choice also leads to Hermitian phase space point operators of order 2. With this alternative choice, the phase space point operators are

$$\tilde{A}_{(0,0)} = \frac{1}{2}(I + X - Y + Z) \qquad \tilde{A}_{(0,1)} = \frac{1}{2}(I + X + Y - Z)$$
$$\tilde{A}_{(1,0)} = \frac{1}{2}(I - X + Y + Z) \qquad \tilde{A}_{(1,1)} = \frac{1}{2}(I - X - Y - Z).$$

Notice that under conjugation by H, the phase space point operators $A_{(0,0)}$, $A_{(0,1)}$, $A_{(1,0)}$, and $A_{(1,1)}$ map to $\tilde{A}_{(0,0)}$, $\tilde{A}_{(0,1)}$, $\tilde{A}_{(1,0)}$, and $\tilde{A}_{(1,1)}$ and vice versa. Instead of taking only four standard phase space point operators, if we take all eight operators to define our phase space we get the so called eight state model [136].

For simplicity we can label the eight phase space point operators as $A_{\pm\pm\pm} = \frac{1}{2}(\mathbb{1} \pm X \pm Y \pm Z)$. These eight phase space points lose some of the properties held by the odd-dimensional phase space. For example, the phase space point operators do not form an operator basis so for a one-qubit density matrix ρ, the decomposition

$$\rho = \sum_{\alpha \in \{\pm\pm\pm\}} W_\rho(\alpha) A_\alpha \tag{2.35}$$

is not unique. However, they regain some of the other properties that made the discrete Wigner function useful for describing quantum computation. For example, phase space

point operators map to phase space point operators under conjugation by all one-qubit Clifford unitaries making the (nonunique) Wigner function defined by Eq. (2.35) Clifford covariant. Further, the Wigner function leads to a simulation algorithm for QCM circuits on one qubit. This simulation algorithm is a special case of the simulation method presented in Chapter 3 so we will not review it in detail. It also goes beyond the odd-dimensional Wigner function in one respect—all one-qubit states are positively representable.[13]

2.7.3 Stabilizer quasimixtures

Let $\mathcal{S}$ denote the set of n-qubit stabilizer states represented by density matrices. Then any state ρ can be decomposed as

$$\rho = \sum_{\sigma \in \mathcal{S}} s_\rho(\sigma)\sigma. \tag{2.36}$$

Since density matrices are Hermitian, the coefficients $s_\rho(\sigma)$ are real for all $\sigma \in \mathcal{S}$, and taking a trace of Eq. 2.36 we find $1 = \sum_{\sigma \in \mathcal{S}} s_\rho(\sigma)$. That is, the coefficients $s_\rho(\sigma)$ describe a quasiprobability representation of the state ρ [21].

This representation lends itself to a classical simulation algorithm for quantum computation with magic states on qubits which is similar in structure to the algorithm for odd-dimensional qudits based on the Wigner function. Classical simulation is efficient when the input state of the QCM circuit is positively represented, i.e. the input state is a mixture of stabilizer states. This algorithm can be seen as a special case of the simulation method described in Chapter 3 so we will not review it in detail.

When the quasiprobability representation takes negative values simulation is still possible but it is generally inefficient. In this case, the complexity of simulation (the number of samples required to achieve a given error) is governed by the amount of negativity. Since the stabilizer states do not form a basis, they are overcomplete, the decomposition Eq. (2.36) is not unique for any given state. The freedom to choose the decomposition allows us to attempt to minimize the complexity of the simulation by minizing the amount of negativity in the coefficients. Then the *robustness of magic*, $\mathfrak{R}_S(\rho)$, defined as the solution to the minimization problem

$$\mathfrak{R}_S(\rho) = \min \left\{ \sum_{\sigma \in \mathcal{S}} |s_\rho(\sigma)| \;\middle|\; \rho = \sum_{\sigma \in \mathcal{S}} s_\rho(\sigma)\sigma \right\} \tag{2.37}$$

is the parameter governing the complexity of simulation in the case where negativity in the expansion Eq. (2.36) is unavoidable [21].

[13] We say positively representable here instead of postively represented because the decomposition Eq. (2.35) is not unique. A state being positively representable simply means there exists a decomposition of the form Eq. (2.35) in which the coefficients are nonnegative.

Chapter 3

Phase-space-simulation method for quantum computation with magic states on qubits

We propose a method for classical simulation of finite-dimensional quantum systems, based on sampling from a quasiprobability distribution, i.e., a generalized Wigner function. Our construction applies to all finite dimensions, with the most interesting case being that of qubits. For multiple qubits, we find that quantum computation by Clifford gates and Pauli measurements on magic states can be efficiently classically simulated if the quasiprobability distribution of the magic states is nonnegative. This provides the so far missing qubit counterpart of the corresponding result [V. Veitch et al., New J. Phys. **14**, 113011 (2012)] applying only to odd dimension. Our approach is more general than previous ones based on mixtures of stabilizer states. Namely, all mixtures of stabilizer states can be efficiently simulated, but for any number of qubits there also exist efficiently simulable states outside the stabilizer polytope. Further, our simulation method extends to negative quasiprobability distributions, where it provides probability estimation. The simulation cost is then proportional to a robustness measure squared. For all quantum states, this robustness is smaller than or equal to robustness of magic.

This chapter has been published in Ref. [1].

3.1 Introduction

How to mark the classical-to-quantum boundary is a question that dates back almost to the beginning of quantum theory. Ehrenfest's theorem [138] provides an early insight, and the Einstein-Podolsky-Rosen paradox [96] and Schrödinger's cat [94] are two early puzzles. The advent of quantum computation [14, 139, 140] added a computational angle: When does it become hard to simulate a quantum mechanical computing device on a classical computer? Which quantum mechanical resource do quantum computers harness to generate a computational speedup?

One instructive computational model is quantum computation with magic states (QCM) [38]. In QCM, both "traditional" indicators of quantumness (developed in the fields of quantum optics and foundations of quantum mechanics) and a computational indicator can be applied. From quantum optics and foundations, the indicators are the negativity of a Wigner function [43, 44], and the breakdown of noncontextual hidden variable models [62, 123, 126, 127]. Computer science is concerned with the breakdown of efficient classical simulation.

In the particular setting of QCM, an important distinction arises between the cases of even and odd local Hilbert space dimension d. If d is odd, then all three of the above indicators for the classical-to-quantum boundary align [19, 20, 32, 51]. This is a very satisfying situation: the physicist, the philosopher and the computer scientist can have compatible notions of what is "quantum".

In even local dimension, the situation differs starkly. Noncontextual hidden variable models for QCM are not viable regardless of computational power [20], which voids the foundational indicator, and furthermore obstructs the view of contextuality as a computational resource. Also, the multiqubit Wigner functions constructed to date do not support efficient classical simulation of QCM by sampling over phase space. Thus, the physics and computer science based criteria for classicality differ, which is an unsatisfactory state of affairs compared to odd d. The purpose of this chapter is to align the perspectives of the physicist and the computer scientist on the classical-to-quantum transition in QCM on qubits.

To prepare for the subsequent discussion, we provide a short summary of QCM, and the role of the Wigner function in it. Quantum computation with magic states operates with a restricted set of instructions, the Clifford gates. These are unitary operations defined by the property that they map all Pauli operators onto Pauli operators under conjugation. Clifford gates are not universal, and, in fact, can be efficiently classically simulated [40, 41, 114]. This operational restriction is compensated for by invoking the "magic" states, which are special quantum states that cannot be created by Clifford gates and Pauli measurements. Suitable magic states restore quantum computational universality; and in fact QCM is a leading paradigm for fault-tolerant universal quantum computation. In sum, computational power is transferred from the quantum gates to the magic states, and one is thus led to ask: Which quantum properties give the magic states their computational power?

One such property is, for odd d at least, negativity in the Wigner function. A quantum speedup can arise only if the Wigner function of the magic states assumes negative values. If, to the contrary, the Wigner function is positive, then the whole quantum computation can be efficiently classically simulated [19, 32]. Further, a positive Wigner function is, for $n \geq 2$ quantum systems, equivalent to the existence of a noncontextual hidden variable model [20, 51]. Both Wigner function negativity and contextuality of the magic states are therefore necessary quantum computational resources.

As we noted, this picture only applies if the local Hilbert space dimension is odd. This excludes the full multiqubit case, which arguably is the most important. Approaches to the qubit case have been made, e.g. through the rebit scenario [57] and multiqubit settings with operational restrictions [59, 61], or by invoking a Wigner function over Grassmann variables [60, 141], or multiple Wigner functions at once [27]. Common to these approaches is that, unlike for odd d [19], they do not efficiently simulate the evolution under general Clifford gates and Pauli measurements by sampling, a.k.a. weak simulation [142–144].

An alternative approach to weak simulation is by defining a quasiprobability function over stabilizer states [21, 38, 49], bypassing Wigner functions. It has the advantage of efficiently simulating all Clifford circuits on positively represented states. For multiqubit systems, it has so far been unknown how the stabilizer method relates to Wigner functions, but we clarify the relation here.

In this chapter, we provide the thus far missing phase space picture for QCM on multiqubit systems. Central to our discussion is a new quasiprobability function defined for all local Hilbert space dimensions d and all numbers of subsystems n. When applied to odd d, it reproduces the known finite-dimensional adaption [29–31, 46, 47] of the original Wigner function [43]; but for even d, in particular $d = 2$, it is different. Then, this quasiprobability function requires a phase space of increased size, in accordance with [71]. Even in $d = 2$, the positivity of this quasiprobability is preserved under all Pauli measurements. This property is crucial for the efficient classical simulation of QCM on positively represented states. Also, this simulation contains the efficient classical simulation [38] of stabilizer mixtures as a special case. We thus reproduce the essential features of the odd-dimensional scenario in $d = 2$.

Starting from the definition of the quasiprobability function W, we treat the following subjects: characterization of phase space for $d = 2$, preservation of positivity of W under Pauli measurements, covariance of W under all Clifford unitaries, efficient classical simulation of QCM for $W \geq 0$, relation to the qubit stabilizer formalism, hardness of classical simulation for $W < 0$, and a monotone under the free operations.

In summary, we arrive at a description that resembles the corresponding scenario in odd local dimension. Namely, negativity in the quasiprobability distribution W for the initial magic state is a necessary precondition for quantum speedup. However, one difference between even and odd d remains. In odd d, every positive Wigner function is also a noncontextual hidden variable model. This is not so for even d, due to the phenomenon of state-independent contextuality among Pauli observables.

Summary of results

This chapter addresses the full n-qubit case of quantum computation with magic states, from the perspectives of the classical-to-quantum transition and quantum computational resources. For the case of local dimension $d = 2$ we closely reproduce the relations between Wigner function and efficient classical simulation existing in odd d. Central to our discussion is a novel quasiprobability function W defined for all local Hilbert space dimensions d. It has the following general properties:

(i) For all n and d, W is Clifford-covariant and positivity-preserving under Pauli measurements.

(ii) If the local Hilbert space dimension d is even, W_ρ is nonunique for any given quantum state ρ. The set of phase space point operators corresponding to W is over-complete.

(iii) If d is odd and $n \geq 2$, then W reduces to the standard Wigner function [29–31] for odd finite dimension.

(iv) For all n and d, the stabilizer formalism is contained as a special case. All stabilizer states can be positively represented by W, and efficiently updated under Clifford operations.

(v) The present description goes beyond the stabilizer formalism. In particular, for

$d = 2$, for every number n of qubits there exist nonmixtures of stabilizer states which are positively represented by W. Furthermore, for any quantum state ρ, the 1-norm of the optimal W_ρ is smaller or equal than the robustness of magic $\mathfrak{R}_S(\rho)$. (Both robustness measures are instances of sum negativity [49].)

The following properties of W for special values of n (and $d = 2$) are also worth noting. (a) The Eight-state model [136] is a special case of W, namely for $n = 1$. (b) For Mermin's square [62], the present simulation algorithm saturates the lower bound [145] on the memory cost of classical simulation. (c) Up to two copies of magic T and H states are positively represented by W.

We establish the following main results: (I) The set of states positively represented by W is closed under Pauli measurement (Theorem 3 in Section 3.4). (II) If a quantum state ρ has a nonnegative function W_ρ, and W_ρ can be efficiently sampled, then, for every Clifford circuit applied to ρ, the corresponding measurement statistics can be efficiently sampled (Theorem 4 in Section 3.5). In this sense, $W \geq 0$ leads to efficient classical simulation of the corresponding quantum computation. (III) For $d = 2, n \geq 2$, the n-system phase space has a more complicated structure than in the case of odd d, reflecting the fact that the phase space point operators are dependent. The points in generalized multiqubit phase space are classified (Theorem 2 in Section 3.3). (IV) There exists a robustness measure $\mathfrak{R}$ which bounds the hardness of classical simulation of quantum computation with magic states, when $W_{\rho_{\text{init}}} < 0$ for the initial state ρ_{init}. $\mathfrak{R}$ is less than or equal to the robustness of magic (Lemma 16), and a monotone under Clifford unitaries and Pauli measurements (Theorem 5 in Section 3.6).

Outline

The remainder of this chapter is organized as follows. In Section 3.2 we define a quasiprobability function W. We show that it reduces to Gross' Wigner function [29] whenever the local Hilbert space dimension d is odd, but, more importantly, is different in even dimension. Specifically, W represents all quantum states redundantly for even d, which enables Clifford covariance and positivity preservation under Pauli measurement. In Section 3.3 we analyze the structure of the phase space on which W lives, for the case of multiple qubits. In particular, we classify the points of phase space. We also clarify the relation to the qubit stabilizer states and their mixtures.

In Sections 3.4 and 3.5 we turn to dynamics. In Section 3.4 we discuss the update of W under Pauli measurement, and in Section 3.5 the efficient classical simulation of QCM for positive W.

In Section 3.6 we address the case of $W_\rho < 0$. We discuss hardness of classical simulation, as well as the elements of a resource theory based on W.

In Section 3.7 we discuss the extent to which the quasiprobility function W satisfies the Stratonovich-Weyl criteria, and its relation to hidden variable models. We conclude in Section 3.8.

3.2 The quasiprobability function

In this section we introduce the generalized n-qudit phase space $\mathcal{V}$, for any local Hilbert space dimension d, and a quasiprobability distribution $W : \mathcal{V} \to \mathbb{R}$ living on it. In Section 3.2.1 we define the phase space point operators corresponding to W, and in Section 3.2.2 identify a minimal set of them. Section 3.2.3 reveals the cohomological underpinning of our construction, which links the present subject to parity proofs of quantum contextuality [69] and contextuality in measurement-based quantum computation [70].

3.2.1 Generalized phase space

We now proceed to the definition of the phase space point operators. We consider a subset Ω of E, and a function $\gamma : \Omega \to \mathbb{Z}_d$, both subject to additional constraints that will be specified in Definitions 5–7 below. The pair (Ω, γ) specifies a corresponding phase space point operator A_Ω^γ,

$$A_\Omega^\gamma := \frac{1}{d^n} \sum_{b \in \Omega} \omega^{-\gamma(b)} T_b, \tag{3.1}$$

with the constraint that

$$\omega^{-\gamma(0)} T_0 = \mathbb{1}. \tag{3.2}$$

When comparing Eq. (3.1) to the phase space point operators of the previously discussed qudit [19], rebit [57] and restricted qubit [61] cases, we note that the overall structure remains the same. In this case, the sets Ω are an additional varying parameter, and the phase space thereby becomes larger.

Based on the phase space point operators A_Ω^γ of Eq. (3.1), we introduce the counterpart to the Wigner function that applies to our setting. The generalized phase space $\mathcal{V}$ consists of all admissible pairs (Ω, γ), to be specified below. Any n-system quantum state ρ can be expanded in terms of a function $W_\rho : \mathcal{V} \to \mathbb{R}$,

$$\rho = \sum_{(\Omega,\gamma) \in \mathcal{V}} W_\rho(\Omega, \gamma) A_\Omega^\gamma. \tag{3.3}$$

The reason for imposing Eq. (3.2) is that it implies $\mathrm{Tr} A_\Omega^\gamma = 1$, for all $(\Omega, \gamma) \in \mathcal{V}$. Hence, W defined in Eq. (3.3) is a quasiprobability distribution. As we see shortly, it generalizes the Wigner function [31] for odd-dimensional qudits to qubits.

We note that when d is even, the quasiprobability distribution W_ρ is nonunique because the set of phase space point operators of Eq. (3.1) is overcomplete.

Definition 4. *An n-qudit quantum state ρ is* positively representable *if it can be expanded in the form of Eq.* (3.3)*, with $W_\rho(\Omega, \gamma) \geq 0$, for all $(\Omega, \gamma) \in \mathcal{V}$.*

The efficient classical simulation algorithm described in Section 3.5 applies to positively representable quantum states ρ. The nonuniqueness of W_ρ allows for more positively representable states than prior quasiprobability representations.

We now turn to the properties of admissible sets Ω and functions γ that define points in the phase space $\mathcal{V}$. We constrain Ω by the following definitions:

Definition 5. *A set $\Omega \subset E$ is* closed under inference *if it holds that*

$$a, b \in \Omega \wedge [a, b] = 0 \Longrightarrow a + b \in \Omega. \tag{3.4}$$

The motivation for this definition is that if T_a and T_b can be simultaneously measured, then the value of T_{a+b} can be inferred from the measurement outcomes, through relation Eq. (2.6). A consequence of the closedness under inference is that $0 \in \Omega$ for all closed sets Ω.

Definition 6. *A set $\Omega \subset E$ is* noncontextual *if there exists a value assignment $\gamma : \Omega \to \mathbb{Z}_d$ that satisfies the condition*

$$\gamma(a) + \gamma(b) - \gamma(a + b) = -\beta(a, b), \tag{3.5}$$

for all $a, b \in \Omega$ with $[a, b] = 0$.

To motivate the nomenclature, if the set $\Omega \subset E$ is noncontextual per the above definition, then it does not admit a parity-based contextuality proof [69]. Namely, Eq. (3.5) represents the constraints on noncontextual value assignments γ that result from the operator constraints Eq. (2.6). If these constraints can be satisfied, then there is no parity-based contextuality proof.

Definition 7. *The generalized phase space $\mathcal{V}$ consists of all pairs (Ω, γ) such that (i) Ω is closed under inference, (ii) Ω is noncontextual, (iii) $\gamma : \Omega \to \mathbb{Z}_d$ satisfies the relation Eq.* (3.5)*, and (iv) Eq.* (3.2) *holds.*

Thus, for the generalized phase space $\mathcal{V}$, the only sets Ω that matter are simultaneously closed and noncontextual. For short, we call such sets "cnc".

3.2.2 Maximal sets Ω

The cnc sets Ω partially specify the points in phase space, and it is thus desirable to eliminate possible redundancies among them. It turns out that, for qubits ($d = 2$), only the "maximal" sets Ω need to be considered for $\mathcal{V}$.

Definition 8. *A cnc set $\Omega \subset E$ is maximal if there is no cnc set $\tilde{\Omega} \subset E$ such that $\Omega \subsetneq \tilde{\Omega}$.*

We denote by $\mathcal{V}_M$ the subset of $\mathcal{V}$ constructed only from the maximal cnc sets Ω. Then, any quantum state ρ has expansions like Eq. (3.3), but with $\mathcal{V}$ replaced by $\mathcal{V}_M$. If one of those expansions is nonnegative, then we say that ρ is positively representable with respect to $\mathcal{V}_M$.

Lemma 8. *For any n and $d = 2$, a quantum state ρ is positively representable with respect to $\mathcal{V}$ if and only if it is positively representable with respect to $\mathcal{V}_M$.*

From the perspective of positive representability, for the case of qubits we may therefore shrink $\mathcal{V}$ to $\mathcal{V}_M$ without loss. We make use of this property in the classification of cnc sets Ω for the multiqubit case in Section 3.3.3. The proof of Lemma 8 is given in Appendix B.2.1.

3.2.3 The cohomological viewpoint

The above Definitions 6 and 7 have a cohomological underpinning, which connects the subject of the present chapter to the topological treatment of parity-based contextuality proofs [69], and of contextuality in measurement-based quantum computation [70].

The cohomological picture arises as follows. The partial value assignments γ and the function β are cochains in a chain complex, with Eqs. (2.7) and (3.5) constraining them. Eq. (2.7) says that β is a special cochain, namely a cocycle. Now, the basic reason for why the case of even d is so much more involved than the case of odd d is that, for even d, the cocycle β is nontrivial whereas for odd d it is trivial [69].

Eqs. (2.7) and (3.5) are frequently used in this chapter, for example in the update rules of the phase space point operators under Pauli measurements (proof of Lemma 12), the closedness of the generalized phase space $\mathcal{V}$ under update by Pauli measurement (proof of Lemma 14), and covariance of the quasiprobability function W under Clifford unitaries (proof of Lemma 17). These are central properties for the phase-space description of quantum computation with magic states. and they are all matters of cohomology.

The cohomological formulation is based on a chain complex $\mathcal{C}_n$ constructed from the n-qubit Pauli operators T_a. The operator labels a define the edges of this complex; the faces of $\mathcal{C}_n$ correspond to commuting pairs (a, b) and volumes (a, b, c) to commuting triples. For details, the interested reader is referred to [69]. Here, we only state two basic topological properties of the present scenario.

As already noted, the cochain β defined in Eq. (2.6) is in fact a 2-cocycle, with the cocycle condition $d\beta = 0$ enforced by Eq. (2.7). For any given volume $v = (a, b, c)$, the coboundary $d\beta$ evaluates on v to

$$d\beta(a, b, c) := \beta(a, b) + \beta(a + b, c) - \beta(b, c) - \beta(a, b + c). \tag{3.6}$$

Thus, Eq. (2.7) says that $d\beta(v) = 0$, for all volumes v.

Eq. (3.5) in Definition 6 also has a cohomological interpretation, namely $d\gamma = -\beta|_{\Omega\times\Omega}$, with

$$d\gamma(a, b) := \gamma(a) + \gamma(b) - \gamma(a + b), \tag{3.7}$$

for any face (a, b) spanned by commuting edges $a, b \in E$.

Subsequently, we use evaluations of $d\beta$ and $d\gamma$, defined in Eqs. (3.6) and (3.7), as a short-hand to express Eqs. (2.7) and (3.5). As outlined above, it is conceptually helpful to remember that $d\beta$ and $d\gamma$ denote coboundaries, but it is not required for the technical results presented in this chapter.

3.3 Properties of the phase space $\mathcal{V}$

In this section, we look at the structure of the phase space $\mathcal{V}$ more closely, and make connections to previous phase space formulations. Namely, in Section 3.3.1 we address the relationship of this phase space to the usual qudit phase space, and in Section 3.3.2 we make clear the connections to the previously addressed rebit case. Further, in Section 3.3.3, we classify the cnc sets Ω, and for every Ω describe the sets $\Gamma(\Omega)$ of value assignments γ. In Section 3.3.4 we clarify the relation to the stabilizer formalism.

3.3.1 Qudits of odd dimension

This is the only place in the present chapter where we consider the case of odd d. The purpose of this section is to show that if d is odd then for $n \geq 2$ qudits the generalized phase space $\mathcal{V}$ includes the standard phase space $V = \mathbb{Z}_d^{2n}$. There, the quasiprobability function W becomes the standard Wigner function [31] for odd finite-dimensional systems. Hence, the present quasiprobability function W is a generalization of the finite odd dimensional Wigner function [31], which in turn is a descendant of the original Wigner function [43].

If d is odd then the whole set E is cnc. First, E is closed under inference by definition. And second, it is known that in odd dimension Pauli observables have noncontextual deterministic value assignments [61, 146]. These yield the functions γ, satisfying the condition Eq. (3.5). E is thus noncontextual.

Hence,

$$A_E^\gamma = \frac{1}{d^n} \sum_{a \in E} \omega^{-\gamma(a)} T_a,$$

are phase space point operators with the functions γ satisfying Eqs. (3.2) and (3.5). The former condition ensures that the identity operator appears with weight $1/d^n$ in the expansion (real and positive). If $n \geq 2$, the latter condition has d^{2n} solutions for the functions γ [51]. For a suitable phase convention for the Pauli operators (see Section 2.3.3), it holds that $\beta \equiv 0$ (odd d only). The solutions for γ then form a vector space

$$V = \mathbb{Z}_d^{2n} \quad \text{(for odd } d\text{)}.$$

We note that the case of a single qudit, $n = 1$, is an exception [51].

3.3.2 Qubits and rebits

The remainder of this chapter is about local Hilbert space dimension $d = 2$. This means mostly qubits, but we will occasionally also consider systems of rebits. The reason is that the major complication of the $d = 2$ case stems from Mermin's square and star [62]—two strikingly simple contextuality proofs. Those settings embed most efficiently in rebits rather than qubits, which warrants the inclusion of rebits here.

We remark that the present discussion of rebits is almost identical to the discussion of qubits, but very different from the earlier discussion of rebits in [57]. In the latter, the physically measurable observables were restricted from real Pauli operators to CSS-type Pauli operators, and the real Clifford unitaries to CSS-ness preserving Clifford unitaries. No such restrictions are imposed here. If the restriction to CSS-ness preserving operations is imposed, then Mermin's square and star, along with all other state-independent contextuality proofs based on Pauli observables, are effectively excised [57]. Here, we retain those contextuality proofs, and consequently have to adjust to their presence. Notably, these contextuality proofs constrain quasiprobability distributions that preserve positivity under Pauli measurement.

We start the exploration of the $d = 2$ case with two examples that illustrate the concept of generalized phases space $\mathcal{V}$. The second example also illustrates the differences between contextuality, negative quasiprobability and quantum computational power for two-level systems.

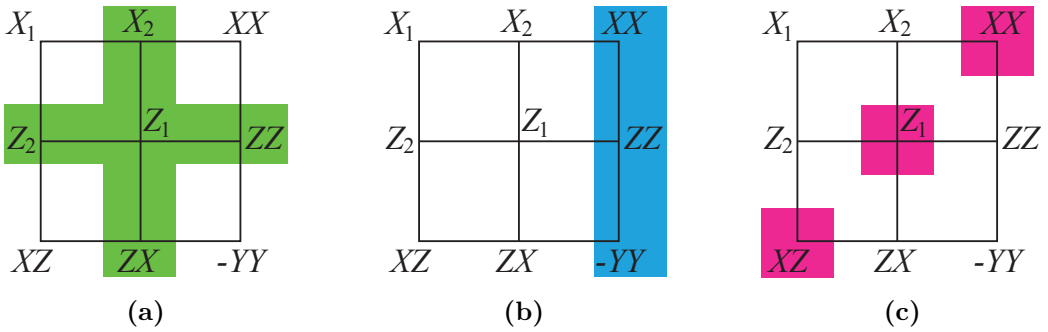


Figure 3.1: Three types of cnc sets Ω for Mermin's square. (a) union of two isotropic subspaces intersecting in one element, (b) isotropic subspace, (c) triple of anti-commuting elements.

Example 1 (Eight-state model). It is known that *every* one-qubit quantum state can be positively represented by the so-called Eight-state-model [136], which consists of two standard 1-qubit Wigner functions tagged together. The Eight-state-model is an instance of the state expansion Eq. (3.3), namely for $d = 2$, $n = 1$, and it contains only one set Ω,

$$\Omega_0 = \{0, x, y, z\},$$

with $T_0 = I$, $T_x = X$, $T_y = Y$ and $T_z = Z$. It is easily checked that Ω_0 is noncontextual and closed under inference (no inference possible). The value assignments γ are constrained by Eq. (3.2), hence $\gamma(0) = 0$, and no constraints arise from Eq. (2.7) due to the lack of nontrivial commuting elements in Ω_0. Thus, $\gamma(x)$, $\gamma(y)$ and $\gamma(z)$ can be freely chosen. There are eight resulting functions, and they define the eight states of the model.

All one-qubit quantum states can be positively represented by this model, which is strictly more than all mixtures of one-qubit stabilizer states.

Example 2 (Mermin's square). Mermin's square is at the very root of the complications that arise for Wigner functions in even dimension. In particular, no n-qubit Wigner function for which the corresponding phase space point operators form an operator basis can preserve positivity under all Pauli measurements [61].

All observables appearing in Mermin's star are real, and can thus be embedded in two rebits. Our formalism is easily adaptable to this slightly simpler scenario. Figure 3.1 shows three distinct types of cnc sets Ω. Type (a) is the union of two nontrivially intersecting isotropic subspaces (9 sets), type (b) is isotropic subspaces (6 sets), and type (c) is triples of anti-commuting elements, i.e., one from each row and column of the square (6 sets). For each cnc set Ω of type (a), (b) and (c) of Figure 3.1, the constraint Eq. (3.5) allows for 2^3, 2^2, 2^3 functions γ, respectively. The number of phase space points of each type therefore is 72, 24, 48.

We make the following numerical observations about the two-rebit case: (i) Random sampling suggests that all 2-rebit states are positively representable; see Table 3.2. (ii) In

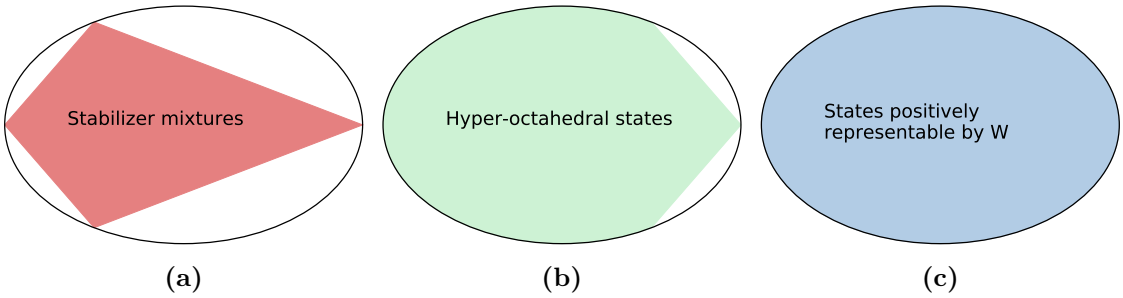


Figure 3.2: Two-dimensional cross section of the two-qubit state space, as parameterized in Eq. (3.8). The shaded regions indicate the positively representable states by various methods; (a) mixtures of stabilizer states, (b) hyper-octahedral states [147], and (c) states positively represented by the present phase space method.

Figure 3.2 the region of positively representable density matrices of the form

$$\rho(x,y) = \frac{1}{4} I_{12} + x(X_1X_2 + Z_1Z_2 - Y_1Y_2) + y(Z_1 + Z_2), \tag{3.8}$$

for $x, y \in \mathbb{R}$, is displayed for three different methods; namely the stabilizer method, the hyper-octrahedral method [147], and the present phase space method. We find that all quantum states in the plane spanned by the parameters x, y are positively represented by the present phase space method, and this is not the case for the stabilizer [38] and hyper-octahedral method [147].

Example 3 (2 qubits). Numerical analysis shows that two copies of the state

$$|H(\phi)\rangle := (|0\rangle + e^{-i\phi}|1\rangle)/\sqrt{2} \tag{3.9}$$

can be positively represented, for all angles ϕ.

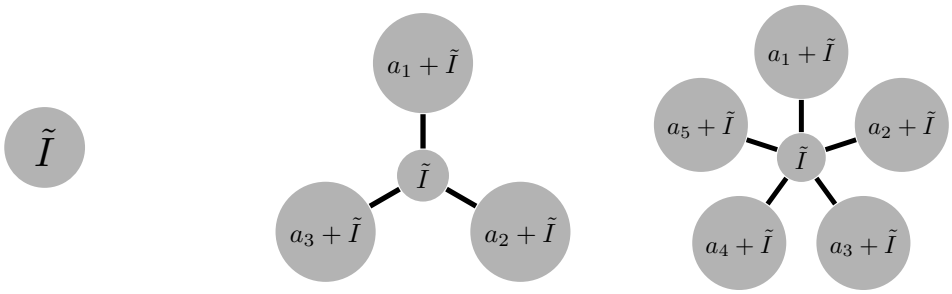
3.3.3 Classification of phase space points

Denote by $\Gamma(\Omega)$ the set of functions $\gamma : \Omega \to \mathbb{Z}_2$ that satisfy the constraints Eqs. (3.2) and (3.5). Then, the following statement holds.

Lemma 9. *For all sets Ω of Definition 7, $\Gamma(\Omega)$ is the coset of a vector space $U(\Omega)$.*

Proof of Lemma 9. Write $\gamma = \gamma_0 + \eta$, where $\gamma_0 \in \Gamma(\Omega)$ is some reference function. Then, the only condition on the functions $\eta \in U(\Omega)$ is $d\eta = 0$. Thus, if $\eta, \eta' \in U(\Omega)$ then $c\eta + c'\eta' \in U(\Omega)$, for all $c, c' \in \mathbb{Z}_2$. □

Lemma 9 reproduces a familiar feature. In infinite and finite odd dimension, the whole phase space is an orbit under the vector space of translations. There is an origin 0 of phase space, and all other phase space points are obtained from it by translation. In our present case of $d = 2$, the phase space $\mathcal{V}$ splinters into many fragments, each of which corresponds to a vector space U attached to a cnc set Ω.



(a) $m = 0$ **(b)** $m = 1$ **(c)** $m = 2$

Figure 3.3: Commutativity graph representation for the cosets of Eq. (3.10) sets. Elements pair-wise commute within each vertex and elements in adjacent vertices pair-wise commute. Elements in nonadjacent vertices anti-commute.

At this point, one question about the structure of $\mathcal{V}$ remains: Can the cnc sets Ω be classified? It is resolved by Lemma 10 and Theorem 2 below.

Lemma 10. *For n qubits, consider an isotropic subspace $\tilde{I} \subset E$ of dimension $n - m$, with $m \leq n$, and $\xi \leq 2m + 1$ elements $a_k \in E$ that pairwise anti-commute but all commute with $\tilde{I}$. Denote $I_k := \langle a_k, \tilde{I} \rangle$ for $k = 1, .., \xi$. For any number n of qubits, the sets*

$$\Omega = \bigcup_{k=1}^{\xi} I_k \tag{3.10}$$

are noncontextual and closed under inference.

Proof of Lemma 10. Existence. The sets Ω of Eq. (3.10) exist for all m, n. To see this, consider the m-qubit Jordan-Wigner transforms of the Majorana fermion operators acting on qubits 1 to m,

$$\begin{array}{rcl} C_{2j-1} & = & I_{1..j-1} X_j Z_{j+1} Z_{j+2} .. Z_{m-1} Z_m, \\ C_{2j} & = & I_{1..j-1} Y_j Z_{j+1} Z_{j+2} .. Z_{m-1} Z_m, \end{array} \tag{3.11}$$

for $j = 1, .., m$, and, if $m > 0$, the further observable

$$C_{2m+1} = Z_1 Z_2 ... Z_{m-1} Z_m. \tag{3.12}$$

Further, be $\tilde{I}$ the isotropic subspace corresponding to a stabilizer state supported on the $n - m$ qubits numbered $m + 1, .., n$. Define a_k via $C_k = T_{a_k}$, for all $k = 1, .., 2m + 1$. These a_k and $a \in \tilde{I}$ have the commutation relations required.

Closedness. Consider a pair $c, d \in \Omega$ such that $[c, d] = 0$. There are two cases. (i) $c, d \in I_k$, for some k. Then, $c + d \in I_k$, hence $c + d \in \Omega$.

(ii) $c \in I_k$ and $d \in I_l$, $k \neq l$. We may write $c = \nu\, x + g$, $d = \mu\, y + g'$, for some $\nu, \mu \in \mathbb{Z}_2$ and $g, g' \in \tilde{I}$. The commutation relation $[c, d] = 0$ then implies that $\nu\mu = 0$, hence either $\nu = 0$ or $\mu = 0$. Wlog. assume that $\nu = 0$. Then, $c \in \tilde{I}$, hence $c, d \in I_l$. Thus, $c+d \in I_l \subset \Omega$.

In both cases, $c, d \in \Omega$ and $[c, d] = 0$ implies that $c + d \in \Omega$. Hence, Ω is closed under inference.

Noncontextuality. There exists a function $\gamma|_{\tilde{I}} : \tilde{I} \to \mathbb{Z}_2$ that satisfies Eq. (3.5) on $\tilde{I}$. We now extend this function to Ω as follows. The values $\gamma(a_k)$, for $k = 1, .., \xi$ can be freely chosen, and for all $a \in \tilde{I}$ and all k, $\gamma(a_k + a) := \gamma(a_k) + \gamma(a) + \beta(a_k, a)$. This fully defines $\gamma : \Omega \to \mathbb{Z}_2$. All commuting triples $c, d, c+d$ lie within one of the isotropic spaces I_k forming Ω, and $d\gamma(a, b) = \beta(a, b)$ thus holds.

This establishes that the sets Ω of Eq. (3.10) exist for the maximum value of ξ, $\xi = 2m+1$. One may always choose ξ smaller, which neither affects closedness nor noncontextuality. □

Theorem 2. *All maximal cnc sets Ω are of the form Eq.* (3.10)*, with $\xi = 2m + 1$ and $1 \leq m \leq n$.*

Proof of Theorem 2. Let $\Omega \subset E$ be closed under inference and noncontextual. We can partition the elements of Ω into two subsets, $\Omega = \{q_1, \ldots, q_\mu | g_1, \ldots, g_\nu\}$, where $\tilde{I} = \{g_1, \ldots, g_\nu\}$ are the elements of Ω which commute with the whole set. $\tilde{I}$ is an isotropic subspace since if two elements, a and b, commute with Ω, then clearly their sum, $a + b$, also commutes with Ω, and $\tilde{I}$ is isotropic by definition.

If all elements of Ω pair-wise commute then $\Omega = \tilde{I}$ is an isotropic subspace. Isotropic subspaces are not maximal cnc sets because they are always contained in Eq. (3.10) sets with parameter $m = 1$. If Ω is not an isotropic subspace then it can be written compactly as

$$\Omega = \bigcup_{k=1}^{\xi} \langle p_k, \tilde{I} \rangle \tag{3.13}$$

where $\xi \geq 2$, the cosets $p_1 + \tilde{I}, \ldots, p_\xi + \tilde{I}$ are distinct and $q_1, \ldots, q_\mu$ are in the cosets $p_1 + \tilde{I}, \ldots, p_\xi + \tilde{I}$. Note that in this form, there can be no element p_j which commutes with all of $p_1, \ldots, p_\xi$ because $\tilde{I}$ is defined to contain all such elements. Now we consider the possible commutation relations that $p_1, \ldots, p_\xi$ can have if Ω is noncontextual.

The Mermin square is generated by products of commuting pairs of the two qubit Pauli operators $\{X_1, X_2, Z_1, Z_2\}$. This is a contextual set. Therefore, any set which is closed under inference and contains four elements p_1, p_2, p_3, p_4 with the commutation relations like those of $\{X_1, X_2, Z_1, Z_2\}$:

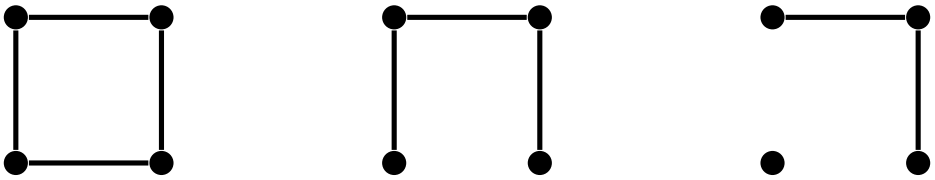
$$\begin{cases} [p_1, p_2] & = [p_1, p_4] = [p_2, p_3] = [p_3, p_4] = 0 \\ [p_1, p_3] & = [p_2, p_4] = 1 \end{cases} \tag{3.14}$$

will necessarily contain the full Mermin square and therefore be contextual.

Another sufficient condition for a closed under inference set to be contextual is that it contains four elements with the commutation relations

$$\begin{cases} [p_1, p_2] & = [p_2, p_3] = [p_3, p_4] = 0 \\ [p_1, p_3] & = [p_1, p_4] = [p_2, p_4] = 1. \end{cases} \tag{3.15}$$

The reason is that since the set is closed under inference, it will necessarily contain the elements $p_1 + p_2$ and $p_3 + p_4$, and the elements $p_1, p_1 + p_2, p_3 + p_4, p_4$ have the commutation relations of Eq. (3.14). Thus, it must contain a Mermin square.



(a) C_4 **(b)** P_4 **(c)** $K_1 \cup P_3$

Figure 3.4: Forbidden induced subgraphs of the commutativity graph, resulting from Mermin's square (also see [149]).

A similar argument shows that another sufficient condition for a closed under inference set to be contextual is that it contains four elements with the commutation relations

$$\begin{cases} [p_1, p_2] &= [p_2, p_3] = 0 \\ [p_1, p_3] &= [p_1, p_4] = [p_2, p_4] = [p_3, p_4] = 1. \end{cases} \tag{3.16}$$

In this case, since the set is closed under inference, it must also contain the elements $p_1 + p_2$ and $p_2 + p_3$ and the elements $p_1 + p_2, p_2, p_2 + p_3, p_4$ have the commutation relations of Eq. (3.14).

To determine the possible commutation relations of the elements $p_1, \ldots, p_\xi$, we will look at their commutativity graph $\mathcal{G}$. That is the undirected graph with a vertex for each of $p_1, \ldots, p_\xi$ and an edge connecting each pair of commuting vertices. Since Ω is noncontextual, the commutation relations of Eq. (3.14), Eq. (3.15), and Eq. (3.16) provide restrictions on the possible commutation relations of the elements $p_1, \ldots, p_\xi$ of Ω. In terms of the commutativity graph $\mathcal{G}$, these are forbidden induced subgraphs[1].

The restriction of Eq. (3.14) says that $\mathcal{G}$ cannot have a four vertex chordless cycle (C_4) as an induced subgraph and the restriction from Eq. (3.15) says that $\mathcal{G}$ cannot have a four vertex path (P_4) as an induced subgraph. These two forbidden induced subgraphs characterize the trivially perfect graphs [148]. I.e. $\mathcal{G}$ must be a trivially perfect graph.

Connected trivially perfect graphs have the the property that they contain a universal vertex [148][2]. If the commutativity graph $\mathcal{G}$ were connected then there would be an element p_j which commutes with all other elements of $\{p_1, \ldots, p_\xi\}$. This is also forbidden. Therefore, the graph $\mathcal{G}$ is disconnected.

Given that $\mathcal{G}$ is disconnected, Eq. (3.16) provides another restriction. Namely that each connected component of $\mathcal{G}$ cannot have a three vertex path (P_3) as an induced subgraph. I.e. each connected component of $\mathcal{G}$ is a clique.

This means we can partition the elements $\{p_1, \ldots, p_\xi\}$ into disjoint subsets

$$\{p_1, \ldots, p_\xi\} = \{p_{1,1}, p_{1,2}, \ldots, p_{1,\xi_1}\} \cup \{p_{2,1}, p_{2,2}, \ldots, p_{2,\xi_2}\} \cup \cdots \cup \{p_{\pi,1}, p_{\pi,2}, \ldots, p_{\pi,\xi_\pi}\} \tag{3.17}$$

where two elements commute if and only if they are in the same subset in the partition. Since the set $\{p_1, \ldots, p_\xi\}$ is closed under inference, each subset in the partition must be closed

[1] An induced subgraph of a graph is the graph obtained by taking a subset of the vertices of the original graph and all of the edges connecting pairs of vertices in the subset

[2] A universal vertex is a vertex that is adjacent to every other vertex in the graph.

under inference. Now suppose a subset in the partition contained at least two elements. Then since the subset is closed under inference it must also contain their sum. But each of the two elements anticommutes with the elements of all other subsets in the partition so their sum must commute with the elements of all other subsets in the partition. This is a contradiction. Therefore, each subset in the partition contains a single element. Thus, the elements $\{p_1, \ldots, p_\xi\}$ of Eq. (3.13) pair-wise anticommute.

Maximal cnc sets are sets of the form Eq. (3.13) for which ξ is maximal for a given isotropic subspace $\tilde{I}$. If the isotropic subspace $\tilde{I}$ has dimension $n - m$ where n is the number of qubits and $1 \leq m \leq n$, then the pair-wise anticommuting elements p_k which complete the set are elements of the symplectic complement $\tilde{I}^\perp$. This is a m dimensional symplectic subspace, therefore the maximal value of ξ is the largest number of pair-wise anticommuting Pauli operators on m qubits. The largest sets of pair-wise anticommuting Pauli operators on m qubits have $2m + 1$ elements. This can be seen as follows. Consider the elements $a_k \in E$ given by $T_{a_k} = C_k$, with C_k defined in Eq. (3.11).

The set $\{a_k |\ 1 \leq k \leq 2m\}$ consists of pairwise anticommuting elements. There is an element c, with $T_c = C_{2m+1}$ (see Eq. (3.12)), that anticommutes with each one of the elements in this set. It is the only element in E to do so, since the set of equations

$$[c, a_k] = 1 \quad 1 \leq k \leq 2m \tag{3.18}$$

has a unique solution. Therefore together with this element we can construct a set of size $2m + 1$.

We would like to show any other set of pairwise anticommuting elements whose size is $2m$ can be mapped bijectively to the set we constructed. Suppose $\{\tilde{a}_k |\ 1 \leq k \leq 2m\}$ is such a set. By Witt's lemma [150, §20] the function that sends a_k to $\tilde{a}_k$ extends to a linear map $f : E \to E$ that satisfies $[f(v), f(w)] = [v, w]$ for all $v, w \in E$ (symplectic transformation). Therefore there is a unique element that anticommutes with all the $\tilde{a}_k$, and it is given by $f(c)$. In particular, $2m + 1$ is the maximal number.

To complete the proof we must show that maximal sets of pair-wise anticommuting elements on m qubits with size less than $2m + 1$ do not lead to maximal cnc sets. To see this note that by Witt's lemma, for any maximal anticommuting set of size $2m' + 1$ ($m' < m$), there is a bijection $f : E \to E$ which maps the set to one of the form Eqs. (3.11) and (3.12). Therefore, we can find $m - m'$ independent elements which commute with the set. For example, if $g_1, g_2, \ldots, g_{m-m'}$ are the vectors corresponding to Pauli operators $X_{m'+1}, X_{m'+2}, \ldots, X_m$, then we could take $f^{-1}(g_1), f^{-1}(g_2), \ldots, f^{-1}(g_{m-m'})$. Therefore, the $n - m$ dimensional isotropic subspace can be extended to one with dimension $n - m'$.

This completes the proof. Therefore, all maximal cnc sets have the form Eq. (3.10), with $\xi = 2m + 1$. □

A result equivalent to the characterization of Eq. (3.17) is given in Theorem 3 of [149].

Tensor products of phase space point operators are not typically phase space point operators. Consider, for example, two phase space point operators with $m = 2$ and $n \geq 2$. Their tensor product does not appear in the classification provided by Theorem 2, as the commutativity graph shows. Physically, such tensor products are not closed under inference, violating Definition 7. Upon closure, they cease to be noncontextual as they then contain a Mermin square. Hence the closures also violate Definition 7.

But there is an exception. If one of the two phase space point operators in the tensor product corresponds to an isotropic subspace, i.e., has $m = 0$, then the tensor product *is* a valid phase space point operator. See Appendix A.2.2 for details.

3.3.4 Relation to the stabilizer formalism

The purpose of this section is to describe the relation between positive representability by the quasiprobability distribution W and qubit stabilizer states. We demonstrate that, for all n, the set of positively W-representable states contains the stabilizer mixtures as a strict subset. This is the content of Lemma 11 below. The lemma is based on two examples.

Example 4. Let $|\mathrm{stab}\rangle$ be an n-qubit stabilizer state, with isotropic subspace $\tilde{I} \subset E$ corresponding to its stabilizer. Then, it is easily verified that $\tilde{I}$ is noncontextual and closed under inference. Namely, $\tilde{I}$ is of form Eq. (3.10), with $m = 0$, $\xi = 1$.

The next example generalizes Example 1 to n-qubit states.

Example 5. Every n-qubit state of the form $\Psi = \rho_1 \otimes |\mathrm{stab}\rangle\langle\mathrm{stab}|_{2,..,n}$, with ρ a general one-qubit state and $|\mathrm{stab}\rangle$ an $n-1$-qubit pure stabilizer state, is positively representable.

To prove this statement, for any number n of qubits, consider an isotropic subspace $\tilde{I} \subset E$ of rank $n-1$ representing the stabilizer state $|\mathrm{stab}\rangle_{2,..,n}$, and three elements $x, y, z \in E$, such that $T_x = X_1$, $T_y = Y_1$ and $T_z = Z_1$. Define the three isotropic subspaces $I_x, I_y, I_z \subset E$,

$$I_x = \langle x, \tilde{I}\rangle,\ I_y = \langle y, \tilde{I}\rangle,\ I_z = \langle z, \tilde{I}\rangle,$$

and $\Omega_{xyz} := I_x \cup I_y \cup I_z$. Ω_{xyz} is of form Eq. (3.10), with $m = 1$, $\xi = 3$ and $n \geq 2$, hence cnc by Lemma 10.

We now apply this result to the state $\Psi = \rho_1 \otimes |\mathrm{stab}\rangle\langle\mathrm{stab}|_{2,..,n}$ above. We can write the constitutents as $\rho = \sum_{\gamma_0} W_\rho(\Omega_0, \gamma_0) A^{\gamma_0}_{\Omega_0}$, with $W_\rho \geq 0$ (see Example 1), and $|\mathrm{stab}\rangle\langle\mathrm{stab}| = A^{\tilde{\gamma}}_{\tilde{I}}$ (see Example 2). We observe that

$$A^{\gamma_0}_{\Omega_0} \otimes A^{\tilde{\gamma}}_{\tilde{I}} = A^{\gamma}_{\Omega_{xyz}},$$

with $\gamma := \gamma_0(a|_1) + \tilde{\gamma}(a|_{2,..,n}) \mod 2$. To see this, recall that $\Omega_0 = \{0, x, y, x\}$, and note that the set Ω_{xyz} can also be written as $\Omega_{xyz} = \tilde{I} \cup (\tilde{I} + x) \cup (\tilde{I} + y) \cup (\tilde{I} + z)$, where the coset $\tilde{I} + x := \{a + x,\ \forall a \in \tilde{I}\}$, etc. Thus, $\Psi = \sum_{\gamma_0} W_\rho(\Omega_0, \gamma_0) A^{\gamma}_{\Omega_{xyz}}$. Since $W_\rho \geq 0$ by Example 1, the states Ψ are all positively representable. Yet not all these states are mixtures of stabilizer states. Stabilizer mixedness is preserved under partial trace. Now assume that Ψ is a stabilizer mixture for all ρ. Then $\mathrm{Tr}_{2..n}\Psi = \rho_1$ is also a stabilizer mixture. Contradiction.

We cast the combined conclusion of Examples 4 and 5 as a Lemma.

Lemma 11. *For all $n \in \mathbb{N}$, all mixtures of n-qubit stabilizer states are positively representable, and furthermore there exist positively representable states that are not mixtures of stabilizer states.*

3.4 Quantum mechanical rules for state update under measurement

In the previous sections we have analyzed the generalized phase space $\mathcal{V}$ on which the quasiprobability function W is defined. We now turn to dynamics.

For our setting of QCM this concerns evolution under the free operations, i.e., the Clifford unitaries and Pauli measurements. As already noted in [57] and [61], the situation simplifies even further. If the goal is to sample from the joint probability distribution of measurement outcomes—which is the case in quantum computation—then only the update under Pauli measurements needs to be considered.

The Clifford unitaries can be propagated forward in time, thereby conjugating the Pauli measurements into other such measurements, past the final measurement and then discarded. (This redundancy notwithstanding, we will visit the update of W under Clifford unitaries in Section 3.6.3, where we prove covariance.) The main results of this section are Theorem 3 and Lemma 12.

Theorem 3. *For any $n \in \mathbb{N}$, the set $\mathcal{P}_n$ of positively representable n-qubit quantum states is closed under Pauli measurement.*

To describe the dynamics under measurement, we need to set up some further notation. For every set Ω we introduce the derived set $\Omega \times a$. Denoting $\mathrm{Comm}(a) := \{b \in E | [a,b] = 0\}$ and $\Omega_a := \Omega \cap \mathrm{Comm}(a)$,

$$\Omega \times a := \Omega_a \cup \{a + b |\, b \in \Omega_a\}, \quad \forall a \notin \Omega. \tag{3.19}$$

Likewise, we define an update on functions γ invoking the measurement outcome s of an observable T_a, namely $(\cdot) \times s : (\gamma : \Omega \to \mathbb{Z}_2) \mapsto (\gamma \times s : \Omega \times a \to \mathbb{Z}_2)$. We define this update only for $(\Omega, \gamma) \in \mathcal{V}$, and only for $a \notin \Omega$[3]. The updated function $\gamma \times s : \Omega \times a \to \mathbb{Z}_2$ is given by

$$\gamma \times s(b) := \gamma(b), \qquad \forall b \in \Omega_a, \tag{3.20a}$$

$$\gamma \times s(b) := \gamma(a+b) + s + \beta(a,b), \qquad \forall a + b \in \Omega_a. \tag{3.20b}$$

The rules of Eq. (3.20) are used to formulate the update rule for phase space point operators of Eq. (3.1) under Pauli measurement.

Remark. Update rules similar to Eq. (3.20) have been used previously [151] to construct a ψ-epistemic model of the multiqubit stabilizer formalism. Those rules update the value assignments in the same way but are applied under different conditions. Specifically, the update in [151] does not refer to general sets Ω satisfying the conditions of Def 7.

Lemma 12. *Denote the projectors $\Pi_a^s := (I + (-1)^s T_a)/2$, and be A_Ω^γ a phase space point operator defined through Eq.* (3.1)*, with $(\Omega, \gamma) \in \mathcal{V}$ satisfying the conditions of Definition 7.*

[3]The definitions of $\Omega \times a$ and $\gamma \times s$ can without modification be extended to $a \in \Omega$. However, in that case the function values $\gamma \times s(b)$ can be determined both through Eqs. (3.20a) and (3.20b), and we need to check consistency. These inferences are indeed consistent, as a consequence of Eq. (3.5). Since we do not need the case of $a \in \Omega$ subsequently, we skip the details of the argument.

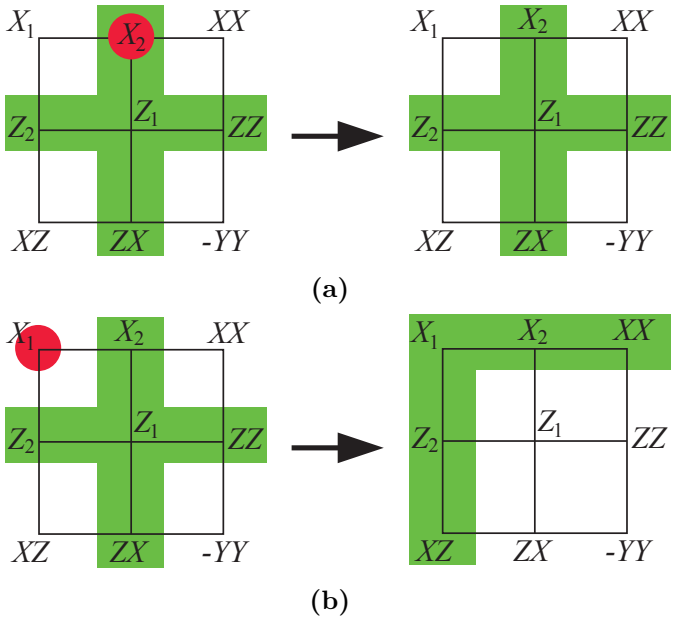


Figure 3.5: Update of a cnc set Ω in Mermin's square, under two Pauli measurements. (a) The measured observable X_2 is such that $a(X_2) \in \Omega$, hence the update proceeds by Eq. (3.21a). (b) The measured observable X_1 is such that $a(X_1) \notin \Omega$; hence the update proceeds by Eq. (3.21b).

Then, the effect of a measurement of the Pauli observable T_a with outcome s on A^γ_Ω is

$$\Pi^s_a A^\gamma_\Omega \Pi^s_a = \delta_{s,\gamma(a)} \frac{A^\gamma_\Omega + A^{\gamma+[a,\cdot]}_\Omega}{2}, \qquad \text{if } a \in \Omega, \tag{3.21a}$$

$$\Pi^s_a A^\gamma_\Omega \Pi^s_a = \frac{1}{2} A^{\gamma\times s}_{\Omega\times a}, \qquad \text{if } a \notin \Omega. \tag{3.21b}$$

Example 2, continued. Eq. (3.21) entails the update of both the sets Ω and the functions γ. Here we only consider the former. Figure 3.5 displays the update of the set Ω shown in Figure 3.1 a, under the measurement of (a) the observable X_1, with $a(X_1) \notin \Omega$, and (b) the observable X_2, with $a(X_2) \in \Omega$.

In preparation for the proof of Lemma 12 it is useful to state two relations of the function β for $d = 2$. With the definition Eq. (2.6) of β and Eq. (3.2), the operator identities $T_a T_a = \mathbb{1}$ and $T_b \mathbb{1} = T_b$ imply that

$$\beta(a, a) = \beta(a, 0) = \gamma(0), \quad \forall a \in E. \tag{3.22}$$

Furthermore, evaluating $d\beta(a, a, 0) = 0$ (see Eqs. (2.7) and (3.6)), and using Eq. (3.22) yields

$$\beta(a, b) = \beta(a, a + b), \quad \forall a, b. \tag{3.23}$$

To prove Lemma 12 we also need the following result.

Lemma 13. *If $\Omega \subset E$ is noncontextual and closed under inference, then so is Ω_a, for all $a \in E$.*

Proof of Lemma 13. First consider closure. Assume that $c, d \in \Omega_a$ and $[c, d] = 0$. Then, $c, d \in \Omega$, and also $c + d \in \Omega$, since Ω is closed by assumption. Further, $[c, a] = [d, a] = 0$ implies $[c + d, a] = 0$, and hence $c + d \in \Omega_a$. Ω_a is thus closed.

Now consider noncontextuality. Since Ω is noncontextual, there exists a function γ such that $d\gamma = \beta$ on Ω. Since Ω_a is closed, β can be properly restricted to $\mathcal{C}(\Omega_a)$, and so can γ. Hence, $d\gamma|_{\mathcal{C}(\Omega_a)} = \beta|_{\mathcal{C}(\Omega_a)}$. Thus, Ω_a is noncontextual. □

Proof of Lemma 12. Under the measurement of T_a with outcome $s \in \mathbb{Z}_2$ we have

$$\begin{aligned}\frac{I + (-1)^s T_a}{2} A^\gamma_\Omega \frac{I + (-1)^s T_a}{2} =& \frac{I + (-1)^s T_a}{2} \frac{1}{2^n} \sum_{b \in \Omega_a} (-1)^{\gamma(b)} T_b \\ =& \frac{1}{2 \cdot 2^n} \sum_{b \in \Omega_a} (-1)^{\gamma(b)} T_b + \frac{(-1)^s}{2 \cdot 2^n} \sum_{b \in \Omega_a} (-1)^{\gamma(b)} T_a T_b. \end{aligned} \tag{3.24}$$

From hereon we need to distinguish two cases, $a \in \Omega$ and $a \notin \Omega$.

Case I: $a \in \Omega$. Focusing on the second term in the expansion Eq. (3.24),

$$\begin{aligned}(-1)^s \sum_{b \in \Omega_a} (-1)^{\gamma(b)} T_a T_b =& (-1)^s \sum_{b \in \Omega_a} (-1)^{\gamma(b) + \beta(a,b)} T_{a+b} \\ =& (-1)^{s+\gamma(a)} \sum_{b \in \Omega_a} (-1)^{\gamma(a+b)} T_{a+b} \\ =& (-1)^{s+\gamma(a)} \sum_{a+b \in \Omega_a} (-1)^{\gamma(a+b)} T_{a+b} \\ =& (-1)^{s+\gamma(a)} \sum_{b \in \Omega_a} (-1)^{\gamma(b)} T_b. \end{aligned}$$

Therein, in the first line we have used Eq. (2.6), in the second line Eq. (3.5), in the third line the completeness of Ω_a under inference (Lemma 13), and the fourth line is just a relabeling of the elements in Ω_a. Inserting this result in the above expansion Eq. (3.24), we find

$$\Pi^s_a A^\gamma_\Omega \Pi^s_a = \delta_{s,\gamma(a)} \frac{1}{2^n} \sum_{b \in \Omega_a} (-1)^{\gamma(b)} T_b, \tag{3.25}$$

and Eq. (3.21a) follows.

Case II: $a \notin \Omega$. Substituting $b \longrightarrow a + b$ in Eq. (3.20b) gives $\gamma \times s(a + b) = \gamma(b) + s + \beta(a, a + b)$, for $b \in \Omega_a$. With Eq. (3.23) we obtain

$$\gamma \times s(a + b) = \gamma(b) + s + \beta(a, b), \quad \forall b \in \Omega_a. \tag{3.26}$$

With this, we now look at the second term in the expansion Eq. (3.24),

$$\begin{aligned}(-1)^s \sum_{b\in\Omega_a} (-1)^{\gamma(b)} T_a T_b =& (-1)^s \sum_{b\in\Omega_a} (-1)^{\gamma(b)+\beta(a,b)} T_{a+b} \\ &= \sum_{b\in\Omega_a} (-1)^{\gamma\times s(a+b)} T_{a+b}.\end{aligned}$$

The first line above follows with Eq. (2.6), and the second with Eq. (3.26).

Considering the first term in the expansion Eq. (3.24), with Eq. (3.20a) we have

$$\sum_{b\in\Omega_a} (-1)^{\gamma(b)} T_b = \sum_{b\in\Omega_a} (-1)^{\gamma\times s(b)} T_b$$

Inserting the above expressions for the two terms in Eq. (3.24), and using the definition Eq. (3.19) of $\Omega \times a$, we obtain Eq. (3.21b). □

We have so far shown how the phase space point operators can be updated under measurement once. We still need to show that this update can be iterated. This requires that the new phase space point operators appearing on the r.h.s. of Eq. (3.21) satisfy the consistency constraints of Definition 7.

Lemma 14. *If $(\Omega, \gamma) \in \mathcal{V}$ then $(\Omega, \gamma + [a, \cdot]) \in \mathcal{V}$, for all $a \in \Omega$, and $(\Omega \times a, \gamma \times s) \in \mathcal{V}$, for all $a \notin \Omega$ and $s \in \mathbb{Z}_2$.*

The proof of Lemma 14 is given in Appendix B.2.2.

Proof of Theorem 3. Consider a state $\rho \in \mathcal{P}_n$, and a measurement of the Pauli observable T_a on it. Assume that the measurement outcome s can occur, $p_a(s) := \mathrm{Tr}(\Pi_a^s \rho) > 0$. We have to show that under these conditions, the post-measurement state

$$\rho' = \frac{\Pi_a^s \rho \Pi_a^s}{p_a(s)}$$

is also contained in the set $\mathcal{P}_n$.

Denote $\overline{\delta}_{a\in\Omega} := 1 - \delta_{a\in\Omega}$. Then, with Lemma 12 and the state expansion Eq. (3.3) of ρ, we have

$$\rho' = \sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{W_\rho(\Omega,\gamma)}{p_a(s)} \Big(\delta_{a\in\Omega}\delta_{s,\gamma(a)} \frac{A_\Omega^\gamma + A_\Omega^{\gamma+[a,\cdot]}}{2} + \frac{1}{2}\overline{\delta}_{a\in\Omega} A_{\Omega\times a}^{\gamma\times s} \Big). \tag{3.27}$$

Thus, ρ' can be represented by a quasiprobability distribution $W_{\rho'}$ with elements

$$\begin{aligned}W_{\rho'}(\Omega',\gamma') = \sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{W_\rho(\Omega,\gamma)}{2p_a(s)} \big(&\delta_{a\in\Omega}\delta_{s,\gamma(a)} \left(\delta_{(\Omega',\gamma'),(\Omega,\gamma)} + \delta_{(\Omega',\gamma'),(\Omega,\gamma+[a,\cdot])} \right) \\ &+ \overline{\delta}_{a\in\Omega} \delta_{(\Omega',\gamma'),(\Omega\times a,\gamma\times s)} \big).\end{aligned} \tag{3.28}$$

The $W_{\rho'}(\Omega',\gamma')$ are thus linear combinations of $W_\rho(\Omega,\gamma)$ with nonnegative coefficients (0 or $1/2p_a(s)$). Since the $W_\rho(\Omega,\gamma)$ are nonnegative by assumption, it follows that $W_{\rho'}(\Omega',\gamma') \geq 0$, for all $(\Omega',\gamma') \in \mathcal{V}$. □

m	0	1	$\{1,2\}$
2 rebits	24	72	120
2 qubits	60	240	432

Table 3.1: Number of points in phase space as a function of $\{m\}$.

3.5 Classical simulation for $W_\rho \geq 0$

3.5.1 Simulation algorithm

We now turn to the question of how hard it is to classically simulate the outcome statistics for a sequence of Pauli measurements on an initial quantum state. In this regard, we show that if the initial quantum state is positively represented and the corresponding probability distribution W can be efficiently sampled from, then the statistics of the measurement outcomes can be efficiently simulated.

The classical simulation procedure in Algorithm 2 describes weak simulation [143, 144], i.e., it outputs one sample from the joint probability distribution $p(s_{a_1}, s_{a_2}, .., s_{a_N})$ of outcomes corresponding to a sequence of measurements of Pauli operators $T_{a_1}, T_{a_2}, \ldots, T_{a_N}$ (T_{a_1} is measured first, T_{a_N} last). If more than one sample are desired, the procedure is just repeated. We note that the measurement can be adaptive. I.e., it is not necessary for the simulation algorithm that a measurement sequence is committed to at the beginning. As a special case of this, the measured observables may depend on earlier measurement outcomes.

We have the following result.

Theorem 4. *If for an initial quantum state ρ it holds that $W_\rho \geq 0$ and furthermore W_ρ can be efficiently sampled from, then the output distribution of all sequences of Pauli measurements, possibly interspersed with Clifford gates, on ρ can be classically efficiently sampled from.*

As a first application of Theorem 4, we return to Example 2, Mermin's square.

Example 6 (Example 2 continued). How much memory capacity is needed to classically simulate measurements of the observables in Mermin's square? We first turn to the state-independent case, which was previously discussed in [145]. The task is to devise a classical algorithm that outputs an outcome sequence for any given sequence of Pauli measurements, which can occur according to quantum mechanics. The measurement sequence can be of any length and the measurements therein may be commuting or anti-commuting. In [145], a lower bound on the memory cost of any such simulation was established, $\log_2 24$ bits; and a specific model was constructed that attains it.

The classical simulation algorithm of Table 2 also saturates this limit. To show this, we use as cnc sets Ω the six maximal isotropic subspaces of two rebits, see Figure 3.1 b. This set of sets Ω is closed under update by Pauli measurement, as described by Eq. (3.21). For each such set Ω, each value assignment γ is specified by two evaluations (the other evaluations then follow via Eq. (3.5)). There are thus four functions γ for each cnc set Ω, hence 24 combinations in total, which is the same as in [145].

1: Draw a sample $(\Omega, \gamma) \in \mathcal{V}$ according to the probability distribution W_ρ representing the initial quantum state ρ.
2: **for all** the observables $T_{a_1}, T_{a_2}, .., T_{a_N}$ measured in this sequence, repeat the following steps. **do**
3: For the i-th measurement, set $a := a_i$.
4: **if** $a \in \Omega$ **then**
5: Ω is unchanged.
6: Output the value $s = \gamma(a)$.
7: Flip a coin.
8: **if** Heads **then**
9: $\gamma \longrightarrow \gamma$,
10: **else**
11: $\gamma \longrightarrow \gamma + [a, \cdot]$.
12: **end if**
13: **else if** $a \notin \Omega$ **then**
14: $\Omega \longrightarrow \Omega \times a$.
15: Flip a coin.
16: **if** Heads **then**
17: $s = 0$,
18: **else**
19: $s = 1$.
20: **end if**
21: Output this value s.
22: Update $\gamma \longrightarrow \gamma \times s$, through Eq. (3.20).
23: **end if**
24: **end for**

Algorithm 2: Classical simulation algorithm for sampling from the joint probability distribution of a sequence of Pauli measurements on a positively represented initial quantum state.

We now turn to the state-dependent version of the problem. How much memory is needed to sample from the correct outcome statistics for arbitrary measurement sequences, for any two-rebit state ρ with $W_\rho \geq 0$, and given the capability to sample from W_ρ? This problem is harder than the former: Not only must the sequence of outcomes be internally consistent for all measurement sequences, but also it needs to represent the state ρ.

Memory cost now depends on the state ρ. If ρ is a mixture of stabilizer states, i.e., the sets Ω can be limited to $m = 0$, then the classical simulation algorithm of Algorithm 2 can still run on $\log_2 24 \approx 4.59$ bits.

If sets Ω with $m = 1$ are included in the expansion, then more two-rebit states ρ can be positively represented (among them, for example, $|T\rangle_1 \otimes |T\rangle_2$) but on the other hand, memory consumption goes up. For $m = 1$, there are $3^2 \times 2^3$ pairs (Ω, γ), see Figure 3.1a. Hence the memory consumption for configurations with $m = 1$ is $\log_2 72 \approx 6.17$ bits. (Note that the sets Ω for $m = 0$ are not maximal. If sets with $m = 1$ are included, then sets with $m = 0$ can be omitted without loss.) The size $|\mathcal{V}(\{m\})|$ of the phase space vs. the

m	0	1	$\{1,2\}$	hy.oct.
V_+/V [pure]	0	1	1	0
V_+/V [mixed]	0.144	1	1	0.924

(a) Two rebits

m	0	1	$\{1,2\}$	hy.oct.
V_+/V [pure]	0	0.980	0.980	0
V_+/V [mixed]	0.009	1	1	0.568

(b) Two qubits

Table 3.2: Volume fraction of state space filled by the positively representable states, as a function of $\{m\}$; (top) two rebits, (bottom) two qubits. The volume fraction V_+/V was obtained numerically, by sampling 10^6 random states according to the Fubini-Study measure for pure states (second row) and the Hilbert-Schmidt measure for mixed states (third row). The first column, $m = 0$, describes mixtures of stabilizer states, and the last column hyper-octahedral states [147] for comparison.

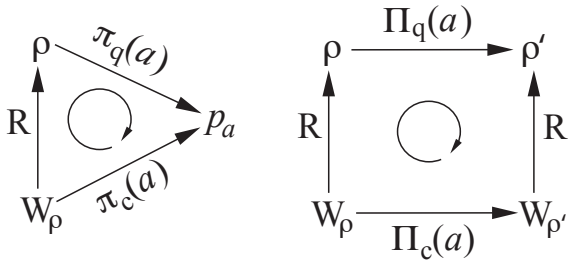


Figure 3.6: Diagrams representing the quantum and the classical way of calculating the probability of measurement outcomes (left) and the post-measurement state (right).

maximum value of m is displayed in Table 3.1. The memory cost is $\log_2 |\mathcal{V}(\{m\})|$. The volume fraction of positively representable two-rebit and two-qubit states is displayed in Table 3.2, for various sets $\{m\}$.

3.5.2 Correctness and efficiency of the classical simulation

In preparation for the proof of correctness of the classical simulation algorithm, we introduce the following notation. Given a probability distribution W_ρ, there are two objects that the classical simulation algorithm needs to reproduce correctly, namely the probability $p_a(s)$ for the outcomes $s \in \mathbb{Z}_2$ of the measurement of any Pauli observable T_a, and the post-measurement state ρ'. There are two ways of obtaining these quantities, a quantum-mechanical one and a classical one using the simulation algorithm of Section 3.5.1.

Regarding the outcome probability $p_a(s)$ given W_ρ, the quantum mechanical way first obtains the corresponding quantum state ρ from W_ρ through Eq. (3.3). This is represented by a map $R : W_\rho \mapsto \rho$. Second, from ρ, the outcome probability $p_a(s)$ is obtained via

the Born rule, $p_a(s) = \mathrm{Tr}(\Pi_a^s \rho)$. This is represented by a map $\pi_q(a) : \rho \mapsto p_a$. The classical way uses the algorithm of Section 3.5.1 to obtain p_a. This is represented by a map $\pi_c(a) : W_\rho \mapsto p_a$.

Likewise, the quantum mechanical way of obtaining the post-measurement state ρ' from W_ρ proceeds by first applying the map R (Eq. (3.3)) to obtain ρ, and second by obtaining ρ' from ρ through the Dirac projection postulate. The second step is represented by a map $\Pi_q(a)$.

The classical way of obtaining ρ' from W_ρ proceeds by first using the simulation algorithm to obtain $W_{\rho'}$, and second by mapping $W_{\rho'}$ to ρ' using the map R. The first step in this procedure is represented by the map $\Pi_c(a)$.

The classical simulation algorithm of Section 3.5.1 is correct only if the quantum and the classical ways of computing $p_a(s)$ and ρ' agree. That is, we require the diagrams in Figure 3.6 to commute.

Lemma 15. *The diagrams of Figure 3.6 commute.*

Proof of Lemma 15. We discuss the outcome probability and the post-measurement state separately.

Outcome probability $p_a(s)$. Then, the quantum mechanical expression for $p_a(s)$ is

$$\begin{aligned} p_a(s) &= \sum_{(\Omega,\gamma)\in\mathcal{V}} W_\rho(\Omega,\gamma)\ \mathrm{Tr}\left(\frac{I+(-1)^s T_a}{2} A_\Omega^\gamma\right) \\ &= \sum_{(\Omega,\gamma)\in\mathcal{V}} W_\rho(\Omega,\gamma)\left(\delta_{a\in\Omega}\delta_{s,\gamma(a)} + \frac{1}{2}\overline{\delta}_{a\in\Omega}\right). \end{aligned} \tag{3.29}$$

The classical expression $p_a^{(c)}(s)$ for $p_a(s)$ obtained through the algorithm of Section 3.5.1 is as follows. If $a \in \Omega$, then the conditional probability for the outcome s given the state (Ω,γ) is $\delta_{s,\gamma(a)}$. If $a \notin \Omega$ then the conditional probability for the outcome s is $1/2$. Thus,

$$p_a^{(c)}(s) = \sum_{(\Omega,\gamma)\in\mathcal{V}} W_\rho(\Omega,\gamma)\left(\delta_{a\in\Omega}\delta_{s,\gamma(a)} + \frac{1}{2}\overline{\delta}_{a\in\Omega}\right).$$

By comparing the two expressions, we find that $p_a(s) = p_a^{(c)}(s)$ for all a, s, and the left diagram of Eq. (3.6) thus commutes.

Post-measurement state ρ'. The quantum mechanical expression for the post-measurement state ρ' has already been given in Eq. (3.27), and we now derive the corresponding expression $\rho'_{(c)}$ that follows from the classical simulation algorithm.

We consider the joint probability $p((\Omega',\gamma')\cap s)$ of obtaining the outcome s in the measurement of T_a and ending up in the state (Ω',γ'). We may invoke conditional probabilities in two ways,

$$p((\Omega',\gamma')\cap s) = \sum_{(\Omega,\gamma)\in\mathcal{V}} p((\Omega',\gamma')\cap s|(\Omega,\gamma))W_\rho(\Omega,\gamma) = p((\Omega',\gamma')|s)p_a(s).$$

Noting that $p((\Omega',\gamma')|s) = W_{\rho'}(\Omega',\gamma')$, and equating the two above expressions we find

$$W_{\rho'}(\Omega',\gamma') = \sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{W_\rho(\Omega,\gamma)}{p_a(s)} p((\Omega',\gamma')\cap s|(\Omega,\gamma)). \tag{3.30}$$

We now infer the conditional probabilities $p((\Omega',\gamma')\cap s|(\Omega,\gamma))$ from the classical simulation algorithm of Section 3.5.1,

$$p((\Omega',\gamma')\cap s|(\Omega,\gamma)) = \begin{cases} \frac{1}{2}\delta_{s,\gamma(a)}\delta_{\Omega',\Omega}\left(\delta_{\gamma',\gamma} + \delta_{\gamma',\gamma+[a,\cdot]}\right), & \text{if } a\in\Omega, \\ \frac{1}{2}\delta_{\Omega',\Omega\times a}\delta_{\gamma',\gamma\times s}, & \text{if } a\notin\Omega. \end{cases}$$

Inserting this into Eq. (3.30), and using the resulting expression in Eq. (3.3), i.e. applying the map R, we obtain

$$\begin{aligned} \rho'_{(c)} &= \sum_{(\Omega',\gamma')\in\mathcal{V}} \sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{W_\rho(\Omega,\gamma)}{2p_a(s)} \left(\delta_{a\in\Omega}\delta_{s,\gamma(a)}\delta_{\Omega',\Omega}\left(\delta_{\gamma',\gamma}+\delta_{\gamma',\gamma+[a,\cdot]}\right) + \overline{\delta}_{a\in\Omega}\delta_{\Omega',\Omega\times a}\delta_{\gamma',\gamma\times s}\right) A^{\gamma'}_{\Omega'} \\ &= \sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{W_\rho(\Omega,\gamma)}{2p_a(s)} \left(\delta_{a\in\Omega}\delta_{s,\gamma(a)}\left(A^\gamma_\Omega + A^{\gamma+[a,\cdot]}_\Omega\right) + \overline{\delta}_{a\in\Omega}A^{\gamma\times s}_{\Omega\times a}\right). \end{aligned}$$

Comparing the last expression with Eq. (3.27), we find that $\rho'_{(c)} = \rho'$ for all a, s, and the right diagram in Figure 3.6 thus commutes. □

Proof of Theorem 4. As explained in Section 3.4, we only need to discuss sequences of Pauli measurements. For those, we show that Algorithm 2 is correct, and, if the initial W_ρ can be efficiently sampled from, it is also computationally efficient. (i) *Correctness.* Denote by $\rho(t)$ the state before the t-th measurement. With Lemma 15, by induction on the right diagram in Eq. (3.6), if $W_{\rho(1)}$ represents the initial state $\rho(1)$, then $W_{\rho(t)}$ represents $\rho(t)$ for all time steps $t = 1,..,N$. Then, by the left diagram in Eq. (3.6), the outcome probabilities $p_{a_t}(s_{a_t}|\mathbf{s}_{\prec t})$, with $\mathbf{s}_{\prec t} = (s_{a_1},..,s_{a_{t-1}})$ the measurement record prior to time t, are also correct. Thus the joint outcome probability sampled from

$$p_{a_1,\ldots,a_n}(s_{a_1},\ldots,s_{a_N}) = \prod_{t=1}^N p_{a_t}(s_{a_t}|\mathbf{s}_{\prec t}),$$

is also correct.

(ii) *Efficiency.* We recall that all cnc sets Ω are unions of $O(n)$ isotropic spaces Ω_i (Theorem 2). Further, each Ω_i defines a stabilizer group

$$T^\gamma_{\Omega_i} := \{T^\gamma_a := (-1)^{\gamma(a)}T_a, a\in\Omega_i\}. \tag{3.31}$$

This allows us to describe $(\Omega,\gamma)\in\mathcal{V}$ using polynomial memory by storing $O(n)$ stabilizer tables of size $O(n^2)$ [40, 114]. Indeed, by Definitions 5–6 and Lemma 10, $T^\gamma_{\Omega_i}$ is a closed commutative group. Furthermore, with Definition 6, it holds that $T^\gamma_a T^\gamma_b = T^\gamma_{a+b}, \forall a,b\in\Omega_i$. This implies the existence of a nontrivial stabilized subspace: $\Pi^\gamma_{\Omega_i} := \sum_{a\in\Omega_i} T^\gamma_a/|\Omega_i|$ is a common $+1$-eigenprojector of every $T_a\in T^\gamma_{\Omega_i}$ as $T^\gamma_a\Pi^\gamma_{\Omega_i} = \sum_{b\in\Omega_i}\frac{T^\gamma_{a+b}}{|\Omega_i|} = \sum_{b'\in\Omega_i}\frac{T^\gamma_{b'}}{|\Omega_i|} =$

$\Pi^{\gamma}_{\Omega_i}, \forall T^{\gamma}_{\Omega_i}$, which also implies ${\Pi^{\gamma}_{\Omega_i}}^2 = \Pi^{\gamma}_{\Omega_i}$.

We now note that the update rules in Algorithm 2, namely (i) checking whether $a \in \Omega$, (ii) evaluating γ on $a \in \Omega$, (iii) updating $\gamma \longrightarrow \gamma + [a, \cdot]$, (iv) $\Omega \longrightarrow \Omega \times a$ and (v) $\gamma \longrightarrow \gamma \times s$, implement tasks that admit efficient classical algorithms in the stabilizer formalism [40, 114]. *Rules (i) and (ii):* To test $a \in \Omega$, we check whether $a \in \Omega_i, i = 1, \ldots, O(n)$. If $a \in \Omega_j$ for some value of j, then $\gamma(a)$ is computed as the bit determining the phase of the stabilizer operator $T^{\gamma}_a \in T^{\gamma}_{\Omega_i}$. Both tasks can be solved classically efficiently via Gaussian elimination given the stabilizer table data [40, 114]. *Rule (iii):* γ is updated to $\gamma' = \gamma + [a, \cdot]$ by (classically efficiently) evaluating $\gamma(\cdot) + [a, \cdot]$ on the generators of every Ω_i. *Rules (iv) and (v):* For all j, $T^{\gamma|_{\Omega_j} \times s}_{\Omega_j \times a}$ is the stabilizer group resulting from the measurement of T_a with outcome s on a state with stabilizer group $T^{\gamma|_{\Omega_j}}_{\Omega_j}$. This update can be efficiently performed using the standard measurement update-rule of Ref. [40, 114] to every stabilizer table in the description of (Ω, γ). Thus, all steps of the algorithm run in polynomial time. □

3.6 The case of $W_\rho < 0$

As we have established in the previous sections, $W_\rho < 0$ is a precondition for quantum speedup. When the initial state is represented by a quasiprobability rather than a true probability function, a standard problem of interest is estimating outcome probabilities for sequences of measurements. An established method for probability estimation is [49], utilizing the Hoeffding bound. Note that probability estimation is a different problem than weak simulation [143, 144], and is not efficiently adaptive.

3.6.1 Robustness

In close analogy to the "robustness of magic" [21] $\mathfrak{R}_S$ (the subscript S is for "stabilizer"), we define a phase space robustness $\mathfrak{R}$, through

$$\mathfrak{R}(\rho) := \min_{W \,|\, \langle \mathcal{A}, W \rangle = \rho} \|W\|_1, \tag{3.32}$$

with $\langle \mathcal{A}, W \rangle := \sum_{\alpha \in \mathcal{V}} W_\alpha A_\alpha$.

Since the definitions of the robustness $\mathfrak{R}$ and of the robustness of magic $\mathfrak{R}_S$ [21] are so similar, one may wonder if there is a relation between them. This is indeed the case; namely, we have the following result.

Lemma 16. *For all quantum states ρ, of any number n of qubits, the phase space robustness $\mathfrak{R}(\rho)$ and the robustness of magic $\mathfrak{R}_S(\rho)$ are related via*

$$\mathfrak{R}(\rho) \leq \mathfrak{R}_S(\rho) \leq (4n + 1)\, \mathfrak{R}(\rho). \tag{3.33}$$

Thus, the phase space robustness $\mathfrak{R}$ is never larger than the robustness of magic, but can only be moderately smaller. The proof of Lemma 16 is given in Appendix B.2.3.

3.6.2 Hardness of classical simulation

The Hoeffding bound says that the number N of samples required to estimate the output probability distribution up to an error ϵ scales as $N \sim \mathcal{M}^2/\epsilon^2$, where $\mathcal{M}$ is a measure of the negativity contained in the quantum process. In our case, the operations are positivity-preserving, and all negativity comes from the initial state. The algorithm of Pashayan et al. [49], when applied to our setting, says that the number N of samples required to estimate the output probability scales as

$$N \sim \frac{\mathfrak{R}(\rho_{\text{init}})^2}{\epsilon^2}.$$

Thus, the robustness $\mathfrak{R}(\rho_{\text{init}})$ of the initial state ρ_{init} is the critical parameter determining the classical hardness of probability estimation.

The same relation, with the robustness $\mathfrak{R}$ replaced by the robustness of magic $\mathfrak{R}_S$ holds for the classical simulation based on quasiprobability distributions over stabilizer states [21]. Lemma 16 above is therefore of interest for relating the operational costs of the two simulation methods.

Classical simulation also requires a quasiprobability function $W_{\mu^{\otimes n}}$ for n copies of the magic state μ. Since the n-qubit phase space is large, the numerical optimization to obtain the least-negative expansion $W_{\mu^{\otimes n}}$ is computationally costly. However, we can apply a similar splitting into smaller blocks of magic states as in the stabilizer case [21]. The computational cost for providing the expansion is then a function of block size rather than total number of copies n. The 1-norm of the resulting expansion is smaller than of the stabilizer expansion, by a factor that is constant in n. Details are given in Appendix A.2.2.

3.6.3 Elements of a resource theory based on W

It is illuminating to discuss QCM within the framework of resource theories. Every resource theory has three main operational components [152], (i) the resource(s), (ii) the nonresources, or free states, (iii) the free operations.

In the physical setting of our interest, the resources are quantum states which cannot be positively represented by W (see Theorem 4). The free operations are Clifford unitaries and Pauli measurements. The free states are those that can be created from the free operations from a completely mixed state, i.e., all mixtures of stabilizer states.

We observe that there is a third class of states which are neither resources nor free, namely the positively representable states which are not mixtures of stabilizer states. Such states are called (iv) bound magic states. We have seen an example of them in Section 3.3.4, the general 1-qubit states tensored with a stabilizer state on arbitrarily many qubits.

The reason for calling those states "bound magic" is that they cannot be distilled into computationally useful ones by free operations. In our setting, by Theorem 3, positive representability is an invariant under the free operations. Hence, bound states can only be converted into other bound states or into free states by the free operations, but never into a resource.

The question of inter-convertibility may more generally be asked for resource states. To facilitate this discussion, one may identify monotones, i.e., real-valued functions on the state space that never increase under the free operations. The main result of this section

is that the robustness $\mathfrak{R}$, defined in Eq. (3.32) and already known to measure hardness of classical simulation by sampling, is a monotone.

Theorem 5. *The robustness $\mathfrak{R}$ is a monotone under all Clifford unitaries and Pauli measurements.*

As part of the proof of Theorem 5, we now discuss an important structural property of the quasiprobability function W, namely its covariance under Clifford unitaries. Be $\mathcal{C}\ell_n$ the n-qubit Clifford group. It acts on the n-qubit Pauli operators via

$$g(T_a) := gT_ag^\dagger = (-1)^{\Phi_g(a)}T_{S_ga}, \ \ \forall g \in \mathcal{C}\ell_n.$$

This implies an action of the Clifford group on the phase space point operators A^γ_Ω, which in turn induces an action on the sets Ω and the functions γ, via

$$g(A^\gamma_\Omega) = \frac{1}{2^n}\sum_{a\in\Omega}(-1)^{\gamma(a)}g(T_a) = \frac{1}{2^n}\sum_{b\in\Omega'}(-1)^{\gamma'(b)}T_b.$$

Therein, the set Ω' is defined as $\Omega' := \{S_ga, \ a \in \Omega\}$, and the function $\gamma' : \Omega' \to \mathbb{Z}_2$ is given by

$$\gamma'(S_ga) := \gamma(a) + \Phi_g(a), \ \forall a \in \Omega.$$

Henceforth we denote Ω' as $g \cdot \Omega$ and γ' as $g \cdot \gamma$, to emphasize the dependence on $g \in \mathcal{C}\ell_n$.

For use in the proof below we quote Lemma 3 from [69] which says that, for any face $(a, b) \in \Omega \times \Omega$,

$$\Phi_g(\partial(a, b)) = \beta(S_ga, S_gb) + \beta(a, b) \mod 2.$$

We then have the following result.

Lemma 17. *$\mathcal{V}$ is mapped to itself under $\mathcal{C}\ell_n$, and the quasiprobability function W transforms covariantly. That is, if the state ρ can be described by W_ρ through Eq. (3.3), then for any $g \in \mathcal{C}\ell_n$ the state $g\rho g^\dagger$ can be described by a quasiprobability function $W_{g\rho g^\dagger}$ defined by*

$$W_{g\rho g^\dagger}(\Omega, \gamma) := W_\rho(g^{-1} \cdot \Omega, g^{-1} \cdot \gamma).$$

Remark 1. We say "the state ρ can be described by W_ρ" rather than "is described" because W_ρ is not unique.

Proof of Lemma 17. First, we show that the phase space $\mathcal{V}$ is closed under the action of $\mathcal{C}\ell_n$, i.e., if $(\Omega, \gamma) \in \mathcal{V}$ then $(\Omega', \gamma') \in \mathcal{V}$. The four items in Definition 7 need to be checked. (i) *Closedness under inference.* Assume that $c, d \in \Omega'$ and $[c, d] = 0$. Then there exist $a, b \in \Omega$ such that $c = S_ga$, $d = S_gb$ and $[a, b] = 0$. Then, $c + d = S_ga + S_gb = S_g(a + b) \in \Omega'$, since $a + b \in \Omega$ by the assumption of closedness. Hence Ω' is closed under inference.

(iii) *γ' satisfies Eq.* (3.5). With the definition of γ' we have (all addition mod 2)

$$\begin{aligned} d\gamma'(S_ga, S_gb) =& d\gamma(a, b) + \Phi_g(\partial(a, b)) \\ =& d\gamma(a, b) + \beta(S_ga, S_gb) + \beta(a, b) \\ =& \beta(S_ga, S_gb). \end{aligned}$$

Therein, in the second line we have used Eq. (3.5). Thus, γ' satisfies Eq. (3.5) on its domain.

(ii) Ω' *is noncontextual.* With γ' we have just proved the existence of a function on Ω' that satisfies Eq. (3.5).

(iv) γ' *satisfies Eq.* (3.2). Since γ satisfies Eq. (3.2), it follows $\mathbb{1} = g(\mathbb{1}) = g\left((-1)^{\gamma(0)}T_0\right) = (-1)^{\gamma(0)+\Phi_g(0)}T_0 = (-1)^{\gamma'(0)}T_0$. Eq. (3.2) is thus satisfied for γ'.

Hence, if $(\Omega,\gamma) \in \mathcal{V}$ then $(\Omega',\gamma') \in \mathcal{V}$, as claimed.

Next we turn to the covariance of W under $\mathcal{C}\ell_n$. We have

$$\begin{aligned} g\rho g^\dagger &= \sum_{(\Omega,\gamma)\in\mathcal{V}} W_\rho(\Omega,\gamma) g(A^\gamma_\Omega) \\ &= \sum_{(\Omega,\gamma)\in\mathcal{V}} W_\rho(\Omega,\gamma) A^{g\cdot\gamma}_{g\cdot\Omega} \\ &= \sum_{(\Omega,\gamma)\in\mathcal{V}} W_\rho(g^{-1}\cdot\Omega, g^{-1}\cdot\gamma) A^\gamma_\Omega. \end{aligned}$$

Comparing the last expression with the expansion Eq. (3.3) for $g\rho g^\dagger$, we find that for all $g \in \mathcal{C}\ell_n$, the quasiprobability distribution $W_{g\rho g^\dagger}$ defined by

$$W_{g\rho g^\dagger}(\Omega,\gamma) = W_\rho(g^{-1}\cdot\Omega, g^{-1}\cdot\gamma) \tag{3.34}$$

describes the state $g\rho g^\dagger$. This is the covariance condition. □

We are now ready to prove the monotonicity of $\mathfrak{R}$, as stated in Theorem 5.

Proof of Theorem 5. (a) Clifford unitaries. With Lemma 17, we have that for any n-qubit Clifford gate g applied to any n-qubit state ρ, the quasiprobability distribution $W_{g\rho g^\dagger}$ can be related to W_ρ via the covariance condition Eq. (3.34). Since W is nonunique, there may a priori be a representation $W'_{g\rho g^\dagger}$ with smaller 1-norm, and thus it holds that

$$\mathfrak{R}(g\rho g^\dagger) \leq \mathfrak{R}(\rho), \quad \forall\rho,\ \forall g \in \mathcal{C}\ell_n. \tag{3.35}$$

(b) Pauli measurements. We consider the measurement of a Pauli observable T_a on a quantum state ρ. Denote by $\rho_{a,s}$ the normalized post-measurement states for the outcomes $s = 0, 1$, respectively. We have to show that, for all n, for all $a \in \mathbb{Z}_2^n \times \mathbb{Z}_2^n$ and all n-qubit states ρ it holds that

$$p_a(0)\,\mathfrak{R}(\rho_{a,0}) + p_a(1)\,\mathfrak{R}(\rho_{a,1}) \leq \mathfrak{R}(\rho). \tag{3.36}$$

With Eq. (3.28), we can write $p_a(0)\,W_{\rho_{a,0}} = W_+ + \overline{W}_+$, and $p_a(1)\,W_{\rho_{a,1}} = W_- + \overline{W}_-$, where

$$\begin{aligned} W_+(\Omega',\gamma') &:= \textstyle\sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{W_\rho(\Omega,\gamma)}{2} \delta_{a\in\Omega}\delta_{\gamma(a),0} \left(\delta_{(\Omega',\gamma'),(\Omega,\gamma)} + \delta_{(\Omega',\gamma'),(\Omega,\gamma+[a,\cdot])}\right), \\ W_-(\Omega',\gamma') &:= \textstyle\sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{W_\rho(\Omega,\gamma)}{2} \delta_{a\in\Omega}\delta_{\gamma(a),1} \left(\delta_{(\Omega',\gamma'),(\Omega,\gamma)} + \delta_{(\Omega',\gamma'),(\Omega,\gamma+[a,\cdot])}\right), \\ \overline{W}_+(\Omega',\gamma') &:= \textstyle\sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{W_\rho(\Omega,\gamma)}{2} \overline{\delta}_{a\in\Omega}\delta_{(\Omega',\gamma'),(\Omega,\gamma\times(s=0))}, \\ \overline{W}_-(\Omega',\gamma') &:= \textstyle\sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{W_\rho(\Omega,\gamma)}{2} \overline{\delta}_{a\in\Omega}\delta_{(\Omega',\gamma'),(\Omega,\gamma\times(s=1))}. \end{aligned} \tag{3.37}$$

From now on, denote by W_ρ the optimal representation for ρ with respect to 1-norm, i.e., $\mathfrak{R}(\rho) = \|W_\rho\|_1$. With the triangle inequality, and the fact that the functions $W_{\rho_{a,s}}$ induced from the optimal W_ρ through Eq. (3.37) need not be optimal for the states $\rho_{a,s}$ with respect to their 1-norm, it holds that $p_{a,0}\,\mathfrak{R}(\rho_{a,0}) \leq \|W_+\|_1 + \left\|\overline{W}_+\right\|_1$, and $p_{a,1}\,\mathfrak{R}(\rho_{a,1}) \leq \|W_-\|_1 + \left\|\overline{W}_-\right\|_1$, hence

$$p_{a,0}\,\mathfrak{R}(\rho_{a,0}) + p_{a,1}\,\mathfrak{R}(\rho_{a,1}) \leq \|W_+\|_1 + \|W_-\|_1 + \left\|\overline{W}_+\right\|_1 + \left\|\overline{W}_-\right\|_1 . \tag{3.38}$$

With Eq. (3.37) we find that

$$\begin{aligned}\|W_+\|_1 + \|W_-\|_1 &= \sum_{(\Omega',\gamma')\in\mathcal{V}} \frac{\delta_{a\in\Omega'}}{2} |W_\rho(\Omega',\gamma') + W_\rho(\Omega',\gamma'+[a,\cdot])| \\ &\leq \sum_{(\Omega',\gamma')\in\mathcal{V}} \delta_{a\in\Omega'} |W_\rho(\Omega',\gamma')|,\end{aligned}$$

where in the second line we used the triangle inequality again. Furthermore, performing the summation over all $(\Omega',\gamma') \in \mathcal{V}$ first, we obtain

$$\left\|\overline{W}_+\right\|_1 = \left\|\overline{W}_-\right\|_1 = \sum_{(\Omega,\gamma)\in\mathcal{V}} \frac{\overline{\delta}_{a\in\Omega}}{2} |W_\rho(\Omega,\gamma)|.$$

Inserting the last two relations into Ineq. (3.38), we arrive at

$$p_{a,0}\,\mathfrak{R}(\rho_{a,0}) + p_{a,1}\,\mathfrak{R}(\rho_{a,1}) \leq \|W_\rho\|_1 .$$

Since $\mathfrak{R}(\rho) = \|W_\rho\|_1$ by assumption, Eq. (3.36) follows. □

3.6.4 Numerical results

In Table 3.3 and Figure 3.7 we present numerical values[4] for the robustness of various magic states, and compare them to robustness of magic as defined by Howard and Campbell [21]. Table 3.3 summarizes the robustness comparisons for the common magic states, as well as the maximal-robustness Hoggar state [21]. In Figure 3.7 we plot the robustness against the stabilizer state robustness for three qubits, as a function of rotation angle. Note the wide and almost flat—though not perfectly flat—plateaus of robustness $\mathfrak{R}$ in the vicinity of stabilizer states.

3.6.5 Curious resurgence of 4^n-dimensional phase space

Numerical calculations of robustness for various quantum states revealed an unexpected feature. Namely, the optimal quasiprobability distribution W_ρ with respect to Eq. (3.32) for a given n-qubit state ρ always was nonzero only on 4^n phase space points, or fewer. 4^n is only a tiny fraction of the whole phase space $\mathcal{V}$, and furthermore the naive expectation if one were completely oblivious of the differences between even and odd d. However, the

[4] Our calculations use the software packages CVXPY [153] and GUROBI [154].

state	$\mathfrak{R}$	$\mathfrak{R}_S$
$\|H\rangle^{\otimes 2}$	1.0	1.7472
$\|T\rangle^{\otimes 2}$	1.0	2.23205
$\|H\rangle^{\otimes 3}$	1.283	2.2189
$\|T\rangle^{\otimes 3}$	1.385	3.09807
$\|\text{Hoggar}\rangle$	1.80	3.8000

Table 3.3: Robustness values of selected magic states. For robustness of magic ($\mathfrak{R}_S$), also see [155].

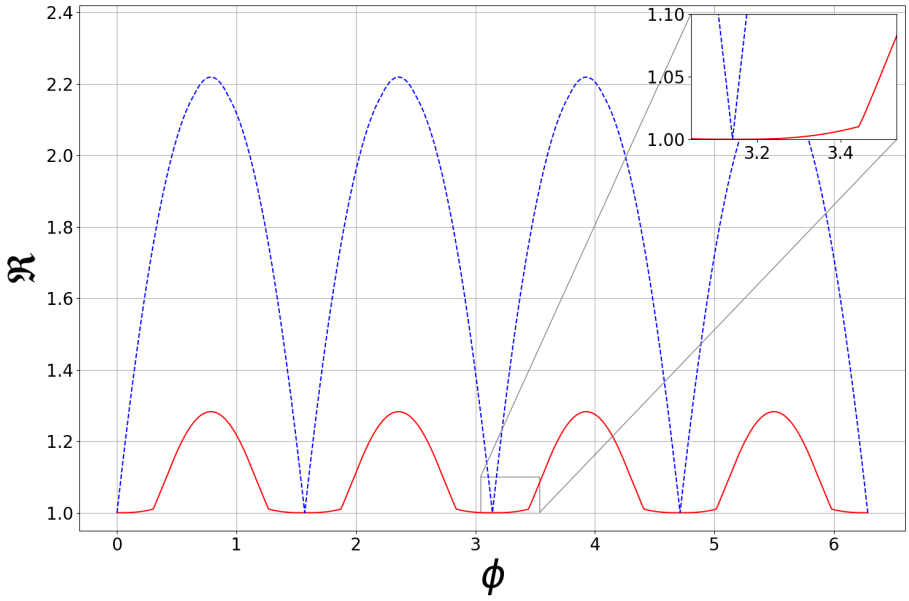


Figure 3.7: Robustness $\mathfrak{R}$ (solid line) and robustness of magic $\mathfrak{R}_S$ (dotted line) for the state $|H(\phi)\rangle^{\otimes 3}$, see Eq. (3.9), as a function of ϕ. Highlighted is the region near a stabilizer state, at $\phi = \pi$.

support of the optimal $W(\rho)$ depends on the state ρ. We can now explain the initially puzzling upper bound on the size of the support, 4^n.

The robustness $\mathfrak{R}$ of a state ρ defined in Eq. (3.32) is the solution to the convex optimization problem

$$\min_{\mathbf{q}} \{||\mathbf{q}||_1 : \ M\mathbf{q} = \mathbf{b}\}, \tag{3.39}$$

where $M_{i,j} = \text{Tr}(A_{\alpha_j} P_i)$, $b_i = \text{Tr}(\rho P_i)$, $\{\alpha_j : 1 \leq j \leq |\mathcal{V}|\}$ is an enumeration of the phase space points and P_i are the n-qubit Pauli operators.

For each variable q_j in Eq. (3.39), define two new variables $q_j^+ := \max(0, q_j)$ and $q_j^- := \max(0, -q_j)$. Then the convex optimization problem of Eq. (3.39) is equivalent to the

standard form linear program

$$\min_{\mathbf{q}} \left\{ \sum_j q_j^+ + q_j^- : \tilde{M}\tilde{\mathbf{q}} = \mathbf{b},\ \tilde{\mathbf{q}} \geq 0 \right\}, \tag{3.40}$$

where $\tilde{M} = \begin{bmatrix} M & -M \end{bmatrix}$ and $\tilde{\mathbf{q}} = \begin{bmatrix} (\mathbf{q}^+)^T & (\mathbf{q}^-)^T \end{bmatrix}^T$. This doubles the number of variables but does not change the number of equality constraints. Since we know this problem is feasible (any physical state can be written as an affine combination of phase space point operators) and bounded (no physical state can have robustness less than 1), by the fundamental theorem of linear programming, for any physical state, Eq. (3.40) has a solution at a vertex of the feasible polytope [156].

Since Eq. (3.40) has an equality constraint for each n-qubit Pauli operator (including the identity), this means any state ρ has a robustness-minimizing expansion in phase space point operators with no more than 4^n nonzero coefficients.

3.7 Discussion

3.7.1 Stratonovich-Weyl correspondence

In the field of quantum optics, what passes for a proper quasiprobability distribution over phase space is codified in the Stratonovich-Weyl (SW) correspondence. Denote by $F_A^{(s)} : X \to \mathbb{R}$ the quasiprobability distribution corresponding to the operator A, with X the phase space and s a real parameter in the interval $[-1, 1]$. In the standard formalism for infinite-dimensional Hilbert spaces, $s = -1, 0, 1$ correspond to the Glauber-Sudarshan P, Wigner, and Husimi Q function, respectively. The SW correspondence is the following set of criteria on the $F_A^{(s)}$ [48]; also see [119],

(0) Linearity: $A \mapsto F_A^{(s)}$ is a one-to-one linear map.

(1) Reality:
$$F_{A^\dagger}^{(s)}(u) = \left(F_A^{(s)}(u)\right)^*, \forall u \in X.$$

(2) Standardization:
$$\int_X d\mu(u) F_A^{(s)}(u) = \operatorname{Tr} A.$$

(3) Covariance:
$$F_{g\cdot A}^{(s)}(u) = F_A^{(s)}(g^{-1}u),\ g \in G,$$
with G the dynamical symmetry group.

(4) Traciality:
$$\int_X d\mu(u) F_A^{(s)}(u) F_B^{(-s)}(u) = \operatorname{Tr} AB.$$

We now investigate to which extent these SW criteria apply to the present quasiprobability function W. There are two deviations. First, the present quasiprobability function W

does not come with a parameter s; there is only a single function W. This will affect the formulation of traciality. Second, the present mapping $A \mapsto W_A$ is one-to-many, as we have noted in Section 3.2.1. The mapping is nonetheless linear, $A + B$ can be represented as $W_A + W_B$.

The remaining SW conditions do apply. (1) Reality of W follows directly from the definition Eq. (3.3), since all A^γ_Ω are Hermitian. (2) Standardization: The definition Eq. (3.1) and property Eq. (3.2) of the phase space point operators imply $\mathrm{Tr}\, A^\gamma_\Omega = 1$, for all Ω and γ; and standardization then follows from Eq. (3.3).

(3) Covariance holds for the entire Clifford group, as stated in Lemma 17. In fact, insisting on Clifford covariance leads to the nonuniqueness of W. Namely, an over-complete set of phase space point operators is necessary to achieve Clifford covariance [58].

(4) Traciality. In the absence of a continuously varying parameter s, we define a dual quasiprobability function $\tilde{W}$ in addition to W, to stand in for $F^{(-s)}$. For all Pauli operators T_a we have

$$\tilde{W}_{T_a}(\Omega, \gamma) := \begin{cases} (-1)^{\gamma(a)}, & a \in \Omega \\ 0, & a \notin \Omega \end{cases} .$$

Since the n-qubit Pauli operators form an operator basis, $\tilde{W}$ can be extended to all n qubit operators by linearity. With Eq. (3.29) we then have

$$\mathrm{Tr}(A\rho) = \sum_{(\Omega,\gamma)\in\mathcal{V}} \tilde{W}_A(\Omega, \gamma) W_\rho(\Omega, \gamma).$$

We thus satisfy the SW criteria (1)–(4).

To conclude, we reiterate that for the present purpose of classically simulating QCM, a crucial property of W is positivity preservation under Pauli measurement. This property has no counterpart in the Stratonovich-Weyl correspondence.

3.7.2 Interpretation as a probabilistic hidden variable model

In the case of odd d [19], there is a third equivalent indicator of classicality, next to positivity of the initial Wigner function and the efficiency of classical simulation of QCM by sampling. Namely, a positive Wigner function is equivalent to a noncontextual hidden variable model (HVM) with deterministic value assignments [20]. This triple coincidence cannot be replicated in $d = 2$, because, for $n \geq 2$ all quantum states—even the completely mixed state—are contextual [20].

One interpretation of this situation is that contextuality, i.e., the unviability of noncontextual HVMs, is not sufficiently tight a criterion to reveal genuine quantumness. A more stringent marker is required, which (i) classifies the present HVM as classical, and (ii) for QCM in odd d reduces to contextuality. At present, we have no suggestion for this more restrictive notion of quantumness. However, we point to a hidden variable model that is illustrative of the shifted quantum-to-classical boundary in the multiqubit case, and we propose it for further study.

Namely, when positive, the quasiprobability distribution W can be considered an HVM. While classified as contextual by the common definitions, it shares many features with ncHVMs and should be considered classical.

This HVM consists of a triple $(\Lambda, \{h^\lambda\}, p_\lambda)$ where $\Lambda = \mathcal{V}$ or $\mathcal{V}_M$, h^λ is a compatible family of distributions on the set of outcomes on contexts and p_λ is a probability distribution on the set Λ of hidden variables. For each $\alpha = (\Omega, \gamma)$ we define h^α by

$$h_I^\alpha(s) = \mathrm{Tr}(\Pi_s A_\Omega^\gamma). \tag{3.41}$$

Therein, I is any isotropic subspace, $s : I \to \mathbb{Z}_2$ is a function, and Π_s is the projector corresponding to the outcome. Note that $P_s = 0$ if $ds \neq \beta$.

It is useful to state the probability distributions $h_I^\alpha(\cdot)$ in their explicit form. Let Π_s denote the projector corresponding to the noncontextual value assignment $s : I \to \mathbb{Z}_2$. Then we have

$$h_I^{(\Omega,\gamma)}(s) = \frac{|I \cap \Omega|}{|I|} \delta_{s|_{I\cap\Omega}, \gamma|_{I\cap\Omega}}. \tag{3.42}$$

From Eq. (3.42) we see that the value assignments in our HVM are generally probabilistic; only in the special case of $I \subset \Omega$ they become deterministic. Further, the $\{h_I^\alpha\}$ form compatible families,

$$h_I^\alpha|_{I\cap I'} = h_{I'}^\alpha|_{I'\cap I}, \ \forall I, I', \ \forall \alpha \in \mathcal{V}.$$

When applicable, this HVM reproduces the predictions of quantum mechanics (see Theorem 4) for measurements of Pauli observables, in single contexts or arbitrary measurement sequences.

The HVM of Eq. (3.41) is (a) contextual in the sense of Abramsky and Brandenburger [68], (b) preparation and transformation contextual, as well as measurement-noncontextual, in the sense of Spekkens [128], and (c) contextual for sequences of transformations in the recently-introduced sense of Mansfield and Kashefi [130].

Its contextuality notwithstanding, we argue that the HVM of Eq. (3.41) should be considered classical. The reason is the following. The only resource the HVM of Eq. (3.41) uses beyond those required by noncontextual HVMs with deterministic value assignments is that of classical uniform randomness (in the evaluation of value assignments). Such use of classical randomness should not render the present HVM genuinely quantum.

Specifically, the HVM of Eq. (3.41) has the following properties: (I) It is conceptually very close to a noncontextual HVM with deterministic value assignments. The difference is that for every state of the HVM, only a subset Ω of the observables has deterministic values, and all other observables are evaluated by coin flipping. (II) It applies to all QCM on states with positive W. (III) It reduces to the standard noncontextual HVM with deterministic value assignments for QCM in odd d.

To summarize, we have described a hidden variable model corresponding to positive quasiprobabilites W. By this correspondence, the HVM is considered classical from the quantum optics perspective. It is also classical from the computational perspective, as it leads to efficient classical simulation of QCM (for applicable magic states). And yet this HVM is contextual. As such, it may serve as a reference point for a refined foundational notion of quantumness that goes beyond contextuality.

3.8 Conclusion

We have introduced a quasiprobability distribution W over generalized phase space, which is defined for any number n of qudits with any number d of levels. For multiqudit systems with odd local dimension d, W reduces to the familiar Wigner function for finite-dimensional systems defined by Gross [31]. For even d, the phase space is enlarged and W becomes nonunique. Importantly, also for $d = 2$ (the multiqubit case), W has the property that a positive quasiprobability function remains positive under all Pauli measurements. This property is crucial for classical simulation algorithms of quantum computation with magic states (QCM) by sampling.

Once this fundamental property is established, it is natural investigate the efficiency (or nonefficiency) of classical simulation in the various regimes, and resource theories characterizing QCM. Here we have treated the canonical questions that arise in this context: we have devised an efficient classical simulation of QCM for $W \geq 0$, and clarified the relation to the qubit stabilizer formalism. Namely, the present method for efficient classical simulation of QCM strictly contains the stabilizer method. It applies to all mixtures of stabilizer states, but in addition to certain states outside the stabilizer polytope. We have further characterized the hardness of classical simulation for $W < 0$ in terms of a robustness measure, and established this robustness is a monotone under the free operations of QCM.

In summary, we arrive at a resource perspective of QCM on qubits that closely resembles the corresponding picture for odd dimension d. However, there are two deviations. First, the phase space on which the quasiprobability function W is defined has a far more intricate structure for $d = 2$ than for odd d. Second, for $d = 2$ the hidden variable model (HVM) induced by any nonnegative quasiprobability function W_ρ is contextual, as a consequence of Mermin's square.

The latter observation leads to a puzzle. The HVM induced for positively representable states ρ is classified as "classical" from the perspectives of quantum optics ($W_\rho \geq 0$) and computer science (classical simulation is efficient), but it is classified as "quantum" from the perspective of contextuality.

In this regard, we have argued (also see [20]) that in multiqubit QCM, contextuality is not suitable as an indicator of genuine quantumness. We have proposed the notion of "HVM with partial noncontextual value assignments" in which classicality and contextuality coexist.

Chapter 4

Hidden variable model for quantum computation with magic states on qubits

We show that every quantum computation can be described by a probabilistic update of a probability distribution on a finite phase space. Negativity in a quasiprobability function is not required in states or operations. Our result is consistent with Gleason's Theorem and the Pusey-Barrett-Rudolph theorem.

This chapter has been published in Ref. [2].

4.1 Introduction

It is often pointed out that the fundamental objects in quantum mechanics are amplitudes, not probabilities [102, 157]. This fact notwithstanding, here we construct a description of universal quantum computation—and hence of all quantum mechanics in finite-dimensional Hilbert spaces—in terms of a probabilistic update of a probability distribution. In this formulation, quantum algorithms look structurally akin to classical diffusion problems.

While this seems implausible, there exists a well-known special instance of it: quantum computation with magic states (QCM) [38] on a single qubit. Compounding two standard one-qubit Wigner functions, a hidden variable model can be constructed in which every one-qubit quantum state is positively represented [136]. This representation is furthermore covariant under all one-qubit Clifford unitaries and "positivity preserving" under all one-qubit Pauli measurements. The update under such operations preserves the probabilistic character of the model, and hence QCM on one qubit can be classically simulated by a probabilistic update of a probability function on eight elements (see Figure 4.1 for illustration).

The prevailing view on the one-qubit example is that it is an exception and that for multiple qubits negativity will inevitably creep into any quasiprobability function of any computationally useful quantum state, rendering classical simulations inefficient [49]. This hypothesis is informed by the study of Wigner functions in finite-dimensional state spaces, which establishes Wigner function negativity as a necessary computational resource, i.e.,

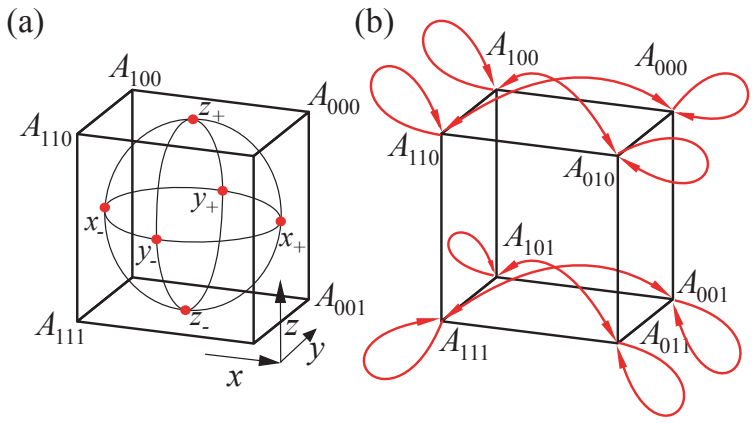


Figure 4.1: One-qubit model. (a) The state space Λ_1 is a cube with eight vertices corresponding to the phase space point operators $A_\alpha = [I + (-1)^{s_x} X + (-1)^{s_y} Y + (-1)^{s_z} Z]/2$, with $\alpha = (s_x, s_y, s_z) \in \mathbb{Z}_2^3$. The physical one-qubit states lie on or in the Bloch sphere which is contained in Λ_1 and touches the boundary of Λ_1 at six points corresponding to the six one-qubit stabilizer states. (b) Update of the phase space point operators A_α under measurement of the Pauli observable Z. Each red arrow represents a transition probability of $1/2$.

there can be no quantum speedup without negativity [1, 19, 21, 27, 28, 31, 51, 57, 58, 60, 61, 71, 137, 155, 158]. A quantum optics notion of quantumness—negativity of Wigner functions [44, 45]—and a computational notion—hardness of classical simulation—thus align.

The viewpoint just summarized requires correction. As we show in this chapter, the one-qubit case is not an exception; rather it is an example illustrating the general case. *Every* quantum state on any number of qubits can be represented by a probability function, and the update of this probability function under Pauli measurement is also probabilistic. This is the content of Theorem 6 below. We emphasize that the states and operations are *both* represented positively, not just one or the other. In Theorem 7, we apply this to quantum computation with magic states, showing that universal quantum computation can be classically simulated by the probabilistic update of a probability distribution.

This looks all very classical, and therein lies a puzzle. In fact, our Theorem 2 is running up against a number of no-go theorems: Theorem 2 in [121] and the Pusey-Barrett-Rudolph (PBR) theorem [159] say that probability representations for quantum mechanics do not exist, and [1, 19, 51, 57, 61] show that negativity in certain Wigner functions is a precondition for speedup in quantum computation. Further, does not Gleason's theorem prove that the proper representation of state in quantum mechanics is density matrices rather than probability distributions?

As we explain in the discussion part of this chapter, there is no contradiction with those works. Rather, the above-quoted theorems make stronger assumptions than we do and establish no-go theorems because of that. However, for describing universal quantum computation—hence all quantum mechanics in finite-dimensional Hilbert spaces—our weaker assumptions suffice.

Outline

The remainder of this chapter is organized as follows. First, in Section 4.2 we define our setting and state our main results, Theorems 6 and 7. Then we prove them in Section 4.3. We conclude with a discussion the above questions in Section 4.5.

4.2 Setting and Results

We focus on systems of n qubits for any $n \in \mathbb{N}$ and consider arbitrary quantum states evolving under sequences of Pauli measurements. The measurements need not commute, and the sequences may be arbitrarily long. This setting comprises universal quantum computation with magic states.

Below we devise a probability representation for this setting. The representation lives on a finite generalized phase space, and its purpose is to correctly reproduce the joint measurements statistics for all quantum states and all sequences of Pauli measurements.

The state space Λ_n of our probabilistic model is defined as follows. We denote by $\mathrm{Herm}_1(\mathcal{H}_n)$ the set of Hermitian operators on n-qubit Hilbert space $\mathcal{H}_n = \mathbb{C}^{2^n}$ with the property that $\mathrm{Tr}(X) = 1$ for all $X \in \mathrm{Herm}_1(\mathcal{H}_n)$, and by $\mathcal{S}_n$ the set of all n-qubit pure stabilizer states [12, 40, 41].

Then, we define the polytope Λ_n as

$$\Lambda_n := \{X \in \mathrm{Herm}_1(\mathcal{H}_n) \mid \mathrm{Tr}(|\sigma\rangle\langle\sigma| X) \geq 0,\ \forall\, |\sigma\rangle \in \mathcal{S}_n\} \tag{4.1}$$

(also see [72]). The elements $X \in \Lambda_n$ are the "states" (though not necessarily proper quantum states) that behave "well" under all sequences of Pauli measurements; namely, the probabilities for the outcome sequences are all nonnegative and add up to unity.

Λ_n is defined as the intersection of a finite number of halfspaces. Furthermore, it is bounded (see Section A.3.5). Therefore, by the Minkowski-Weyl theorem [160, 161], Λ_n can equivalently be described as the convex hull of finitely many extreme points (vertices). Denote by $\mathcal{A}_n$ the set of vertices of Λ_n, and the vertices by $A_\alpha \in \mathcal{A}_n$. These are our generalized phase space point operators, and the corresponding index set $\{\alpha\} =: \mathcal{V}_n$ is the generalized phase space.

We now have the following result.

Theorem 6. *For all numbers of qubits $n \in \mathbb{N}$, (i) each n-qubit quantum state ρ can be represented by a probability function $p_\rho : \mathcal{V}_n \longrightarrow \mathbb{R}_{\geq 0}$,*

$$\rho = \sum_{\alpha \in \mathcal{V}_n} p_\rho(\alpha)\, A_\alpha. \tag{4.2}$$

(ii) For the state update under Pauli measurements it holds that

$$\Pi_a^s A_\alpha \Pi_a^s = \sum_{\beta \in \mathcal{V}_n} q_{\alpha,a}(\beta, s)\, A_\beta. \tag{4.3}$$

For all $a \in E_n$, $\alpha \in \mathcal{V}_n$, the $q_{\alpha,a} : \mathcal{V}_n \times \mathbb{Z}_2 \to \mathbb{R}_{\geq 0}$ are probability functions,

(iii) Denote by $P_{\rho,a}(s)$ the probability of obtaining outcome s for a measurement of T_a on the state ρ. Then, the Born rule $P_{\rho,a}(s) = \mathrm{Tr}(\Pi_a^s \rho)$ takes the form

$$\mathrm{Tr}(\Pi_a^s \rho) = \sum_{\alpha \in \mathcal{V}_n} p_\rho(\alpha) Q_a(s|\alpha), \tag{4.4}$$

where $Q_a(s|\alpha)$ is given by

$$Q_a(s|\alpha) := \sum_{\beta \in \mathcal{V}_n} q_{\alpha,a}(\beta, s). \tag{4.5}$$

Hence $0 \leq Q_a(s|\alpha) \leq 1$, for all a, s, α.

The above theorem describes a hidden variable model (HVM) [96, 98, 99, 123, 126, 127, 162]. For any fixed number of qubits, any quantum state can be described by a probability function with finitely many elements. This property distinguishes it from the HVM of Beltrametti and Bugajski [162], which also applies to all quantum states but requires an infinite state space. A further distinguishing property is the probabilistic state update under a dynamical process, Pauli measurement. Theorem 6 is illustrated in Figure 4.1 through the example of a single qubit and in Section A.3.4 for two qubits.

Because of its capability to describe Pauli measurements, the above HVM has bearing on a model of universal quantum computation, namely quantum computation with magic states (QCM) [38]. QCM is closely related to the circuit model (see the SM, Sec. IV, for background). The difference is that in QCM the set of operations is restricted to a sequence of Clifford unitaries interspersed with Pauli measurements. These operations are applied to an initial "magic" state. The only property of QCM relevant for the present discussion is its quantum computational universality [12, 38, 106].

To apply the above probabilistic representation to QCM, we need to consider all its operational primitives—the Pauli measurements, the Clifford unitaries, and the magic states. Magic states (like all other quantum states) and Pauli measurements are positively represented by our HVM, cf. Theorem 6. This leaves the Clifford gates. The easiest way of dealing with them is to observe that they are redundant, i.e., no computational power is lost if we consider sequences of Pauli measurements only. The reason is that the Clifford unitaries may be propagated past all measurements, thereby conjugating the Pauli measurements into (other) Pauli measurements. After forward propagation, the unitaries can be dropped since they do not affect the statistics of the (now earlier) measurements (see e.g. [57, 61]).

With the Pauli measurements as the only essential dynamical element, QCM matches the setting described in Theorem 6. This leads to the following result.

1: Sample from the probability distribution p_ρ. Obtain a phase space point $\alpha_0 \in \mathcal{V}_n$.
2: For all Pauli measurements T_{a_t}, $a_t \in E_n$, from $t = 1$ to $t = t_{\max}$, sample from $q_{\alpha,a=a_t}$ to obtain the new phase space point $\beta \in \mathcal{V}_n$ and measurement outcome s. Output s as the outcome for the measurement of T_{a_t}, update the phase space point $\alpha_{t-1} \to \alpha_t = \beta$, and increment $t \to t + 1$.

Algorithm 3: Classical procedure to simulate a single run of a given QCM.

Theorem 7. *For any $n \in \mathbb{N}$ and all n-qubit quantum states ρ the classical algorithm of Algorithm 3 for sampling the outcomes of any sequence of Pauli measurements on ρ agrees with the predictions of quantum mechanics.*

Thus, the HVM of Theorem 6 describes all of universal quantum computation and hence arbitrarily closely approximates all quantum mechanical dynamics in finite-dimensional Hilbert spaces.

Theorem 7 does not imply that the classical simulation algorithm of Algorithm 3 is efficient. Intuition derived from previous classical simulation algorithms for quantum computation [1, 18, 19, 163] suggests that it is inefficient. However, at present we can neither prove the efficiency nor the inefficiency of this algorithm.

4.3 Proofs

We now turn to the proofs of Theorems 6 and 7. The proof of Theorem 6 requires a lemma.

Lemma 18. *The set Λ_n has the following properties.*

(1) Λ_n contains all n-qubit quantum states; i.e., for all n-qubit density operators ρ it holds that $\rho \in \Lambda_n$.

(2) Λ_n is closed under Pauli measurement, i.e., for all Π_a^s it holds that

$$X \in \Lambda_n \ \wedge \ \mathrm{Tr}(\Pi_a^s X) > 0 \Longrightarrow \frac{\Pi_a^s X \Pi_a^s}{\mathrm{Tr}(\Pi_a^s X)} \in \Lambda_n.$$

Proof of Lemma 18. All quantum states ρ satisfy the conditions $\mathrm{Tr}(|\sigma\rangle\langle\sigma|\,\rho) \geq 0$, for all n-qubit stabilizer states $|\sigma\rangle$ (as well as all other pure states), and $\mathrm{Tr}(\rho) = 1$; hence all quantum states ρ are in Λ_n.

Regarding Property 2, we observe that for all stabilizer states $|\sigma\rangle \in \mathcal{S}_n$ and all Pauli observables T_a it holds that

$$\Pi_a^s|\sigma\rangle\langle\sigma|\Pi_a^s = c\,|\sigma'\rangle\langle\sigma'|, \text{ where } |\sigma'\rangle \in \mathcal{S}_n, \ c \geq 0. \tag{4.6}$$

Namely, $c = 1$ if $(-1)^s T_a|\sigma\rangle = |\sigma\rangle$, $c = 0$ if $(-1)^s T_a|\sigma\rangle = -|\sigma\rangle$, and $c = 1/2$ otherwise [12]. Combining Eq. (4.6) and the definition of Λ_n, Eq. (4.1), $\mathrm{Tr}(|\sigma\rangle\langle\sigma|\,\Pi_a^s X \Pi_a^s) = \mathrm{Tr}[(\Pi_a^s\,|\sigma\rangle\langle\sigma|\,\Pi_a^s)X] = c\,\mathrm{Tr}(|\sigma'\rangle\langle\sigma'|\,X) \geq 0$. Therefore, whenever $\mathrm{Tr}(\Pi_a^s X) > 0$, the post-measurement state $X'_{a,s} := \Pi_a^s X \Pi_a^s / \mathrm{Tr}(\Pi_a^s X)$ also has the property that

$$\mathrm{Tr}\big(|\sigma\rangle\langle\sigma|X'_{a,s}\big) \geq 0, \ \ \forall a \in E_n, \forall s \in \mathbb{Z}_2, \forall |\sigma\rangle \in \mathcal{S}_n.$$

Furthermore, $\mathrm{Tr}\big(X'_{a,s}\big) = 1$. Therefore, $X'_{a,s} \in \Lambda_n$. □

Proof of Theorem 6. With Property 1 in Lemma 18, any n-qubit quantum state ρ is in Λ_n. Hence it can be expressed as a convex combination of the vertices A_α, as in Eq. (4.2). Taking the trace of Eq. (4.2) yields $\sum_\alpha p_\rho(\alpha) = 1$, i.e., p_ρ is a probability function. This proves the first statement of Theorem 6.

With Property 2 of Lemma 18, for all phase space point operators A_α and all projectors Π_a^s with $\mathrm{Tr}(\Pi_a^s A_\alpha) > 0$ it holds that $\Pi_a^s A_\alpha \Pi_a^s / \mathrm{Tr}(\Pi_a^s A_\alpha) \in \Lambda_n$. Therefore,

$$\Pi_a^s A_\alpha \Pi_a^s = \sum_{\beta \in \mathcal{V}_n} q_{\alpha,a}(\beta, s) A_\beta,$$

with $q_{\alpha,a}(\beta, s) \geq 0$ for all $\beta \in \mathcal{V}_n$, and $s \in \mathbb{Z}_2$. Now fixing α, a and adding the corresponding equations for $s = 0$ and $s = 1$, and then taking the trace, we find

$$\sum_{s \in \mathbb{Z}_2} \sum_{\beta \in \mathcal{V}_n} q_{\alpha,a}(\beta, s) = 1. \tag{4.7}$$

Hence, $q_{\alpha,a} : \mathcal{A}_n \times \mathbb{Z}_2 \to \mathbb{R}_{\geq 0}$ is a probability distribution for all $\alpha \in \mathcal{V}_n$, $a \in E_n$. This demonstrates Eq. (4.3).

Regarding $Q_a(s|\alpha)$ as defined in Eq. (4.5), since the $q_{\alpha,a}(\beta, s)$ are all positive, it holds that $Q_a(s|\alpha) \geq 0$ for all a, s, α. Furthermore, with Eq. (4.7) it follows that $Q_a(0|\alpha) + Q_a(1|\alpha) = 1$ for all a, α, and therefore

$$0 \leq Q_a(s|\alpha) \leq 1, \quad \forall a, s, \alpha.$$

Combining Eq. (4.2) and the already established Eq. (4.3),

$$\begin{aligned} \mathrm{Tr}(\Pi_a^s \rho) &= \sum_{\alpha \in \mathcal{V}_n} p_\rho(\alpha) \, \mathrm{Tr}(\Pi_a^s A_\alpha \Pi_a^s) \\ &= \sum_{\alpha \in \mathcal{V}_n} p_\rho(\alpha) \sum_{\beta \in \mathcal{V}_n} q_{\alpha,a}(\beta, s) \\ &= \sum_{\alpha \in \mathcal{V}_n} p_\rho(\alpha) Q_a(s|\alpha). \end{aligned}$$

This proves the formulation Eq. (4.4) of the Born rule. □

Proof of Theorem 7. Consider a Pauli measurement T_a on input state ρ. Using the classical simulation algorithm, the conditional probability of obtaining outcome s given the state $\alpha \in \mathcal{V}_n$ is given by Eq. (4.5). Therefore, the probability of obtaining outcome s given a measurement of T_a on state ρ as predicted by the classical simulation algorithm is

$$P_{\rho,a}^{(Sim)}(s) = \sum_{\alpha \in \mathcal{V}_n} p_\rho(\alpha) Q_a(s|\alpha). \tag{4.8}$$

The outcome probability predicted by the Born rule, $P_{\rho,a}^{(QM)}$, is given by Eq. (4.4).

Comparing Eq. (4.8) and Eq. (4.4), we see that the classical simulation algorithm reproduces the outcome probabilities predicted by the Born rule for a single Pauli measurement.

Now we turn to the post-measurement state ρ'. Quantum mechanics predicts it to be

$$\rho'^{(QM)} = \frac{\Pi_a^s \rho \Pi_a^s}{\mathrm{Tr}(\Pi_a^s \rho)}.$$

Here the numerator is

$$\begin{aligned}\Pi_a^s \rho \Pi_a^s &= \sum_{\alpha \in \mathcal{V}_n} p_\rho(\alpha) \Pi_a^s A_\alpha \Pi_a^s \\ &= \sum_{\alpha \in \mathcal{V}_n} p_\rho(\alpha) \sum_{\beta \in \mathcal{V}_n} q_{\alpha,a}(\beta, s) A_\beta,\end{aligned}$$

and so

$$\rho'^{(QM)} = \frac{\sum_\alpha p_\rho(\alpha) \sum_\beta q_{\alpha,a}(\beta, s) A_\beta}{\sum_\alpha p_\rho(\alpha) Q_a(s|\alpha)}. \tag{4.9}$$

Using the classical simulation algorithm, the probability of obtaining outcome s and state β given a measurement of T_a on state ρ is $P_{\rho,a}(\beta, s) = P_{\rho,a}(\beta|s) P_{\rho,a}(s)$. But $P_{\rho,a}(\beta, s) = \sum_\alpha p_\rho(\alpha) P_a(\beta, s|\alpha) = \sum_\alpha p_\rho(\alpha) q_{\alpha,a}(\beta, s)$ and $P_{\rho,a}(\beta|s) = p_{\rho'}(\beta)$. Therefore, the post-measurement state according to the classical simulation algorithm is

$$\begin{aligned}\rho'^{(Sim)} &= \sum_{\beta \in \mathcal{V}_n} p_{\rho'}(\beta) A_\beta = \sum_{\beta \in \mathcal{V}_n} \frac{P_{\rho,a}(\beta, s)}{P_{\rho,a(s)}} A_\beta \\ &= \sum_{\beta \in \mathcal{V}_n} \frac{\sum_\alpha p_\rho(\alpha) q_{\alpha,a}(\beta, s)}{\sum_\alpha p_\rho(\alpha) Q_a(s|\alpha)} A_\beta.\end{aligned} \tag{4.10}$$

This agrees with Eq. (4.9) above. Therefore, the classical simulation algorithm also reproduces the post-measurement state predicted by quantum mechanics for a single Pauli measurement.

Now let $\rho(t)$ denote the state before the tth measurement. Then the above shows that the classical simulation algorithm correctly reproduces the Born rule probabilities $P_{\rho,a_t}(s_t|s_1, s_2, \ldots, s_{t-1})$ as well as the post-measurement state $\rho(t+1)$. Therefore, by induction the simulation algorithm correctly reproduces the outcome probabilities predicted by the Born rule for any sequence of Pauli measurements. □

4.4 Multiqubit CNC phase space points are vertices of Λ

In the present work, there is no negativity anywhere in the classical simulation of QCM. The shifting of the cause for computational hardness away from negativity to other potential sources is a major disruption with the prior works [1, 19, 51, 57, 61].

But underneath this discontinuity lies an element of continuity. Namely, the direct precursor to the present work is Ref. [1]; and the phase space point operators of the multiqubit quasiproability function defined therein are also extremal vertices of the present state polytope Λ_n. This is the content of Lemma 19 below, the main result of this section. It shows that the multiqubit phase space defined in [1] is a subset of the phase space of the present model, describing a sector of it in which the update rules under Clifford unitaries and Pauli measurements are guaranteed to be computationally efficient.

Recall from the previous chapter a couple of definitions. We call a set $\Omega \subset E_n$ closed under inference if for all $a, b \in \Omega$ with the property that $[a, b] = 0$ it holds that $a + b \in \Omega$. We call a set $\Omega \subset E_n$ noncontextual if it supports a noncontextual value assignment. Sets Ω

which are both closed under inference and noncontextual are called "cnc" [1]. Of particular interest in are maximal cnc sets, which are cnc sets that are not strictly contained in any other cnc set. They give rise to the following multiqubit phase space point operators

$$A_\Omega^\gamma = \frac{1}{2^n} \sum_{a \in \Omega} (-1)^{\gamma(a)} T_a, \tag{4.11}$$

where Ω is a maximal cnc set, and $\gamma : \Omega \to \mathbb{Z}_2$ is a noncontextual value assignment.

We have the following result (also see [72] for an independent proof).

Lemma 19. *For any number n of qubits, the phase space point operators A_Ω^γ of Eq.* (4.11) *corresponding to maximal cnc sets Ω are vertices of Λ_n.*

The proof of this Lemma is given in Section B.3.1.

4.5 Discussion

We now return to the questions posed in the introduction.

(i) Existence of a probability representation. It is stated in Theorem 2 of [121] that "a quasiprobability representation of quantum theory must have negativity in either its representation of states or measurements (or both)."

This appears to contradict Theorem 6. However, there is no contradiction here, only a difference in assumptions. Through the definitions made prior to it, the above quoted theorem refers to frame representations. This requires, in particular, that the quasiprobability representation for every quantum state is unique. Clearly, our probability distribution p does not satisfy this condition.

(ii) Contextuality. Given the history of the subject [1, 19–21, 27, 28, 31, 51, 57, 61, 62, 71, 149, 155, 164], an interesting question is whether the present HVM is contextual or noncontextual. The Kochen-Specker notion of contextuality does not apply, because the present value assignments are not deterministic. This leaves us with Spekkens' notion [128, 165] to consider. In this regard, our HVM is preparation contextual and measurement-noncontextual. The former reflects the aforementioned nonuniqueness of p.

(iii) Negativity vs efficiency of simulation. Negativity in quasiprobability representations has been identified as a cause for slowing down the classical simulation of quantum systems by sampling. A general result has been obtained in [49] stating that a quantum system described by a quasiprobality function W with negativity $\mathcal{M} = \|W\|_1$ can be simulated by sampling at a multiplicative cost that scales like $\mathcal{M}^2$.

There are simulation schemes for QCM on qudits [19], on rebits [57], and on qubits [1, 21, 61], where negativity is the only source for the computational hardness of classical simulation. Negativity is therefore singled out as precondition for quantum speedup.

We do not contradict the results [1, 19, 21, 51, 57, 61] but now find that they are an artifact of the particular quasiprobability functions chosen. Our result lies at the opposite end of the spectrum. There is no negativity but, presumably, still computational hardness.

The absence of negativity notwithstanding, there also is continuity with prior works. The probability distribution p satisfies the four criteria of the Stratonovich-Weyl (SW) correspondence [48] (also see [119]; see Section A.3.2 for details). It is thus very closely related

to the original Wigner function [43] and to previously defined discrete Wigner functions for finite-dimensional systems. From the SW perspective, the only condition p doesn't satisfy is uniqueness.

Furthermore, the phase space point operators identified in the multiqubit setting of [1] (also see [149]) are special cases of the phase space point operators discussed here (see the SM, Section IV). And thus, the present approach provides a broader and yet conceptually simpler framework for the classical simulation of quantum computation by sampling, subsuming earlier ones as special cases.

(iv) The PBR theorem. The hidden variable model presented here is ψ-epistemic [166]. The PBR theorem [159] asserts that (with certain assumptions) no ψ-epistemic model can reproduce the predictions of quantum mechanics. Our result does not contradict the PBR theorem for two reasons. First, we consider only sequences of Pauli measurements rather than general measurements (this is sufficent for universal quantum computation). Second, our model does not satisfy the assumption of preparation independence required for the theorem to hold. That is, in general, $p_{\rho_1 \otimes \rho_2} \neq p_{\rho_1} \cdot p_{\rho_2}$.

The assumption of preparation independence is less relevant for quantum computation with magic states, where, in the language of resource theories, the free sector is formed by stabilizer states and stabilizer operations, not local states and local operations. Further, the memory lower bound of Karanjai, Wallman, and Bartlett [71] shows that a classical simulation algorithm like that of Algorithm 3 is incompatible with this assumption.

(v) Gleason's theorem. Gleason's theorem [167] says that in Hilbert spaces $\mathcal{H}$ of dimension 3 or greater, the only way to assign probabilities $p(h)$ to all subspaces of $h \subset \mathcal{H}$, represented by corresponding projectors Π_h, is via $p(h) = \mathrm{Tr}(\Pi_h \rho)$, for some valid density matrix ρ.

That is, the only consistent way to assign probabilities to measurement outcomes is the Born rule involving density matrices. Our Theorem 6 does not contradict this; rather it reproduces the Born rule, cf. Eq. (4.4).

However, Gleason's theorem is sometimes interpreted as a mathematical proof that density operators are the fundamental notion of state in quantum mechanics. In short, density operators are for quantum mechanics, probability distributions for classical statistical mechanics. Theorem 6 escapes this interpretation. It shows that every quantum state *can* be described by a probability distribution, and yet the Born rule for measurement is reproduced. This is possible because we have restricted measurement to Pauli observables. Note though that this restriction does not affect the universality of quantum computation with magic states!

To summarize, we have constructed a probability function over a finite set capable of positively representing all quantum states on any number of qubits, as well as their update under all Pauli measurements, local and nonlocal. All prior quasiprobability representations invoked in the discussion of quantum computation with magic states, such as the Wigner function for qudits [19, 31] or for rebits [57], and the quasiprobability over stabilizer states [21], require negativity to represent universal quantum computation.

In view of the seeming classicality of the hidden variable model for universal quantum computation constructed here, an important open question is: Where is quantumness hiding?—In this regard, we propose the polytopes Λ_n, and in particular the algebraic structure of their extremal vertices as a subject for further study.

Part II

Research Chapters

Chapter 5

The role of cohomology in quantum computation with magic states

A web of cohomological facts relates quantum error correction, measurement-based quantum computation, symmetry protected topological order and contextuality. Here we extend this web to quantum computation with magic states. In this computational scheme, the negativity of certain quasiprobability functions is an indicator for quantumness. However, when constructing quasiprobability functions to which this statement applies, a marked difference arises between the cases of even and odd local Hilbert space dimension. At a technical level, establishing negativity as an indicator of quantumness in quantum computation with magic states relies on two properties of the Wigner function: their covariance with respect to the Clifford group and positive representation of Pauli measurements. In odd dimension, Gross' Wigner function—an adaptation of the original Wigner function to odd-finite-dimensional Hilbert spaces—possesses these properties. In even dimension, Gross' Wigner function doesn't exist. Here we discuss the broader class of Wigner functions that, like Gross', are obtained from operator bases. We find that such Clifford-covariant Wigner functions do not exist in any even dimension, and furthermore, Pauli measurements cannot be positively represented by them in any even dimension whenever the number of qudits is $n \geq 2$. We establish that the obstructions to the existence of such Wigner functions are cohomological.

This chapter has been published in Ref. [4].

5.1 Introduction

Homology and cohomology find widespread uses in physics [64], starting with Gauss' theorem in electromagnetism. They have more recently entered the fields of foundations of quantum mechanics and quantum computation. Regarding the former, Kochen-Specker contextuality [127] has recently been given a cohomological underpinning [53, 68, 69]. Regarding the latter, Wigner function negativity—a traditional indicator of nonclassicality in quantum optics [45]—has been shown to be equivalent to contextuality in certain cases [50,

51, 168, 169], and both have been linked to the possibility of quantum computational advantage [19, 20, 27, 28, 146].

A short list of cohomological phenomena in the field of quantum computation is as follows: (i) quantum error correction with the Kitaev surface code [65, 66], and its measurement-based counterpart with 3D cluster states [67]; (ii) proofs of contextuality of quantum mechanics [53, 68, 69, 127], in particular parity-based proofs such as Mermin's square and star [62]; (iii) the contextuality of measurement-based quantum computation (MBQC) [22] in the limit of flat temporal order[1] [70]; (iv) computational phases of quantum matter [87, 89, 90, 170–174] which relate MBQC to symmetry protected topological order.

The above facts are not isolated but form an inter-related web. For example, fact (iii) on MBQC is a consequence of fact (ii) on parity-based contextuality proofs. Furthermore, recently, a connection between the phenomenologically rather distant fields of contextual MBQCs and computational phases of matter has been established [175]. Cohomology emerges as a language to navigate this web of fundamental facts about quantum computation, and to explain those facts in a unified fashion.

The purpose of this chapter is to extend this 'web of cohomology' to a further scheme of quantum computation, namely quantum computation with magic states (QCM) [38]. In QCM, the gate operations are reduced from a universal set to the so-called Clifford gates, which form a finite group for any number of qudits. By the Gottesman-Knill theorem, the reduced gate set alone cannot generate a quantum speedup. To retain the quantum computational power of the circuit model, the Clifford operations are supplemented by so-called 'magic' states.—Which properties must those magic states have to enable universality and quantum speedup?

It has been found that discrete Wigner functions and related quasiprobability functions are indicators of quantumness for QCM [19]. A quantum speedup can exist only if those functions take negative values on the magic states. We show that, depending on local Hilbert space dimension and number of qudits, whether or not a Wigner function exists that can serve as an indicator of quantumness in QCM is a question of cohomology.

For odd-dimensional qudits, the phenomenology is well understood: A precondition for quantum speedup is negativity in the Wigner function [43, 47] of the magic states [19, 32], as well as their contextuality [20]. At a technical level, the first result rests on two facts about Gross' Wigner function [29, 31], namely (a) that it is covariant under all Clifford transformations, and (b) that its positivity is preserved under all Pauli measurements. The second result rests on the first, and further the fact that (c) noncontextual value assignments for Pauli observables exist in odd dimension. Our discussion will reveal that the facts (a), (b), (c) are cohomological.

Our greater phenomenological interest is with the trickier case of *even* local dimension (qubits etc.); see [1, 2, 21, 28, 49, 52, 57, 58, 60, 61, 72, 149]. Do the above results on Wigner function negativity and state-dependent contextuality cary over?—Mermin's square prevents the latter, leaving the former for discussion. Indeed, it turns out that negativity as precondition for speedup can be reproduced in even dimension *if* one admits more general quasiprobability functions; specifically quasiprobability functions that do not stem from an operator basis and are not even unique [1, 21]. If one is not prepared to consider such

[1]The MBQC-contextuality connection itself is not restricted to flat temporal order [26].

generalizations, one is confronted with no-go theorems. For example, it is known that Wigner functions satisfying the assumption of diagram preservation cannot represent the stabilizer sub-theory of quantum mechanics positively [52]. Furthermore, for qubits, it has been shown that Clifford-covariant Wigner functions from operator bases do not exist [58].

In this chapter, we extend the known no-go results for Wigner functions derived from operator bases to all even dimensions, and also to the question of positivity preservation under Pauli measurement. The latter is central for QCM; see Section 5.2.2. Our technical contributions are two-fold: First, we formalize the obstructions to the existence of Wigner functions with the above "nice" properties (a) and (b). We demonstrate that these obstructions are cohomological in nature. Second, we apply these general results to the case of even dimension. We show that, in all even dimensions, Wigner functions constructed from operator bases cannot be Clifford covariant and cannot represent Pauli measurement positively. Our main results are stated as Theorems 10–14.

Outline

In Section 5.2 we provide the necessary background on the Pauli and Clifford groups, QCM, Wigner functions, cohomology, and contextuality. In Section 5.3 we define the Wigner functions of present interest. Section 5.4 is on the possibility of Clifford covariance of Wigner functions constructed from operator bases. Theorem 10 identifies a necessary and sufficient cohomological condition for the existence of Clifford-covariant Wigner functions, and Theorem 11 applies this condition to the case of even dimension. Section 5.5 discusses positive representation of Pauli measurement by Wigner functions that are constructed from operator bases and fulfill the Stratonovich-Weyl criteria [48, 119]. Theorem 13 provides a necessary and sufficient cohomological criterion for the existence of Wigner functions that represent Pauli measurement positively, and Theorem 14 applies this criterion to even dimension. Section 5.6 is the discussion.

5.2 Background

5.2.1 The Pauli group and the Clifford group

Pauli observables and Clifford unitaries are of central importance for this chapter. Here we provide the definitions for reference.

The Pauli group

In constructing the Pauli group and the Clifford group, there is a distinction between the odd- and the even-dimensional cases. Let $\mu = \omega$ and $Z_\mu = \mathbb{Z}_d$ ($\mu = \sqrt{\omega}$ and $Z_\mu = \mathbb{Z}_{2d}$) if d is odd (even). Pauli operators are defined by

$$T_a = \mu^{\gamma(a)} Z(a_z) X(a_x) \tag{5.1}$$

where $a = (a_z, a_x) \in \mathbb{Z}_d^n \times \mathbb{Z}_d^n =: E$, $Z(a_z) := \bigotimes_{k=1}^n Z^{a_z[k]}$, $X(a_x) := \bigotimes_{k=1}^n X^{a_x[k]}$, and $\gamma : E \to Z_\mu$ is a function chosen such that all operators T_a satisfy $(T_a)^d = \mathbb{1}$. For even d, the requirement that $(T_a)^d = \mathbb{1}$ restricts γ to $\gamma(a) \equiv \langle a_z | a_x \rangle \mod 2$.

The n-qudit Pauli group P_n is generated by the operators T_a where $a \in E$, yielding $P_n = \{\mu^\lambda T_a : \lambda \in Z_\mu, a \in E\}$.

Structural properties. We have the following structural results about the multiplication table of Eq. (2.6).

Observation 1 ([31])**.** *If the dimension d is odd, then for any number n of qudits the phases $\gamma(a)$ in Eq.* (5.1) *can be chosen such that $\beta \equiv 0$.*

The proof of Observation 1 is constructive; choose

$$\gamma(a) = -2^{-1} \langle a_z | a_x \rangle . \tag{5.2}$$

Among the phenomenological implications of Observation 1 are (i) the fact that there is no parity-based contextuality proof on Pauli observables [61, 146] (i.e., no counterpart to Mermin's square and star) when d is odd; and (ii) the multiqudit Wigner function defined in [31] is positivity-preserving under all Pauli measurements. Both properties are important for identifying Wigner function negativity and state-dependent contextuality as preconditions for quantum speedup in QCM, see [19],[20].

The even-dimensional counterpart of Observation 1 will be discussed in Section 5.5.3. For the special case of $d = 2$ the following is known.

Observation 2 ([62])**.** *If $d = 2$ then for all $n \geq 2$ and any choice of the function γ in Eq.* (5.1) *it holds that $\beta \not\equiv 0$.*

The Proof of Observation 2 is provided by Mermin's square [62].

The Clifford group

The n-qudit Clifford group $\mathcal{C}\ell_n$ is the normalizer of the Pauli group P_n in the unitary group $U(2^n)$, with the phases modded out,

$$\mathcal{C}\ell_n = N(P_n)/U(1). \tag{5.3}$$

Of central interest to us is how the Clifford group acts on Pauli observables by conjugation,

$$g(T_a) := g T_a g^\dagger = \omega^{\tilde{\Phi}_g(a)} T_{S_g a}, \;\; \forall g \in \mathcal{C}\ell_n, \; \forall a \in E, \tag{5.4}$$

where S_g is a symplectic transformation acting on E, i.e. an element of $\mathrm{Sp}(E)$. Here we are using two observations: (1) S_g is a group homomorphism $E \to E$ since g respects products of Pauli operators, i.e. $g(T_a T_b) = g(T_a) g(T_b)$, and (2) S_g is symplectic since g respects commutators in the sense that $[g(T_a), g(T_b)] = g([T_a, T_b])$. The Pauli group with the phases modded out, $\mathcal{P}_n := P_n/(U(1) \cap P_n) \cong \mathbb{Z}_d^n \times \mathbb{Z}_d^n$, is a normal subgroup of $\mathcal{C}\ell_n$, and it holds that $\mathrm{Sp}(E) \cong \mathcal{C}\ell_n/\mathcal{P}_n$.

The phase function $\tilde{\Phi}$ plays an important role in the subsequent discussion. $U(1)$-phases in $N(P_n)$ do not affect $\tilde{\Phi}$ in Eq. (5.4), which is why we may mod out those phases to begin with, cf. the definition of $\mathcal{C}\ell_n$ in Eq. (5.3).

5.2.2 A variant of QCM

We observe that in QCM, the Clifford unitaries can be eliminated without loss of computational power [57, 111]. Given the magic states, the computational power rests with the Pauli measurements.

This can be seen as follows. Consider the most general QCM, consisting of a sequence of Clifford gates interspersed with Pauli measurements, both potentially conditional on the outcomes of prior Pauli measurements. Now, starting with the last and ending with the first, the Clifford unitaries may be propagated forward in time, past the last Pauli measurement. Since the computation ends with the last measurement (all measurement outcomes have been gathered), after propagation the unitaries may be dropped without loss.

The only effect of the Clifford unitaries is that, in propagation, they change the measured observables by conjugation. But, by the very definition of the Clifford group, Pauli measurements remain Pauli measurements under conjugation by Clifford unitaries. Thus, for every QCM circuit consisting of Clifford unitaries and Pauli measurements, there is a computationally equivalent circuit that consists of Pauli measurements only.

This observation impacts the interpretation of the results of this chapter. Theorems 10 and 11 below deal with the question of when Clifford-covariant Wigner functions exist, and Theorems 13 and 14 with the question of when Wigner functions exist that represent Pauli measurements positively. With the observation just made, the latter two are more important for QCM. However, we still address Clifford covariance, as it has traditionally been invoked in the discussion of QCM [19], and as it is of general interest.

5.2.3 Cohomology

The purpose of the section is to explain that the function β defined in Eq. (2.6) and the phase function $\tilde{\Phi}$ defined in Eq. (2.9) are cohomological objects. As it will turn out, β governs the existence of Wigner functions that represent Pauli measurement positively, and $\tilde{\Phi}$ governs the existence of Wigner functions that are Clifford covariant. Thus, Clifford covariance of Wigner functions and positive representation of Pauli measurement by Wigner functions are cohomological properties of the Clifford and Pauli groups. Also see Ref. [69] for a more detailed exposition.

Motivation

Here we lay out a short path to recognizing β and $\tilde{\Phi}$ as cohomological objects. A problem gets us started. The Pauli operators contain in their definition an arbitrary phase γ, cf. Eq. (5.1), with no physical significance. Respecting the constraint on γ imposed by the condition $(T_a)^d = I$, $\forall a \in E$, the phases γ can be changed by any $\nu : E \to \mathbb{Z}_d$ as

$$\begin{array}{rcll} \gamma & \longmapsto & \gamma + \nu, & \text{if } d \text{ is odd,} \\ \gamma & \longmapsto & \gamma + 2\nu, & \text{if } d \text{ is even.} \end{array} \tag{5.5}$$

The corresponding effect on β is, irrespective of whether d is even or odd,

$$\beta(a,b) \longrightarrow \beta(a,b) + \nu(a) + \nu(b) - \nu(a+b) \mod d. \tag{5.6}$$

An object can be of physical significance only if it is invariant under the gauge transformations of Eq. (5.5). Can we construct such objects out of the function β?

The prepared eye recognizes Eq. (5.6) as a cohomological equivalence transformation $\beta \longrightarrow \beta + d\nu$, where $d\nu : (a, b) \mapsto \nu(a) + \nu(b) - \nu(a + b)$ is the coboundary of ν. This gives a clue: the cohomology classes $[\beta]$ are invariant objects!

Indeed, as we demonstrate in Section 5.5, those cohomology classes determine the existence of Wigner functions that represent Pauli measurement positively. To provide the foundation for the cohomological formulation, in Section 5.2.3 we describe the chain complex where β lives.

The motivation for the cohomological description of $\tilde{\Phi}$ is analogous. With Eq. (5.4), the effect of the equivalence transformation Eq. (5.5) on $\tilde{\Phi}$ is

$$\tilde{\Phi}_g(a) \longrightarrow \tilde{\Phi}_g(a) + \nu(a) - \nu(S_g a) \mod d. \tag{5.7}$$

Again, this looks like an equivalence transformation in group cohomology, $\tilde{\Phi} \longrightarrow \tilde{\Phi} - d^h\nu$. Therein, d^h is the coboundary operator in group cohomology. The bi-complex where $\tilde{\Phi}$ lives is introduced in Section 5.2.3. As we demonstrate in Section 5.4, the existence of Clifford-covariant Wigner functions hinges on a cohomological invariant extracted from $\tilde{\Phi}$.

β is a cocycle

Let $\mathcal{C}_* = (C_0, C_1, C_2, C_3)$ denote the chain complex for which C_k is defined to be the free $\mathbb{Z}_d$-module with basis $[v_1|v_2|\cdots|v_k]$, where $v_i \in E$ and

$$[v_i, v_j] = 0, \ \forall i, j = 1, \cdots, k. \tag{5.8}$$

Note that $C_0 = \mathbb{Z}_d$. The boundary map ∂ is given by the formula

$$\partial[v_1|v_2|\cdots|v_k] = [v_2|\cdots|v_k] + \left(\sum_{i=1}^{k-1}(-1)^i[v_1|\cdots|v_i + v_{i+1}|\cdots|v_k]\right) + (-1)^k[v_1|\cdots|v_{k-1}]. \tag{5.9}$$

There is also an associated cochain complex $\mathcal{C}^*$ whose k-cochains C^k are given by $\mathbb{Z}_d$-module maps (i.e. $\mathbb{Z}_d$-linear functions) $f : C_k \to \mathbb{Z}_d$. The coboundary map δ is defined by the formula $\delta f(-) = f(\partial -)$.

With the identification $\beta([a|b]) := \beta(a, b)$, the above definitions make β a 2-cochain in $\mathcal{C}^*$, i.e. $\beta \in C^2$. But we can say more; β is in fact a cocycle, $d\beta = 0$. Namely, associativity of operator multiplication, $T_a(T_bT_c) = (T_aT_b)T_c$, implies, for all $a, b, c \in E$ such that $[a, b] = [a, c] = [b, c] = 0$,

$$\beta(b, c) - \beta(a + b, c) + \beta(a, b + c) - \beta(a, b) = 0.$$

The surface $F := [b|c] - [a + b|c] + [a|b + c] - [a|b]$ is the boundary of the volume $V = [a|b|c]$, $F = \partial V$. Thus, $0 = \beta(\partial V) = d\beta(V)$, for all volumes V. Hence,

$$d\beta = 0, \tag{5.10}$$

as claimed. The equivalence class of β is now defined by

$$[\beta] := \{\beta + d\nu, \nu \in C^1\}. \tag{5.11}$$

The equivalence classes of 2-cocycles form the second cohomology group $H^2(\mathcal{C}, \mathbb{Z}_d)$. They are independent of the choice of phase γ in Eq. (5.1).

To provide a first application of the cohomological formulation, recall that in Section 5.2.1 we discussed whether the phase factors in the Pauli operator multiplication table of Eq. (2.6) can be eliminated by a clever choice of the phase convention γ. We now find that this is a topological question. Namely, with Eqs. (5.5) and (5.11), the phase factor ω^β can be removed if and only if $[\beta] = 0$. Observations 1 and 2 may thus be reformulated in a cohomological fashion as (i) If d is odd then for any n it holds that $[\beta] = 0$, and (ii) If $d = 2$ then for any $n \geq 2$ it holds that $[\beta] \neq 0$.

The group cocycles $\tilde{\Phi}$ and $\Phi_{\mathbf{cov}}$

In this section we will regard $\tilde{\Phi}$, introduced in Eq. (5.4), as a cocycle and relate its cohomology class to certain properties of Wigner functions. For this we need to introduce a chain complex which is slightly different than the one used above. The main difference is that we remove the commutativity constraint imposed on tuples constituting the basis of the chain complexes. We define a chain complex $\tilde{\mathcal{C}}_* = (\tilde{C}_0, \tilde{C}_1, \tilde{C}_2, \tilde{C}_3)$. Here $\tilde{C}_k$ is the free $\mathbb{Z}_d$-module with basis consisting of the tuples $[v_1|v_2|\cdots|v_k]$ where $v_i \in E$.

The boundary map is the same as in Eq. (5.9). The associated cochain complex is denoted by $\tilde{\mathcal{C}}^* = (\tilde{C}^0, \tilde{C}^1, \tilde{C}^2, \tilde{C}^3)$ and the coboundary map $\delta : \tilde{C}^k \to \tilde{C}^{k+1}$ is induced by the boundary map as before. Note that by definition $\tilde{C}_1 = C_1$ and $\tilde{C}^1 = C^1$, thus in this case we remove the extra decoration for simplicity of notation. We also note that the (co)chain complex defined here is the standard complex (which is called the bar construction) that computes the group (co)homology of the abelian group E.

For a symmetry group specified by a subgroup $G \subseteq \mathcal{C}\ell_n$, we consider the bicomplex $C^p(G, \tilde{C}^q)$; see also [69, Section 5.2]. The bicomplex $C^p(G, \tilde{C}^q)$ comes with two types of coboundaries: group cohomological $d^h : C^p(G, \tilde{C}^q) \to C^{p+1}(G, \tilde{C}^q)$, and $d^v : C^p(G, \tilde{C}^q) \to C^p(G, \tilde{C}^{q+1})$ induced by δ.

The phase function $\tilde{\Phi} : G \to C^1$ is per definition a 1-cochain in group cohomology. Its coboundary, a 2-cochain, is

$$(d^h\tilde{\Phi})_{g,h}(a) := \tilde{\Phi}_h(a) - \tilde{\Phi}_{gh}(a) + \tilde{\Phi}_g(S_h a). \tag{5.12}$$

In fact, $\tilde{\Phi}$ is not only a 1-cochain but a 1-cocycle, i.e., $d^h\tilde{\Phi} = 0$. Namely, with the associativity of matrix multiplication, it holds that $(gh)(T_a) = g(h(T_a))$, $\forall g, h \in G$, $\forall a \in E$. Evaluating both sides using Eq. (5.4) yields $\tilde{\Phi}_{gh}(a) = \tilde{\Phi}_h(a) + \tilde{\Phi}_g(S_h a)$, $\forall g, h \in G$ and $\forall a \in E$. Thus, with Eq. (5.12), $(d^h\tilde{\Phi})_{g,h}(a) = 0$, for all $g, h \in G$ and all $a \in E$. $\tilde{\Phi}$ is indeed a group cocycle.

This group cocycle may be trivial or nontrivial. Consider a 0-cochain ν in group cohomology. Its coboundary, a 1-cocycle, is

$$(d^h\nu)_g(a) := \nu(S_g a) - \nu(a). \tag{5.13}$$

We define the group cohomology classes

$$[\tilde{\Phi}] = \{\tilde{\Phi} + d^h\nu,\ \forall \nu \in C^1\}.$$

A cocycle $\tilde{\Phi}$ is trivial iff it can be written in the form $\tilde{\Phi} = d^h\nu$, for some $\nu \in C^1$. A class $[\tilde{\Phi}]$ is trivial, $[\tilde{\Phi}] = 0$, if and only if it contains $\tilde{\Phi} \equiv 0$.

The cocycle $\tilde{\Phi}$ may not only be evaluated on edges $a \in E$ but, by linear extension, on all 1-chains: A 1-chain is a $\mathbb{Z}_d$-linear sum of elements in E and we define

$$\tilde{\Phi}_g\left(\sum_j \alpha_j[a_j]\right) = \sum_j \alpha_j \tilde{\Phi}_g(a_j), \quad \alpha_j \in \mathbb{Z}_d, a_j \in E,$$

where each $\tilde{\Phi}_g(a_j)$ is determined by Eq. (5.4). Subsequently, we will have occasion to evaluate $\tilde{\Phi}$ on boundaries ∂f, for $f \in C_2$. When $G = \mathcal{C}\ell_n$ it is easily verified that, for boundaries ∂f,

$$\tilde{\Phi}_g(\partial f) = \tilde{\Phi}_{[g]}(\partial f),\ \forall g \in \mathcal{C}\ell_n,$$

where $[g] \in \mathcal{C}\ell_n/\mathcal{P}_n$. That is, when $\tilde{\Phi}$ is evaluated on a boundary, it depends on its first argument g only through the equivalence class $[g]$.

To formalize this property in general, let $N \subset G$ denote the subgroup of symmetries g such that S_g is the identity transformation. The quotient group $Q = G/N$ is the essential part of the symmetries acting on the complex. Let $\tilde{B}_1$ denote the image of the boundary map $\partial : \tilde{C}_2 \to C_1$. We write U_{cov} for the set of $\mathbb{Z}_d$-module maps $\tilde{B}_1 \to \mathbb{Z}_d$. We choose a set-theoretic section $\theta : Q \to G$ of the quotient map $\pi : G \to Q$; i. e., $\theta(\pi(g)) \in \pi(g)$. Then $\Phi_{\text{cov}} \in C^1(Q, U_{\text{cov}})$ is defined to be the composite

$$\Phi_{\text{cov}} : Q \xrightarrow{\theta} G \xrightarrow{\tilde{\Phi}} C^1 \xrightarrow{d^v} U_{\text{cov}} \tag{5.14}$$

where the last map can also be thought of as the restriction of a $\mathbb{Z}_d$-module map $C_1 \to \mathbb{Z}_d$ to the boundaries $\tilde{B}_1$. More explicitly we have

$$\Phi_{\text{cov}}(q, \partial f) = d^v\tilde{\Phi}_{\theta(q)}(f) = \tilde{\Phi}_{\theta(q)}(\partial f)$$

for any $q \in Q$ and $f \in \tilde{C}_2$. Although $\tilde{C}^1 = C^1$, the object U_{cov} is different from its counterpart U introduced in [69, Eq. (39)] as part of the symmetry discussion in contextuality. See also Theorem 9 below.

Like $\tilde{\Phi}$, Φ_{cov} is also a group cocycle, $d^h\Phi_{\text{cov}} = 0$. The cocycle class of Φ_{cov} is given by

$$[\Phi_{\text{cov}}] = \{\Phi_{\text{cov}} + d^h\nu,\ \nu \in U_{\text{cov}}\}.$$

$[\Phi_{\text{cov}}]$ is the object of interest for Wigner function covariance.

As a first application of the group cohomological formalism, recall Theorem 37 from [31], on the structure of the Clifford group in odd dimension. Items (i) and (iii) of that theorem read: (i) For any symplectic S, there is a unitary operator $\theta(S)$ such that $\theta(S)T_a\theta(S)^\dagger = T_{Sa}$; (iii) Up to a phase, any Clifford operation is of the form $U = T_b\,\theta(S)$, for a suitable $b \in E$ and symplectic transformation S.

Comparing the relation in item (i) with the general relation Eq. (5.4), we find that $\tilde{\Phi}_{\theta(S)} = 0$ for all symplectic S, with respect to the phase convention γ chosen in [31]. Thus, in particular, $\tilde{\Phi}_{\theta(S)}(\partial f) = 0$ for all $f \in \tilde{C}_2$. Since Pauli flips don't change $\tilde{\Phi}$ on boundaries, $\tilde{\Phi}_{T_b\,\theta(S)}(\partial f) = 0$ for all $f \in \tilde{C}_2$ and all $b \in E$. But with item (iii) this covers the entire Clifford group. I.e., there is a phase convention γ such that $\Phi_{\text{cov}} \equiv 0$. The phase-convention independent version of this statement is

Observation 3. *For the Clifford group $\mathcal{C}\ell_n$ in any odd dimension d it holds that $[\Phi_{\text{cov}}] = 0$.*

Splitting and the group cocycle ζ

For a subgroup $G \subset \mathcal{C}\ell_n$ define $N = G \cap \mathcal{P}_n$ as before. Since N is a normal subgroup we can consider the quotient group $Q = G/N$. A structural question about G is how it is put back together from its two parts N and Q. A particularly simple composition is the semi-direct product, $G \cong Q \ltimes N$. If it applies, then G is said to 'split'. But the semi-direct product is only one among a number of ways to compose Q and N. Those ways are classified by a cohomology group [176].

The notion of splitting is of interest here because, as we shall prove in Section 5.4.4, it governs the existence of the faithful group action which is required in Eq. (2.26) and is, thus, a precondition for covariance.

Let us formulate splitting as a cohomological property. We identify Q with the subgroup of $\mathrm{Sp}(E)$ to which it is isomorphic by Eq. (5.4). For convenience, we denote T_a modulo phase by t_a, such that $\mathcal{P}_n = \{t_a, a \in E\} \cong E$. Any set-theoretic section $\theta : Q \to G$ has the following properties:

$$\theta(S) t_a (\theta(S))^\dagger = t_{Sa},\ \forall t_a \in E, \tag{5.15a}$$

$$g = t_{\alpha_g} \theta(S_g),\ \forall g \in G, \tag{5.15b}$$

$$\theta(S_1)\theta(S_2) = t_{\zeta(S_1,S_2)} \theta(S_1 S_2),\ \zeta : Q \times Q \longrightarrow E \tag{5.15c}$$

Herein $t_{\alpha_g}, t_{\zeta(S_1,S_2)} \in N$. The section θ is not unique. From any given θ we may switch to new θ' via

$$\theta'(S) = t(S)\theta(S),\ t(S) \in N,\ \forall S \in Q. \tag{5.16}$$

The function ζ changes under the transformation in Eq. (5.16), and the set $[\zeta]$ of functions ζ' that can be reached from ζ via Eq. (5.16) is an element of the second cohomology group $H^2(Q, E)$. The group G splits if and only if ζ vanishes for a suitable choice of θ. In this case we say that $[\zeta] = 0 \in H^2(Q, E)$. See Eq. (A.17) in the appendix for a particular choice of the cocycle ζ.

5.2.4 Contextuality

Unlike the previous parts of this background section, the material on contextuality discussed here is not necessary to understand the main results of this chapter, Theorems 10–14. However, it is helpful for connecting to the broader picture. Namely, two previously established theorems on state-independent contextuality [69], restated below as Theorems 8 and 9, are structurally akin to Theorems 10 and 13, on the existence of Wigner functions with "nice"

properties. They invoke the same cohomological conditions. This background portion prepares for the discussion in Section 5.6.

Contextuality is a foundational property that distinguishes quantum mechanics from classical physics. A priori, one may attempt to describe quantum phenomena by so-called hidden variable models (HVMs) in which all observables have predetermined outcomes that are merely revealed upon measurement. A probability distribution over such predetermined outcomes is then intended to mimic the randomness of quantum measurement. An additional constraint on HVMs is the assumption of noncontextuality: the value assigned to any given observable just depends on that particular observable, and not on any compatible observable that may be measured in conjunction. The Kochen-Specker theorem says that, in Hilbert spaces of dimension greater than two, no noncontextual hidden variable model can reproduce the predictions of quantum mechanics. Since noncontextual hidden variable models fail in this realm, quantum mechanics is said to be "contextual".

The original proof of the Kochen-Specker theorem is intricate. However, when sacrificing a modest amount of generality, namely the case of Hilbert space dimension 3, a very simple proof can be given—Mermin's square [62]. Here we review two types of proofs of the Kochen-Specker theorem, the parity-based proofs (Mermin's square is the simplest example), and the symmetry-based proofs.

Parity-based proofs. The main examples of parity-based proofs are the well-known Mermin's square in dimension 4 and Mermin's star in dimension 8 [62]. In those examples, the proof is based on a cleverly chosen set of Pauli observables. However, for parity proofs to work, the observables don't need to be of Pauli type; it suffices that all their eigenvalues are kth roots of unity for some fixed $k \in \mathbb{N}$.

Parity proofs have a cohomological formulation. With the cocycle β and corresponding cohomology class $[\beta]$ in $H^2(\mathcal{C})$ defined as in Section 5.2.3, but this time for a set $\mathcal{O}$ of observables whose eigenvalues are all powers of ω (not necessarily, but possibly, Pauli observables), we have the following result.

Theorem 8 ([69])**.** *For a set of observables $\mathcal{O}$ with all eigenvalues of form $e^{i2\pi\, m/k}$, for $m, k \in \mathbb{N}$ and k fixed, a parity-based contextuality proof exists if and only if $[\beta] \neq 0$.*

Symmetry-based proofs. Proofs of the Kochen-Specker theorem may also be based on the transformation behavior of a set $\mathcal{O}$ of observables under a symmetry group G [69]. To be a symmetry group, G (i) needs to map the set $\mathcal{O}$ to itself up to phases that preserve the constraint on the eigenvalue spectrum, and (ii) needs to preserve algebraic relations among the transformed observables. Again, the symmetry-based contextuality proofs have a cohomological formulation.

Theorem 9 ([69])**.** *For a given set $\mathcal{O}$ of observables as above, and a corresponding symmetry group G, if $[\Phi] \neq 0 \in H^1(Q, U)$ then $\mathcal{O}$ exhibits state-independent contextuality.*

There is a small difference between Φ in Theorem 9 and Φ_{cov} from Section 5.2.3. Both phase functions are defined only on boundaries ∂f where $f = [a|b]$. However, for Φ all f are constraint to $[a, b] = 0$, whereas for Φ_{cov} as defined in Section 5.2.3 this extra condition is not imposed.

5.3 Wigner functions from operator bases

In this section we define the Wigner functions we are concerned with in this chapter, namely Wigner functions based on operator bases, and derive elementary properties of them. Simultaneously, we also define matching effect functions. This lays the groundwork for Sections 5.4 and 5.5, where we discuss Clifford covariance of Wigner functions and positive representation of Pauli measurement, respectively.

Quasiprobability functions derived from operator bases are a natural choice: the original Wigner function [43], its finite-dimensional adaptions [47], and thereof in particular Gross' Wigner function in odd dimension [29, 31] are all of this form[2].

We remark that, for even local dimension $d = 2$, if we move beyond operator bases and admit nonunique probability functions defined on larger (generalized) phase spaces then Theorem 1 is known to extend [1, 21]. However, moving into this territory, we also meet *probability* functions that represent universal QCM [2]. No negativity is needed, neither in the states nor the operations! For those scenarios, since negativity never occurs, it cannot be a precondition for speedup.

Quasiprobability functions derived from operator bases represent a calmer and more standard scenario, and therefore we settle their case here. For the remainder of this chapter, we impose the condition

(OB) For any linear operator Y, a corresponding Wigner function W_Y satisfies

$$Y = \sum_{v \in V} W_Y(v) A_v, \tag{5.17}$$

where the operators $\{A_v,\, v \in V\}$ form an operator basis and the phase space is $V = \mathbb{Z}_d^n \times \mathbb{Z}_d^n$.

As before in Section 2.4, in addition to the quasiprobability function W representing quantum states, we define a dual object, the effect function Θ, representing observables. For any linear operator Y, $\Theta_Y : V \longrightarrow \mathbb{C}$.

E vs. V. We briefly comment on the distinction between the sets E of Pauli labels and the phase space V. Both are isomorphic to the group $\mathbb{Z}_d^n \times \mathbb{Z}_d^n$, but they represent different physical objects. E, as introduced right after Eq. (5.1), is the set of Pauli operators modulo phase. The multiplication of Pauli operators induces an addition in E, endowing it with a group structure. V, introduced right above through Eq. (5.17), is the set of points in phase space. Once an origin is fixed, points in phase space may also be added; but in the present chapter we make no use of this operation. Finally, the property (SW3) of Pauli covariance implies an action of E on V,

$$E \ni a : v \mapsto a(v) = (a + v) \in V, \;\; \forall v \in V. \tag{5.18}$$

Some of our results refer to the SW criteria, in the following formulation adapted to finite-dimensional Hilbert spaces:

[2]Note that in the infinite-dimensional case of the original Wigner function, the set of phase space point operators is not a basis in the standard sense, i.e., it is not a Hamel basis, but it is a Hilbert basis for the ambient space that the density operators live in viewed as an inner product space (see e.g. Ref. [177, §9] for details on this distinction). In finite dimensions there is no such distinction.

(SW1) (Reality):

$$W_{Y^\dagger}(u) = (W_Y(u))^* \quad \forall u \in V,$$

(SW2) (Standardization):

$$\sum_{u \in V} W_Y(u) = \mathrm{Tr}(Y),$$

(SW3) (Pauli covariance):

$$W_{T_a(Y)}(u + a) = W_Y(u) \quad \forall u \in V \; \forall a \in E,$$

(SW4) (Traciality):

$$\sum_{u \in V} W_{Y_1}(u) \Theta_{Y_2}(u) = \mathrm{Tr}(Y_1 Y_2).$$

Condition (OB) implies (sw0) but is not implied by it. (SW2) and (SW4) are obtained from (sw2) and (sw4) by choosing a natural measure μ. As opposed to (sw3), (SW3) refers to a specific symmetry group G—the Pauli group—and a particular action of the symmetry group on the phase space.

Among the Wigner functions permitted by the constraints (OB) and (SW1)–(SW4), we are interested in those that transform covariantly under all Clifford gates and represent Pauli measurement positively. Note that we do not require diagram preservation—the assumption that the representation of a composition of processes is identical to the composition of the representations of the processes [135]. Full diagram preservation requires this for both parallel and sequential composition of states and operations. In particular, for parallel composition of states, diagram preservation requires that for a product state $\rho \otimes \sigma$ we have $W_{\rho\otimes\sigma}(u \oplus v) = W_\rho(u) W_\sigma(v)$. This is a strong assumption which, though reasonable for some local realist models, is not satisfied by several models which are useful for describing quantum computation, e.g. [1, 2, 7, 21, 151]. In fact, this assumption is incompatible with the ability of a model to simulate contextuality as in quantum computation on multiple qubits [71]. The conditions imposed on quasiprobability distributions in [135] include diagram preservation and imply (OB), (SW1), (SW2), and (SW4) [52, 132]. If a quasiprobability distribution from [135] positively represents the entire stabilizer subtheory, it additionally fulfills (SW3) [52] and, thus, all of our conditions.

Regarding the effect function Θ, we observe that all such functions admissible by (OB) and (SW4) take a simple form,

$$\Theta_Y(v) = \mathrm{Tr}(Y A_v), \;\; \forall v \in V. \tag{5.19}$$

Namely, for $Y = A_u$ in Eq. (5.17) the condition (OB) implies that $W_{A_u}(v) = \delta_{uv}$. Using this in the traciality condition (SW4) for $Y_1 = A_v$ implies Eq. (5.19). Note furthermore that Eq. (5.19) implies (SW4).

We further observe that for the projectors Π_a^s corresponding to the measurement of the Pauli observable T_a with outcome s, for all Θ, W satisfying (OB), (SW2), (SW4), it holds that

$$\sum_s \Theta_{\Pi_a^s}(v) = 1, \;\; \forall v \in V. \tag{5.20}$$

Namely, $\sum_s \Theta_{\Pi_a^s}(v) = \sum_s \mathrm{Tr}(\Pi_a^s A_v) = \mathrm{Tr}(A_v)$. The last expression equals $\sum_{u\in V} W_{A_v}(u) = \sum_{u\in V} \delta_{uv} = 1$, with (SW2), yielding Eq. (5.20).

To prepare for the subsequent discussion, we parametrize the phase space point operators. Pauli covariance (SW3) holds if and only if

$$A_v = \frac{1}{d^n} \sum_b \omega^{-[v,b]} c_b T_b^\dagger, \ \forall v \in V, \tag{5.21}$$

where $c_b \in \mathbb{C}$ for all b, $V = \mathbb{Z}_d^n \times \mathbb{Z}_d^n$.

Additionally imposing Standardization (SW2) and the fact that the operators A_v span an operator basis (OB) is equivalent to the following two conditions on the coefficients c_b in Eq (5.21), respectively:

$$c_0 = \mu^{\gamma(0)}, \tag{5.22a}$$

$$c_b \neq 0, \ \forall b. \tag{5.22b}$$

With these properties established, we are now ready to address the questions of Clifford covariance and positive representation of Pauli measurement.

5.4 Clifford covariance

In this section we establish a necessary and sufficient condition for the existence of Clifford covariance of Wigner functions constructed from operator bases, for varying local Hilbert space dimension and number of particles. There are cohomological obstructions to the existence of such Wigner functions, and in even local dimension these obstructions don't vanish.

5.4.1 When are Wigner functions Clifford covariant?

We start out from the following

Definition 9. *For any subgroup $G \subseteq \mathcal{C}\ell_n$, a Wigner function W is called G-covariant if*

$$W_{g(Y)}(S_g u + a_g) = W_Y(u) \quad \forall u \in V \ \forall g \in G \tag{5.23}$$

where S_g is the symplectic transformation defined in Eq. (5.4) and $a_g \in E$. When $G = \mathcal{C}\ell_n$, we say that W is Clifford covariant.

This definition, in the most general case of full Clifford covariance, is analogous to the one given in [31] for odd dimension d; cf. Theorem 41 therein. From the perspective of classical simulation of QCM, the usefulness of the above covariance property is that the map $g: \ u \mapsto S_g u + a_g$ can be efficiently computed, for all $g \in \mathcal{C}\ell_n$. We remark that the addition is between an element of V and one of E, cf. Eq. (5.18). For consistency, we observe that in the limiting case of $G = P_n$, $S_g = I$ for all $g \propto T_u \in P_n$, and Eq. (5.23) reduces to the Pauli covariance condition (SW3).

It is useful to restate the covariance condition Eq. (5.23) in terms of phase space point operators. With Eq. (5.17) it holds that $W_{A_w} = \delta_w(v)$. Now, for any $g \in G$ we have

$$\begin{aligned} g(A_w) &= \textstyle\sum_v W_{g(A_w)}(v) A_v \\ &= \textstyle\sum_v W_{g(A_w)}(S_g v + a_g) A_{S_g v + a_g} \\ &= \textstyle\sum_v W_{A_w}(v) A_{S_g v + a_g} \\ &= \textstyle\sum_v \delta_w(v) A_{S_g v + a_g} \\ &= A_{S_g w + a_g} \end{aligned} \tag{5.24}$$

Therein, in the second line we have relabeled the summation index, and in the third line we have used the covariance condition Eq. (5.23). Thus, Eq. (5.23) implies Eq. (5.24).

Now we show the reverse. Assuming Eq. (5.24), the operator $g(Y)$ can be expanded in two ways

$$g(Y) = \sum_v W_Y(v) g(A_v) = \sum_v W_Y(v) A_{S_g v + a_g},$$

and

$$g(Y) = \sum_v W_{g(Y)}(v) A_v = \sum_v W_{g(Y)}(S_g v + a_g) A_{S_g v + a_g}$$

Comparing the r.h.s.-es, and noting that by (OB) the expansions are unique, Eq. (5.23) follows.

Both arguments combined show that a Wigner function W satisfying (OB) is G-covariant if and only if

$$g(A_v) = A_{S_g v + a_g}, \quad \forall v \in V, \ \forall g \in G. \tag{5.25}$$

5.4.2 Existence of Clifford-covariant Wigner functions

Given a particular Wigner function in terms of phase space point operators, Clifford covariance may be verified using the criterion Eq. (5.25). Now we are concerned with the question of whether for a system of n qudits of d levels a Clifford-covariant Wigner function *exists*. The result of this section is a cohomological criterion for the existence of a Clifford-covariant Wigner function.

Theorem 10. *For any integers $n \geq 1$, $d \geq 2$, a Clifford-covariant Wigner function according to (OB) exists if and only if $[\Phi_{\mathrm{cov}}] = 0 \in H^1(Q, U_{\mathrm{cov}})$.*

Proof of Theorem 10. "Only if": Suppose W is Clifford covariant. Then, with Eq. (5.25), $T_a(A_v) = A_{v+y(a)}$, $\forall a \in E$, where for readability we write $y(a)$ instead of a_g with $g = T_a$. First, we want to show that the function $y : E \cong \mathbb{Z}_d^{2n} \to V \cong \mathbb{Z}_d^{2n}$ is invertible. $T_a T_b(A_v) = T_{a+b}(A_v)$ implies that y is linear. Therefore, y is invertible if and only if its kernel is trivial, $\mathrm{Ker}(y) = \{0\}$. We expand $A_v = \sum_{b \in E} c_b(v) T_b^\dagger$. Note that, in contrast to Ansatz (5.21), this expansion does not restrict A_v. Now assume that $\mathrm{Ker}(y) \ni a \neq 0$. Then $\sum_b c_b(v) T_b^\dagger = A_v = T_a(A_v) = \sum_b c_b(v) \omega^{[b,a]} T_b^\dagger \ \forall v$ or, equivalently,

$$c_b(v) = 0 \ \ \forall b \in E : [b, a] \neq 0, \ \forall v \in V.$$

For any $a \neq 0$ there is a $b \in E$ such that $[b, a] \neq 0$, and hence a $b \in E$ such that $c_b(v) = 0$,

$\forall v$. Since this contradicts (OB), y must be invertible. We may thus write

$$T_{y^{-1}(w)}(A_0) = A_w,\ \forall w \in V. \tag{5.26}$$

Because of Clifford covariance, for any $g \in \mathcal{C}\ell_n$ it holds that $g(A_0^\dagger) = A_{a_g}^\dagger$. Therein, we have $g(A_0^\dagger) = \sum_b c_b^*(0)\, g(T_b) = \sum_b c_b^*(0)\omega^{\tilde{\Phi}_g(b)}\, T_{S_g b}$, and, with Eq. (5.26), $A_{a_g}^\dagger = \left(T_{y^{-1}(a_g)}(A_0)\right)^\dagger = \sum_b \omega^{[y^{-1}(a_g),b]} c_b^*(0)\, T_b$. For any $g \in \mathcal{C}\ell_n$ and any pair $b, S_g b \in E$ it follows that

$$c_{S_g b}(0) = c_b(0)\, \omega^{[y^{-1}(a_g),S_g b]-\tilde{\Phi}_g(b)}. \tag{5.27}$$

Thus, for the entire Clifford orbit $\langle b\rangle$ of any $b \in E$, the expansion coefficients $c_b(0)$ have the same magnitude, and differ only by phase factors ω^m, $m \in \mathbb{N}$. They can thus be written in the form

$$c_b(0) = c_{\langle b\rangle}\, \omega^{\nu(b)}, \forall b \in E, \tag{5.28}$$

where $c_{\langle b\rangle} \in \mathbb{C}$ only depends on the Clifford orbit $\langle b\rangle$ of b, and $\nu : E \longrightarrow \mathbb{Z}_d$.

Further inspecting Eq. (5.27), the l.h.s. depends on g only through $S_g b \in E$, whereas the r.h.s. has the a priori more general dependence through $\tilde{\Phi}_g$. Comparing Eqs. (5.27) and (5.28), we find that the functions ν and $\tilde{\Phi}$ are mutually constrained by the relation

$$\nu(S_g b) - \nu(b) = [y^{-1}(a_g), S_g b] - \tilde{\Phi}_g(b),\ \ \forall b \in E,\ \forall g \in \mathcal{C}\ell_n.$$

Now consider a face $f \in \tilde{C}_2$ with boundary $\partial f = [a] + [b] - [a+b]$, and add up the above relations for the edges in the boundary. Because of its linearity, the commutator term then vanishes, and we obtain

$$\nu(g\partial f) - \nu(\partial f) = -\tilde{\Phi}_g(\partial f) = -\Phi_{\mathrm{cov}}([g], \partial f),\ \forall f \in C_2(\mathcal{C}'),\ \forall g \in \mathcal{C}\ell_n,$$

where $g\partial f := [S_g a] + [S_g b] - [S_g(a+b)]$. Thus, $\Phi_{\mathrm{cov}} = d^h(-\nu)$, i.e., $[\Phi_{\mathrm{cov}}] = 0$.

"If": Assume that $[\Phi_{\mathrm{cov}}] = 0$. Thus, there is a phase convention γ such that $\tilde{\Phi}^{(\gamma)}|_{\tilde{B}_1} \equiv 0$. But this means that $\tilde{\Phi}_g^{(\gamma)}(\cdot)$ is a linear function for any $g \in \mathcal{C}\ell_n$, and we may write it as

$$\tilde{\Phi}_g^{(\gamma)}(a) = [x_g, a],$$

for suitable x_g, $g \in \mathcal{C}\ell_n$. Let $W^{(\gamma)}$ be the Wigner function defined by the phase space point operators $A_v^{(\gamma)} = \frac{1}{d^n} \sum_b \omega^{-[v,b]} \left(T_b^{(\gamma)}\right)^\dagger$.

Since the $A_v^{(\gamma)}$ constitute a special case of Ansatz (5.21) with Condition (5.22b), we already know that they span an operator basis (OB). Moreover, $W^{(\gamma)}$ is Clifford covariant.

Namely,

$$\begin{aligned} g\Big(A_v^{(\gamma)}\Big) &= \tfrac{1}{d^n}\sum_b \omega^{-[v,b]}\, g\Big[\Big(T_b^{(\gamma)}\Big)^\dagger\Big] \\ &= \tfrac{1}{d^n}\sum_b \omega^{-[v,b]-\tilde{\Phi}_g^{(\gamma)}(b)}\, \Big(T_{S_g b}^{(\gamma)}\Big)^\dagger \\ &= \tfrac{1}{d^n}\sum_b \omega^{-[v+x_g,b]}\, \Big(T_{S_g b}^{(\gamma)}\Big)^\dagger \\ &= \tfrac{1}{d^n}\sum_b \omega^{-[S_g v+S_g x_g,b]}\, \Big(T_b^{(\gamma)}\Big)^\dagger \\ &= A_{S_g v+S_g x_g}^{(\gamma)} \end{aligned}$$

This is Eq. (5.25) with $a_g = S_g x_g$. □

Theorem 10 has the following generalization, which we will refer to in Section 5.4.4.

Corollary 2. *For any number $n \geq 1$ of qudits of local dimension $d \geq 2$, let $G \subseteq \mathcal{C}\ell_n$ be a subgroup of the Clifford group such that $\mathcal{P}_n \subset G$. A G-covariant Wigner function according to (OB) exists if and only if $[\Phi_{\rm cov}] = 0 \in H^1(Q, U_{\rm cov})$.*

The proof is exactly the same as for Theorem 10.

5.4.3 Even vs. odd dimension

For odd dimension d, and all numbers n of qudits, a Clifford-covariant Wigner function satisfying the conditions (OB), (SW1)–(SW4) has been explicitly constructed in [31]. In the present formalism, the existence of a Clifford-covariant Wigner function satisfying (OB) follows by Observation 3 and Theorem 10.

In even dimension we have the following result.

Theorem 11. *If the local dimension d is even, then, for any number n of local systems, Clifford-covariant Wigner functions satisfying (OB) do not exist.*

Remark: There is a related no-go theorem in [52], addressing the impossibility of positive representation of the stabilizer sub-theory in even dimension. Here, we assume much less about Wigner functions than the related no-go theorem in [52], namely only (OB). We point out specifically that the assumption of diagram preservation is not required.

Furthermore, a special instance of Theorem 11 was proved in [178], namely for the multiqubit case of $n > 1$, $d = 2$. This result is based on the fact—proved in the same paper—that the multiqubit Clifford groups are unitary 3-designs. However, this property does not extend to other even dimension except powers of 2 [63].

Remark: From the perspective of QCM, Theorem 11 can be bypassed. Namely, as explained in Section 5.2.2, no computational power is lost in QCM if the Clifford gates are dispensed with. This holds even though such gates are by convention part of QCM [19, 38]. If no Clifford gates occur, then the breakdown of Clifford covariance is no problem for the classical simulation. The computational power, and hence the hardness of classical simulation, rests with the Pauli measurements.

To prove Theorem 11, we first establish the following statement about the cohomology class $[\Phi_{\rm cov}]$ for the n-qudit Clifford group.

Lemma 20. *For all even local dimensions d and all qudit numbers n, it holds that $[\Phi_{\mathrm{cov}}] \neq 0$.*

Proof of Lemma 20. The basic proof strategy is to identify a group element $g \in \mathcal{C}\ell_n$ and a 2-chain $f \in \tilde{C}_2$ such that

$$g\partial f = \partial f, \text{ and } \tilde{\Phi}_g(\partial f) \neq 0. \tag{5.29}$$

Assume Eq. (5.29) holds and $[\Phi_{\mathrm{cov}}] = 0$. From the latter, $\tilde{\Phi}_g(\partial f) = \Phi_{\mathrm{cov}}([g], \partial f) = \nu(\partial f) - \nu(g\partial f)$, for some $\nu \in C^0(G, U_{\mathrm{cov}})$. Therein, with Eq. (5.29), $0 \neq 0$—Contradiction. Thus, the existence of a pair g, f satisfying Eq. (5.29) implies that $[\Phi_{\mathrm{cov}}] \neq 0$.

In the following, we show that g, f can be chosen in accordance with Eq. (5.29), for all even d. We focus on $n = 1$; replacing g by $g \otimes I^{\otimes n-1}$ and all $a_z, a_x \in \mathbb{Z}_d$ by $(a_z, 0, \ldots, 0), (a_x, 0, \ldots, 0) \in \mathbb{Z}_d^n$ immediately generalizes the proof to all $n \in \mathbb{N}$. For concreteness, we set $\gamma(a) = a_z a_x$, but our final statement is independent of the phase convention γ.

For $d = 4m + 2$, $m \in \mathbb{N}_0$, we consider the Fourier transform $g = \frac{1}{\sqrt{d}} \sum_{k,l=0}^{d-1} \omega^{kl} |k\rangle\langle l|$, which is a Clifford unitary that acts by conjugation as

$$g \cdot g^\dagger : X \longrightarrow Z \longrightarrow X^{-1} \longrightarrow Z^{-1} \longrightarrow X. \tag{5.30}$$

Further, $f := [u|v]$, with $u = (u_z, u_x)$ and $v = (v_z, v_x)$ by $u_x = v_z = 0$ and $u_z = v_x = 2m+1$. Then

$$\begin{aligned} T_u &= Z^{2m+1}, & T_v &= X^{2m+1} & T_{u+v} &= \sqrt{\omega}^{(2m+1)^2} Z^{2m+1} X^{2m+1}, \\ gT_u g^\dagger &= T_v, & gT_v g^\dagger &= T_u, & gT_{u+v} g^\dagger &= \omega^{-(2m+1)} T_{u+v}. \end{aligned}$$

Hence, $g\partial[u|v] = \partial[u|v]$ but $\tilde{\Phi}_g(\partial[u|v]) = 0 + 0 + (2m+1) = d/2 \neq 0 \in \mathbb{Z}_d$. Thus, Eq. (5.29) applies.

For $d = 4m$, $m \in \mathbb{N}$, we introduce the unitary $g = \frac{1}{d} \sum_{k,l=0}^{d-1} \omega^{mk^2} \left(\sum_{j=0}^{d-1} \omega^{mj^2 + (k-l)j} \right) |k\rangle\langle l|$, which is in the Clifford group since

$$gZg^\dagger = \sqrt{\omega}^{2m} ZX^{2m}, \quad gXg^\dagger = \sqrt{\omega}^{-2m} Z^{2m} X. \tag{5.31}$$

Choosing $f := [u|v]$, with $u_z = v_z = 1$, $u_x = 0$, and $v_x = 2m$ yields

$$\begin{aligned} T_u &= Z, & T_v &= \sqrt{\omega}^{2m} ZX^{2m} & T_{u+v} &= \sqrt{\omega}^{d} Z^2 X^{2m}, \\ gT_u g^\dagger &= T_v, & gT_v g^\dagger &= \omega^{2m} T_u, & gT_{u+v} g^\dagger &= T_{u+v}. \end{aligned}$$

Again, $\partial f = g\partial f$, and $\tilde{\Phi}_g(\partial[u|v]) = 0 + 2m - 0 = d/2 \neq 0 \in \mathbb{Z}_d$. Eq. (5.29) thus applies. □

Proof of Theorem 11. The statement is the combined conclusion of Lemma 20 and Theorem 10. □

5.4.4 First vs. second cohomology group

The Clifford group splits when d is odd and $n \geq 1$, or $d = 2$ and $n = 1$. It does not split for d even and $n \geq 2$ [31, 179, 180]. Splitting is also a notion of group cohomology, living

at the second level. *Is there a connection with Clifford covariance?*

In this section we clarify that such a connection does indeed exist. However, it is not one-to-one. Namely, splitting of the symmetry group G is necessary for G-covariance, but not sufficient.

A notion of particular interest for this discussion is the specific faithful group action of the group G (the Clifford group, or a subgroup thereof) on the phase space V required by Definition 9. The existence of such a group action is a precondition for covariance. We establish that splitting is necessary and sufficient for the existence of the faithful group action.

Recall that splitting is equivalent to $[\zeta] = 0$ with ζ defined by Eq. (5.15c), and Q and U_{cov} are defined above Eq. (5.14). In summary we arrive at the following picture:

$$\boxed{\begin{array}{ccc} G\text{-covariance} & \Rightarrow & \begin{array}{c}\exists \text{ faithful group} \\ \text{action of } G \text{ on } V\end{array} \\ \Updownarrow & & \Updownarrow \\ [\Phi_{\text{cov}}] = 0 \in H^1(Q, U_{\text{cov}}) & \Rightarrow & [\zeta] = 0 \in H^2(Q, E) \end{array}} \tag{5.32}$$

The vertical arrow on the left is the content of Corollary 2 in Section 5.4.2. The remaining three arrows are established in Sections 5.4.4 and 5.4.4 below.

Splitting and faithful group action

With ζ and Q as defined in Section 5.2.3, we have the following result.

Theorem 12. *For any given number n of d-level systems, be $G \subseteq \mathrm{Cl}_n$ a subgroup of the Clifford group such that $\mathcal{P}_n \subset G$. A faithful action of G on the phase space V of the form $\tau : G \times V \to V$, $\tau_g(v) = S_g v + a_g$ exists if and only if $[\zeta] = 0 \in H^2(Q, E)$.*

Proof of Theorem 12. "Only if": Assume that a faithful group action of the form $\tau_g(v) = S_g v + a_g$ does exist.

First, we examine τ_g for $g \in \mathcal{P}_n = \{t_a, a \in E\}$, where t_a denotes T_a with the phase modded out. For readability, we write $y(a)$ instead of a_g with $g = t_a$. Note that $t_a t_b = t_{a+b}$ and $t_a^\dagger = t_{-a}$. Therefore, $\forall g \in \{t_a\}$, S_g as defined in Eq. (5.4) is the identity transformation, i. e., $\tau_{t_a}(v) = v + y(a)$.

Using that τ_g is a group action, we find that $y : E \cong \mathbb{Z}^{2d} \to V \cong \mathbb{Z}^{2d}$ is linear:

$$y(a+b) = \tau_{t_a t_b}(v) - v = \tau_{t_a}(\tau_{t_b}(v)) - v = y(a) + y(b) \quad \forall a, b \in E$$

Moreover, y is injective and, thus, invertible:

$$y(a) = y(b) \;\Rightarrow\; \tau_{t_a} = \tau_{t_b} \;\Rightarrow\; t_a = t_b \;\Leftrightarrow\; a = b,$$

where the second step relies on τ being faithful.

Thanks to the invertibility of y, we can now choose a section $\theta : Q \to G$, see Eq. (5.15), such that $\tau_{\theta(S)}(v) = Sv$. In general, $\tau_{\theta'(S)}(v) = Sv + a_{\theta'(S)}$. However, by Eq. (5.16), we

can always transition from $\theta'(S)$ to $\theta(S) = t_{y^{-1}[-a_{\theta'(S)}]}\theta'(S)$. This ensures that $\tau_{\theta(S)}(v) = \tau_{t_{y^{-1}[-a_{\theta'(S)}]}}(\tau_{\theta'(S)}(v)) = Sv$, as desired. By fixing θ we single out a specific representative ζ from the equivalence class $[\zeta] \in H^2(Q, E)$.

At last, we show that, for this choice of θ, $\zeta(S_1, S_2) = 0\ \forall S_1, S_2 \in Q$. Using Eq. (5.15) to go from the first to the second line, we find

$$\begin{aligned} S_1S_2v &= \tau_{\theta(S_1)}(\tau_{\theta(S_2)}(v)) = \tau_{\theta(S_1)\theta(S_2)}(v) \\ &= \tau_{t_{\zeta(S_1,S_2)}\theta(S_1S_2)}(v) = \tau_{t_{\zeta(S_1,S_2)}}(\tau_{\theta(S_1S_2)}(v)) = S_1S_2v + y(\zeta(S_1, S_2)),\ \ \forall S_1, S_2 \in Q. \end{aligned}$$

Hence $y(\zeta(S_1, S_2)) = 0$ and, since y is linear and invertible, $\zeta(S_1, S_2) = 0\ \forall S_1, S_2 \in Q$. This, finally, yields $[\zeta] = 0 \in H^2(Q, E)$.

"If": Assume that $[\zeta] = 0 \in H^2(Q, E)$. Then we can choose θ such that $\zeta(S_1, S_2) = 0$ $\forall S_1, S_2 \in Q$. According to Eq. (5.15), any $g \in G$ can be cast in the form $g = t_{\alpha_g}\theta(S_g)$.

Let us show that $\tau_g(v) = S_gv + \alpha_g$ is a faithful group action of G on V. Observing that $hg = t_{\alpha_h}\theta(S_h)t_{\alpha_g}\theta(S_g) = t_{\alpha_h}\theta(S_h)t_{\alpha_g}(\theta(S_h))^\dagger\theta(S_h)\theta(S_g) = t_{\alpha_h+S_h\alpha_g}\theta(S_hS_g)$, we find $\tau_{hg}(v) = \tau_{t_{\alpha_h+S_h\alpha_g}\theta(S_hS_g)}(v) = S_hS_gv + \alpha_h + S_h\alpha_g = S_h(S_gv + \alpha_g) + \alpha_h = \tau_h(\tau_g(v))$, i. e., τ is a group action.

Moreover, τ is faithful. Let $\tau_g(v) = \tau_h(v)$, $\forall v \in V$. Then $\alpha_g = \alpha_h$, $S_g = S_h$, and, thus, $g = t_{\alpha_g}\theta(S_g) = t_{\alpha_h}\theta(S_h) = h$. □

This establishes the vertical arrow on the right of the diagram Eq. (5.32).

Covariance implies splitting

We have the following result, as a corollary to Theorem 12:

Corollary 3. *For any given number n of d-level systems, be $G \subseteq \mathcal{C}\ell_n$ a subgroup of the Clifford group such that $\mathcal{P}_n \subset G$. A G-covariant Wigner function satisfying (OB) exists only if $[\zeta] = 0 \in H^2(Q, E)$.*

Proof of Corollary 3. We first show that G-covariance as defined in Definition 9 implies the existence of a faithful group action of G on V of the form $\tau : G \times V \to V$, $\tau_g(v) = S_gv + a_g$.

A Wigner function satisfying (OB) is Clifford covariant if and only if $g(A_v) = A_{S_gv+a_g} = A_{\tau_g(v)}\ \forall v \in V$, $\forall g \in G$, cf. Eq. (5.25). Therefore

$$A_{\tau_{hg}(v)} = (hg)(A_v) = h(g(A_v)) = A_{\tau_h(\tau_g(v))}\ \ \forall v \in V,\ \forall g, h \in G. \tag{5.33}$$

Since the A_v constitute an operator basis, $A_v = A_{v'}$ implies $v = v'$. Thus, $\tau_{hg}(v) = \tau_h(\tau_g(v))$ $\forall v \in V$, $\forall g, h \in G$, and τ is, indeed, a group action.

If τ were not faithful, there would exist $g, h \in G$ with $g \neq h$ such that

$$g(A_v) = h(A_v)\ \ \forall v \in V \ \ \Leftrightarrow \ \ h^{-1}gA_v = A_vh^{-1}g\ \ \forall v \in V. \tag{5.34}$$

Using, again, that $\{A_v, v \in V\}$ is a basis, we observe that $[h^{-1}g, Y] = 0$ for any linear operator Y and, thus, the unitary $h^{-1}g$ is, up to phase, the identity operator. This contradicts that $g \neq h$ in G. Hence, τ is a faithful group action.

With Theorem 12, it follows that $[\zeta] = 0$. □

Corollary 3 demonstrates the upper horizontal arrow in the diagram Eq. (5.32). In sum, we have established the two vertical arrows and the upper horizontal one, implying the remaining lower horizontal arrow. This completes the diagram Eq. (5.32).

The lower horizontal arrow, so far derived through reasoning about the physics concept of Wigner function, relates the cohomology groups $H^1(Q, U_{\mathrm{cov}})$ and $H^2(Q, E)$, which are purely mathematical objects. In Appendix A.4.1 we independently establish this relation as a mathematical fact, without recurring to physics concepts.

Splitting does not imply covariance

Splitting of a Clifford subgroup G does not guarantee the existence of a G-covariant Wigner function. This is demonstrated with the following 1-qubit example: We consider the covariance group G generated by the Hadamard gate H and the one-qubit Pauli group. This group splits, $G = \langle H \rangle \ltimes P_1$. However, $[\Phi_{\mathrm{cov}}] \neq 0$. Namely, consider the Pauli operators $T_{a_1} = X$, $T_{a_2} = Y$, $T_{a_3} = Z$. Then, $\tilde{\Phi}_H(a_1) = \tilde{\Phi}_H(a_3) = 0$, and $\tilde{\Phi}_H(a_2) = 1$. Thus, $\Phi_{\mathrm{cov}}(H, \partial[a_1|a_3]) = \tilde{\Phi}_H(\partial[a_1|a_3]) = 1$ and $H\partial[a_1|a_3] = \partial[a_1|a_3]$. These relations imply that $[\Phi_{\mathrm{cov}}] \neq 0$ (see the proof of Lemma 20). Thus, with Corollary 2, no G-covariant Wigner function satisfying (OB) exists.

5.5 Positive representation of Pauli measurement

In this section we first define "positive representation of Pauli measurement", and then establish a necessary and sufficient condition for it, applicable to all Wigner functions that satisfy (OB), (SW1)–(SW4). Finally, we apply this result to even and odd dimensions, making use of the respective structure theorems for Pauli observables.

5.5.1 When are Pauli measurements positively represented?

Denote by Π_a^s the projector associated with the outcome $s \in \mathbb{Z}_d$ in the measurement of the Pauli observable T_a, i. e., Π_a^s is the projector onto the eigenspace of T_a with eigenvalue ω^s. The probability to obtain s when measuring T_a on the state ρ is $\mathrm{Tr}(\Pi_a^s \rho)$. We then have the following definition of "positive representation of Pauli measurement":

Definition 10. *A pair (W, Θ) of a Wigner function W and an effect function Θ satisfying (SW4), i. e., $\mathrm{Tr}(\Pi_a^s \rho) = \sum_{v \in V} \Theta_{\Pi_a^s}(v) W_\rho(v)$, represents Pauli measurements positively if the following two properties hold.*

(a) For all Pauli measurements, the corresponding effect functions $\Theta_{\Pi_a^s} : V \longrightarrow \mathbb{R}$ satisfy

$$\Theta_{\Pi_a^s}(v) \geq 0, \ \forall v \in V, \forall a, s. \tag{5.35}$$

(b) For all Pauli measurements, the nonnegativity of the Wigner function is preserved under measurement, i.e.,

$$W_\rho \geq 0 \Longrightarrow W_{\Pi_a^s \rho \Pi_a^s} \geq 0, \forall a, s. \tag{5.36}$$

The above definition of "positive representation of Pauli measurement" is intuitive. Condition (a) says that the effects $\Theta_{\Pi_a^s}$ associated with all Pauli measurements are nonnegative, and if the Wigner function W_ρ is nonnegative as well, then the outcome probabilities for Pauli measurements can be obtained by sampling from the phase space V. Condition (b) says that if a state ρ is represented by a nonnegative Wigner function, then for any Pauli measurement with any outcome, the post-measurement state is also represented by a nonnegative Wigner function.

5.5.2 Cohomological condition for positive representation

Here we show the following.

Theorem 13. *For any system of n qudits with d levels, $n, d \in \mathbb{N}$, a pair of Wigner and effect function satisfying (OB), (SW1)–(SW4) that represents Pauli measurement positively exists if and only if $[\beta] = 0 \in H^2(\mathcal{C}, \mathbb{Z}_d)$.*

The proof of Theorem 13 proceeds in several steps. To begin, we note that $(T_b)^k \sim T_{bk}$, and define phases $\varphi_b : \mathbb{Z}_d \longrightarrow \mathbb{Z}_d$, for all b, such that

$$(T_b)^k = \omega^{\varphi_b(k)} T_{bk}, \quad \forall b, \ \forall k. \tag{5.37}$$

With Ansatz (5.21), Reality (SW1) becomes equivalent to $c_b^* T_b = c_{-b} T_{-b}^\dagger$, and hence to both

$$c_b^* = \omega^{\varphi_{-b}(-1)} c_{-b} \ \text{ and } \ c_b^* = \omega^{\varphi_b(-1)} c_{-b}. \tag{5.38}$$

To prepare for subsequent applications, we observe that $(T_b)^{kl} = \big((T_b)^k\big)^l$, for all $k, l \in \mathbb{Z}_d$ and all b. For the phases φ_b defined in Eq. (5.37) this entails

$$\varphi_b(kl) = \varphi_{kb}(l) + l\varphi_b(k).$$

We will later make use of this relation for the special case of $l = -1 \mod d$,

$$\varphi_b(-k) = \varphi_{kb}(-1) - \varphi_b(k). \tag{5.39}$$

As the next step towards proving Theorem 13, we have the following two results.

Lemma 21. *For all pairs of Wigner and effect functions satisfying (OB), (SW3), (SW4), the functions $\Theta_{\Pi_a^s}(\cdot)$ are, for all a, s, of the form*

$$\Theta_{\Pi_a^s}(v) = \frac{1}{d} \sum_k \omega^{-ks'} c_a(k) \tag{5.40}$$

with $s' := s + [v, a]$, $c_a(k) := \omega^{\varphi_a(k)} c_{ka}$, and

$$c_a(k)^* = c_a(-k). \tag{5.41}$$

In particular, all $\Theta_{\Pi_a^s}(\cdot)$ are real-valued.

Further,

Lemma 22. *For any system of n qudits with d levels, $n, d \in \mathbb{N}$, a pair of Wigner and effect function satisfying (OB), (SW1)–(SW4) can represent measurement positively only if $|c_b| = 1$, $\forall b$.*

Proof of Lemma 21. For any a, the projector Π_a^s can be represented as

$$\Pi_a^s = \frac{1}{d} \sum_{i=0}^{d-1} \omega^{-si} (T_a)^i.$$

Starting from the expression for Θ in Eq. (5.19) and the expansion of A_v in Eq. (5.21), which invoke the assumptions (OB), (SW3), and (SW4) of the lemma, we obtain by direct computation

$$\Theta_{\Pi_a^s}(v) = \frac{1}{d} \sum_{k=0}^{d-1} \omega^{-k(s+[v,a])+\varphi_a(k)} c_{ka}. \tag{5.42}$$

We have thus established all $\Theta_{\Pi_a^s}$ as functions from V to $\mathbb{C}$. We still need to show that all $\Theta_{\Pi_a^s}(v)$ are real-valued. To this end, we use the above definitions $s' = s + [v, a]$ and $c_a(k) = \omega^{\varphi_a(k)} c_{ka}$, which simplify $\Theta_{\Pi_a^s}(v)$ to Eq. (5.40).

The $\Theta_{\Pi_a^s}(v)$ are thus real-valued if $c_a(k)^* = c_a(-k)$, for all a and all k. This property we now demonstrate.

$$\begin{aligned} c_a(k)^* &= c_{ka}^* \, \omega^{-\varphi_a(k)} \\ &= c_{-ka} \, \omega^{\varphi_{ka}(-1)-\varphi_a(k)} \\ &= c_a(-k) \, \omega^{-\varphi_a(-k)+\varphi_{ka}(-1)-\varphi_a(k)} \\ &= c_a(-k) \end{aligned}$$

Therein, the first and fourth line follow by the definition of $c_a(\cdot)$, and the second line by Eq. (5.38). Finally, using Eq. (5.39) in the last relation, we obtain Eq. (5.41). □

Proof of Lemma 22. The proof consists of two parts. Part (a) invokes the condition Eq. (5.35) in Def. 10, and leads to Eq. (5.43). Part (b) invokes the condition Eq. (5.36) in Def. 10 of positive representation of Pauli measurement, and leads to Eq. (5.44) below. Both constraints together imply the statement of the lemma.

(a). First we exclude the possibility of $|c_a| > 1$ for any a. With the expansion of A_v in Eq. (5.21), for any $v \in V$, $|c_a| = |\mathrm{Tr}(T_a A_v)| = |\sum_s \omega^s \, \mathrm{Tr}(\Pi_a^s A_v)| \leq \sum_s |\mathrm{Tr}(\Pi_a^s A_v)| = \sum_s \Theta_{a,s}(v) = 1$. Herein, the third relation follows by the triangle inequality, and $|\mathrm{Tr}(\Pi_a^s A_v)| = \mathrm{Tr}(\Pi_a^s A_v) = \Theta_{\Pi_a^s}(v)$ relies on Eq. (5.19) and the assumption of positive representation; cf. Eq. (5.35) in Def. 10. The last equality is by Eq. (5.20). Thus we arrive at

$$|c_a| \leq 1, \quad \forall a. \tag{5.43}$$

(b). Consider the expectation value $\langle T_a \rangle_\rho$. With Eqs. (5.17) and (5.21),

$$\langle T_a \rangle_\rho = \sum_{v \in V} W_\rho(v) \omega^{-[v,a]} c_a.$$

Now assume that $|c_a| < 1$, and consider quantum states ρ that are positively represented

by W, i.e., $W_\rho \geq 0$. Then,

$$\begin{aligned} |\langle T_a \rangle_\rho| &\leq \textstyle\sum_{v \in V} |W_\rho(v)||c_a| \\ &< \textstyle\sum_{v \in V} |W_\rho(v)| \\ &= \textstyle\sum_{v \in V} W_\rho(v) \\ &= 1. \end{aligned}$$

In short, $|\langle T_a \rangle_\rho| < 1$, for all $a \in E$. Above, the first line holds by the triangle inequality, the second line invokes the assumption $|c_a| < 1$, the third line invokes the other assumption $W_\rho \geq 0$, and the fourth line follows from the Standardization condition (SW2).

In view of this constraint, consider any eigenstate $\rho(a, s)$ of the Pauli observable T_a, with an eigenvalue ω^s. Hence, $|\langle T_a \rangle_{\rho(a,s)}| = 1$. With the above, $|c_a| < 1$ implies that $W_{\rho(a,s)} < 0$, $\forall\, s \in \mathbb{Z}_d$.

With Eqs. (5.17)–(5.22), we observe that the completely mixed state I/d^n has a positive Wigner function. However, the post-measurement states resulting from measuring T_a on I/d^n are of the type $\rho(a, s)$. Thus, if $|c_a| < 1$ then W is not positivity preserving under measurement of the Pauli observable T_a, contradicting the condition Eq. (5.36) in Def. 10. Hence, a Wigner function represents Pauli measurement positively only if

$$|c_a| \geq 1, \ \ \forall a. \tag{5.44}$$

Combining Eqs. (5.44) and (5.43) leaves $|c_a| = 1$ as the only option, for all a. □

As the final step in proving Theorem 13, we need to constrain the phases of the coefficients $c_a(k)$, cf. Lemma 21. This requires a discrete version of Bochner's theorem.

Lemma 23 (Variation on Bochner's theorem)**.** *For a given function $f : \mathbb{Z}_d \longrightarrow \mathbb{C}$ the Fourier transform $\hat{f}$ of f is nonnegative if and only if the matrix M with coefficients*

$$M_y^x = f(x - y), \ \ \forall x, y \in \mathbb{Z}_d, \tag{5.45}$$

is positive semidefinite.

The proof of Lemma 23 is the same as in [31] (Theorem 44 therein). To explicitly demonstrate that it applies in both odd and even dimensions, we restate it in Appendix B.4.1.

We are now ready to prove the main result of this section, Theorem 13.

Proof of Theorem 13. "Only if": We assume that a given pair of Wigner function W and effect function Θ satisfying (OB), (SW1)–(SW4) represents Pauli measurement positively. Our first goal is to show that the coefficients $c_a(k)$ may then be expressed in the form

$$c_a(k) = (\omega^{r_a})^k, \text{ with } r_a \in \mathbb{Z}_d, \ \forall a. \tag{5.46}$$

The set $\{r_a, \forall a\}$ characterizes the phase space point operators $\{A_v\}$ via $c_a = c_a(1) = \omega^{r_a}$.

With Lemma 22, we can express the coefficients $c_a(k)$ as $c_a(k) = e^{i\chi_a(k)}$, $\chi_a(k) \in \mathbb{R}$. We furthermore observe that, with Eqs. (5.22a) and (5.37), it holds that $c_a(0) = 1$, $\chi_a(0) = 0$, for all $a \in E$. If $d = 2$, then Eq. (5.46) follows directly from Eq. (5.41); namely the $c_a(k)$ are all real.

For $d > 2$, we consider the submatrix $M(a)|_{J\times J}$ for the set of rows (and columns) $J = \{1, 2, k+1\}$, which, using Eq. (5.41), reads

$$M(a)|_{J\times J} = \begin{pmatrix} 1 & e^{i\chi_a(1)} & e^{i\chi_a(k)} \\ e^{-i\chi_a(1)} & 1 & e^{i\chi_a(k-1)} \\ e^{-i\chi_a(k)} & e^{-i\chi_a(k-1)} & 1 \end{pmatrix}$$

By Lemmas 21 and 23, we require the determinant of this matrix to be nonnegative, which leads to the constraint

$$e^{i(\chi_a(1)+\chi_a(k-1)-\chi_a(k))} + \text{c.c.} \geq 2.$$

The only solution of that constraint is $\chi_a(k) = \chi_a(k-1) + \chi_a(1)$, which we may use as a recursion relation for computing the angles $\chi_a(k)$, for all k. With $\chi_a(0) = 0$,

$$\chi_a(k) = k\,\chi_a(1).$$

With the relation $c_a(k)^* = c_a(-k \mod d)$ we further find

$$d\chi_a(1) = 0 \mod 2\pi, \ \ \forall a.$$

Eq. (5.46) follows from the last two relations, where $\omega^{r_a} = e^{i\chi_a(1)}$. With Eq. (5.40) this implies

$$\Theta_{\Pi_a^s}(v) = \delta_{r_a, s+[v,a]}. \tag{5.47}$$

Now consider the simultaneous measurement of the commuting observables T_a, T_b with outcomes $s(a)$, $s(b)$ on the completely mixed state. We denote the resulting state by ρ. A further measurement of T_a or T_b on ρ must produce outcomes $s(a)$, $s(b)$ with certainty, and a measurement of T_{a+b} must produce the outcome $s_{a+b} = s(a) + s(b) - \beta(a,b)$ with certainty. Since the completely mixed state is positively represented and, by assumption, Pauli measurements are positivity preserving, it holds that $W_\rho \geq 0$. Using Standardization (SW2), we observe that for any phase space point v in the support of W_ρ, of which there is at least one, it must hold that $\Theta_{\Pi_a^{s(a)}}(v) = \Theta_{\Pi_b^{s(b)}}(v) = \Theta_{\Pi_{a+b}^{s(a)+s(b)-\beta(a,b)}}(v) = 1$. Therefore, with Eq. (5.47),

$$\begin{aligned} s(a) &= r_a - [v,a], \\ s(b) &= r_b - [v,b], \\ s(a) + s(b) - \beta(a,b) &= r_{a+b} - [v, a+b]. \end{aligned}$$

Adding the first two equations and subtracting the third, we obtain

$$r_a + r_b - r_{a+b} = \beta(a,b); \ \ \forall a, b \ \text{ with } [a,b] = 0. \tag{5.48}$$

Note that in this condition the dependence on the particular phase space point v has disappeared. In cohomological notation, Eq. (5.48) reads $\beta = dr$; hence $[\beta] = 0$.

"If": Assume that $[\beta] = 0$ holds. We show that then Gross' Wigner function, see Section 2.4.4, has all the desired properties. We can choose a gauge such that $\beta \equiv 0$. In this gauge it holds that $\varphi_a \equiv 0$ for all a, cf. Eq. (5.37). Note that $\varphi_a \equiv 0$, in turn, implies

$(T_0)^0 = T_0$ and, thus, $\mu^{\gamma(0)} = 1$. We assert

$$c_a \equiv 1, \tag{5.49}$$

and for this choice verify the Stratonovich-Weyl criteria as well as conditions Eq. (5.35) and (5.36) of Def. 10.

(i) *Stratonovich-Weyl criteria.* (OB), (SW2), and (SW3) hold since we use Ansatz (5.21) and fulfill Conditions (5.22a) and (5.22b). With $\varphi_a \equiv 0$, (SW1) in Ansatz (5.21) becomes equivalent to $c_a^* = c_{-a}$, cf. Eq. (5.38), which is clearly satisfied by $c_a = 1$ in Eq. (5.49). Finally, we ensure (SW4) by setting $\Theta_Y(v) = \mathrm{Tr}(YA_v)$, cf. Eq. (5.19).

(ii) *Condition Eq. (5.35).* With Eq. (5.49), we obtain $\Theta_{\Pi_a^s} = \delta_{s,[a,v]}$. Eq. (5.35) is thus satisfied.

(iii) *Condition Eq. (5.36).* It suffices to show that $\Pi_a^s A_v \Pi_a^s = \sum_{w\in V} q_v(w) A_w$, with all $q_v(w)$ real and nonnegative. We have

$$\begin{aligned}
\Pi_a^s A_v \Pi_a^s &= \Pi_a^s \sum_{b|\,[a,b]=0} \frac{1}{d^n} \omega^{[b,v]} T_b^\dagger \\
&= \left(\frac{1}{d} \sum_{k=0}^{d-1} \omega^{-ks} (T_a)^k \right) \left(\sum_{b|\,[a,b]=0} \frac{1}{d^n} \omega^{[b,v]} T_b^\dagger \right) \\
&= \frac{1}{d^{n+1}} \sum_{k=0}^{d-1} \omega^{-ks} \sum_{b|[a,b]=0} \omega^{[b,v]} T_{b-ka}^\dagger \\
&= \frac{1}{d^{n+1}} \sum_{k=0}^{d-1} \omega^{k([a,v]-s)} \sum_{b|[a,b]=0} \omega^{[b,v]} T_b^\dagger \\
&= \frac{\delta_{s,[a,v]}}{d^n} \sum_b \delta_{[a,b],0}\, \omega^{[b,v]} T_b^\dagger \\
&= \frac{\delta_{s,[a,v]}}{d^{n+1}} \sum_{k=0}^{d-1} \sum_b \omega^{[b,v]+k[b,a]} T_b^\dagger \\
&= \delta_{s,[a,v]} \frac{1}{d} \sum_{k=0}^{d-1} A_{v+ka}
\end{aligned} \tag{5.50}$$

Therein, in the first line we have used the ansatz Eq. (5.49). To remove the noncommuting elements ($[a,b] \neq 0$), we have used the following argument: $\Pi_a^s T_a = T_a \Pi_a^s = \omega^s \Pi_a^s$, and hence $\Pi_a^s T_b \Pi_a^s = \omega^{-s} \Pi_a^s T_a T_b \Pi_a^s = \omega^{-s+[a,b]} \Pi_a^s T_b T_a \Pi_a^s = \omega^{[a,b]} \Pi_a^s T_b \Pi_a^s$. Thus, if $[a,b] \neq 0$ then $\Pi_a^s T_b \Pi_a^s = 0$. On the other hand, if $[a,b] = 0$ then $\Pi_a^s T_b \Pi_a^s = \Pi_a^s \Pi_a^s T_b = \Pi_a^s T_b$. In the third line we have used the phase convention that yields $\beta \equiv 0$, and in the fourth line we re-organized the sum over b. Thus, all nonzero coefficients in the expansion of $\Pi_a^s A_v \Pi_a^s$ are positive. □

5.5.3 Even vs. odd dimension

For odd local dimension, a Wigner function which represents Pauli measurement positively has been explicitly constructed [31]. In the framework established here, the existence of

$T_a = Z^{-1} \otimes I$	$\times$	$T_b = I \otimes Z$	$\times$	$T_{a+b}^{-1} = Z \otimes Z^{-1}$	$= I \Rightarrow$	$\beta(a,b) = 0$
$\times$		$\times$		$\times$		
$T_c = I \otimes \tilde{X}$	$\times$	$T_d = \tilde{X}^{-1} \otimes I$	$\times$	$T_{c+d}^{-1} = \tilde{X} \otimes \tilde{X}^{-1}$	$= I \Rightarrow$	$\beta(c,d) = 0$
$\times$		$\times$		$\times$		
$T_{a+c}^{-1} = Z \otimes \tilde{X}^{-1}$	$\times$	$T_{b+d}^{-1} = \tilde{X} \otimes Z^{-1}$	$\times$	$T_{a+b+c+d} = \tilde{Y}^{-1} \otimes \tilde{Y}$	$= I \Rightarrow$	$\beta(a+c,b+d) = 0$
$\shortparallel$		$\shortparallel$		$\shortparallel$		
I		I		$-I$		
$\Downarrow$		$\Downarrow$		$\Downarrow$		
$\beta(a,c) = 0$		$\beta(b,d) = 0$		$\beta(a+b,c+d) = d/2$		

Table 5.1: Mermin's square (shaded cells) generalized to arbitrary even dimension for the proof of Lemma 24. In each row and column of Mermin's square, the Pauli observables commute and imply a value of β as stated. For the definition of $\tilde{X}$ and $\tilde{Y}$ see text.

such a Wigner function follows by Observation 1 and Theorem 13.

For even local dimension, we have the following result.

Theorem 14. *For any system of $n \geq 2$ qudits with an even number d of levels, a pair of Wigner and effect function satisfying (OB), (SW1)–(SW4) that represents all Pauli measurements positively does not exist.*

To prove Theorem 14, we first establish the following fact.

Lemma 24. *If the dimension d is even, then for any number $n \geq 2$ of local systems it holds that $[\beta] \neq 0$.*

The proof of Lemma 24 uses the existence of a Mermin square for all even d and $n \geq 2$, yielding an alternate route to proving the nonexistence of noncontextual ontological representations [52, 181].

Proof of Lemma 24. The proof proceeds by a construction generalizing Mermin's square to arbitrary even dimension. For any even d, define

$$\tilde{X} = X^{d/2}, \tilde{Y} = \sqrt{\omega}^{d/2} X^{d/2} Z.$$

Now consider the generalized Mermin square shown in Tab. 5.1 and its equivalent topological reformulation depicted in Fig. 5.1. The overall strategy of the proof is to identify a 2-cycle F in the chain complex corresponding to Fig. 5.1 such that $\partial F = 0$ but $\beta(F) \neq 0$. Any such surface F implies that $[\beta] \neq 0$. Namely, assume $[\beta] = 0$, i.e., $\beta = d\gamma$ for some $\gamma \in C^1$. Then, for the above surface F, $\beta(F) = d\gamma(F) = \gamma(\partial F) = \gamma(0) = 0$, which contradicts the assumption $\beta(F) \neq 0$.

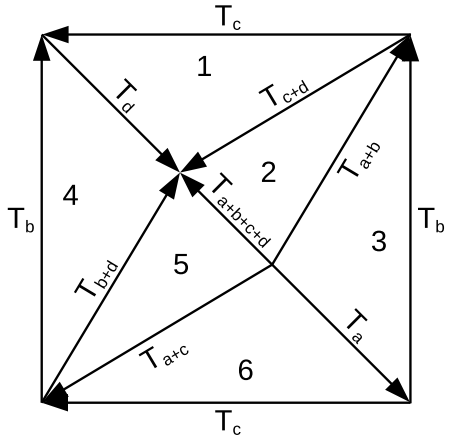


Figure 5.1: Topological reformulation of Mermin's square: each row and column of Mermin's square corresponds to the boundary of an elementary face f_j with $j \in \{1, \ldots, 6\}$ as indicated. The exterior edges are identified as shown. The arrows give an orientation to the edges. For the explicit expressions for the Pauli observables appearing in this figure, see Table 5.1.

The overall surface $F \in C_2$ of the torus in Fig. 5.1 has the required properties. More precisely, orienting F such that all boundaries point counterclockwise and labeling the elementary faces f_j as in Fig. 5.1, we observe that $F = \sum_{j=1}^{6} f_j$ with $f_1 = [c|d]$, $f_2 = [a + b|c + d]$, $f_3 = [a|b]$, $f_4 = -[b|d]$, $f_5 = -[a + c|b + d]$, and $f_6 = -[a|c]$, and $\partial F = 0$. It remains to be shown that $\beta(F) \neq 0$. First, the top row of the square in Tab. 5.1 reads $T_a T_b T_{a+b}^{-1} = I$; or equivalently, in the form matching Eq. (2.6), $T_a T_b = \omega^0 T_{a+b}$, such that $\beta(f_3) = 0$. Now we turn to the rightmost column of Tab. 5.1, which is $T_{a+b}^{-1} T_{c+d}^{-1} T_{a+b+c+d} = -I$. Again we transform this into the normal form of Eq. (2.6), which yields $\omega^{d/2} T_{a+b+c+d} = T_{a+b} T_{c+d}$ and $\beta(f_2) = d/2$. In the same fashion, we find $\beta(f_1) = \beta(f_4) = \beta(f_5) = \beta(f_6) = 0$. Indeed, $\partial F = 0$ and $\beta(F) \mod d = d/2$. Hence $[\beta] \neq 0$. □

Proof of Theorem 14. The statement is the combined conclusion of Lemma 24 and Theorem 13. □

5.6 Discussion

Here we have addressed the following question about quantum computation with magic states (QCM). For qudits in odd dimension, negativity in the Wigner function of the initial magic state is a precondition for quantum speedup [19].—Does the same hold in even dimension, e.g. for qubits?

This question has many facets. For example, it is known that if the notion of Wigner function is suitably generalized, namely to *nonunique* quasiprobability representations, then negativity in those representations can again be established as a precondition for quantum speedup [1, 21].

In this chapter, we have approached the question from a different angle. Namely, we have

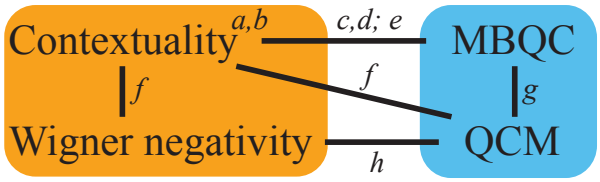


Figure 5.2: Canvas of cohomological properties in quantum computation with magic states (QCM) and measurement-based quantum computation (MBQC), relating to contextuality and Wigner function negativity. a=[53], b=[69]: cohomological formulation of contextuality; c=[25], d=[26]: MBQC is contextual; e=[70]: contextuality in temporally flat MBQCs is cohomological; f=[20]: Wigner function negativity and state-dependent contextuality with respect to Pauli observables are the same in odd dimension; contextuality of the magic states is a precondition for quantum speedup; g=[61]: computational power only resides in states and measurements, for both QCM and MBQC. h=[19]: in odd local dimension, quantum speedup in QCM requires Wigner function negativity. This result relies on the Clifford covariance of Gross' Wigner function, and its positive representation of Pauli measurement. [This work]: In even local dimension, cohomological invariants obstruct the existence of Wigner functions that are Clifford covariant and represent Pauli measurement positively.

investigated the more conventional Wigner functions derived from operator bases, a class to which the original Wigner function [43] and Gross' Wigner function in odd dimension [31] belong. In the case of odd dimension relying on Gross' Wigner function (and likewise in [1, 21]), the key to establishing negativity as a precondition for quantum speedup are two structural properties of the quasiprobability representations involved—Clifford covariance and positive representation of Pauli measurement. Here we have shown that, in even dimension, no Wigner function constructed from operator bases has these important structural properties. The obstructions to the existence of such Wigner functions are cohomological.

Specifically, the results of this chapter are two-fold: First, for Wigner functions constructed from operator bases (and, in some cases, satisfying the Stratonovich-Weyl criteria in addition), we have formalized the obstructions to Clifford covariance and positive representation of Pauli measurement. These obstructions are cohomological in nature. Our most general result is Theorem 10, on obstructions to Clifford covariance. The only assumption made by the theorem is (OB), i.e., that the Wigner function in question is constructed from an operator basis.

Theorem 13 on the positive representation of Pauli measurement is the critical one from the perspective of QCM. Again, a cohomological obstruction is identified, but this time the theorem requires the Stratonovich-Weyl criteria as additional assumptions.

As the second set of our results, we have applied the above theorems to the case of even local dimension, where the cohomological obstructions don't vanish. Thereby we have extended the existing no-go results [52, 58, 61] on the positive representation of the stabilizer sub-theory of quantum mechanics in even dimensions. The most general result so far is [52], which shows that, under the assumption of diagram preservation, the stabilizer sub-theory

cannot be positively represented. From this result, we have removed the assumption of diagram preservation. Specifically, in Theorem 11 we have shown that, in all even dimensions, Clifford-covariant Wigner functions from operator bases do not exist. In Theorem 14 we have shown that—whenever $n \geq 2$—Wigner functions from operator bases that also satisfy the Stratonovich-Weyl criteria cannot represent Pauli measurement positively.

Taking a step back, we observed in the introduction that there is a web of cohomological facts connecting measurement-based quantum computation to contextuality[3], and beyond. This web of cohomological facts we have extended to a further computational scheme, quantum computation with magic states. Fig. 5.2 shows a map of these cohomological facts and their relations.

[3]Any specific MBQC is contextual, i.e, cannot be described by a noncontextual hidden variable model, if it computes a nonlinear Boolean function with sufficiently high probability of success [25, 26]

Chapter 6

On the extremal points of the Λ-polytopes and classical simulation of quantum computation with magic states

We investigate the Λ-polytopes, a convex-linear structure recently defined and applied to the classical simulation of quantum computation with magic states by sampling. There is one such polytope, Λ_n, for every number n of qubits. We establish two properties of the family $\{\Lambda_n, n \in \mathbb{N}\}$, namely (i) Any extremal point (vertex) $A_\alpha \in \Lambda_m$ can be used to construct vertices in Λ_n, for all $n > m$. (ii) For vertices obtained through this mapping, the classical simulation of quantum computation with magic states can be efficiently reduced to the classical simulation based on the preimage A_α. In addition, we describe a new class of vertices in Λ_2 which is outside the known classification. While the hardness of classical simulation remains an open problem for most extremal points of Λ_n, the above results extend efficient classical simulation of quantum computations beyond the presently known range.

This chapter has been published in Ref. [3].

6.1 Introduction

The question of precisely which property of quantum mechanics is responsible for the speedup in quantum computation to date remains open. Various candidates have been put forward, such as superposition and interference [182], entanglement [18, 183], and largeness of Hilbert space [14], but there is always a counterexample that stops broad generalization; see e.g. [40, 184–186]. To identify new candidates, or perhaps even refine the question, one may study classical simulation algorithms for quantum systems, and specifically how the efficiency of such algorithms breaks down when pushed towards the regime of universal quantum computation.

The classical simulation method of present interest is based on sampling from the Λ-polytope [2]. Before turning to it, to provide context, we briefly discuss other classical simulation methods for quantum computation, and the general structure that unifies them. Classical simulation techniques for quantum computation typically seek to exploit proximity to a simple, "effectively classical" reference point. The characteristic simplifying feature of such reference points varies; examples are locality [187], stabilizerness [40], matchgates [74, 188], and the positivity of Wigner functions [19, 20]. Furthermore, often a natural measure exists for the distance between the setting of interest and the classical reference point. This can be, in case of locality as the simplifying feature, the bond dimension in MPS states [18], or the graph-theoretic measure of rankwidth in classical simulation of measurement based quantum computation on graph states [183]. If stabilizer-ness is the simplifying feature, measures of state magic [19–21, 42, 49, 155, 189, 190] such as Wigner function negativity or stabilizer rank quantify the hardness of classical simulation.

Irrespective of the concrete approach, these classical simulation methods share four general features: (i) A sector in which the classical simulation is efficient, (ii) a physical property that characterizes this sector, (iii) a physical property of universal quantum computation that obstructs efficiency of the simulation method in the general case, and (iv) a distance measure from the setting of interest to its closest efficiently simulable setting, governing the hardness of classical simulation.

Here we are concerned with the recently introduced method of simulating quantum computation based on sampling from the Λ-polytope [2]. Being a sampling algorithm, it is closely related to the sampling algorithms invoking Wigner functions and other quasiprobability distributions [1, 19, 27, 28, 49, 57]. However, there is also an important difference: in this method, no negativity ever occurs! This prompts the question of where quantumness is hiding in this scenario. If quantumness cannot be attributed to negativity, then to what else? To summarize the state of knowledge about the classical simulation by sampling from the Λ-polytope [2], of the above general characteristics (i)–(iv), we presently only have a glimpse of (i). Namely, it is known that in the subpolytope spanned by the so-called cnc vertices, classical simulation is efficient [1]; and this includes the domain of the stabilizer formalism [40] as a strict subset. But it is presently unknown how far the efficiently simulable region extends, which physical property this efficiency is to be attributed to, and which opposing physical property would render the simulation of the general case inefficient.

This chapter is a first approach to the systematic study of the Λ-polytopes. We investigate the structure of the extremal points of the polytopes Λ_n for any number n of qubits, asking: *Can the vertices of Λ_n be constructed from smaller parts? And if so, can the classical simulation of quantum computation based on such vertices be reduced to the classical simulation of those parts?*—Here, we give affirmative first answers to these questions, extending the classically efficiently simulable sector of the model [2].

Summary of main results

The main results of this chapter are:

- Theorem 15 (Section 6.3): For every vertex $X \in \Lambda_m$, every projector onto an $(n-m)$-qubit stabilizer state $|\sigma\rangle$, $\Pi_\sigma = |\sigma\rangle\langle\sigma|$, and every n-qubit Clifford unitary U, it holds that $U(X \otimes \Pi_\sigma)U^\dagger$ is a vertex of Λ_n.

- Theorem 16 (Section 6.4): The classical simulation of quantum computation with magic states (QCM), with the n-qubit initial state $\tilde{X} = U(X \otimes \Pi_\sigma)U^\dagger$, can be efficiently reduced to the classical simulation of QCM with the m-qubit initial state X.

- Theorem 17 (Section 6.5): We describe a new family of vertices $X \in \Lambda_2$, i.e., of two qubits, which is outside the closed and noncontextual (cnc) classification [1]. It is the first explicitly known family, including the update under Pauli measurements, beyond cnc.

The significance of the first result is that whenever we learn about a new vertex of Λ_m, it implies an infinite family of vertices on $n > m$ qubits. The second result means that, for the entire infinite family of vertices resulting from X, the hardness of classical simulation of QCM is essentially given by the hardness of simulating QCM on X. This result, together with the new class of efficiently updateable vertices provided by the third result, extends the efficiently classically simulable sector for the present method.

Outline

This chapter is organized as follows. In Section 6.2 we review background material. We state the definition of the state polytopes Λ_n and provide a motivation for studying them. In Section 6.3 we show how to construct vertices of the state polytope Λ_n from vertices of the polytope Λ_m, for $n > m$. Section 6.4 is concerned with the efficiency of classical simulation of quantum computation with magic states. For the cases where a vertex under consideration is the result of the map defined by Theorem 15, we describe a reduction to the classical simulation of the preimage of the map. In Section 6.5, we describe a new family of two-qubit vertices that is outside the presently known cnc classification. We conclude in Section 6.6.

6.2 Background

6.2.1 The Λ polytopes

Cnc vertices. There is an infinite family of vertices of the Λ-polytopes, the so-called cnc vertices, which is completely understood. Any cnc vertex in $\mathcal{V}_n$ has the form

$$A_\Omega^\gamma = \frac{1}{2^n} \sum_{v \in \Omega} (-1)^{\gamma(v)} T_v \tag{6.1}$$

where $\Omega \subset E_n$ is a maximal cnc set, to be defined below. The set of n-qubit cnc-type vertices are denoted by $\mathcal{V}_n^{\text{cnc}}$.

"Cnc" in 'cnc set Ω' stands for "closed noncontextual", and means the following. (Closure) A set $\Omega \subset E_n$ is closed if $v, w \in \Omega$ and $[v, w] = 0$ implies that $v + w \in \Omega$. (Noncontextuality) Ω admits a consistent noncontextual value assignment $\gamma : \Omega \longrightarrow \mathbb{Z}_2$. That is, the eigenvalue found in the measurement of the Pauli observable T_v is $(-1)^{\gamma(v)}$. The consistency requirement is that for a triple of commuting Pauli observables T_v, T_w, T_{v+w}, with $T_{v+w} = (-1)^{\beta(v,w)} T_v T_w$, it holds that $\gamma(v + w) = \gamma(v) + \gamma(w) + \beta(v, w) \mod 2$.

The reason for imposing the consistency condition is that if the three observables T_v, T_w, T_{v+w} are simultaneously measured, on any quantum state, then the measurement outcomes of these observables satisfy the above constraint with certainty. Hence we require the same of all considered value assignments γ. The consistency condition also explains the need for the closure property. If $v, w \in \Omega$ then the expectation values for T_v, T_w are extremal. The expectation value for T_{v+w} must then also be extremal, hence we require $v + w \in \Omega$ for A^γ_Ω to reproduce this.

The following is known about the cnc vertices:

1. For any $n \in \mathbb{N}$, the convex hull of $\mathcal{V}^{\text{cnc}}_n$ is closed under Pauli measurement. That is, for any projector Π^s_a onto the eigenspace of T_a corresponding to the eigenvalue $(-1)^s$ it holds that if $\text{Tr}\big(\Pi^s_a A^\gamma_\Omega\big) > 0$ then $\Pi^s_a A^\gamma_\Omega \Pi^s_a$ is a probabilistic linear combination of cnc vertices (Theorem 2 of [1].)

2. The update of the pair (Ω, γ) under Pauli measurement is computationally efficient (Theorem 3 of [1]).

3. The phase space point operators A^γ_Ω are extreme points of the respective polytopes Λ_n [2, 72].

4. The cnc sets Ω are classified for all $n \in \mathbb{N}$ (Theorem 1 of [1]; also see Theorem 2 of [149]).

Item 3 provides an element of continuity between the simulation method [2] of present interest and the earlier simulation schemes based on Wigner functions [1, 19, 57]. Namely, the phase space point operators A^γ_Ω were initially devised for a classical simulation scheme [1] of quantum computation with magic states in which negativity can arise. The latter is the multiqubit counterpart to the initial work on the role of Wigner function negativity for quantum computation [19], which applies only to odd Hilbert space dimension.

Duality. Each polytope Λ_n is the polar dual of the n-qubit stabilizer polytope SP_n [72]. Therefore, every vertex of $\Lambda_n, \forall n \in \mathbb{N}$, corresponds to a linear inequality that bounds SP_n.

This has an implication for $n = 2$. Namely, the linear inequalities bounding SP_2 have been described in [105], see Table II and Eq. (22) therein. This immediately provides a classification of the vertices of Λ_2.

The vertex enumeration problem, namely the problem of computing the vertices of a polytope given a description as the intersection of the set of halfspaces, is a fundamental problem in computational geometry. Algorithms exist for this problem (e.g. lrslib [191] and cddlib [192]) with computational complexity which typically depends on the input size (the number of bounding halfspace inequalities), the output size (the number of vertices), and the dimension of the space. In our case, the dimension of the space $\text{Herm}_1(\mathbb{C}^{2^n})$ grows exponentially with the number n of qubits, and the input size grows like $2^{[1/2+o(1)]n^2}$ [41]. This makes numerical computation of the vertices of Λ_n difficult even for small n.

So far a complete enumeration has only been possible for up to $n = 2$[105]. There are 60 bounding inequalities of Λ_2 and 22320 vertices of Λ_2 which fall into 8 orbits under the action of the Clifford group.

It is possible that the Clifford group symmetry of Λ_n could be used to reduce the vertex enumeration problem to the problem of enumerating a representative vertex from each Clifford orbit [193]. It is unknown to the authors whether this reduction is sufficient to enumerate the vertices of Λ_n for $n \geq 3$.

6.2.2 Status of the tensor product

Given two extremal points $A_\alpha \in \Lambda_n$ and $A_\beta \in \Lambda_m$, is the tensor product $A_\alpha \otimes A_\beta$ an extremal point of Λ_{n+m}?—This is not generally the case; in fact, $A_\alpha \otimes A_\beta$ may not even be contained in Λ_n. We can show this by example. Consider the one-qubit vertex [2]

$$A_0 = \frac{I + X + Y + Z}{2}.$$

Then it holds that $A_0 \otimes A_0 \notin \Lambda_2$. To see this, consider the stabilizer (Bell) state $|B_{11}\rangle$ defined by the stabilizer relations $Z_1 \otimes Z_2 |B_{11}\rangle = -|B_{11}\rangle$, $X_1 \otimes X_2 |B_{11}\rangle = -|B_{11}\rangle$. With these relations, $\langle B_{11}|A_0 \otimes A_0|B_{11}\rangle = -1/2$. Hence, with Eq. 4.1, $A_0 \otimes A_0 \notin \Lambda_2$. In the terminology of the Pusey-Barrett-Rudolph (PBR) theorem [159], the HVM of Theorem 6 does not satisfy the condition of "preparation independence", thereby evading the consequences of the PBR theorem.

From the above discussion it appears that the tensor product does not play a particular role in the HVM described by Theorem 6. Or does it?

6.3 Mapping vertices of Λ_m to Λ_n

In the introduction we asked whether the extremal points of the state polytopes Λ_n can be constructed from smaller parts, and Section 6.2.2 concluded with the question of whether there is any significant role for the tensor product in the present formalism. It turns out that the two questions are related through a common answer.

In this section we construct a map $\Lambda_m \to \Lambda_n$, for any $m < n$, that sends a vertex of Λ_m to a vertex of Λ_n. Moreover, under this map $\mathcal{V}_m$ maps injectively into $\mathcal{V}_n$. For the statement of our result we recall the definition of stabilizer projectors. Let Π_J^s denote the projector corresponding to the pair (J, s) consisting of an isotropic subspace $J \subset E_n$ and a value assignment $s : J \to \mathbb{Z}_2$. It has the form

$$\Pi_J^s = \frac{1}{|J|} \sum_{v \in J} (-1)^{s(v)} T_v$$

where $|\,|$ denotes the number of elements. For each stabilizer state $|\sigma\rangle \in \mathcal{S}_n$ the projector $\Pi_\sigma = |\sigma\rangle\langle\sigma|$ can be uniquely written as a projector of the form Π_I^s for some maximal isotropic subspace $I \subset E_n$ and a value assignment $s : I \to \mathbb{Z}_2$.

Theorem 15. *Let $J \subset E_n$ be an isotropic subspace of dimension $d = n - m$ and $S \in Sp(\mathbb{Z}_2^{2n})$ such that $S(J_0) = J$ where $J_0 = Span(x_{m+1}, \cdots, x_n)$. Given value assignments $r_0 : J_0 \to \mathbb{Z}_2$ and $r : J \to \mathbb{Z}_2$ define a linear map*

$$\Phi_{J,r} : Herm(\mathbb{C}^{2^m}) \to Herm(\mathbb{C}^{2^n})$$

by the formula

$$\Phi_{J,r}(X) = U(X \otimes \Pi_{J_0}^{r_0})U^\dagger \tag{6.2}$$

where U is a Clifford unitary that implements S and satisfies $U\Pi_{J_0}^{r_0}U^\dagger = \Pi_J^r$. Then the map $\Phi_{J,r}$ satisfies the following properties:

1. *The image of $\Phi_{J,r}$ is given by $\{(\Pi_J^r Y \Pi_J^r)/Tr(Y\Pi_J^r) |\ Y \in \Lambda_n$ and $\mathrm{Tr}(Y\Pi_J^r) \neq 0\}$.*
2. *$\Phi_{J,r}$ is injective and maps a vertex $X \in \Lambda_m$ to a vertex of Λ_n.*
3. *A vertex X is of cnc-type (see Eq. (6.1)) if and only if $\Phi_{J,r}(X)$ is of cnc-type.*

Presently, very little is known about the vertices of the polytopes Λ_n, for $n \geq 3$; and Theorem 15 extends that knowledge base. Essentially, the only fact already known is that the cnc construction [1] provides vertices of Λ_n for all n. Cnc vertices have the property that they consist of a core tensored with a stabilizer tail, cf. Lemma 11 in [1]. Therein, the cores are related to Majorana fermions and the Jordan-Wigner transformation. Now, Theorem 15 establishes that *all* vertices of Λ_n, for any n, have the same overall structure: they consist of a core tensored with a stabilizer tail.

We now prepare for the proof of Theorem 15. We first consider the case $(J, r) = (J_0, r_0)$ so that S is the identity matrix, and also the Clifford unitary U is chosen to be the identity matrix. For $X \in \mathrm{Herm}(\mathbb{C}^{2^m})$ we can write

$$X = \frac{1}{2^m} \sum_{v \in E_m} \alpha_v T_v^{(m)}$$

and Eq. (6.2) implies that

$$\Phi_{J,r}(X) = X \otimes \Pi_J^r = \frac{1}{2^n} \sum_{v+u \in E_m + J} (-1)^{r(u)} \alpha_v T_{v+u}^{(n)}. \tag{6.3}$$

The first step is to show that $\Phi_{J,r}$ maps Λ_m into Λ_n. Given a value assignment $s : I \to \mathbb{Z}_2$ and a subspace $J \subset E_n$ we write $s|_{I\cap J}$ for the restriction of s to the intersection $I \cap J$. Given another value assignment $s' : I \to \mathbb{Z}_2$ the delta function $\delta_{s',s}$ takes the value 1 if $s' = s$ as functions on I and 0 otherwise.

Lemma 25. *Let I be a maximal isotropic subspace of E_n. Then*

$$\mathrm{Tr}(\Phi_{J,r}(X)\Pi_I^s) = \delta_{r|_{I\cap J}, s|_{I\cap J}} \frac{|I \cap J^\perp|}{2^n} \mathrm{Tr}\Big(X \Pi_{I\cap E_m}^{s|_{I\cap E_m}}\Big).$$

In particular, if $X \in \Lambda_m$ then $\Phi_{J,r}(X) \in \Lambda_n$.

Proof of Lemma 25. We have $\mathrm{Tr}(\Phi_{J,r}(X)) = 1$ since $\mathrm{Tr}(X) = 1$. We calculate

$$\begin{aligned}
\mathrm{Tr}(\Phi_{J,r}(X)\Pi_I^s) &= \frac{1}{2^n}\sum_{v+u\in E_m+J}(-1)^{r(u)}\alpha_v\,\mathrm{Tr}(T_{v+u}\Pi_I^s)\\
&= \frac{1}{2^n}\sum_{v+u\in E_m+J}(-1)^{r(u)}\alpha_v\frac{1}{|I|}\sum_{w\in I}(-1)^{s(w)}\,\mathrm{Tr}(T_{v+u}T_w)\\
&= \frac{|I\cap J|}{2^n}\sum_{v\in I\cap E_m}\alpha_v(-1)^{s(v)}\left(\frac{1}{|I\cap J|}\sum_{u\in I\cap J}(-1)^{r(u)}(-1)^{s(u)}\right)\\
&= \delta_{r|_{I\cap J},s|_{I\cap J}}\frac{|I\cap J|}{2^n}\sum_{v\in I\cap E_m}\alpha_v(-1)^{s(v)}\\
&= \delta_{r|_{I\cap J},s|_{I\cap J}}\frac{|I\cap J||I\cap E_m|}{2^n}\,\mathrm{Tr}\Big(X\Pi_{I\cap E_m}^{s|_{I\cap E_m}}\Big)\\
&= \delta_{r|_{I\cap J},s|_{I\cap J}}\frac{|I\cap J^\perp|}{2^n}\,\mathrm{Tr}\Big(X\Pi_{I\cap E_m}^{s|_{I\cap E_m}}\Big)\geq 0.
\end{aligned}$$

As a consequence of this calculation and the definition of Λ_n the image $\Phi_{J,r}(X)$ belongs to Λ_n for any $X\in\Lambda_m$. □

Next we show that if $X\in\Lambda_m$ is a vertex then $\Phi_{J,r}(X)$ is a vertex of Λ_n. To achieve this we will use the following characterization of vertices of a polytope:

Let P be a polytope in $\mathbb{R}^N$. Then a point $p\in P$ is a vertex if and only if there exists no nonzero $x\in\mathbb{R}^N$ such that $p\pm x\in P$ [194, page 18].

Our polytope Λ_n lives inside $\mathrm{Herm}_1(\mathbb{C}^{2^n})$ and we can identify $\mathrm{Herm}_1(\mathbb{C}^{2^n})$ with $\mathbb{R}^{2^{2n}-1}$ by choosing $\mathbb{1}_{2^n}/2^n$ to be the origin. Therefore to prove that $\Phi_{J,r}(X)\in\Lambda_n$ is a vertex we need to show that $\{\Phi_{J,r}(X)\pm Y\}\not\subset\Lambda_n$ for any $Y\in\mathrm{Herm}(\mathbb{C}^{2^n})$ of trace zero. Let us write

$$Y = \frac{1}{2^n}\sum_{0\neq v\in E_n}\beta_v T_v^{(n)} \tag{6.4}$$

and define another trace zero operator

$$\tilde{Y} = \frac{1}{2^m}\sum_{0\neq v\in E_m}\tilde{\beta}_v T_v^{(m)}. \tag{6.5}$$

The coefficients $\tilde{\beta}_v$ are given by $\mathrm{Tr}(Y_v\Pi_J^r)$ where $Y_v = 1/2^n\sum_{u\in J}\beta_{u+v}T_u^{(n)}\in\mathrm{Herm}(\mathbb{C}^{2^n})$. Furthermore, we have the relation

$$\tilde{\beta}_v = \frac{1}{|J|}\sum_{u\in J}\beta_{u+v}(-1)^{r(u)}. \tag{6.6}$$

Lemma 26. *Let $I'\subset E_m$ be a maximal isotropic subspace and $s':I'\to\mathbb{Z}_2$ be a value assignment. Then*

$$\mathrm{Tr}\Big(Y\Pi_{J+I'}^{r*s'}\Big) = \mathrm{Tr}\Big(\tilde{Y}\Pi_{I'}^{s'}\Big) \tag{6.7}$$

where $r * s' : (J + I') \to \mathbb{Z}_2$ is the value assignment defined by $r * s'(u + v) = r(u) + s'(v)$.

Proof of Lemma 26. We calculate

$$\begin{aligned}\mathrm{Tr}\Big(Y\Pi^{r*s'}_{J+I'}\Big) &= \frac{1}{|J||I'|}\sum_{v\in I'}\sum_{u\in J}\beta_{u+v}(-1)^{r(u)+s(v)}\\ &= \frac{1}{|I'|}\sum_{v\in I'}(-1)^{s(v)}\frac{1}{|J|}\sum_{u\in J}\beta_{u+v}(-1)^{r(u)}\\ &= \frac{1}{|I'|}\sum_{v\in I'}(-1)^{s(v)}\tilde{\beta}_v\\ &= \mathrm{Tr}\Big(\tilde{Y}\Pi^{s'}_{I'}\Big).\end{aligned}$$

For the third equality we use Eq. (6.6). □

We are now ready to give a proof of our main result, Theorem 15.

Proof of Theorem 15. We begin by proving Theorem 15 for $(J, r) = (J_0, r_0)$ in which case the map is given by Eq. (6.3). Part (1) follows immediately from comparing $\Pi^{r_0}_{J_0} Y \Pi^{r_0}_{J_0}$ after normalization with Eq. (6.3). For part (2) note that Lemma 25 says that $\Phi_{J,r}(\Lambda_m) \subset \Lambda_n$. Let X be a vertex of Λ_m. We will show that $\{\Phi_{J,r}(X) \pm Y\} \not\subset \Lambda_n$ for any $Y \in \mathrm{Herm}(\mathbb{C}^{2^n})$ of trace zero. Since X is a vertex either $X + \tilde{Y}$ or $X - \tilde{Y}$ lies outside of Λ_m, where $\tilde{Y} \in \mathrm{Herm}(\mathbb{C}^{2^m})$ is defined in Eq. (6.5). Without loss of generality assume that $X+\tilde{Y} \notin \Lambda_m$. This means that there is a pair (I', s') such that $\mathrm{Tr}\Big((X + \tilde{Y})\Pi^{s'}_{I'}\Big) < 0$. Lemma 25 with $I = J + I'$ and $s = r * s'$ gives

$$\mathrm{Tr}\Big(\Phi_{J,r}(X)\Pi^{r*s'}_{J+I'}\Big) = \mathrm{Tr}\Big(X\Pi^{s'}_{I'}\Big). \tag{6.8}$$

Using Eq. (6.8) and Eq. (6.5) we obtain

$$\mathrm{Tr}\Big((\Phi_{J,r}(X) + Y)\Pi^{r*s'}_{J+I'}\Big) = \mathrm{Tr}\Big((X + \tilde{Y})\Pi^{s'}_{I'}\Big) < 0,$$

thus $X + Y$ lies outside of Λ_n. This completes the proof of part (2). Next we prove part (3). If $X = A^\gamma_\Omega$ for some cnc set (Ω, γ) where $\Omega \subset E_m$ then Eq. (6.3) implies that

$$\Phi_{J,r}(A^\gamma_\Omega) = \frac{1}{2^n}\sum_{u+v\in J+E_m}(-1)^{r(u)+\gamma(v)}T_{u+v} = A^{r*\gamma}_{J+\Omega}.$$

Therefore we obtain a vertex of cnc-type. Conversely, if $\Phi_{J,r}(X)$ is a vertex of cnc-type given by $A^{\tilde{\gamma}}_{\tilde{\Omega}}$ for some cnc set $(\tilde{\Omega}, \tilde{\gamma})$ then again from Eq. (6.3) we see that $X = A^\gamma_\Omega$ where $\Omega = \tilde{\Omega} \cap E_m$ and γ is given by the restriction $\tilde{\gamma}|_\Omega$.

For the general case let $J \subset E_n$ be an arbitrary isotropic subspace and $r : J \to \mathbb{Z}_2$ be a value assignment. We will write J_0 for $\mathrm{Span}(x_{m+1}, \cdots, x_n)$ and $r_0 : J_0 \to \mathbb{Z}_2$ for the value assignment defined by $r_0(v) = 0$ for all $v \in J_0$. Let U be a Clifford unitary as described in

the statement of the Theorem. We think of $\Phi_{J,r}$ as the composite of two maps:

$$\Phi_{J,r}(X) = U\Phi_{J_0,r_0}(X)U^\dagger. \tag{6.9}$$

Part (1) follows from the relation $U\Pi^{r_0}_{J_0}U^\dagger = \Pi^r_J$. Part (2) holds since U acts on Λ_n by permuting its vertices [2]. Also this action maps a cnc-type vertex to a cnc-type vertex, which implies part (3).

There is a gap in this proof that remains to be filled. Denote the operator $\tilde{Y}$ above by $Y_{J,r}$, defined as

$$Y_{J,r} = \frac{1}{2^m}\sum_{0\neq v\in E_m}\beta^{J,r}_v T^{(m)}_v \tag{6.10}$$

where

$$\beta^{J,r}_v = \frac{1}{|J|}\sum_{u\in J}\beta_{u+v}(-1)^{r(u)}. \tag{6.11}$$

If $Y_{J,r}\neq 0$ then there exists (I',s') such that $\mathrm{Tr}\Big((X+Y_{J,r})\Pi^{s'}_{I'}\Big) < 0$ (without loss of generality we assume $X+Y_{J,r}\notin\Lambda_m$). Then using Lemmas 25 and 26 with $I = I'+J$ and $s = s'*r$ we have

$$\mathrm{Tr}((\Phi_{J,r}(X)+Y)\Pi^s_I) = \mathrm{Tr}\Big((X+Y_{J,r})\Pi^{s'}_{I'}\Big) < 0, \tag{6.12}$$

which lets us conclude that $\Phi_{J,r}(X)$ is a vertex.

The argument breaksdown for Y's such that $Y_{J,r}=0$. For such cases, the existence of (I',s') satisfying $\mathrm{Tr}\Big((X+Y_{J,r})\Pi^{s'}_{I'}\Big)<0$ is not guaranteed.

To fill this gap, first of all note that it suffices to choose $J=\langle x_n\rangle$ and $r=r_0$ where $r_0(x_n)=0$. That is, we take $n=m+1$. An arbitrary pair (J,r) can be first mapped to (J,r_0) where $J=\langle x_{m+1},\ldots,x_n\rangle$ and $r_0(x_i)=0$ for $m+1\leq i\leq n$ with a symplectic transformation implemented by a unitary. Then the Φ-map for $J=\langle x_{m+1},\ldots,x_n\rangle$ can be obtained by sequentially applying the Φ-maps for the cases $J=\langle x_{m+1}\rangle,\langle x_{m+2}\rangle,\langle x_n\rangle$. Let us write $\Phi=\Phi_{j,r}$ for simplicity of notation where $(J,r)=(\langle x_n\rangle,r_0)$. Assume that $0\neq Y\in\mathrm{Herm}_0(\mathbb{C}^{2^n})$ is such that $Y_{J,r}=0$. This implies that the coefficients in Eq. (6.4) satisfy

$$\beta_{x_n}=0 \quad\text{and}\quad \beta_{v+x_n}=-\beta_v, \quad \forall v\notin\{0,x_n\}.$$

Let r_1 denote the value assignment $r_1: J\to\mathbb{Z}_2$ defined by $r_1(x_n)=1$. First we observe that $Y_{J,r_1}\neq 0$. The coefficients of Y_{J,r_1} in the Pauli basis (see Eq. (6.11)) are given by

$$\beta^{J,1}_{x_n} = -\beta_{x_n}/2 = 0 \quad\text{and}\quad \beta^{J,1}_v = (\beta_v-\beta_{v+x_n})/2 = \beta_v.$$

Since $Y\neq 0$ at least one of $\beta_v\neq 0$ which implies that $Y_{J,r_1}\neq 0$. Now let (I',s') be a pair such tha

$$\mathrm{Tr}\Big((X+Y_{J,r_1})\Pi^{s'}_{I'}\Big)<0 \tag{6.13}$$

holds, such a pair exists since $Y_{J,r_1}\neq 0$ and $X+Y_{J,r_1}\notin\Lambda_n$ (again without loss of generality we assume $X+Y_{J,r_1}$ lies outside the polytope instead of $X-Y_{J,r_1}$). Finally we verify

Eq. (6.12) with $(I, s) = (I' + J, s' * r_1)$:

$$\begin{aligned}
\mathrm{Tr}((\Phi(X) + Y)\Pi_I^s) &= \mathrm{Tr}\Big((X \otimes \Pi_J^{r_0} + Y)\Pi_{I'}^{s'} \otimes \Pi_J^{r_1}\Big) \\
&= \mathrm{Tr}\Big(Y(\Pi_{I'}^{s'} \otimes \Pi_J^{r_1})\Big) \\
&= \mathrm{Tr}\Big(Y_{J,r_1}\Pi_{I'}^{s'}\Big) \\
&\leq \mathrm{Tr}\Big((X + Y_{J,r_1})\Pi_{I'}^{s'}\Big) < 0
\end{aligned}$$

In the second line we use $\Pi_{J,r_0}\Pi_{J,r_1} = 0$, in the third line we use Lemma 26 and in the fourth line we use $\mathrm{Tr}\big(X\Pi_{I',s'}\big) \geq 0$ and Eq. (6.13). Therefore $\Phi(X) + Y$ lies outside of Λ_n. □

6.4 Reduction of the classical simulation

We now turn to the second question posed in the introduction, namely whether the classical simulation of the update of the composite vertices—whose existence was established in Theorem 15—under the operations of QCM can be reduced to the classical simulation of their constituent parts. This is indeed the case, as Theorem 16 below demonstrates.

Theorem 16. *Any quantum computation in the magic state model (QCM) that operates on an initial state $U(X_A \otimes (\Pi_\sigma)_B)U^\dagger$, where $X \in \Lambda_m$, is an m-qubit vertex and $\Pi_\sigma := |\sigma\rangle\langle\sigma|$ is the projector on an $(n-m)$-qubit stabilizer state $|\sigma\rangle$ and U is an n-qubit Clifford unitary, can be efficiently reduced to a QCM on initial state X alone.*

The theorem says that supplementing a vertex X with a stabilizer state does not increase the computational power of QCM. A proof of a similar result is given in Ref. [111, §V] for the case where X_A in the statement of Theorem 16 is replaced by the state $|T\rangle\langle T|^{\otimes m}$.

Proof of Theorem 16. We start from the version of QCM where the quantum computation consists of a sequence of Pauli measurements. All Clifford unitaries can be propagated forward past the last measurement (conjugating the measured observables in the passing), and then discarded. Thus, without loss of generality we consider initial states of form $\tilde{X} = X_A \otimes (\Pi_\sigma)_B$.

Let $\mathcal{T}$ denote a finite sequence of Pauli observables on $A \otimes B$. We give an explicit procedure to replace the sequence $\mathcal{T}$ by an equivalent sequence $\tilde{\mathcal{T}}^{(A)}$ of observables that act only on the subsystem A. The proof is by induction, and the induction hypothesis is that, at time t, the sequence $\mathcal{T}_{\leq t}$ of measurements has been replaced by a computationally equivalent sequence $\tilde{\mathcal{T}}_{\leq t}^{(A)}$ of Pauli measurements on the register A only. This statement is true for $t = 0$, i.e., the empty measurement sequence. We now show that the above statement for time t implies the analogous statement for time $t + 1$.

At time t, the state of the quantum register evolved under the computationally equivalent measurement sequence $\tilde{\mathcal{T}}_{\leq t}^{(A)}$ is $\tilde{Y}(t) = Y(t)_A \otimes (\Pi_\sigma)_B$. We now consider the Pauli observable $T(t+1) \in \mathcal{T}$ to be measured next, and write $T(t+1) = R_A(t+1) \otimes S_B(t+1)$. There are two cases:

Case I: $T(t+1)$ *commutes with the entire stabilizer* $\mathcal{S}$ *of* $|\sigma\rangle$. Hence, also $S_B(t+1)$ commutes with $\mathcal{S}$. But then, either $S_B(t+1)$ or $-S_B(t+1)$ is in $\mathcal{S}$, and $S_B(t+1)$ may be replaced by its eigenvalue ± 1 in the measurement. Hence, the measurement of $T(t+1)$ is equivalent to the measurement of $\pm R_A(t+1)$.

Case II: $T(t+1)$ *does not commute with the entire stabilizer* $\mathcal{S}$ *of* $|\sigma\rangle$. Then, the measurement outcome s_{t+1} is completely random. Further, there exists a Clifford unitary V such that

$$V\mathcal{S}(t+1)V^\dagger = \langle X_{B:1}, X_{B:2}, .., X_{B:m}\rangle, \;\; VT(t+1)V^\dagger = Z_{B:1}.$$

Therefore, the state resulting from the measurement of $T(t+1)$, with outcome s_{t+1} on the state $\tilde{Y}(t)$ is the same state as the one resulting from the following procedure:

1. Apply the Clifford unitary V to $\tilde{Y}(t) = Y(t)_A \otimes (\Pi_\sigma)_B$, leading to

   $$V\tilde{Y}(t)V^\dagger = Y'(t) \otimes |\overline{+}\rangle\langle\overline{+}|_B,$$

   where $|\overline{+}\rangle_B := \bigotimes_{i\in B} |+\rangle_{B:i}$.

2. Measure $Z_{B:1}$ on $Y'(t) \otimes |\overline{+}\rangle\langle\overline{+}|_B$, with outcome s_{t+1}.

3. Apply $V^\dagger$.

Now, note that the measurement in Step 2, of the Pauli observable $Z_{B:1}$ is applied to the stabilizer state $|\overline{+}\rangle_B$. The result is $|\sigma'(t+1)\rangle = |s_{t+1}\rangle_{B:1} \bigotimes_{j=2}^m |+\rangle_{B:j}$; that is the first qubit of subsystem B is now in a Z-eigenstate, and the other qubits are unchanged. Therefore, after normalization, the effect of the measurement can be replaced by the unitary $(X_{B:1})^{s_{t+1}} H_{B:1}$.

Thus, the whole procedure may be replaced by the Clifford unitary $V^\dagger \; (X_{B:1})^{s_{t+1}} H_{B:1} V$. But Clifford unitaries don't need to be implemented. They are just propagated past the last measurement, thereby affecting the measured observables by conjugation whereby their Pauli-ness is preserved. In result, in Case II, the measurement of $T(t+1)$ doesn't need to be performed at all. It is replaced by a coin flip, and efficient classical post-processing of the subsequent measurement sequence.

We conclude that in both the cases I and II, given the induction assumption, the original measurement sequence $\mathcal{T}_{\leq t+1}$ can be replaced by a computationally equivalent measurement sequence $\tilde{\mathcal{T}}^{(A)}_{\leq t+1}$ acting on register A only. By induction, the complete measurement sequence $\mathcal{T}$ can be replaced by a computationally equivalent sequence $\tilde{\mathcal{T}}^{(A)}$ acting on A only.

Since the measurements $\tilde{\mathcal{T}}^{(A)}$ are applied to an unentangled initial state $X_A \otimes (\Pi_\sigma)_B$, the register B may be dropped without loss of information. Thus, any sequence of Pauli measurements on the initial state $X_A \otimes (\Pi_\sigma)_B$ is efficiently reduced to a Pauli measurement sequence of at most the same length on X alone. □

Discussion. Theorem 16 extends the previously known range of efficient classical simulation of QCM by sampling from the polytopes Λ_n. The prior result is Theorem 3 from [1]. It says that if a quantum state ρ can be represented as a probabilistic linear combination of

cnc vertices, and the probability distribution defining this expansion can be efficiently sampled from, then any QCM on ρ can be efficiently classically simulated. Classical simulation of probability distributions over stabilizer states is contained therein as a limiting case.

To apply the above Theorem 16, we define vertex classes

$$\mathcal{W}_K := \{U(A_\alpha \otimes \Pi_\sigma)U^\dagger |\, A_\alpha \in \Lambda_K, |\sigma\rangle \in \mathcal{S}_{n-K}, U \in \mathcal{C}\ell_n, n \geq K\}$$

consisting of a union of certain n-qubit vertices for $n \geq K$. Therein, Π_σ is the projector corresponding to a $(n-K)$-qubit stabilizer state $|\sigma\rangle$ and $\mathcal{C}\ell_n$ is the n-qubit Clifford group. Every class $\mathcal{W}_K$ contains of vertices on arbitrarily many qubits for $n \geq K$. The present extension of Theorem 3 of [1], based on Theorem 16 above, is the following.

Corollary 4. *Consider an n-qubit quantum state that can be expanded into a probabilistic linear combination of vertices from the set $\mathcal{W}_K$, for a given value of K. Then, the computational cost of simulating any QCM on ρ is polynomial in n.*

Note that the simulation is (likely to be) inefficient in K.

6.5 Beyond vertices of cnc-type

In the previous section we established that the complexity of simulating QCM for a composite vertex $U(X_A \otimes (\Pi_\sigma)_B)U^\dagger$ efficiently reduces to the simulation of QCM on X_A. But which vertices X can we actually put explicitly into this reduction?—To date, the only family of vertices which is explicitly described, including the update under Pauli measurements, are the vertices of cnc-type.

For a single qubit, all eight vertices are equivalent under Clifford transformations, and cnc. For two qubits, there are 8 Clifford-equivalence classes of vertices [72, 105], two of which are of cnc-type.

In this section, we give a complete characterization of the update rules under Pauli measurement for one class of two-qubit vertices that are not cnc. A distinguishing feature of the vertices considered is that the expectations $\langle T_v \rangle$ take values in $\{0, \pm 1/2, \pm 1\}$ whereas in the cnc case these expectations belong to $\{0, \pm 1\}$.

Specifically, we are concerned with the Clifford orbit of the vertex denoted by A_{α_0} whose coordinates in the Pauli basis (i.e. the expectations $\langle T_v \rangle$) are given as follows

II	IX	XI	XX	IZ	IY	XZ	XY	ZI	ZX	YI	YX	ZZ	ZY	YZ	YY
1	$-1/2$	$1/2$	0	$-1/2$	$-1/2$	-1	0	$1/2$	-1	$-1/2$	0	0	0	0	1

We know that there are 1920 vertices in this orbit (by computer calculation). The set of vertices in the Clifford orbit of A_{α_0} will be denoted by

$$\mathcal{O} = \{U A_{\alpha_0} U^\dagger |\, U \in \mathcal{C}\ell_n\}. \tag{6.14}$$

For our construction we will consider noncontextual subsets of E_2 that are not necessarily closed. Let Ω be a subset of E_n and $\gamma : \Omega \to \mathbb{Z}_2$ be a value assignment. Associated to (Ω, γ) we can define the operator A_Ω^γ as in Eq. (6.1).

Construction. The vertices in the Clifford orbit $\mathcal{O}$ given in Eq. (6.14) have the following form

$$A^{\gamma}_{I,\Omega} = A^{\gamma}_{I} + \frac{1}{4}(A^{\gamma'}_{\Omega} - A^{\gamma''}_{\Omega}). \tag{6.15}$$

where

C.1 $I \subset E_2$ is a maximal isotropic subspace.

C.2 $\gamma : I \to \mathbb{Z}_2$ is a value assignment.

C.3 The set Ω is constructed from a collection $\mathcal{C}$ described in (C.4) by the following formula

$$\Omega = E_2 - \Omega^{\perp} \quad \text{where} \quad \Omega^{\perp} = \left(\bigcup_{J \in \mathcal{C}} J \right) - \{0\}. \tag{6.16}$$

C.4 The collection $\mathcal{C} = \mathcal{C}(I)$ of maximal isotropics (see Fig. 6.1) is defined using the following rules

R.1 $\mathcal{C}$ contains I,

R.2 for each $J \in \mathcal{C}$ and $0 \neq v \in J$ exactly one of the two subspaces $J' \in \mathcal{I}(E_2) - \{J\}$ containing v is contained in $\mathcal{C}$.

C.5 The value assignments γ' and γ'' on Ω are uniquely specified by the requirements

V.1 $\gamma'(0) = \gamma''(0) = 0$,

V.2 $\gamma'(v) = 1 + \gamma''(v)$ for $v \in \Omega - \{0\}$,

V.3 $\gamma'(v) + \gamma'(w) + \beta(v,w) = \gamma(v+w)$ for $v, w \in \Omega$ such that $[v,w] = 0$ and $v + w \in I$. Note that (V.2) implies that γ'' also satisfies this property.

In summary a vertex in the Clifford orbit $\mathcal{O}$ is specified by $(I, \gamma, \mathcal{C})$. Counting these components: 15 maximal isotropics, 2^2 value assignments, 2^5 collections $\mathcal{C}$ we obtain a total of $1920 = 15 \times 2^7$ vertices. This covers all the vertices in the Clifford orbit of A_{α_0}.

The vertex A_{α_0} has the following parameters: $I = \langle x_1 + z_2, z_1 + x_2 \rangle$, $\mathcal{C}$ consists of the following collection

$$\{\langle x_1 + z_2, z_1 + x_2 \rangle, \langle x_1 + x_2, z_1 + z_2 \rangle, \langle x_1 + y_2, y_1 + z_2 \rangle, \langle z_1 + y_2, y_1 + x_2 \rangle, \langle y_1 + x_2, x_1 + y_2 \rangle, \langle z_1 + y_2, y_1 + z_2 \rangle\},$$

and the value assignments are given by

	II	IX	XI	XX	IZ	IY	XZ	XY	ZI	ZX	YI	YX	ZZ	ZY	YZ	YY
$(-1)^{\gamma}$	1						-1			-1						1
$(-1)^{\gamma'}$	1	-1	1		-1	-1			1		-1					
$(-1)^{\gamma''}$	1	1	-1		1	1			-1		1					

The set Ω can be calculated from $\mathcal{C}$ using Eq. (6.16) and it turns out to be $\{0, x_1, y_1, z_1, x_2, y_2, z_2\}$.

Our strategy for computing the update rules for $A^{\gamma}_{I,\Omega}$ is to consider the update of A^{γ}_{I} and $\frac{1}{4}(A^{\gamma'}_{\Omega} - A^{\gamma''}_{\Omega})$ in Eq. (6.15) separately. First one can be updated using the update rule [1]

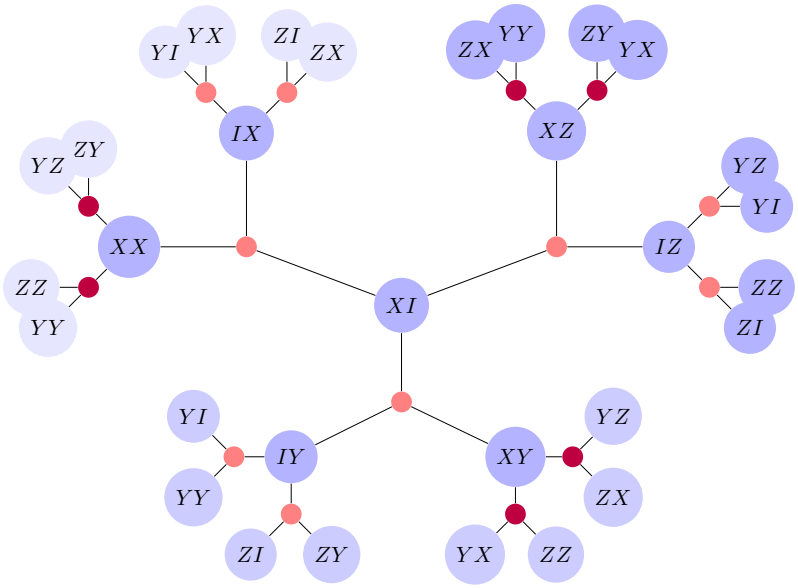


Figure 6.1: The poset of isotropic subspaces of E_2. Large blue nodes correspond 1-dimensional subspaces and small red nodes correspond to 2-dimensional subspaces. Each node at the boundary repeats 3 times and they are identified. The dark red colored nodes represent the maximal isotropic subspaces in the collection $\mathcal{C}$ corresponding to the vertex A_{α_0}.

of cnc-type vertices:

$$\Pi_a^s A_I^\gamma \Pi_a^s = \begin{cases} \delta_{s_a,\gamma(a)} A_I^\gamma & a \in I \\ \frac{1}{2} A_{I\times a}^{\gamma\times a} & a \notin I. \end{cases} \tag{6.17}$$

For the second one we first define a cnc set $\tilde{\Omega}$ obtained as the closure of $\Omega \cap \langle a\rangle^\perp$ under inference. We also extend γ' and γ'' on this cnc set in the expected way. The resulting value assignments are denoted by $\tilde{\gamma}'$ and $\tilde{\gamma}''$, respectively. Then the update of $\frac{1}{4}(A_\Omega^{\gamma'} - A_\Omega^{\gamma''})$ coincides with the update of $\frac{1}{4}(A_{\tilde{\Omega}}^{\tilde{\gamma}'} - A_{\tilde{\Omega}}^{\tilde{\gamma}''})$ by the properties of γ' and γ''. By construction $\tilde{\Omega}$ is contained in $\langle a\rangle^\perp$, in particular it contains a. Eq. (6.17) gives us

$$\Pi_a^s A_{\tilde{\Omega}}^{\tilde{\gamma}'} \Pi_a^s = \delta_{s_a,\tilde{\gamma}'(a)} A_{\tilde{\Omega}}^{\tilde{\gamma}'}.$$

Replacing $\tilde{\gamma}'$ with $\tilde{\gamma}''$ gives the update rule for $A_{\tilde{\Omega}}^{\tilde{\gamma}''}$. Next we give the update rules for $A_{I,\Omega}^\gamma$. The proof of this result is given in Section B.5.1.

Theorem 17. *The update rules for $A_{I,\Omega}^\gamma$ under Pauli measurements consist of the following three cases:*

Case I. Suppose that $a \in I$ and $a \notin \Omega$.

(1.a) If $\gamma(a) \neq s$ *then* $\Pi_a^s A_{I,\Omega}^{\gamma} \Pi_a^s = 0$.

(1.b) If $\gamma(a) = s$ *then*

$$\frac{\Pi_a^s A_{I,\Omega}^{\gamma} \Pi_a^s}{\mathrm{Tr}\Big(A_{I,\Omega}^{\gamma} \Pi_a^s\Big)} = \frac{2A_{\langle a\rangle^\perp}^{\tilde{\alpha}_0} + A_{\langle a\rangle^\perp}^{\tilde{\alpha}_1} + A_{\langle a\rangle^\perp}^{\tilde{\alpha}_2}}{4}$$

where the value assignments are given in Eq. (B.53).

Case II. Suppose that $a \notin I$ *and* $a \in \Omega$.

(2.a) If $\tilde{\gamma}'(a) \neq s$ *then*

$$\frac{\Pi_a^s A_{I,\Omega}^{\gamma} \Pi_a^s}{\mathrm{Tr}\Big(A_{I,\Omega}^{\gamma} \Pi_a^s\Big)} = A_{\langle a\rangle^\perp}^{\tilde{\alpha}'}$$

where the value assignments are given in Eq. (B.54).

(2.b) If $\tilde{\gamma}'(a) = s$ *then*

$$\frac{\Pi_a^s A_{I,\Omega}^{\gamma} \Pi_a^s}{\mathrm{Tr}\Big(A_{I,\Omega}^{\gamma} \Pi_a^s\Big)} = \frac{2A_{\langle a\rangle^\perp}^{\tilde{\alpha}} + A_{\langle a\rangle^\perp}^{\tilde{\alpha}'}}{3}$$

where the value assignments are given in Eq. (B.55).

Case III. Suppose that $a \notin I$ *and* $a \notin \Omega$.

(3.a) If $\tilde{\gamma}'(a) \neq s$ *then*

$$\frac{\Pi_a^s A_{I,\Omega}^{\gamma} \Pi_a^s}{\mathrm{Tr}\Big(A_{I,\Omega}^{\gamma} \Pi_a^s\Big)} = \frac{A_{\langle a\rangle^\perp}^{\tilde{\alpha}_0} + A_{\langle a\rangle^\perp}^{\tilde{\alpha}_1}}{2}$$

where the value assignments are given in Eq. (B.56).

(3.b) If $\tilde{\gamma}'(a) = s$ *then*

$$\frac{\Pi_a^s A_{I,\Omega}^{\gamma} \Pi_a^s}{\mathrm{Tr}\Big(A_{I,\Omega}^{\gamma} \Pi_a^s\Big)} = \frac{A_{\langle a\rangle^\perp}^{\tilde{\alpha}_0} + A_{\langle a\rangle^\perp}^{\tilde{\alpha}_1}}{2}$$

where the value assignments are given in Eq. (B.57).

6.6 Conclusion

This chapter is an expedition into the state polytopes Λ_n [2, 72], which are presently largely uncharted territory. The study of these polytopes is motived by the fact that they form the structural basis of a hidden-variable description of universal quantum computation [2].

Here we have shown that certain extremal points of the Λ_n can be built from "smaller parts", namely a vertex of state polytope Λ_m on a smaller number $m < n$ of qubits, and

a stabilizer state. Further, we have shown that the classical simulation of the evolution of such composite vertices can be reduced to the evolution of their parts. We have also described a new class of vertices outside the known cnc classification, together with their update rules under Pauli measurements.

A possible next question is as follows. The map Φ takes vertices of Λ_m to vertices of Λ_n, $n > m$, by tensoring on projectors onto stabilizer states. Are there generalizations of this map that use more general constructs than stabilizer states as parameters of the mapping, and more general notions of composition than the tensor product?

To conclude with a broader comment, the state polytopes Λ_n are a novel object in the theory of quantum computation and foundations of quantum mechanics. It is presently unknown how easy or hard their study is going to be, and which techniques will be useful. With the present investigation, we have made a first dent into the subject.

Chapter 7

Hidden variable model for quantum computation with magic states on qudits of any dimension

It was recently shown that a hidden variable model can be constructed for universal quantum computation with magic states on qubits. Here we show that this result can be extended, and a hidden variable model can be defined for quantum computation with magic states on qudits with any Hilbert space dimension. This model leads to a classical simulation algorithm for universal quantum computation.

This chapter has been published in Ref. [7].

7.1 Introduction

The field of quantum computation has seen an explosion of interest in recent years. It is widely believed that the era of quantum advantage is upon us and that we are entering the realm of so-called Noisy Intermediate-Scale Quantum (NISQ) computation. This view is evidenced by the impressive performance of quantum devices in recent hardware demonstrations.

But in spite of the age of the field and the recent surge in interest, a key question at the heart of quantum computation remains without an entirely satisfying answer: what is the essential quantum resource that provides the speedup of quantum computation over classical computation? This is clearly an important question as its resolution could inform the development of quantum hardware and the design of quantum computer architectures.

One inroad to approaching this question comes from quantum computation with magic states (QCM) [37, 38]. QCM—a universal model of quantum computation closely related to the circuit model—is one of the leading candidates for scalable fault-tolerant quantum computation [39]. In QCM, the allowed operations are restricted to a subset of unitary gates forming the Clifford group, as well as arbitrary Pauli measurements. These operations by themselves are not universal for quantum computation. In fact, any quantum circuit consisting of only these operations can be simulated efficiently on a classical computer [40,

41], and so with these operations alone no quantum computational speedup is possible. Universality is restored in QCM through the inclusion of additional nonstabilizer quantum states to the input of the circuit. Therefore, this model allows us to refine the question posed above. Instead of asking broadly "which nonclassical resources are required for a quantum computational speedup?", we can focus on the quantum states and ask "which states could provide a speedup in QCM?"

A partial answer to this question is provided by the study of quasiprobability representations like the Wigner function [43]. The Wigner function is the closest quantum mechanical counterpart to the classical notion of a probability distribution over a phase space, but unlike a probability distribution it can take negative values making it a *quasi*probability function. Accordingly, negativity in the Wigner function has traditionally been considered an indicator distinguishing classically behaving quantum states from those that exhibit genuinely quantum features [44, 45]. When adapted to finite-dimensional quantum mechanics, the setting relevant for quantum computation, quantum states are represented by a discrete Wigner function [27–31, 46, 47]—a quasiprobability function over a finite set (a generalized phase space) usually satisfying certain constraints [4, 48, 119].

Veitch et al. [19] showed that a necessary condition for a quantum computational speedup in QCM on odd-prime-dimensional qudits—quantum systems with odd prime Hilbert space dimension—is that the discrete Wigner function of the input state of the quantum circuit must take negative values (this result is easily extended to QCM on qudits with any odd dimension [9]). In particular, the amount of negativity in the discrete Wigner function quantifies the cost of classical simulation of a quantum computation [49] with simulation being efficient if the Wigner function is nonnegative everywhere. Since nonnegativity of the discrete Wigner function also implies the existence of a classical (noncontextual) hidden variable model (HVM) describing the computation [50, 51], this proves that two traditional notions of nonclassicality for quantum systems—Wigner negativity and failure of a classical HVM description—herald a quantum computational advantage over classical computation. This aligns with work that shows contextuality is required for quantum advantage in other settings [20, 26, 55].

The usual discrete Wigner function cannot be used to extend this result to QCM on even-dimensional qudits, including arguably the most important case—QCM on qubits [4, 52, 58, 71]. That said, similar necessary conditions for quantum advantage in QCM have been proven based on other quasiprobability representations [1, 21, 57, 59–61]. In all cases, negativity is required in the representation of states or measurements in order to describe universal quantum computation.

Recently, a hidden variable model was defined which bucks this trend by representing all quantum states, operations, and measurements relevant for QCM on qubits using only classical (nonnegative) probabilities [2]. This model is structurally similar to previous quasiprobability representations (modulo absence of negativity) and leads to a classical simulation method for universal quantum computation based on sampling from the defining probability distributions. In this chapter we show that this result can be significantly extended in that a nonnegative hidden variable model can be constructed for quantum computation with magic states on qudits of any dimension. We also show that many of the properties of the qubit HVM also apply in the qudit case, for example, this model leads to a classical simulation algorithm for quantum computation, and it subsumes previously

defined quasiprobability representations. Note that, although this model can simulate any quantum computation, the simulation is presumably inefficient in general.

Outline

The remainder of this chapter is structured as follows. In Section 7.2 we define the hidden variable model alluded to above. In Section 7.3 we present a classical simulation algorithm for quantum computation with magic states based on sampling from the probability distributions that define the model. In Section 7.4 we characterize a subset of the hidden variables of the model. Finally, in Section 7.5 we consider some properties of the qubit HVM [3] which extend to the qudit case. We conclude with a discussion of the significance of these results in Section 7.6.

7.2 Hidden variable model

In this section we define a hidden variable model that represents all components of quantum computation with magic states by a family of probability distributions. This is in contrast to previous quasiprobability representations which required negativity in the representation of either the states or the operations of QCM in order to represent universal quantum computation. The main result of this section is Theorem 18.

Let $\mathrm{Herm}(\mathcal{H})$ be the space of Hermitian operators on n-qudit Hilbert space $\mathcal{H} \cong \mathbb{C}^{d^n}$, $\mathrm{Herm}_1(\mathcal{H})$ be the affine subspace of this space obtained by fixing the trace of the operators to be 1, and let $\mathrm{Herm}_1^{\succeq 0}(\mathcal{H})$ be the subset of $\mathrm{Herm}_1(\mathcal{H})$ consisting of positive semidefinite operators. Let $\mathcal{S}$ denote the set of pure n-qudit stabilizer states. The state space of the hidden variable model is based on the set

$$\Lambda = \{X \in \mathrm{Herm}_1(\mathcal{H}) | \, \mathrm{Tr}(|\sigma\rangle \langle\sigma| \, X) \geq 0 \; \forall \, |\sigma\rangle \in \mathcal{S}\}. \tag{7.1}$$

The elements of Λ are much like density operators in that they are Hermitian operators with unit trace, but unlike density operators they are not necessarily positive semidefinite. In order to define the hidden variable model we first need to establish some basic properties of Λ.

Lemma 27. *For any number of qudits $n \in \mathbb{N}$ of any dimension $d \in \mathbb{N}$, (i) Λ is convex, and (ii) Λ is compact.*

The proof of Lemma 27 is in Section B.6.1.

Λ can be interpreted as a subset of a real affine space defined by the intersection of a finite number of linear inequalities, i.e., it is a polyhedral set. Since Λ is compact, it is a polytope, and so by the Minkowski-Weyl theorem [161] it can equivalently be described as the convex hull of finitely many vertices. Let $\{A_\alpha \mid \alpha \in \mathcal{V}\}$ denote the (finite) set of vertices of Λ. Then we have the following result, which is a generalization of [9, Theorem 1] to qudits of arbitrary Hilbert space dimension d.

Theorem 18. *For any number $n \in \mathbb{N}$ of qudits with any Hilbert space dimension $d \in \mathbb{N}$,*

1. *For any quantum state $\rho \in Herm_1^{\succeq 0}(\mathcal{H})$, there is a probability function $p_\rho : \mathcal{V} \to \mathbb{R}_{\geq 0}$ such that*
$$\rho = \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) A_\alpha. \tag{7.2}$$

2. *For any vertex A_α of Λ and any Clifford operation $g \in \mathcal{C}\ell$, $gA_\alpha g^\dagger =: A_{g\cdot\alpha}$ is a vertex of Λ.*

3. *For update under Pauli measurements it holds that for any isotropic subgroup $I \subset E$, any noncontextual value assignment $r : I \to \mathbb{Z}_d$, and any vertex A_α,*
$$\Pi_I^r A_\alpha \Pi_I^r = \sum_{\alpha' \in \mathcal{V}} q_{\alpha,I}(\alpha', r) A_{\alpha'}, \tag{7.3}$$
where $q_{\alpha,I}(\alpha', r) \geq 0$ for all $\alpha' \in \mathcal{V}$ and $\sum_{\alpha',r} q_{\alpha,I}(\alpha', r) = 1$.

4. *The Born rule takes the form*
$$\mathrm{Tr}(\Pi_I^r \rho) = \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) Q_I(r \mid \alpha) \tag{7.4}$$
where Q is given by
$$Q_I(r \mid \alpha) = \sum_{\alpha' \in \mathcal{V}} q_{\alpha,I}(\alpha', r). \tag{7.5}$$

This theorem defines a hidden variable model which represents all of the primitives of quantum computation with magic states using only probability distributions for states and measurements, and probabilistic update rules for dynamics. It has a similar structure to previous hidden variable models based on quasiprobability representations [1, 2, 9, 21, 29–31, 49, 51, 57, 137], with a key difference being that in this model every state can be represented by a probability distribution, no negativity is required. Unlike the model of Beltrametti and Bugajski [162] which requires a hidden variable for each pure quantum state, this model has a finite number of hidden variables for any number of qudits.

The proof of Theorem 18 requires the following lemma.

Lemma 28. *The polytope Λ has the following properties:*

1. *For any density operator $\rho \in Herm_1^{\succeq 0}(\mathcal{H})$ representing a physical n-qudit quantum state, $\rho \in \Lambda$.*

2. *For any $X \in \Lambda$ and any Clifford operation $g \in \mathcal{C}\ell$, $gXg^\dagger \in \Lambda$,*

3. *For any $X \in \Lambda$, any isotropic subgroup $I \subset E$, and any noncontextual value assignment $r : I \to \mathbb{Z}_d$, if $\mathrm{Tr}(\Pi_I^r X) > 0$ then*
$$\frac{\Pi_I^r X \Pi_I^r}{\mathrm{Tr}\big(\Pi_I^r X\big)} \in \Lambda. \tag{7.6}$$

Proof of Lemma 28. We will prove the three properties of the lemma in order. For the first property note that for any density operator $\rho \in \mathrm{Herm}_1^{\succeq 0}(\mathcal{H})$, ρ is positive semidefinite. Therefore, for any pure quantum state $|\psi\rangle$, $\mathrm{Tr}(|\psi\rangle\langle\psi|\,\rho) \geq 0$. This holds in particular for any pure stabilizer state. Therefore, ρ satisfies all of the defining inequalities of the polytope in Eq. (7.1) and so $\rho \in \Lambda$.

For the second property, let $X \in \Lambda$ and $g \in \mathcal{C}\ell$. Then for any stabilizer state $|\sigma\rangle \in S$,

$$\mathrm{Tr}\Big(|\sigma\rangle\langle\sigma|\,(gXg^\dagger)\Big) = \mathrm{Tr}\Big((g^\dagger|\sigma\rangle\langle\sigma|\,g)X\Big) \tag{7.7}$$

$$= \mathrm{Tr}\big(|\sigma'\rangle\langle\sigma'|\,X\big) \geq 0. \tag{7.8}$$

Here the first equality follows from the cyclic property of the trace, the second equality from the fact that Clifford operations map stabilizer states to stabilizer states (see Lemma 4), and the last inequality from the assumption $X \in \Lambda$.

Now we can prove the third property of Lemma 28. Let $I, J \subset E$ be isotropic subgroups with noncontextual value assignments $r : I \to \mathbb{Z}_d$ and $s : J \to \mathbb{Z}_d$. By Lemma 5,

$$\Pi_I^r \Pi_J^s \Pi_I^r = \delta_{r|_{I\cap J}, s|_{I\cap J}} \frac{|J \cap I^\perp|}{|J|} \Pi^{r\star s}_{I+(J\cap I^\perp)} \tag{7.9}$$

where $\delta_{r|_{I\cap J}, s|_{I\cap J}}$ is equal to one if r and s agree on the intersection $I \cap J$, and it is zero otherwise. Here $r \star s$ is the unique noncontextual value assignment on the set $I + (J \cap I^\perp)$ such that $r \star s|_I = r$ and $r \star s|_{J\cap I^\perp} = s|_{J\cap I^\perp}$. For any $X \in \Lambda$ and any Pauli projector Π_I^r, if $\mathrm{Tr}(\Pi_I^r X) > 0$, then for any projector onto a stabilizer state Π_J^s,

$$\mathrm{Tr}\left(\Pi_J^s \frac{\Pi_I^r X \Pi_I^r}{\mathrm{Tr}\big(\Pi_I^r X\big)}\right) = \frac{\mathrm{Tr}((\Pi_I^r \Pi_J^s \Pi_I^r)X)}{\mathrm{Tr}\big(\Pi_I^r X\big)} \tag{7.10}$$

$$= \delta_{r|_{I\cap J}, s|_{I\cap J}} \frac{|J \cap I^\perp|}{|J|} \frac{\mathrm{Tr}\Big(\Pi^{r\star s}_{I+(J\cap I^\perp)} X\Big)}{\mathrm{Tr}\big(\Pi_I^r X\big)} \geq 0. \tag{7.11}$$

Here the first line follows from linearity and the cyclic property of the trace. The last inequality follows from the fact that by Lemma 3, $\Pi^{r\star s}_{I+(J\cap I^\perp)}$ can be written as a conic combination of projectors onto stabilizer states, and from the assumption $X \in \Lambda$. Therefore, for any $X \in \Lambda$, any Pauli projector Π_I^r, and any stabilizer state $|\sigma\rangle \in \mathcal{S}$, if $\mathrm{Tr}(\Pi_I^r X) > 0$, then

$$\mathrm{Tr}\left(|\sigma\rangle\langle\sigma|\,\frac{\Pi_I^r X \Pi_I^r}{\mathrm{Tr}\big(\Pi_I^r X\big)}\right) \geq 0 \tag{7.12}$$

and so $\Pi_I^r X \Pi_I^r / \mathrm{Tr}(\Pi_I^r X) \in \Lambda$. This proves the third statement of the lemma. □

We can now prove the main result of this section.

Proof of Theorem 18. We will prove the four statements of the theorem in order. First, as shown in Lemma 28, Λ contains all density matrices corresponding to physical n-qudit quantum states. Therefore, by the Krein-Milman theorem [161] any state can be written as a convex combination of the vertices of Λ. This is the first statement of the theorem.

The second property from Lemma 28 shows that for any Clifford operation $g \in \mathcal{C}\ell$ and any vertex A_α of Λ, we have $A_{g\cdot\alpha} := gA_\alpha g^\dagger \in \Lambda$. It remains to show that $A_{g\cdot\alpha}$ is a vertex of Λ.

Let $\mathcal{S}_\alpha = \{|\sigma\rangle \in \mathcal{S} \mid \operatorname{Tr}(|\sigma\rangle\langle\sigma| A_\alpha) = 0\}$ be the set of stabilizer states with projectors orthogonal to vertex A_α with respect to the Hilbert-Schmidt inner product. By Theorem 18.1 of Ref. [160], since A_α is a vertex of Λ, A_α is the unique solution in $\mathrm{Herm}_1(\mathcal{H})$ of the system

$$\begin{cases} \operatorname{Tr}(|\sigma\rangle\langle\sigma| X) = 0 & \forall\, |\sigma\rangle \in \mathcal{S}_\alpha \\ \operatorname{Tr}(|\sigma\rangle\langle\sigma| X) \geq 0 & \forall\, |\sigma\rangle \in \mathcal{S} \setminus \mathcal{S}_\alpha. \end{cases} \tag{7.13}$$

For any stabilizer state $|\sigma\rangle \in \mathcal{S}_\alpha$,

$$\operatorname{Tr}(|\sigma\rangle\langle\sigma| X) = \operatorname{Tr}\Big(g|\sigma\rangle\langle\sigma| g^\dagger g X g^\dagger\Big). \tag{7.14}$$

Therefore, under conjugation by $g \in \mathcal{C}\ell$, solutions to the system Eq. (7.13) are mapped bijectively to solutions of the system

$$\begin{cases} \operatorname{Tr}(|\sigma\rangle\langle\sigma| X) = 0 & \forall\, |\sigma\rangle \in \mathcal{S}_{g\cdot\alpha} \\ \operatorname{Tr}(|\sigma\rangle\langle\sigma| X) \geq 0 & \forall\, |\sigma\rangle \in \mathcal{S} \setminus \mathcal{S}_{g\cdot\alpha} \end{cases} \tag{7.15}$$

where $S_{g\cdot\alpha} := \{g|\sigma\rangle \mid |\sigma\rangle \in S_\alpha\}$. In particular, $A_{g\cdot\alpha} := gA_\alpha g^\dagger$ is the unique solution to this system, so by Theorem 18.1 of Ref. [160], it is a vertex of Λ.

For the third statement of Theorem 18, let A_α be a vertex of Λ and Π_I^r be any Pauli projector. We have two cases. (I) First, if $\operatorname{Tr}(\Pi_I^r A_\alpha) = 0$, then $\Pi_I^r A_\alpha \Pi_I^r$ is zero as an operator. To see this, consider the inner product $\operatorname{Tr}(T_a \Pi_I^r A_\alpha \Pi_I^r)$ for any Pauli operator T_a, $a \in E$. Here we have three subcases: (i) if $a \in I$, then

$$\begin{aligned} T_a \Pi_I^r &= \frac{1}{|I|} \sum_{b \in I} \omega^{-r(b)} T_a T_b && (7.16) \\ &= \frac{1}{|I|} \sum_{b \in I} \omega^{-r(b) - \beta(a,b)} T_{a+b} && (7.17) \\ &= \frac{1}{|I|} \sum_{b \in I} \omega^{r(a) - r(a+b)} T_{a+b} && (7.18) \\ &= \omega^{r(a)} \Pi_I^r. && (7.19) \end{aligned}$$

Therefore,

$$\operatorname{Tr}(T_a \Pi_I^r A_\alpha \Pi_I^r) = \omega^{r(a)} \operatorname{Tr}(\Pi_I^r A_\alpha) = 0. \tag{7.20}$$

(ii) If $a \in I^\perp \setminus I$, then by Lemma 3 we have

$$\Pi_I^r = \sum_{\gamma \in \Gamma_{I,r}^{\langle a, I\rangle}} \Pi_{\langle a, I\rangle}^\gamma \tag{7.21}$$

where $\Gamma_{I,r}^{\langle a,I\rangle}$ is the set of noncontextual value assignments on $\langle a, I\rangle$ satisfying $\gamma|_I = r$. Multiplying this equation on the right by A_α and taking a trace we get

$$\operatorname{Tr}(\Pi_I^r A_\alpha) = \sum_{\gamma\in\Gamma_{I,r}^{\langle a,I\rangle}} \operatorname{Tr}\Big(\Pi_{\langle a,I\rangle}^{\gamma} A_\alpha\Big). \tag{7.22}$$

With Lemma 3, each projector on the right hand side can be written as a sum of projectors onto stabilizer states. Therefore, since $A_\alpha \in \Lambda$ each term on the right hand side is nonnegative. But the left hand side is zero by assumption. Therefore, each term on the right hand side is zero. I.e. $\operatorname{Tr}\Big(\Pi_{\langle a,I\rangle}^{\gamma} A_\alpha\Big) = 0$ for every $\gamma \in \Gamma_{I,r}^{\langle a,I\rangle}$.

We can write the spectral decomposition of the operator T_a as

$$T_a = \sum_{\gamma\in\Gamma_{I,r}^{\langle a,I\rangle}} \omega^{\gamma(a)} \Pi_{\langle a\rangle}^{\gamma|_{\langle a\rangle}}. \tag{7.23}$$

Then

$$\operatorname{Tr}(T_a\Pi_I^r A_\alpha \Pi_I^r) = \sum_{\gamma\in\Gamma_{I,r}^{\langle a,I\rangle}} \omega^{\gamma(a)} \operatorname{Tr}\Big(\Pi_{\langle a\rangle}^{\gamma|_{\langle a\rangle}} \Pi_I^r A_\alpha \Pi_I^r\Big) \tag{7.24}$$

$$= \sum_{\gamma\in\Gamma_{I,r}^{\langle a,I\rangle}} \omega^{\gamma(a)} \operatorname{Tr}\Big(\Pi_{\langle a,I\rangle}^{\gamma} A_\alpha\Big) = 0. \tag{7.25}$$

(iii) If $a \notin I^\perp$, then

$$\Pi_I^r T_a \Pi_I^r = \frac{1}{|I|^2} \sum_{b,c\in I} \omega^{-r(b)-r(c)} T_b T_a T_c \tag{7.26}$$

$$= \frac{1}{|I|^2} \sum_{b,c\in I} \omega^{-r(b)-r(c)+[b,a]-\beta(b,c)} T_a T_{b+c} \tag{7.27}$$

$$= \frac{1}{|I|^2} T_a \sum_{b,c\in I} \omega^{-r(b+c)+[b,a]} T_{b+c} \tag{7.28}$$

$$= \frac{1}{|I|} T_a \Pi_I^r \sum_{b\in I} \omega^{[b,a]}. \tag{7.29}$$

By character orthogonality, the sum in the final expression vanishes. Therefore,

$$\operatorname{Tr}(T_a\Pi_I^r A_\alpha \Pi_I^r) = \operatorname{Tr}(\Pi_I^r T_a \Pi_I^r A_\alpha) = 0. \tag{7.30}$$

We have $\operatorname{Tr}(T_a\Pi_I^r A_\alpha \Pi_I^r) = 0$ for every Pauli operator T_a, $a \in E$. Therefore, $\Pi_I^r A_\alpha \Pi_I^r$ is zero as an operator.

(II) Second, if $\operatorname{Tr}(\Pi_I^r A_\alpha) > 0$, then by the third statement of Lemma 28 we have $\Pi_I^r A_\alpha \Pi_I^r / \operatorname{Tr}(\Pi_I^r A_\alpha) \in \Lambda$, and so there exists a decomposition of $\Pi_I^r A_\alpha \Pi_I^r / \operatorname{Tr}(\Pi_I^r A_\alpha)$ as a convex combination of the vertices of Λ. Therefore, there exist nonnegative coefficients

$q_{\alpha,I}(\alpha', r)$ such that

$$\Pi_I^r A_\alpha \Pi_I^r = \sum_{\alpha' \in \mathcal{V}} q_{\alpha,I}(\alpha', r) A_{\alpha'}. \tag{7.31}$$

Taking a trace of this equation and adding the corresponding equations for all noncontextual value assignments r of I we have on the left hand side

$$\sum_r \mathrm{Tr}(\Pi_I^r A_\alpha \Pi_I^r) = \mathrm{Tr}\left[\left(\sum_r \Pi_I^r\right) A_\alpha\right] = \mathrm{Tr}(A_\alpha) = 1 \tag{7.32}$$

and on the right hand side

$$\sum_{\alpha',r} q_{\alpha,I}(\alpha', r) \,\mathrm{Tr}(A_{\alpha'}) = \sum_{\alpha',r} q_{\alpha,I}(\alpha', r). \tag{7.33}$$

Therefore, $\sum_{\alpha',r} q_{\alpha,I}(\alpha', r) = 1$. This proves the third statement of the Theorem.

Finally, we calculate

$$\begin{aligned} \mathrm{Tr}(\Pi_I^r \rho) &= \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) \,\mathrm{Tr}(\Pi_I^r A_\alpha) && (7.34)\\ &= \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) \sum_{\alpha' \in \mathcal{V}} q_{\alpha,I}(\alpha', r) && (7.35)\\ &= \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) Q_I(r \mid \alpha) && (7.36) \end{aligned}$$

and we obtain the fourth statement of the theorem. □

7.3 Classical simulation algorithm

Theorem 18 shows that all of the components of quantum computation with magic states can be described by a hidden variable model which represents all relevant states and dynamical operations by probabilities. This leads to a classical simulation algorithm for quantum computation with magic states, Algorithm 4, based on sampling from these probability distributions.

In short, a vertex label $\alpha \in \mathcal{V}$ is sampled according to the probability distribution of Eq. (7.2) representing the input state of the quantum circuit. This vertex is then propagated through the circuit. When a Clifford gate $g \in \mathcal{C}\ell$ is encountered, we have a deterministic update rule: $\alpha \to g \cdot \alpha$, according to the second statement of Theorem 18. When a Pauli measurement $a \in E$ is encountered, the third and fourth statements of Theorem 18 give a way of determining probabilities for measurement outcomes, as well as a probabilistic update rule. That is, we sample a pair (α', r) according to the probability distribution $q_{\alpha,\langle a\rangle}$ where $\alpha' \in \mathcal{V}$ and $r : \langle a\rangle \to \mathbb{Z}_d$ is a noncontextual value assignment. Then $r(a)$ is returned as the measurement outcome and the vertex is updated as $\alpha \to \alpha'$. This process continues until the end of the circuit is reached.

Input: p_{ρ_0}

sample a point $\alpha \in \mathcal{V}$ according to the probability distribution p_{ρ_0}
while end of circuit has not been reached **do**
  if a Clifford gate $g \in \mathcal{C}\ell$ is encountered **then**
    update $\alpha \leftarrow g \cdot \alpha$
  end if
  if a Pauli measurement $T_a,\ a \in E$ is encountered **then**
    sample (α', r) according to the probability distribution $q_{\alpha,\langle a\rangle}$
    Output: $r(a)$ as the outcome of the measurement
    update $\alpha \leftarrow \alpha'$
  end if
end while

Algorithm 4: One run of the classical simulation algorithm for quantum computation with magic states based on the hidden variable model of Theorem 18. The algorithm provides samples from the joint probability distribution of the Pauli measurements in a quantum circuit consisting of Clifford unitaries and Pauli measurements applied to an input state ρ_0.

A proof of the correctness of this simulation algorithm is given below.

Theorem 19. *The classical simulation algorithm, Algorithm 4, correctly reproduces the predictions of quantum theory.*

Proof of Theorem 19. Without loss of generality, a QCM circuit can be represented as a sequence $g_1, I_1, g_2, I_2, \ldots$ with $g_1, g_2, \cdots \in \mathcal{C}\ell$ specifying the Clifford unitaries to be applied, and $I_1, I_2, \cdots \subset E$ isotropic subgroups specifying the Pauli measurements to be performed. First, consider a single layer of this circuit consisting of a Clifford gate $g \in \mathcal{C}\ell$ followed by Pauli measurements corresponding to an isotropic subgroup $I \subset E$.

Using the classical simulation algorithm, Algorithm 4, the conditional probability of obtaining measurement outcomes specified by the noncontextual value assignment $r : I \to \mathbb{Z}_d$ for the measurements given the state $\alpha \in \mathcal{V}$ is $Q_I(r \mid g \cdot \alpha)$. Therefore, the probability of obtaining outcomes r given the gate g is applied to the state ρ followed by the measurements of the Pauli observables in I is given by

$$P^{(Sim)}_{\rho,g,I}(r) = \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) Q_I(r \mid g \cdot \alpha). \tag{7.37}$$

The corresponding outcome probability predicted by the Born rule, $P^{(QM)}_{\rho,g,I}(r)$, is

$$\operatorname{Tr}\left(\Pi^r_I g \rho g^\dagger\right) = \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) \operatorname{Tr}\left(\Pi^r_I g A_\alpha g^\dagger\right) \tag{7.38}$$

$$= \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) \operatorname{Tr}(\Pi^r_I A_{g\cdot\alpha}) \tag{7.39}$$

$$= \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) Q_I(r \mid g \cdot \alpha). \tag{7.40}$$

Here in the first line we use the expansion of ρ in the vertices of Λ, Eq. (7.2), in the second line we use the second statement of Theorem 18, and in the last line we use the fourth statement of Theorem 18. This agrees with the outcome probability predicted by the classical simulation algorithm.

Now consider the postmeasurement state ρ'. According to quantum mechanics, the postmeasurement state is

$$\rho'^{(QM)} = \frac{\Pi_I^r g\rho g^\dagger \Pi_I^r}{\mathrm{Tr}\big(\Pi_I^r g\rho g^\dagger\big)}. \tag{7.41}$$

Here the numerator is

$$\Pi_I^r g\rho g^\dagger \Pi_I^r = \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha)\Pi_I^r g A_\alpha g^\dagger \Pi_I^r \tag{7.42}$$

$$= \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha)\Pi_I^r A_{g\cdot\alpha}\Pi_I^r \tag{7.43}$$

$$= \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha) \sum_{\alpha'\in\mathcal{V}} q_{g\cdot\alpha,I}(\alpha', r)A_{\alpha'} \tag{7.44}$$

and the denominator is

$$\mathrm{Tr}\Big(\Pi_I^r g\rho g^\dagger \Pi_I^r\Big) = \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha) \sum_{\alpha'\in\mathcal{V}} q_{g\cdot\alpha,I}(\alpha', r) \tag{7.45}$$

$$= \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha) Q_I(r \mid g\cdot\alpha), \tag{7.46}$$

so the postmeasurement state predicted by quantum theory is

$$\rho'^{(QM)} = \frac{\sum_\alpha p_\rho(\alpha) \sum_{\alpha'} q_{g\cdot\alpha,I}(\alpha', r)A_{\alpha'}}{\sum_\alpha p_\rho(\alpha) Q_I(r \mid g\cdot\alpha)}. \tag{7.47}$$

Using the classical simulation algorithm, the probability of obtaining outcomes r and state $A_{\alpha'}$ given a Clifford gate g followed by measurements of the Pauli observables I on state ρ is $P_{\rho,g,I}(\alpha', r) = P_{\rho,g,I}(\alpha'|r)P_{\rho,g,I}(r)$. But $P_{\rho,g,I}(\alpha', r) = \sum_\alpha p_\rho(\alpha)P_{g,I}(\alpha', r|\alpha) = \sum_\alpha p_\rho(\alpha) q_{g\cdot\alpha,I}(\alpha', r)$, and $P_{\rho,g,I}(\alpha'|r) = p_{\rho'}(\alpha')$. Therefore, the postmeasurement state predicted by the classical simulation algorithm is

$$\rho'^{(Sim)} = \sum_{\alpha'\in\mathcal{V}} p_{\rho'}(\alpha')A_{\alpha'} \tag{7.48}$$

$$= \sum_{\alpha'\in\mathcal{V}} \frac{P_{\rho,g,I}(\alpha', r)}{P_{\rho,g,I}(r)} A_{\alpha'} \tag{7.49}$$

$$= \sum_{\alpha'\in\mathcal{V}} \frac{\sum_\alpha p_\rho(\alpha) q_{g\cdot\alpha,I}(\alpha', r)}{\sum_\alpha p_\rho(\alpha) Q_I(r \mid g\cdot\alpha)} A_{\alpha'}. \tag{7.50}$$

This agrees with the postmeasurement state predicted by quantum mechanics. Therefore, the classical simulation algorithm correctly reproduces the outcome probabilities and the postmeasurement state predicted by quantum mechanics for a single layer of a QCM circuit.

Now let $\rho(t)$ denote the state after t layers of the circuit. Then the argument above shows that the classical simulation algorithm correctly reproduces the Born rule probabilities $P_{\rho_0,g_{t+1},I_{t+1}}(r_{t+1} \mid r_1, r_2, \ldots, r_t)$ as well as the postmeasurement state $\rho(t+1)$. Therefore, by induction, the simulation algorithm correctly reproduces the outcome probabilities predicted by the Born rule for any QCM circuit. □

7.4 Partial characterization of vertices of Λ

In Ref. [1] a classical simulation algorithm for quantum computation with magic states is introduced based on a quasiprobability representation. Points in the generalized phase space over which the quasiprobability function is defined are associated with pairs (Ω, γ) where $\Omega \subset E$ is a cnc set and $\gamma : \Omega \to \mathbb{Z}_d$ is a noncontextual value assignment for Ω. For each point in phase space there is a corresponding phase space point operator defined as

$$A_\Omega^\gamma = \frac{1}{d^n} \sum_{b \in \Omega} \omega^{-\gamma(b)} T_b. \tag{7.51}$$

For qubits, if Ω is a maximal cnc set then phase space point operators A_Ω^γ of the form Eq. (7.51) are vertices of Λ [2, 72]. Vertices of the type Eq. (7.51) we call cnc vertices.

A similar result holds for odd-prime-dimensional qudits. Namely, phase space point operators of the form A_E^γ define facets of the stabilizer polytope [19, 28], or equivalently, by polar duality, A_E^γ are vertices of Λ. Here we show that this holds for qudits with any odd Hilbert space dimension. This is the result of the following theorem.

Theorem 20. *For any number n of qudits with any odd Hilbert space dimension d, phase space point operators of the form*

$$A_E^\gamma = \frac{1}{d^n} \sum_{b \in E} \omega^{-\gamma(b)} T_b \tag{7.52}$$

where $\gamma : E \longrightarrow \mathbb{Z}_d$ is a noncontextual value assignment for E are vertices of Λ.

This theorem provides a partial characterization of the vertices of Λ which define the hidden variable model of Theorem 18. When $n \geq 2$, the operators A_E^γ are exactly the phase space point operators of Gross' Wigner function for odd-dimensional qudits [29, 31, 51], therefore, the standard multiqudit phase space point operators are a subset of the hidden variables of our model (when $n = 1$, Gross' phase space point operators still have this form, but there are also other operators with this form [51]). Note that the phase space point operators of the form Eq. (7.52) only exist when the qudit Hilbert space dimension d is odd since noncontextual value assignments on E exist only when d is odd [69]. In the case of qubits, some additional classes of vertices have been characterized [69], but the vertices with the simplest description are the cnc type vertices of Eq. (7.51).

The proof of Theorem 20 requires the following lemma.

Lemma 29. *For any vertex A_α of Λ and any Pauli operator T_a, $|\operatorname{Tr}(T_a A_\alpha)| \leq 1$.*

Proof of Lemma 29. For any $\alpha \in \mathcal{V}$ and any $a \in E$,

$$\operatorname{Tr}(T_a A_\alpha) = \sum_{s \in \Gamma^{\langle a \rangle}} \operatorname{Tr}\Big(\Pi^s_{\langle a \rangle} A_\alpha\Big) \omega^{s(a)} \tag{7.53}$$

where $\Gamma^{\langle a \rangle}$ is the set of noncontextual value assignments on $\langle a \rangle$ and $\Pi^s_{\langle a \rangle}$ is the projector onto the eigenspace of the Pauli observable T_a with eigenvalue $\omega^{s(a)}$. $\{\Pi^s_{\langle a \rangle} | s \in \Gamma^{\langle a \rangle}\}$ is a projection-valued measure, i.e. $\sum_s \Pi^s_{\langle a \rangle} = \mathbb{1}$. Therefore,

$$\sum_{s \in \Gamma^{\langle a \rangle}} \operatorname{Tr}\Big(\Pi^s_{\langle a \rangle} A_\alpha\Big) = \operatorname{Tr}\left[\left(\sum_{s \in \Gamma^{\langle a \rangle}} \Pi^s_{\langle a \rangle}\right) A_\alpha\right] = 1 \tag{7.54}$$

and so

$$\begin{aligned} |\operatorname{Tr}(T_a A_\alpha)| &\leq \sum_{s \in \Gamma^{\langle a \rangle}} \left|\operatorname{Tr}\Big(\Pi^s_{\langle a \rangle} A_\alpha\Big) \omega^{s(a)}\right| && (7.55)\\ &\leq \sum_{s \in \Gamma^{\langle a \rangle}} \left|\operatorname{Tr}\Big(\Pi^s_{\langle a \rangle} A_\alpha\Big)\right| && (7.56)\\ &= \sum_{s \in \Gamma^{\langle a \rangle}} \operatorname{Tr}\Big(\Pi^s_{\langle a \rangle} A_\alpha\Big) = 1 && (7.57) \end{aligned}$$

which proves the claimed bound. □

We can now prove Theorem 20.

Proof of Theorem 20. First we need to show that the phase space point operators of Eq. (7.52) are in Λ. This requires checking that the Hilbert-Schmidt inner product of A^γ_E with the projector onto any stabilizer state is nonnegative. For any maximal isotropic subgroup I with any noncontextual value assignment $r : I \to \mathbb{Z}_d$,

$$\begin{aligned} \operatorname{Tr}\big(\Pi^r_I A^\gamma_E\big) &= \frac{1}{d^{2n}} \sum_{a \in I} \sum_{b \in E} \omega^{-r(a)-\gamma(b)} \operatorname{Tr}(T_a T_b) && (7.58)\\ &= \frac{1}{d^n} \sum_{a \in I} \omega^{-r(a)-\gamma(-a)} && (7.59)\\ &= \frac{1}{d^n} \sum_{a \in I} \omega^{-r(a)+\gamma(a)}. && (7.60) \end{aligned}$$

To obtain the last equality, note that with the phase convention chosen in Eq. (2.11) we have $T_{-a} = T_a^\dagger = T_a^{-1}$ and $T_0 = \mathbb{1}$. Therefore, $T_a T_{-a} = T_a T_a^{-1} = \mathbb{1} = T_{a+(-a)}$ and so by definition $\beta(a, -a) = 0$. Further, Eq. (3.5) applied to the case $a = b = 0$ implies $\gamma(0) = -\beta(0, 0) = 0$. Thus, with Eq. (3.5), we have $\gamma(a) + \gamma(-a) - \gamma(0) = -\beta(a, -a)$, and so $\gamma(-a) = -\gamma(a)$.

Here we have two cases.

(I) If $r|_I = \gamma|_I$ then we have

$$\mathrm{Tr}\big(\Pi_I^r A_\Omega^\gamma\big) = \frac{1}{d^n}\sum_{a\in I}\omega^{-r(a)+\gamma(a)} = \frac{|I|}{d^n} = 1. \tag{7.61}$$

since $|I| = d^n$ by [115, Theorem 1].

(II) If $r|_I \neq \gamma|_I$ then by orthogonality of twisted characters [195], we have

$$\mathrm{Tr}\big(\Pi_I^r A_E^\gamma\big) = \frac{1}{d^n}\sum_{a\in I}\omega^{-r(a)+\gamma(a)} = 0. \tag{7.62}$$

Therefore, in both cases $\mathrm{Tr}\big(\Pi_I^r A_E^\gamma\big) \geq 0$. This proves that the phase space point operators of Eq. (7.52) are in Λ.Since $A_E^\gamma \in \Lambda$, there exists an expansion of A_E^γ as a convex combination of the vertices of Λ:

$$A_E^\gamma = \sum_{\alpha\in\mathcal{V}} p(\alpha) A_\alpha. \tag{7.63}$$

Consider a Pauli operator $T_a,\ a \in E$. We have $\mathrm{Tr}\big(T_a A_E^\gamma\big) = \omega^{\gamma(a)}$ so multiplying Eq. (7.63) by T_a and taking a trace we get

$$\omega^{\gamma(a)} = \sum_{\alpha\in\mathcal{V}} p(\alpha)\,\mathrm{Tr}(T_a A_\alpha) \tag{7.64}$$

Taking the absolute value of this equation we get on the left hand side $\left|\omega^{\gamma(a)}\right| = 1$ and on the right hand side

$$\left|\sum_{\alpha\in\mathcal{V}} p(\alpha)\,\mathrm{Tr}(T_a A_\alpha)\right| \leq \sum_{\alpha\in\mathcal{V}} p(\alpha)|\,\mathrm{Tr}(T_a A_\alpha)| \tag{7.65}$$

$$\leq \sum_{\alpha\in\mathcal{V}} p(\alpha) = 1. \tag{7.66}$$

Here the first inequality is the triangle inequality and the second inequality follows from Lemma 29. The second inequality is strict if $|\,\mathrm{Tr}(T_a A_\alpha)| < 1$ for any A_α with $p(\alpha) > 0$. If this were the case then this would lead to a contradiction: $1 < 1$. Therefore, $|\,\mathrm{Tr}(T_a A_\alpha)| = 1$ for all A_α with $p(\alpha) > 0$.

Now consider the real part of the equation above:

$$\mathrm{Re}\left[\omega^{\gamma(a)}\right] = \sum_{\alpha\in\mathcal{V}} p(\alpha)\mathrm{Re}\left[\mathrm{Tr}(T_a A_\alpha)\right]. \tag{7.67}$$

Since the coefficients are nonnegative and sum to one, this implies that $\mathrm{Re}\left[\mathrm{Tr}(T_a A_\alpha)\right] = \mathrm{Re}\left[\omega^{\gamma(a)}\right]$ for every α with $p(\alpha) > 0$. The same argument holds for the imaginary part. Thus, $\mathrm{Tr}(T_a A_\alpha) = \omega^{\gamma(a)}$ for all α with $p(\alpha) > 0$. I.e. each A_α that appears with nonzero weight in the expansion of A_E^γ in Eq. (7.63) must agree with A_E^γ on the expectation of T_a for all $a \in E$. There is exactly one such operator in $\mathrm{Herm}_1(\mathcal{H})$, namely A_E^γ. Therefore, A_E^γ is a vertex of Λ. □

7.5 Mapping vertices of Λ_m to Λ_n

In this section we introduce a version of the Φ-map [3], that embeds the m-qudit polytope Λ_m as a subpolytope of the n-qudit polytope Λ_n where $n \geq m$, that works over qudits for an arbitrary $d \geq 2$. However, for $d \neq 2$ this map fails to map vertices of Λ_m to vertices of Λ_n.

We will regard E_m as a subgroup of E_n by identifying it with $\langle x_1, \cdots, x_m, z_1, \cdots, z_m \rangle$. We will write $E^{(n)}_{n-m}$ for the subgroup $\langle x_{m+1}, \cdots, x_n, z_{m+1}, \cdots, z_n \rangle$. These two subgroups intersect at the zero element and they generate the whole group. In this decomposition E_{n-m} is identified with $E^{(n)}_{n-m}$ via the map $x_i \mapsto x_{m+i}$ and $z_i \mapsto z_{m+i}$. Any isotropic subgroup $J \subset E_{n-m}$ will be identified with its image in $E^{(n)}_{n-m}$.

Theorem 21. *Let Π^r_J denote an $(n-m)$-qudit stabilizer projector and g denote an n-qudit Clifford unitary. The linear map*

$$\Phi^r_{g,J} : Herm\left(\mathbb{C}^{d^m}\right) \to Herm\left(\mathbb{C}^{d^n}\right), \quad X \mapsto g(X \otimes \Pi^r_J)g^\dagger \tag{7.68}$$

is injective and sends Λ_m to the subpolytope given by

$$\{(\Pi^{r'}_{J'} Y \Pi^{r'}_{J'})/Tr(Y\Pi^{r'}_{J'}) |\ Y \in \Lambda_n \text{ and } \mathrm{Tr}\left(Y\Pi^{r'}_{J'}\right) \neq 0\} \tag{7.69}$$

where $\Pi^{r'}_{J'} = g(\mathbb{1} \otimes \Pi^r_J)g^\dagger$.

This partially generalizes to qudits of arbitrary dimension a stronger result which applies only to the case $d = 2$ [69].

Proof of Theorem 21. To be able to distinguish Pauli operators we will write $T^{(n)}_a$ to indicate an n-qudit Pauli operator. With the choice of phase function ϕ in Eq. (2.11) we have (1) $T_0 = \mathbb{1}$, (2) $\beta(a, ka) = 0$ for all $a \in E_n$ and $k \in \mathbb{Z}_d$ and (3) $\beta(a, b) = 0$ for all $a \in E_m$ and $b \in E^{(n)}_{n-m}$. It suffices to prove the theorem for the case where (g, J, r) is given by $(\mathbb{1}, J_0, r_0)$. Here $\mathbb{1}$ is the identity operator, $J_0 = \langle x_{m+1}, x_{m+2}, \cdots, x_n \rangle$, and r_0 is the value assignment defined by $r_0(x_i) = 0$ for all $1 \leq i \leq n-m$. This is because we can write

$$\Phi^r_{g,J}(X) = g\Phi^r_{\mathbb{1},J}(X)g^\dagger \tag{7.70}$$

$$= g(\mathbb{1} \otimes W)\Phi^{r_0}_{\mathbb{1},J_0}(X)(\mathbb{1} \otimes W^\dagger)g^\dagger \tag{7.71}$$

$$= V\Phi^{r_0}_{\mathbb{1},J_0}(X)V^\dagger \tag{7.72}$$

where W is the Clifford unitary such that $W\Pi^{r_0}_{J_0}W^\dagger = \Pi^r_J$ and $V = g(\mathbb{1} \otimes W)$. For the rest we will take $(g, J, r) = (\mathbb{1}, J_0, r_0)$. Let us write $\Phi = \Phi^{r_0}_{\mathbb{1},J_0}$ for simplicity of notation.

The map Φ sends X to the tensor product $X \otimes \Pi^{r_0}_{J_0}$, and therefore it is linear because of the properties of the tensor product operation. For $X \in \mathrm{Herm}(\mathbb{C}^{d^m})$ we can write

$$X = \frac{1}{d^m} \sum_{a \in E_m} \alpha_a T^{(m)}_a \tag{7.73}$$

which gives

$$\Phi(X) = X \otimes \Pi_J^r = \frac{1}{d^n} \sum_{a+b \in E_m + J} \alpha_a T_{a+b}^{(n)}. \tag{7.74}$$

In particular, $\Phi(X) = 0$ implies that $X = 0$, hence Φ is injective. Next we show that Φ maps Λ_m into Λ_n. First note that $\mathrm{Tr}(\Phi(X)) = \alpha_0 = 1$ since $\mathrm{Tr}(X) = 1$. For a maximal isotropic subgroup $I \subset E_n$ we compute

$$\mathrm{Tr}(\Pi_I^s \Phi(X)) = \frac{1}{d^{2n}} \sum_{c \in I} \sum_{a+b \in E_m + J} \alpha_a \omega^{s(c)} \underbrace{\mathrm{Tr}\Big(T_c^\dagger T_{a+b}\Big)}_{d^n \delta_{c,a+b}} \tag{7.75}$$

$$= \frac{1}{d^n} \sum_{a+b \in (E_m+J) \cap I} \alpha_a \omega^{s(a+b)} \tag{7.76}$$

$$= \frac{|K \cap J|}{d^n} \sum_{a \in K \cap E_m} \alpha_a \omega^{s(a)} \left(\frac{1}{|K \cap J|} \sum_{b \in K \cap J} \omega^{s(b)} \right) \tag{7.77}$$

$$= \delta_{r|_{K\cap J}, s|_{K\cap J}} \frac{|K \cap J|}{d^n} \sum_{a \in K \cap E_m} \alpha_a \omega^{s(a)} \tag{7.78}$$

$$= \delta_{r|_{K\cap J}, s|_{K\cap J}} \frac{|K|}{d^n} \mathrm{Tr}\Big(\Pi_{K \cap E_m, s|_{K\cap E_m}} X\Big) \geq 0 \tag{7.79}$$

where $K = (E_m + J) \cap I$. Therefore $\Phi(X) \in \Lambda_n$. This image is a convex polytope since the image of a convex polytope under a linear map is also a convex polytope. We want to show that the image of Φ is given by

$$\{\, \Pi_J^r Y \Pi_J^r / \,\mathrm{Tr}(Y \Pi_J^r) \,|\, Y \in \Lambda_n \;\text{ and }\; \mathrm{Tr}(Y \Pi_J^r) \neq 0\}. \tag{7.80}$$

First observe that this set is contained in Λ_n by Theorem 18. Writing Y in the Pauli basis $Y = (\sum_{c \in E_n} \beta_c T_c^{(n)})/d^n$ we obtain

$$\Pi_J^r Y \Pi_J^r = \frac{1}{d^n} \sum_{c \in E_n} \beta_c \Pi_J^r T_c^{(n)} \Pi_J^r \tag{7.81}$$

$$= \frac{1}{d^n} \sum_{a+b \in E_m + J} \beta_{a+b} \Pi_J^r T_{a+b}^{(n)} \tag{7.82}$$

$$= \frac{1}{d^n} \sum_{a+b \in E_m + J} \beta_{a+b} \Pi_J^r (T_a^{(m)} \otimes T_b^{(n-m)}) \tag{7.83}$$

$$= \frac{1}{d^n} \sum_{a+b \in E_m + J} \beta_{a+b} T_a^{(m)} \otimes \Pi_J^r \tag{7.84}$$

$$= \frac{1}{d^m} \sum_{a \in E_m} \left(\frac{1}{d^{n-m}} \sum_{b \in J} \beta_{a+b} \right) T_a^{(m)} \otimes \Pi_J^r. \tag{7.85}$$

Defining X by setting $\alpha_a = \mathrm{Tr}(Y \Pi_J^r)(\sum_{b \in J} \beta_{a+b})/d^{n-m}$ in Eq. (7.73) this computation shows that $\Phi(X) = (\Pi_J^r Y \Pi_J^r)/\,\mathrm{Tr}(Y \Pi_J^r)$. Moreover, X belongs to Λ_m since for any isotropic

subgroup $I' \subset E_m$ and a value assignment s' defined on it we have that $\mathrm{Tr}\Big(X\Pi_{I'}^{s'}\Big) = \mathrm{Tr}\Big((X\Pi_{I'}^{s'}) \otimes \Pi_J^r\Big) = \mathrm{Tr}\Big(\Phi(X)(\Pi_{I'}^{s'} \otimes \Pi_J^r)\Big) \geq 0$ since $\Pi_{I'}^{s'} \otimes \Pi_J^r$ is a stabilizer projector and $\Phi(X) \in \Lambda_n$. This shows that the set in Eq. (7.80) is contained in the image of Φ. For the converse, we define Y by setting $\beta_{a+b} = \alpha_a$ for all $b \in J$. Then by the computation above we find that $\Pi_J^r Y \Pi_J^r = \Phi(X)$ with $\mathrm{Tr}(Y\Pi_J^r) = \mathrm{Tr}(\Phi(X)) = 1$. □

In contrast to the qubit case [3], the map Φ does not necessarily map a vertex of Λ_m to a vertex of Λ_n. For example, this is the case for odd qudit dimension d vertices of the form $A_{E_m}^{\gamma}$, as in (7.52).

To see this, consider $\Phi_{g,J}^r$ with $g = \mathbb{1}$. Analogously to Lemma 3, the operator

$$\Phi_{\mathbb{1},J}^r(A_{E_m}^{\gamma}) = A_{E_m}^{\gamma} \otimes \Pi_J^r \tag{7.86}$$

can be written as a proper convex combination of vertices of Λ_n:

$$A_{E_m}^{\gamma} \otimes \Pi_J^r = A_{E_m}^{\gamma} \otimes \left(\frac{1}{\left|\Gamma_{J,r}^{E_{n-m}}\right|} \sum_{\gamma' \in \Gamma_{J,r}^{E_{n-m}}} A_{E_{n-m}}^{\gamma'} \right) \tag{7.87}$$

$$= \frac{1}{\left|\Gamma_{J,r}^{E_{n-m}}\right|} \sum_{\gamma' \in \Gamma_{J,r}^{E_{n-m}}} A_{E_m}^{\gamma} \otimes A_{E_{n-m}}^{\gamma'} \tag{7.88}$$

$$= \frac{1}{\left|\Gamma_{J,r}^{E_{n-m}}\right|} \sum_{\gamma' \in \Gamma_{J,r}^{E_{n-m}}} A_{E_n}^{\gamma \star \gamma'}, \tag{7.89}$$

where $\Gamma_{J,r}^{E_{n-m}}$ is the set of noncontextual value assignments r on E_{n-m} such that $r_{|J} = r$ and $\gamma \star \gamma_0$ is the unique value assignment on E_n satisfying $(\gamma \star \gamma_0)_{|E_m} = \gamma$ and $(\gamma \star \gamma_0)_{|E_{n-m}} = \gamma_0$.

7.6 Discussion

In this chapter, we have presented a hidden variable model which represents all components of quantum computation with magic states using only classical (nonnegative) probabilities. In this model, magic states are represented by a probability distribution over a finite set according to Eq. (7.2). Clifford gates are represented by a deterministic update rule—a map from the set of hidden variables to itself. Pauli measurements are represented by a probabilistic update rule—a map from hidden variables to probability distributions over the set of hidden variables according to Eq. (7.3). This model is similar in form to many previously defined quasiprobability representations of quantum computation with magic states [1, 19, 21, 29–31, 57], but with the distinguishing feature that in our model all states can be represented by a probability distribution. No "negative probabilities" are required. This is the generalization of the hidden variable model of Ref. [2] to qudits of arbitrary local Hilbert space dimension.

Since everything is represented probabilistically, the model leads to a classical simulation algorithm for quantum computation with magic states based on sampling from the defining

probability distributions. This is Algorithm 4 presented in Section 7.3. This algorithm is similar in structure to simulation algorithms based on sampling from quasiprobability distributions like the discrete Wigner function [1, 9, 19, 21, 57], except that those algorithms are limited in their scope. They can only simulate quantum circuits for which the input state of the circuit is represented by a probability distribution. Since in our model any state can be represented by a probability distribution, we have no such limitation. Note that although Algorithm 4 can simulate any quantum computation, we make no claims that the simulation is efficient in general. In fact, if a quantum computational speedup over classical computation is possible at all, as many believe, then efficient classical simulation must fail.

There are, however, some important cases where the simulation is efficient. For example, in the qubit case, it is known that the (efficient) classical simulation algorithm of Ref. [1] is a special case of this more general model/simulation algorithm. This is a result of the fact that the phase space point operators of Eq. (7.51) are vertices of Λ [2, 72], these are the cnc-type vertices. In Section 7.4 we show that for higher-dimensional qudits as well there are vertices of Λ of cnc-type. We conjecture that the update of these vertices under Clifford gates and Pauli measurements will be efficiently computable classically. This would result in the simulation algorithm being efficient whenever the following two conditions are met: (i) the support of the probability distribution representing the input state of the circuit is restricted to vertices of this type and (ii) samples from the distribution representing the input state can be obtained efficiently.

Of course not all vertices of Λ are cnc-type. For example, Ref. [105] characterizes all vertices of the two-qubit Λ polytope. Under the action of the Clifford group there are eight orbits of vertices, only two of which are cnc-type vertices. Characterizing the remaining non-cnc-type vertices of Λ for arbitrary n and d is an open problem which could expand the scope of efficiency of Algorithm 4. This has already been partially achieved in the qubits case: Ref. [3] provides an efficient description, along with update rules under Pauli measurements, for one of the non-cnc-type orbits of the two qubit Λ polytope. As a result of the Φ-map, which maps vertices of the m-qubit polytope to vertices of the n-qubit polytope with $m < n$, this also provides a description of a non-cnc-type orbit of vertices for the Λ polytope on any number of qubits. Ref. [3] also provides a reduction of the simulation algorithm that shows that the Φ-map does not significantly increase the cost of classical simulation. Therefore, the scope of efficiency of Algorithm 4 in the qubit case goes beyond states supported on cnc-type vertices. In section 7.5 we show that a version of the Φ-map also holds for qudits of arbitrary dimension. We conjecture that a version of the corresponding reduction of the simulation algorithm also holds in this more general setting. This would align the model and simulation algorithm with the resource theory perspective in which preparation of stabilizer states is considered a "free" operation, along with Clifford gates and Pauli measurements [42].

We have seen evidence that the Λ polytopes could prove to be of independent interest. A subset of vertices, namely the cnc-type vertices, have already found a handful of applications [1, 72, 149]. The remaining vertices are less well studied but have proven useful in certain areas as well [3, 105]. Therefore, we conclude by proposing the family of Λ polytopes for arbitrary n and d as a subject of further study.

The main open questions regarding the Λ polytopes are "where does the efficiency of

classically simulating quantum computation end?" (assuming it does), and now that we have found it's not always negativity in a quasiprobability representation, which physical property is responsible for the breakdown of efficient classical simulation?". In this chapter we have shown that those questions do not only arise in the multiqubit case, but rather in quantum computation on qudits of any dimension.

Chapter 8

Efficient classical simulation of quantum computation beyond Wigner positivity

We present the generalization of the CNC formalism, based on closed and noncontextual sets of Pauli observables, to the setting of odd-prime-dimensional qudits. By introducing new CNC-type phase space point operators, we construct a quasiprobability representation for quantum computation which is covariant with respect to the Clifford group and positivity preserving under Pauli measurements, and whose nonnegative sector strictly contains the subtheory of quantum theory described by nonnegative Wigner functions. This allows for a broader class of magic state quantum circuits to be efficiently classically simulated than those covered by the stabilizer formalism and Wigner function methods.

This chapter has previously appeared in Ref. [8].

8.1 Introduction

Quasiprobability representations have long played an important role in physics bridging the gap between classical and quantum. Originally conceived by Wigner [43] as a means to adapt phase space methods of statistical mechanics to quantum physics, they have since found many applications in diverse contexts [45, 48, 116–118, 121, 131]. Notably, they have proven useful in describing and understanding quantum computation [19, 27–32].

Gross' Wigner function [29–31]—a Wigner function for systems of odd-dimensional qudits—has proven particularly effective for describing quantum computation. It defines a noncontextual (thus classical) model of the stabilizer subtheory [4, 20, 50, 52, 69], a realm where the function remains nonnegative. Negativity in this function has been proposed as the source of quantum computational power [27]. This aligns with the traditional view where, as negativity in the Wigner function is the feature that distinguishes it from a classical probability distribution over a phase space, it is considered an indicator of nonclassical behaviour [45].

Veitch et al [19] showed that negativity in this Wigner function is a necessary condition for a quantum computational advantage in the model of quantum computation with magic states (QCM) on odd-prime dimensional qudits. This is achieved by introducing an efficient classical simulation algorithm for QCM circuits that applies whenever the Wigner representation of the input state of the quantum circuit in question is nonnegative. Further, negativity in the representation of the input state can be used to quantify the cost of classical simulation, since in that case another sampling-based classical simulation algorithm can be applied where the number of samples required to achieve a given error scales with the amount of negativity [49].

These result can be easily extended to quantum computation on qudits of any odd dimension [9], but issues arise when attempting to extend them to computation on even-dimensional qudits. It was widely believed that no Wigner function exists which describes quantum computation on even-dimensional qudits, and this was proven in a few special cases [58, 63, 178]. This issue has since been settled for the general case [4, 52], but in the meantime a range of alternative quasiprobability representations have been defined [1, 21, 57, 60, 137, 147, 196]. These representations relax some of the assumptions that typically define a Wigner function [4, 48, 131], but they retain the properties required for classical simulation of quantum computation, namely, covariance with respect to the Clifford group and preservation of nonnegativity under Pauli measurements.

One of these recently introduced quasiprobability representations is the so-called CNC construction [1]. In the case of odd-dimensional qudits, phase space points of the Wigner function are identified with noncontextual value assignment functions on the set of Pauli observables [51], however, in the case of even dimensions, such noncontextual value assignments do not exist [4, 62, 69]. The CNC construction circumvents this issue by allowing the set of Pauli observables over which the value assignment is defined to also vary. That is, phase space points are identified with pairs (Ω, γ), where Ω is a noncontextual subset of the Pauli observables and γ is a noncontextual value assignment on Ω. "CNC" refers to these sets Ω, CNC stands for Closed under inference and NonContextual, two technical constraints we impose on the sets that we will define later. The result is a much larger phase space than a typical Wigner function would have, as required for consistency with a proven memory lower bound for simulating quantum contextuality [71], but the properties of the phase space required for simulating quantum computations are retained. This construction applies to any Hilbert space dimension, though so far it has mostly been studied for systems of multiple qubits [1].

In this chapter, we address the CNC construction for the case of odd-prime dimensional qudits. In Ref. [1], it was claimed that in this case the CNC construction is equivalent to Gross' Wigner function [29–31], but this was later shown to be incorrect [10]. Here we show that it can actually outperform Gross' Wigner function. That is, although the CNC phase space contains all of the points of the Wigner function phase space, our approach also involves the introduction of new CNC-type phase space points that are not equivalent to Wigner function phase space points, or convex mixtures thereof. The resulting quasiprobability representation obtained by introducing these new phase space points remains closed under Clifford operations and Pauli measurements, ensuring that the representation is capable of simulating quantum computations in the magic state model.

We provide a characterization of the CNC-type phase space points for odd-prime-

dimensional qudits. One of the primary differences between this case and the qubit case stems from the fact that there are no state-independent proofs of contextuality among Pauli observables on odd-dimensional qudits [4, 69]. As a result, only the closure under inference condition is required for the characterization, not the noncontextuality condition. Despite this difference, the classification of CNC-type phase space point operators in odd-prime-dimensional qudits looks very similar to the qubit case.

We also compare this construction with the Λ polytope model [2, 7], a recently introduced hidden variable model for QCM. In the multiqubit case, CNC-type phase space point operators are vertices of the Λ polytopes, and so they are identified with some of the hidden variables of that model [2, 72]. This is partially mirrored in the odd-prime-dimensional case where at least some of the CNC-type phase space point operators, including the phase space points of the Wigner function, are vertices of the qudit version of the Λ polytope model [7].

The primary advantage of this new construction lies in the expanded set of states that can be positively represented by the new quasiprobability function. Specifically, our construction goes beyond the capabilities of the previous Wigner function [19, 29–31] and stabilizer [40, 41, 101, 114] methods in identifying states that can be efficiently classically simulated. This expanded scope provides a more precise delineation of the classical-to-quantum transition in QCM on odd-prime-dimensional qudits, shedding light on the properties of magic states that enable a computational advantage in these systems and more generally offering new insights into the nature of computational resources in quantum computation with magic states.

Outline

The rest of this chapter is organized as follows. In Section 8.2, we define the CNC phase space, and we determine its relation to the phase space of the Wigner function for odd-prime-dimensional qudits. We also consider a simplification that removes redundancy in the phase space points for the case of multiple qubits, and we show why this reduction fails for odd-prime-dimensional qudits, thus proving that the CNC construction does not reduce to the Wigner function in odd dimensions. In Section 8.3, we provide a characterization of the CNC phase space for odd-prime-dimensional qudits, this characterization partially mirrors the structure of the Wigner function phase space [51], as well as that of the multiqubit CNC phase space [1, §IV].

In Section 8.4, we demonstrate that this new phase space is closed under Clifford gates and Pauli measurements, and we show how these facts can be exploited to define a classical simulation algorithm for quantum computation. We also demonstrate that this algorithm is efficient whenever the function representing the input state is nonnegative and samples from this distribution can be obtained efficiently. In Section 8.5, we show how the CNC construction relates to the Λ-polytope model [2, 7], a probabilistic representation of quantum computation with magic states. Finally, in Section 8.6, we conclude with a discussion of the implications of our results for the classical-to-quantum transition in QCM and propose directions for future research.

8.2 The CNC phase space

In this section we define a generalized phase space picture that can be used for describing quantum computation with magic states on qudits of any dimension. This phase space was first introduced in Ref. [1], and there it was characterized for the case of qubits. In this chapter we primarily consider the case of odd-prime dimensional qudits.

8.2.1 Definition of the phase space

Recall from Ref. [1] a few definitions.

Definition 11. *A set $\Omega \subset E$ is* closed under inference *if for every pair of elements $a, b \in \Omega$ satisfying $[a, b] = 0$, it holds that $a + b \in \Omega$.*

The closure under inference of a set $\Omega \subset E$, denoted $\overline{\Omega}$, is the smallest subset of E which is closed under inference and contains Ω. When using the language of the symplectic vector space E, we also refer to the closure under inference of a set Ω as the orthogonal closure. This is related to what is called a partial closure in Ref. [197].

Definition 12. *A* noncontextual value assignment *for a set $\Omega \subset E$ is a function $\gamma : \Omega \to \mathbb{Z}_d$ that satisfies*

$$\gamma(a) + \gamma(b) - \gamma(a + b) = -\beta(a, b) \tag{8.1}$$

for all pairs $a, b \in \Omega$ such that $[a, b] = 0$. A set $\Omega \subset E$ is called noncontextual *if there exists a noncontextual value assignment for the set.*[1]

As a result of the Mermin square proof of contextuality [62], the set E of all Pauli observables does not satisfy Definition 12 when the qudit Hilbert space dimension is $d = 2$ and the number of qudits is $n \geq 2$. Similar proofs of contextuality can be constructed using the multiqudit Pauli observables on qudits of any even dimension, but no such proofs exist for Pauli observables on odd-dimensional qudits [4, 69].

A set which is both closed under inference and noncontextual we call CNC for short. The CNC phase space $\mathcal{V}$ consists of all pairs (Ω, γ) where $\Omega \subset E$ is a CNC set and $\gamma : \Omega \to \mathbb{Z}_d$ is a noncontextual value assignment. For any CNC set Ω and any noncontextual value assignment $\gamma : \Omega \to \mathbb{Z}_d$, we define a phase space point operator

$$A^\gamma_\Omega = \frac{1}{d^n} \sum_{b \in \Omega} \omega^{-\gamma(b)} T_b. \tag{8.2}$$

These operators satisfy (QR1)–(QR3), i.e., they define a quasiprobability function where the representation of states is given by the coefficients in the expansion

$$\rho = \sum_{(\Omega,\gamma) \in \mathcal{V}} W_\rho(\Omega, \gamma) A^\gamma_\Omega. \tag{8.3}$$

In Ref. [1] it was shown that for multiple qubits, this representation also satisfies (QR4) and (QR5). For the case of odd dimensions, we return to the question of how the dynamics

[1] For odd d, $\beta \equiv 0$, so the noncontextuality condition simplifies to $\gamma(a) + \gamma(b) = \gamma(a + b)$ for all a, b such that $[a, b] = 0$.

of QCM, Clifford gates and Pauli measurements, are represented in the CNC model in Section 8.4.

8.2.2 Extremal phase space points

By including the CNC set Ω over which the noncontextual value assignments are defined as an extra varying parameter, the phase space becomes much larger than one would expect for a Wigner function. As a result, the representation of states in Eq. (8.3) is not unique. For the case of multiple qubits, when all pairs (Ω, γ) are included, the phase space contains redundancy. It is convenient to reduce the size of the phase space by eliminating this redundancy. This is achieved by the following lemma.

Lemma 30. *A CNC set $\Omega \subset E$ is called maximal if it is not strictly contained in a larger CNC set. Let $\mathcal{V}_M$ denote the multiqubit phase space consisting of pairs (Ω, γ) where Ω is a maximal CNC set, and γ is a noncontextual value assignment on Ω, satisfying Definitions 11 and 12. Then for any $(\tilde{\Omega}, \tilde{\gamma}) \in \mathcal{V}$ where $\tilde{\Omega}$ is not maximal, there are nonnegative coefficients $c(\Omega, \gamma) \geq 0 \; \forall (\Omega, \gamma) \in \mathcal{V}_M$ such that*

$$A_{\tilde{\Omega}}^{\tilde{\gamma}} = \sum_{(\Omega,\gamma)\in\mathcal{V}_M} c(\Omega, \gamma) A_{\Omega}^{\gamma}. \tag{8.4}$$

Further, a multiqubit state ρ is positively representable with respect to $\mathcal{V}$ if and only if it is positively representable with respect to $\mathcal{V}_M$.

The proof of this lemma was given first in Ref. [9, §3.3]. We include it here for completeness.

Proof of Lemma 30. For any CNC set $\tilde{\Omega}$ and value assignment $\tilde{\gamma}$ on $\tilde{\Omega}$, we have a phase space point operator $A_{\tilde{\Omega}}^{\tilde{\gamma}}$. If $\tilde{\Omega}$ is not a maximal set, then as a result of the classification of [1, Theorem 1], $\tilde{\Omega}$ has the form

$$\tilde{\Omega} = \bigcup_{k=1}^{\zeta} \langle a_k, I \rangle \tag{8.5}$$

where I is an isotropic subspace of E, all a_k commute with all elements of I, and all a_k pair-wise anti-commute. Also, $\tilde{\Omega}$ is contained in a maximal CNC set

$$\Omega = \bigcup_{k=1}^{\xi} \langle a_k, I \rangle \tag{8.6}$$

with the same commutation structure and with $\xi > \zeta$. Define two value assignments γ_0 and γ_1 on Ω as follows: $\gamma_0(b) = \gamma_1(b) = \gamma(b)$ for each $b \in \tilde{\Omega}$, and for each $b \in \{a_{\zeta+1}, \ldots, a_\xi\}$, define $\gamma_0(b) = 0$ and $\gamma_1(b) = 1$. The values of γ_0 and γ_1 on the remaining elements of $\Omega \setminus \tilde{\Omega}$

are determined by Eq. (8.1). Then,

$$\frac{1}{2}\left(A_\Omega^{\gamma_0} + A_\Omega^{\gamma_1}\right) = \frac{1}{2^{n+1}} \sum_{b\in\Omega} \left((-1)^{\gamma_0(b)} + (-1)^{\gamma_1(b)}\right) T_b \tag{8.7}$$

$$= \frac{1}{2^n} \sum_{b\in\tilde{\Omega}} (-1)^{\gamma(b)} T_b + \frac{1}{2^{n+1}} \sum_{b\in\Omega\setminus\tilde{\Omega}} \left((-1)^{\gamma_0(b)} + (-1)^{\gamma_1(b)}\right) T_b \tag{8.8}$$

For each $b \in \Omega \setminus \tilde{\Omega}$, we have one of two cases:
Case 1: If $b \in \{a_{\zeta+1}, \ldots a_\xi\}$ then by definition $\gamma_0(b) = 0$ and $\gamma_1(b) = 1$.
Case 2: If $b \notin \{a_{\zeta+1}, \ldots a_\xi\}$ then $b = a_j + g$ for some $j \in \{\zeta+1, \ldots, \xi\}$ and $g \in I$. Then, with all addition mod 2, we have

$$\gamma_0(b) = \beta(a_j, g) + \gamma_0(a_j) + \gamma(g) = \beta(a_j, g) + \gamma(g) \tag{8.9}$$

and

$$\gamma_1(b) = \beta(a_j, g) + \gamma_1(a_j) + \gamma(g) = \beta(a_j, g) + 1 + \gamma(g). \tag{8.10}$$

In both cases, $\gamma_1(b) \equiv \gamma_0(b) + 1 \mod 2$. Therefore, each term in the second sum in the expression above vanishes and we have

$$\frac{1}{2}\left(A_\Omega^{\gamma_0} + A_\Omega^{\gamma_1}\right) = \frac{1}{2^n} \sum_{b\in\Omega} (-1)^{\gamma(b)} T_b = A_\Omega^{\tilde{\gamma}}. \tag{8.11}$$

If a state has a nonnegative representation with respect to $\mathcal{V}_M$, then clearly it has a nonnegative representation with respect to $\mathcal{V}$ since $\mathcal{V}_M \subseteq \mathcal{V}$. Conversely, if a state ρ has a nonnegative expansion with respect to $\mathcal{V}$ with a positive coefficient on $A_\Omega^{\tilde{\gamma}}$, substituting the right hand side of Eq. (8.4) for $A_\Omega^{\tilde{\gamma}}$ does not introduce any negativity. Therefore, if a state is positively representable with respect to $\mathcal{V}$ then it is positively representable with respect to $\mathcal{V}_M$. □

This lemma relies on the fact that any noncontextual value assignment on $\tilde{\Omega}$ can be extended to a noncontextual value assignment on the larger CNC set Ω. This is true for qubits, but in general this extension is not possible for higher dimensional qudits. See Ref. [10] for an example of where this fails.

The only maximal CNC set for odd-dimensional qudits is the full set of Pauli observables E. If we were to restrict the phase space to include only maximal CNC sets, for $n \geq 2$ we obtain exactly the phase space of the Wigner function [51]. But since Lemma 30 fails for odd-dimensional qudits, we cannot restrict to maximal CNC sets without losing expressibility in terms of the quantum states that are postitively representable. In the next section we characterize all CNC phase space points for odd-prime-dimensional qudits. The phase space includes the phase space points of the Wigner function, but it also includes pairs (Ω, γ) where $\Omega \subsetneq E$ where γ cannot be extended to a noncontextual value assignment on E.

8.3 Characterization of phase space points

Our goal in this section is to characterize the CNC phase space for odd-prime-dimensional qudits, as was achieved for the case of qubits in Ref. [1, §IV]. This requires characterizing all pairs (Ω, γ), where $\Omega \subset E$ is a CNC set and $\gamma : \Omega \to \mathbb{Z}_d$ is a noncontextual value assignment for Ω, since points in the phase space are identified with these pairs. To start, we characterize the CNC sets.

8.3.1 Characterization of CNC sets

Note that there are no state-independent proofs of contextuality like the Mermin square [62] using only odd-dimensional Pauli observables [4, 52, 69]. Therefore, for the purpose of characterizing CNC sets, we can ignore the noncontextuality condition of Definition 12, and focus only on the closure under inference condition, Definition 11. We now characterize the sets $\Omega \subset E$ that are closed under inference. A similar characterization for the case $d = 2$ has already been found [1, 72, 149]. For the case of odd prime d, this is achieved by the following theorem.

Theorem 22. *For any number of qudits n of any odd prime dimension d, a set $\Omega \subset E$ is closed under inference if and only if*

(i) Ω is a subspace of E (i.e. Ω is closed under the addition), or

(ii) Ω has the form

$$\Omega = \bigcup_{k=1}^{\xi} \langle a_k, I \rangle \tag{8.12}$$

where $I \subset E$ is an isotropic subspace and the generators $a_1, \ldots, a_\xi$ satisfy $[a_i, a_j] \neq 0$ for all $i \neq j$, and $a_i \in I^\perp$ for all $i \in \{1, \ldots, \xi\}$.

The proof of this theorem requires a couple of lemmas.

Lemma 31. *Let $\Omega \subset E$ be a subspace and $v \in E$. Then $\overline{\Omega \cup \{v\}}$ has the form Eq. (8.12) if and only if $\Omega \cap v^\perp$ is isotropic. Otherwise, $\overline{\Omega \cup \{v\}} = \langle v, \Omega \rangle$.*

Lemma 32. *Suppose $\Omega \subset E$ has the form Eq. (8.12), and $v \in E$. Then $\overline{\Omega \cup \{v\}}$ has the form Eq. (8.12), or it is a subspace.*

The proofs of these lemmas are in Appendix B.7.1.

Proof of Theorem 22. It is easy to verify that subspaces and sets of the form Eq. (8.12) are closed under inference. The proof for the $d = 2$ case is given by Lemma 3 of Ref. [1]). The proof is identical for the case of odd prime d. Here we focus on the other direction, that every set that is closed under inference is either a subspace or has the form described by Eq. (8.12).

The overall proof strategy is induction. For the base case, start with one element $u \in E$. Obviously $\overline{\{u\}} = \langle u \rangle$, which is simultaneously a subspace and of the form Eq. (8.12). In the induction step, suppose that a set $\Omega \subset E$ is a subspace or of the form (8.12). Then we show that for all $v \in E$ the orthogonal closure $\overline{\Omega \cup \{v\}}$ is again a subspace or of the form (8.12). The induction step follows immediately from Lemmas 31 and 32. □

8.3.2 Maximal size of cnc sets

For a set Ω of the form Eq. (8.12), once the isotropic subspace I is fixed, the elements $a_1, \ldots, a_\xi$ are chosen from $I^\perp/I \simeq \mathbb{Z}_d^{2(n-\dim(I))}$. Then the size of the set is $(\xi d - \xi + 1) \cdot d^{\dim(I)}$. That is, it depends on the value of ξ, which is bounded by the maximal number of mutually nonorthogonal elements in $\mathbb{Z}_d^{2(n-\dim(I))}$. The maximal number of mutually nonorthogonal elements in $\mathbb{Z}_d^{2n}$ is unknown [198], but we can prove that it is at least $dn+1$.[2] This can be seen as a generalization of the corresponding qubit construction in [1, Theorem 1].

Let $e_1, \ldots, e_n, f_1, \ldots f_n$ be the standard basis of the symplectic vector space $E_n \simeq \mathbb{Z}_d^{2n}$ ($[e_i, e_j] = 0$, $[f_i, f_j] = 0$, and $[e_i, f_j] = \delta_{i,j}$). The space E_1 can be decomposed into $d+1$ mutually nonorthogonal lines as

$$E_1 = \langle c_1 \rangle \cup \cdots \cup \langle c_{d+1} \rangle, \quad [c_i, c_j] \neq 0 \text{ for all } i \neq j. \tag{8.13}$$

For example, such a decomposition if given by

$$E_1 = \langle e_1 \rangle \cup \langle f_1 \rangle \cup \langle e_1 + f_1 \rangle \cup \langle e_1 + 2f_1 \rangle \cup \cdots \cup \langle e_1 + (d-1)f_1 \rangle. \tag{8.14}$$

This gives rise to the following construction for any $n \in \mathbb{N}$ (which is in a sense a generalization of the construction given in [1, eqs. (16), (17)]). Labeling the elements of E_n by $(u_1, \ldots, u_n)$ with $u_i \in E_1$ for $i = 1, \ldots, n$. We have the following construction of $dn+1$ mutually nonorthogonal lines,

$$\begin{cases} (c_1, 0, \ldots, 0), & (c_2, 0, \ldots, 0), & \ldots, & (c_d, 0, \ldots, 0), \\ (c_{d+1}, c_1, 0 \ldots, 0), & (c_{d+1}, c_2, 0, \ldots, 0), & \ldots, & (c_{d+1}, c_d, 0, \ldots, 0), \\ (c_{d+1}, c_{d+1}, c_1, 0 \ldots, 0), & (c_{d+1}, c_{d+1}, c_2, 0, \ldots, 0), & \ldots, & (c_{d+1}, c_{d+1}, c_d, 0, \ldots, 0), \\ \vdots & & & \\ (c_{d+1}, c_{d+1} \ldots, c_{d+1}, c_1), & (c_{d+1}, c_{d+1} \ldots, c_{d+1}, c_2), & \ldots, & (c_{d+1}, c_{d+1} \ldots, c_{d+1}, c_d), \\ (c_{d+1}, c_{d+1} \ldots, c_{d+1}, c_{d+1}). & & & \end{cases} \tag{8.15}$$

It is straightforward to check that all of the given generators are mutually nonorthogonal. The symplectic inner product of two generators in the same row above is proportional to $[c_i, c_j] \neq 0$ for all $i \neq j$. The symplectic inner product of two elements that lie in different rows above will be proportional to $[c_i, c_{d+1}] \neq 0$ for some $i \in \{1, \ldots, d\}$.

8.3.3 Characterization of noncontextual value assignments on CNC sets

Theorem 22 gives a characterization of all CNC sets Ω. With a characterization of the noncontextual value assignments γ on Ω, we obtain a characterization of the phase space $\mathcal{V}$. The proof of [1, Lemma 2] can be adapted to show that the value assignments are a coset

[2] In order to prove efficiency of the classical simulation algorithm in Section 8.4, we need the number ξ of the noncommuting generators of the CNC sets to be polynomial in the number n of qudits. We expect the maximal number of pair-wise noncommuting elements in $\mathbb{Z}_d^{2n}$ to be bounded by a polynomial in n, but as noted this has not been proven. If it turns out that this conjecture is false, then in order for the algorithm to be efficient we must explicitly impose a polynomial bound on the value ξ for the admissible CNC phase space points used in the simulation. See Section 8.4 for details.

of a vector space, and in fact, since $\beta \equiv 0$ in odd dimensions, they are a vector space. This characterization can be made more explicit.

First, for the case $\Omega = E$ with $n \geq 2$, Ref. [51, Lemma 1] gives a characterization of the value assignments. The value assignments $\gamma : \Omega \to \mathbb{Z}_d$ are linear functions (equivalently, $\omega^{\gamma(\cdot)}$ are characters of Ω). When Ω is a subspace of E, the restriction of these characters to Ω are still value assignments. The single-qudit case is an exception, with more value assignments than characters (see [51, §3.4]). Second, in the case Ω has the form Eq. (8.12) with $\xi \geq 2$, we can describe the value assignments as follows. The value assignment γ can be chosen freely on $a_1, \ldots, a_\xi$, as well as on a basis for I. Then the value of γ on the rest of Ω is determined through Eq. (8.1).

If a value assignment on Ω happens to be linear, then it can be extended by linearity to a value assignment on E of all Pauli operators. In this case, Lemma 30 can be adapted to show that the corresponding phase space point (Ω, γ) is redundant (the proof given in [1, Lemma 1] holds). But if γ is not linear on Ω, then it cannot be extended to a value assignment on E, and the phase space point (Ω, γ) is not necessarily redundant. See Ref. [10] for an example of a phase space point (Ω, γ) with a nonlinear value assignment γ. Therefore, there exist phase space points in the CNC construction that do not correspond to phase space points of Gross' Wigner function or convex mixtures thereof, namely, those with $\Omega \subsetneq E$ and nonlinear value assignments $\gamma : \Omega \to \mathbb{Z}_d$.

8.4 Extended classical simulation algorithm

In this section we show that the CNC construction provides a classical simulation algorithm for quantum computation with magic states. To start we have show how CNC-type phase space point operators are updated under Clifford gates and Pauli measurements, and from these update rules, the classical simulation algorithm follows. It is given explicitly in Algorithm 5. Furthermore, with the characterization of the phase space given in Section 8.3, we are able to show that the update of phase space points under Clifford gates and Pauli measurements can be computed efficiently classically. This is the result of Theorem 23.

8.4.1 State update under Clifford gates

Using the action of the Clifford group on the Pauli operators in Eq. (2.9), for every CNC set Ω, each noncontextual value assignment $\gamma : \Omega \to \mathbb{Z}_d$, and every Clifford group element $g \in \mathcal{C}\ell$, we introduce the derived objects

$$g \cdot \Omega := \{S_g a \mid a \in \Omega\} \tag{8.16}$$

and $g \cdot \gamma : g \cdot \Omega \to \mathbb{Z}_d$ defined by

$$g \cdot \gamma(S_g b) = \gamma(b) - \Phi_g(b), \quad \forall b \in \Omega. \tag{8.17}$$

Lemma 33. *If $\Omega \subset E$ is CNC and $\gamma : \Omega \to \mathbb{Z}_d$ is a noncontextual value assignment for Ω, then for any Clifford group element $g \in \mathcal{C}\ell$,*

$$g A^\gamma_\Omega g^\dagger = A^{g\cdot\gamma}_{g\cdot\Omega}, \tag{8.18}$$

where $g \cdot \Omega$ is CNC and $g \cdot \gamma$ is a noncontextual value assignment for $g \cdot \Omega$.

Proof of Lemma 33. For any $a, b \in g \cdot \Omega$ with $[a, b] = 0$, there exist $c, d \in \Omega$ such that $a = S_g c$ and $b = S_g d$, and since $S_g \in \mathrm{Sp}(E)$, $[c, d] = 0$. Since Ω is closed under inference, $c + d \in \Omega$, therefore, $S_g(c + d) = a + b \in g \cdot \Omega$. Thus, $g \cdot \Omega$ is closed under inference. $g \cdot \gamma$ satisfies Eq. (8.1) by Lemma 4 in Ref. [69]. Therefore, $g \cdot \gamma : g \cdot \Omega \to \mathbb{Z}_d$ is a noncontextual value assignment for $g \cdot \Omega$.

Finally, we calculate

$$gA^{\gamma}_{\Omega}g^{\dagger} = \frac{1}{d^n} \sum_{b \in \Omega} \omega^{-\gamma(b)} g T_b g^{\dagger} \tag{8.19}$$

$$= \frac{1}{d^n} \sum_{b \in \Omega} \omega^{-\gamma(b) + \Phi_g(b)} T_{S_g b} \tag{8.20}$$

$$= \frac{1}{d^n} \sum_{b \in g \cdot \Omega} \omega^{-g \cdot \gamma(b)} T_b = A^{g \cdot \gamma}_{g \cdot \Omega} \tag{8.21}$$

and we obtain the final statement of the lemma. □

This lemma shows that the CNC phase space is closed under the action of the Clifford group. As a result, the corresponding quasiprobability function is Clifford covariant.

Corollary 5. *For any quantum state ρ and any Clifford group element $g \in \mathcal{C}\ell$, there exists a representation of the state $g\rho g^{\dagger}$ satisfying $W_{g\rho g^{\dagger}}(g \cdot \Omega, g \cdot \gamma) = W_{\rho}(\Omega, \gamma)$ for all $(\Omega, \gamma) \in \mathcal{V}$.*

The proof of this corollary follows immediately by conjugating the expression Eq. (8.3) by a Clifford group element $g \in \mathcal{C}\ell$. This also implies that the negativity in the representation does not increase under Cliffords.

Corollary 6. *Suppose a state ρ has an expansion as in Eq. (8.3) such that $W_{\rho}(\Omega, \gamma) \geq 0$ for all $(\Omega, \gamma) \in \mathcal{V}$. Then for any Clifford group element $g \in \mathcal{C}\ell$, there exists an expansion of $g\rho g^{\dagger}$ such that $W_{g\rho g^{\dagger}}(\Omega, \gamma) \geq 0$ for all $(\Omega, \gamma) \in \mathcal{V}$.*

8.4.2 State update under Pauli measurements

Lemma 34. *Let $\Omega \subset E$ be a CNC set and $\gamma : \Omega \to \mathbb{Z}_d$ be a noncontextual value assignement for Ω. Then for any isotropic subspace $I \subset E$ with any noncontextual value assignment $r : I \to \mathbb{Z}_d$, either*

(I) $\mathrm{Tr}\big(\Pi^r_I A^{\gamma}_{\Omega}\big) = 0$ *and* $\Pi^r_I A^{\gamma}_{\Omega} \Pi^r_I = 0$, *or*

(II) $\mathrm{Tr}\big(\Pi^r_I A^{\gamma}_{\Omega}\big) > 0$ *and*

$$\frac{\Pi^r_I A^{\gamma}_{\Omega} \Pi^r_I}{\mathrm{Tr}\big(\Pi^r_I A^{\gamma}_{\Omega}\big)} = A^{\gamma \times r}_{I + \Omega \cap I^{\perp}} \tag{8.22}$$

where $I + \Omega \cap I^{\perp}$ is CNC and $\gamma \times r : I + \Omega \cap I^{\perp} \to \mathbb{Z}_d$ is the unique noncontextual value assignment for $I + \Omega \cap I^{\perp}$ satisfying $\gamma \times r|_I = r$ and $\gamma \times r|_{\Omega \cap I^{\perp}} = \gamma|_{\Omega \cap I^{\perp}}$.

Proof of Lemma 34. The proof is similar to the third part of the proof of Theorem 1 in Ref. [7] and to the proof of Lemma 7 in Ref. [7], the difference being here A_Ω^γ is not in general a stabilizer code projector. We have

$$\begin{aligned}\operatorname{Tr}\big(\Pi_I^r A_\Omega^\gamma\big) =& \frac{1}{d^n|I|}\sum_{a\in I}\sum_{b\in\Omega}\omega^{-r(a)-\gamma(b)}\operatorname{Tr}(T_aT_b) && (8.23)\\ =& \frac{1}{d^n|I|}\sum_{a\in I}\sum_{b\in\Omega\cap I^\perp}\omega^{-r(a)-\gamma(b)}\operatorname{Tr}(T_aT_b)+\sum_{a\in I}\sum_{b\in\Omega\setminus(\Omega\cap I^\perp)}\omega^{-r(a)-\gamma(b)}\operatorname{Tr}(T_aT_b) && (8.24)\end{aligned}$$

The second sum vanishes because I is isotropic so I and $\Omega\setminus(\Omega\cap I^\perp)$ are disjoint, therefore, $\operatorname{Tr}(T_aT_b)=0$ for all $a\in I$, $b\in\Omega\setminus(\Omega\cap I^\perp)$. Continuing with the first sum we have

$$\begin{aligned}\operatorname{Tr}\big(\Pi_I^r A_\Omega^\gamma\big) =& \frac{1}{d^n|I|}\sum_{a\in I}\sum_{b\in\Omega\cap I^\perp}\omega^{-r(a)-\gamma(b)}\operatorname{Tr}(T_aT_b) && (8.25)\\ =& \frac{1}{|I|}\sum_{a\in\Omega\cap I}\omega^{-r(a)-\gamma(-a)} && (8.26)\\ =& \frac{1}{|I|}\sum_{a\in\Omega\cap I}\omega^{-r(a)+\gamma(a)}. && (8.27)\end{aligned}$$

Here we have two cases. First, if $r|_{\Omega\cap I}\neq\gamma|_{\Omega\cap I}$, then by orthogonality of twisted characters [195], we have $\operatorname{Tr}\big(\Pi_I^r A_\Omega^\gamma\big)=0$. Second, if $r|_{\Omega\cap I}=\gamma|_{\Omega\cap I}$ then

$$\operatorname{Tr}\big(\Pi_I^r A_\Omega^\gamma\big)=\frac{1}{|I|}\sum_{a\in\Omega\cap I}\omega^{-r(a)+\gamma(a)}=\frac{|\Omega\cap I|}{|I|}>0. \tag{8.28}$$

Therefore, either $\operatorname{Tr}\big(\Pi_I^r A_\Omega^\gamma\big)=0$, or $\operatorname{Tr}\big(\Pi_I^r A_\Omega^\gamma\big)>0$.[3] Here we have two cases.

(I) In the first case, we must show that $\Pi_I^r A_\Omega^\gamma \Pi_I^r = 0$. For any Pauli operator $T_a, a\in E$, consider the inner product $\operatorname{Tr}\big(T_a\Pi_I^r A_\Omega^\gamma\Pi_I^r\big)$. Here we have three subcases. (i) First, if $a\in I$ then

$$\begin{aligned}T_a\Pi_I^r =& \frac{1}{|I|}\sum_{b\in I}\omega^{-r(b)}T_aT_b && (8.29)\\ =& \frac{1}{|I|}\sum_{b\in I}\omega^{-r(b)-\beta(a,b)}T_{a+b} && (8.30)\\ =& \frac{1}{|I|}\sum_{b\in I}\omega^{r(a)-r(a+b)}T_{a+b} && (8.31)\\ =& \omega^{r(a)}\Pi_I^r. && (8.32)\end{aligned}$$

Therefore,

$$\operatorname{Tr}\big(T_a\Pi_I^r A_\Omega^\gamma\Pi_I^r\big)=\omega^{r(a)}\operatorname{Tr}\big(\Pi_I^r A_\Omega^\gamma\big)=0. \tag{8.33}$$

[3] This implies the operators A_Ω^γ are contained in the Λ polytopes of Refs. [2, 7].

(ii) Second, if $a \in I^\perp \setminus I$, then by Lemma 5 of Ref. [7]

$$\Pi_I^r = \sum_{r' \in \Gamma_{I,r}^{\langle I,a\rangle}} \Pi_{\langle I,a\rangle}^{r'} \tag{8.34}$$

where $\Gamma_{I,r}^{\langle I,a\rangle}$ is the set of noncontextual value assignments on $\langle I, a\rangle$ satisfying $r'|_I = r$. Multiplying this equation by A_Ω^γ and taking a trace we get

$$\mathrm{Tr}\big(\Pi_I^r A_\Omega^\gamma\big) = \sum_{r' \in \Gamma_{I,r}^{\langle I,a\rangle}} \mathrm{Tr}\Big(\Pi_{\langle a,I\rangle}^{r'} A_\Omega^\gamma\Big). \tag{8.35}$$

The left hand side is zero by assumption. As shown above, each term on the right hand side is nonnegative, therefore each term on the right hand side is zero, i.e. $\mathrm{Tr}\Big(\Pi_{\langle a,I\rangle}^{r'} A_\Omega^\gamma\Big) = 0$ for all $r' \in \Gamma_{I,r}^{\langle a,I\rangle}$.

We can write the spectral decomposition of T_a as

$$T_a = \sum_{r' \in \Gamma_{I,r}^{\langle a,I\rangle}} \omega^{r(a)} \Pi_{\langle a\rangle}^{r'|_{\langle a\rangle}}. \tag{8.36}$$

Then

$$\mathrm{Tr}\big(T_a \Pi_I^r A_\Omega^\gamma \Pi_I^r\big) = \sum_{r' \in \Gamma_{I,r}^{\langle a,I\rangle}} \omega^{r'(a)}\, \mathrm{Tr}\Big(\Pi_{\langle a\rangle}^{r'|_{\langle a\rangle}} \Pi_I^r A_\Omega^\gamma \Pi_I^r\Big) = \sum_{r' \in \Gamma_{I,r}} \omega^{r'(a)}\, \mathrm{Tr}\Big(\Pi_{\langle a,I\rangle}^{r'} A_\Omega^\gamma\Big) = 0. \tag{8.37}$$

(iii) Third, if $a \notin I^\perp$ then

$$\Pi_I^r T_a \Pi_I^r = \frac{1}{|I|^2} \sum_{b,c \in I} \omega^{-r(b)-r(c)} T_b T_a T_c \tag{8.38}$$

$$= \frac{1}{|I|^2} \sum_{b,c \in I} \omega^{-r(b)-r(c)+[b,a]-\beta(b,c)} T_a T_{b+c} \tag{8.39}$$

$$= \frac{1}{|I|^2} T_a \sum_{b,c \in I} \omega^{-r(b+c)+[b,a]} T_{b+c} \tag{8.40}$$

$$= \frac{1}{|I|} T_a \Pi_I^r \sum_{b \in I} \omega^{[b,a]}. \tag{8.41}$$

By character orthogonality, the sum in the final expression vanishes. Therefore,

$$\mathrm{Tr}\big(T_a \Pi_I^r A_\Omega^\gamma \Pi_I^r\big) = \mathrm{Tr}\big(\Pi_I^r T_a \Pi_I^r A_\Omega^\gamma\big) = 0. \tag{8.42}$$

Thus, $\mathrm{Tr}\big(T_a \Pi_I^r A_\Omega^\gamma \Pi_I^r\big) = 0$ for all Pauli operators $T_a, a \in E$, and so $\Pi_I^r A_\Omega^\gamma \Pi_I^r = 0$.

(II) In the second case, $\mathrm{Tr}\big(\Pi_I^r A_\Omega^\gamma\big) > 0$, as shown above we have $r|_{\Omega\cap I} = \gamma|_{\Omega\cap I}$. Then

$$\Pi_I^r A_\Omega^\gamma \Pi_I^r = \frac{1}{|I|d^n} \sum_{b\in\Omega}\sum_{c\in I} \omega^{-\gamma(b)-r(c)} \Pi_I^r T_b T_c \tag{8.43}$$

$$= \frac{1}{|I|d^n} \sum_{b\in\Omega}\sum_{c\in I} \omega^{-\gamma(b)-r(c)+[b,c]} \Pi_I^r T_c T_b \tag{8.44}$$

$$= \frac{1}{|I|d^n} \sum_{b\in\Omega}\sum_{c\in I} \omega^{-\gamma(b)+[b,c]} \Pi_I^r T_b \tag{8.45}$$

$$= \frac{1}{|I|d^n} \Pi_I^r \sum_{b\in\Omega} \left[\sum_{c\in I} \omega^{[b,c]}\right] \omega^{-\gamma(b)} T_b \tag{8.46}$$

$$= \frac{1}{d^n} \Pi_I^r \sum_{b\in\Omega\cap I^\perp} \omega^{-\gamma(b)} T_b \tag{8.47}$$

$$= \frac{1}{|I|d^n} \sum_{a\in I}\sum_{b\in\Omega\cap I^\perp} \omega^{-r(a)-\gamma(b)} T_a T_b \tag{8.48}$$

$$= \frac{1}{|I|d^n} \sum_{a\in I}\sum_{b\in\Omega\cap I^\perp} \omega^{-r(a)-\gamma(b)-\beta(a,b)} T_{a+b} \tag{8.49}$$

Here each Pauli operator in the sum has the same multiplicity. For any $c \in I + \Omega\cap I^\perp$, let $\mu(c)$ denote the number of pairs $(a,b) \in I\times\Omega\cap I^\perp$ such that $a+b=c$. We will show that $\mu(c) = \mu(0)$ for all $c \in I+\Omega\cap I^\perp$. First, suppose $(a_1,b_1),(a_2,b_2),\ldots,(a_{\mu(c)},b_{\mu(c)})$ are $\mu(c)$ distinct pairs in $I\times\Omega\cap I^\perp$ such that $a_j+b_j=c$. Then the pairs $(a_j-a_1, b_j-b_1) \in I\times\Omega\cap I^\perp$ for $j=2,3,\ldots,\mu(c)$, together with the pair $(0,0)$ show that $\mu(0)\geq\mu(c)$ for any $c\in I+\Omega\cap I^\perp$. Now let $(a_1,b_1),(a_2,b_2),\ldots,(a_{\mu(0)},b_{\mu(0)})$ be distinct pairs in $I\times\Omega\cap I^\perp$ such that $a_j+b_j=0$, and let (a,b) be such that $a+b=c$. Then the pairs (a_j+a, b_j+b) for $j=1,2,\ldots,\mu(0)$ show that $\mu(c)\geq\mu(0)$ for any $c\in I+\Omega\cap I^\perp$. Thus, $\mu(c)=\mu(0)$ for any $c\in I+\Omega\cap I^\perp$. We have $\mu(0) = |\Omega\cap I|$.

The noncontextual value assignment $\gamma\times r : I+\Omega\cap I^\perp \to \mathbb{Z}_d$ satisfying $\gamma\times r|_I = r$ and $\gamma\times r|_{\Omega\cap I^\perp} = \gamma|_{\Omega\cap I^\perp}$ is uniquely defined. Then

$$\Pi_I^r A_\Omega^\gamma \Pi_I^r = \frac{|\Omega\cap I|}{|I|d^n} \sum_{a\in I+\Omega\cap I^\perp} \omega^{-\gamma\times r(a)} T_a = \frac{|\Omega\cap I|}{|I|} A_{I+\Omega\cap I^\perp}^{\gamma\times r} \tag{8.50}$$

Thus

$$\frac{\Pi_I^r A_\Omega^\gamma \Pi_I^r}{\mathrm{Tr}\big(\Pi_I^r A_\Omega^\gamma\big)} = A_{I+\Omega\cap I^\perp}^{\gamma\times r} \tag{8.51}$$

which concludes the proof of the lemma. □

For a single Pauli measurement, $I = \langle a\rangle$ for some $a\in E$, this lemma simplifies with the following corollary.

Corollary 7. *Let $(\Omega,\gamma)\in\mathcal{V}$ be a CNC pair. Then for any $a\in E$,*

(I) if $a \in \Omega$, *then* $\begin{cases} \mathrm{Tr}\big(\Pi_a^s A_\Omega^\gamma\big) = 1 \text{ and } \Pi_a^s A_\Omega^\gamma \Pi_a^s = A_{\Omega\cap a^\perp}^{\gamma|_{\Omega\cap a^\perp}} & \text{if } s = \gamma(a), \\ \mathrm{Tr}\big(\Pi_a^s A_\Omega^\gamma\big) = 0 \text{ and } \Pi_a^s A_\Omega^\gamma \Pi_a^s = 0 & \text{if } s \neq \gamma(a). \end{cases}$

(II) if $a \notin \Omega$, *then* $\mathrm{Tr}\big(\Pi_a^s A_\Omega^\gamma\big) = 1/d$ *for each* $s \in \mathbb{Z}_d$, *and* $\Pi_a^s A_\Omega^\gamma \Pi_a^s = \frac{1}{d} A_{\langle a\rangle + \Omega\cap a^\perp}^{\gamma\times s}$.

8.4.3 Classical simulation algorithm

The update rules of the phase space points under Clifford gates and Pauli measurements given in Sections 8.4.1 and 8.4.2 allow us to introduce a classical simulation algorithm for quantum computation with magic states. This simulation is given explicitly in Algorithm 5. The proof of correctness of this algorithm is analogous to the proofs of [1, Theorem 3], [2, Theorem 2], and [7, Theorem 2].[4]

Input: $W_\rho(\Omega, \gamma)$; a Clifford+Pauli circuit $\mathcal{C}$
1: sample a point $(\Omega, \gamma) \in \mathcal{V}$ according to the probability distribution p_ρ
2: **while** end of circuit has not been reached **do**
3: **if** a Clifford unitary $g \in \mathcal{C}\ell$ is encountered **then**
4: update $(\Omega, \gamma) \leftarrow\!\!\prec (g \cdot \Omega, g \cdot \gamma)$
5: **end if**
6: **if** a Pauli measurements T_a, $a \in E$ is encountered **then**
7: **if** $a \in \Omega$ **then**
8: set $s = \gamma(a)$
9: update $(\Omega, \gamma) \leftarrow\!\!\prec (\Omega \cap a^\perp, \gamma|_{\Omega\cap a^\perp})$
10: **else if** $a \notin \Omega$ **then**
11: choose a random $s \in \mathbb{Z}_d$ distributed uniformly
12: update $(\Omega, \gamma) \leftarrow\!\!\prec (\langle a\rangle + \Omega \cap a^\perp, \gamma \times s)$
13: **end if**
14: **Output:** s as the outcome of the measurement
15: **end if**
16: **end while**

Algorithm 5: One run of the classical simulation algorithm for quantum computation with magic states. The algorithm provides samples from the joint probability distribution of the Pauli measurements in a quantum circuit consisting of Clifford gates and Pauli measurements applied to an input state ρ.

8.4.4 Efficiency of classical simulation

In the case of the multiqubit CNC construction, the classical simulation algorithm [1, Table 1] is efficient on states with a nonnegative quasiprobability function W_ρ. The proof

[4] One difference between the simulation Algorithm 5 and that of Ref. [1] is there only extremal (in the sense of Lemma 30) CNC phase space points are used. Clifford gates preserve the property of extremality and after each Pauli measurement the resulting post-measurement state $\Pi_a^s A_\Omega^\gamma \Pi_a^s / \mathrm{Tr}(\Pi_a^s A_\Omega^\gamma)$ is decomposed as a convex combination of extremal CNC phase space point operators, c.f. [1, Equation 26]. On the other hand, in our case we allow nonextremal CNC points to appear in the simulation, this does not affect the proof of correctness of the algorithm.

of this fact relies on the structure of the multiqubit CNC phase space points found in [1, §IV]. With the structure of the phase space found in Section 8.3, an adaptation of the proof of efficiency should apply to our case, with one caveat. Namely, for multiple qubits, the maximum number of pair-wise noncommuting Pauli operators on n qubits is known to be $2n + 1$. In particular, it is linear in the number of qubits. For odd-prime-dimensional qudits, a polynomial bound on this number is not known, the best known upper bound is exponential in the number of qudits n [198]. If there were to exist sets of noncommuting Pauli operators of size exponential in n, the classical updates of the CNC phase space points would not be efficiently computable classically.

Let $\tilde{\mathcal{V}}$ denote a subset of the CNC phase space defined by the additional constraint that the CNC sets have a set of generators (consisting of a basis in case of a subspace of E, or a basis of I together with the noncommuting generators $a_1, \ldots, a_\xi$ in the case of Eq. (8.12), c.f. Theorem 22) that has size polynomial in the number of qudits. The state updates under Clifford gates and Pauli measurements described in Lemmas 33 and 34 (lines 4, 9, and 12 of Algorithm 5) do not increase the number of generators needed to specify a CNC set, so $\tilde{\mathcal{V}}$ is closed under these operations. With respect this this phase space $\tilde{\mathcal{V}}$, we have the following theorem.

Theorem 23. *For any number of qudits n of any odd prime Hilbert space dimension d,*

1. *The update of phase space points $\tilde{\mathcal{V}}$ under Clifford gates given by Lemma 33 can be computed in time polynomial in the number of qudits.*

2. *The update of phase space points $\tilde{\mathcal{V}}$ under Pauli measurements given by Lemma 34 can be computed in time polynomial in the number of qudits.*

The proof of Theorem 23 is analogous to the proof of Theorem 3 in Ref. [1]. Basically, with the characterization of phase space points in Section 8.3, for any phase space points (Ω, γ), the set Ω can be identified by a small number of vectors in $\mathbb{Z}_d^{2n}$. Namely, if Ω is a subspace of E then we identify Ω by a basis of vectors, and if Ω has the form Eq. (8.12) then we use the vectors $a_1, \ldots, a_\xi$ in Eq. (8.12) as well as a basis for I. Further, the value assignment γ can be specified by specifying its value on these generating vectors, and then its value on the rest of Ω is determined through Eq. (8.1). Then all of the steps in the update can be performed by linear algebra in $\mathbb{Z}_d^{2n}$ on this small number of vectors.

Corollary 8. *For any quantum circuit consisting of a sequence of polynomially many Clifford gates and Pauli measurements on an n-qudit input state ρ, satisfying the following conditions:*

- $\rho = \sum_{(\Omega,\gamma)\in\tilde{\mathcal{V}}} W_\rho(\Omega,\gamma) A_\Omega^\gamma$ *where* $W_\rho(\Omega,\gamma) \geq 0$ *for all* $(\Omega,\gamma) \in \tilde{\mathcal{V}}$,

- *samples from the distribution W_ρ can be obtained efficiently,*

the outcome distributions of the measurements can be efficiently sampled from using Algorithm 5.

Of course, it could turn out that the maximal number of noncommuting Pauli operators on n qudits is bounded by a polynomial in the number of qudits n, in which case Theorem 23 and Corollary 8 apply to the full CNC phase space $\mathcal{V}$, but as mentioned above this has not been proven. This is an open problem.

8.5 Relation to the Λ-polytope model

The Λ polytope models [2, 7] are a family of quasiprobability representations for quantum computation with magic states satisfying (QR1)–(QR5) in which no negativity is needed in the representation of any state. That is, all elements of this model of quantum computation—all states, Clifford gates, and Pauli measurements are described probabilistically.

The qubit version of the model was extensively studied in [2, 3, 72]. There it was shown that for the multiqubit case ($d = 2$), for maximal CNC sets Ω, the phase space point operators A^{γ}_{Ω} are vertices of the Λ polytopes, and thus define phase space points of the Λ polytope models. Here we consider a similar relation between the odd-prime-dimensional qudit extension of the CNC construction and the Λ polytope models.

Let $\mathcal{S}$ denote the set of pure n-qudit stabilizer states. The Λ polytope models are based on the set

$$\Lambda = \{X \in \mathrm{Herm}_1(\mathbb{C}^{d^n}) \mid \mathrm{Tr}(|\sigma\rangle\langle\sigma| X) \geq 0 \ \ \forall\, |\sigma\rangle \in \mathcal{S}\}. \tag{8.52}$$

The set Λ is a polytope for any dimension d (see [7, Lemma 1]). One can easily verify that $A^{\gamma}_{\Omega} \in \Lambda$ for every A^{γ}_{Ω} of the form Eq. (8.2) (see the proof of Lemma 34). More generally, we want to specify Hermitian operators in Λ whose expansion coefficients in the generalized Pauli basis are *arbitrary* coefficients on the complex unit circle. That is, we want to consider Hermitian operators of the form

$$A^{\eta}_{\Omega} = \frac{1}{d^n} \sum_{u \in \Omega} \omega^{\eta(u)} T_u, \quad \Omega \subset E, \ \eta : \Omega \to [0, \pi] \tag{8.53}$$

that are contained in Λ. As it will turn out, the condition $A^{\eta}_{\Omega} \in \Lambda$ implies that A^{η}_{Ω} is precisely of the form (8.2), i.e. Ω is closed under inference, $\eta(u) \pmod d \in \mathbb{Z}_d$, and

$$\eta(a + b) \equiv \eta(a) + \eta(b) \pmod d \text{ if } [a, b] = 0. \tag{8.54}$$

For a linear subspace $L \subset E$ let

$$L^* := \{\gamma : L \to \mathbb{Z}_d, \gamma(a + b) = \gamma(a) + \gamma(b)\} \tag{8.55}$$

be its dual space, i.e. the space of linear functions on L.

Theorem 24. *If $A^{\eta}_{\Omega} \in \Lambda$, then*

(i) If I is an isotropic subspace and $I \subset \Omega$, then $\eta_{|I} \in I^*$.

(ii) The set Ω is closed under inference, i.e., if $a, b \in \Omega$ and $[a, b] = 0$, then $a + b \in \Omega$.

For $M \subset E$ let $\mathrm{pr}_M : \mathrm{Herm}(\mathbb{C}^{d^n}) \to \mathrm{span}\{T_b : b \in M\}$ be the projection that acts via

$$\sum_{u \in E_n} c_b T_b \overset{\mathrm{pr}_M}{\longmapsto} \sum_{b \in M} c_b T_b. \tag{8.56}$$

To prove the lemma, we will use the fact that if $X \in \Lambda$, then $\mathrm{pr}_I(X) \in \mathrm{pr}_I(\Lambda)$. Lemma 9

in [7] allows us to add redundant inequalities to Λ, so that we can write

$$\Lambda = \{X \in \mathrm{Herm}_1(\mathbb{C}^{d^n}) \mid \mathrm{Tr}\big(\Pi_I^\gamma X\big) \geq 0 \; \forall I, \forall \gamma \in I^*\}. \tag{8.57}$$

where we range over all isotropic subspaces I and all $\gamma \in I^*$. We will be mainly interested in $\mathrm{pr}_I(\Lambda)$ for I being an isotropic subspace.

Lemma 35. *If I is an isotropic subspace, then $\mathrm{pr}_I(\Lambda)$ is a self dual simplex with vertices $\Pi_I^\gamma, \gamma \in I^*$, in particular $\mathrm{pr}_I(\Lambda) = \mathrm{conv}\{\Pi_I^\gamma \mid \gamma \in I^*\}$.*

Proof of Lemma 35. Due to $\mathrm{pr}_I(\Pi_I^\gamma) = \Pi_I^\gamma$ it follows that

$$\mathrm{pr}_I(\Lambda) \subseteq \{X = \sum_{b \in I} c_b T_b \mid \mathrm{Tr}(X) = 1, \, \mathrm{Tr}\big(X\Pi_I^\gamma\big) \geq 0 \text{ for all } \gamma \in I^*\}. \tag{8.58}$$

Analogously, to Lemma 15 in [199], character orthogonality implies that the right hand side of (8.58) is a self-dual simplex, so

$$\{X = \sum_{b \in I} c_b T_b \mid \mathrm{Tr}(X) = 1, \, \mathrm{Tr}\big(X\Pi_I^\gamma\big) \geq 0 \text{ for all } \gamma \in I^*\} \tag{8.59}$$

$$= \mathrm{conv}\{\Pi_I^\gamma \mid \gamma \in I^*\} \tag{8.60}$$

and the vertices of $\mathrm{conv}\{\Pi_I^\gamma \mid \gamma \in I^*\}$ are precisely the operators Π_I^γ. On the other hand $\Pi_I^\gamma \in \Lambda$ and therefore $\Pi_I^\gamma \in \mathrm{pr}_I(\Lambda)$ for all $\gamma \in I^*$ and therefore also $\mathrm{conv}\{\Pi_I^\gamma \mid \gamma \in I^*\} \subseteq \mathrm{pr}_I(\Lambda)$. $\square$

We will use the characterization Lemma 35 to prove Lemma 24. The overall strategy will be to argue that if Ω and η violate one of the conditions of the theorem, then there is an isotropic subspace $I \subset E$ such that $\mathrm{pr}_I(A_\Omega^\gamma) \notin \mathrm{pr}_I(\Lambda)$ implying $A_\Omega^\gamma \notin \Lambda$. To show $\mathrm{pr}_I(A_\Omega^\gamma) \notin \mathrm{pr}_I(\Lambda)$, we will construct a hyperplane that separates the point $\mathrm{pr}_I(A_\Omega^\gamma)$ from the polytope $\mathrm{pr}_I(\Lambda)$, that is, we construct some explicit Hermitian operator $Y \in \mathrm{span}\{T_b \mid b \in I\}$ such that

$$\mathrm{Tr}\big(\mathrm{pr}_I(A_\Omega^\gamma)Y\big) > \max_{X \in \mathrm{pr}_I(\Lambda)} \mathrm{Tr}(XY) = \max_{\gamma \in I^*} \mathrm{Tr}\big(\Pi_I^\gamma Y\big), \tag{8.61}$$

where the last equation is a consequence of Lemma 35.

Proof of Theorem 24. Let $A_\Omega^\eta = \frac{1}{d^n} \sum_{u \in \Omega} \omega^{\eta(u)} T_u$ be Hermitian with $\eta : \Omega \to [0, \pi]$ and $\mathrm{Tr}\big(A_\Omega^\eta\big) = 1$ (the last condition implies $0 \in \Omega$ and $\eta(0) = 0$). As A_Ω^η is Hermitian, we have

$$\sum_{u \in \Omega} \omega^{\eta(u)} T_u = A_\Omega^\eta = (A_\Omega^\eta)^\dagger = \sum_{u \in \Omega} \omega^{-\eta(u)} T_{-u}, \tag{8.62}$$

implying $-\eta(u) = \eta(-u)$ for all $u \in \Omega$.

(i) Let I be an isotropic subspace contained in Ω. We will show that if $A_\Omega^\gamma \in \Lambda$, then $\eta_{|I} \equiv \gamma$ for some linear function $\gamma : I \to \mathbb{Z}_d$: Therefore, assume that $\eta_{|I} \not\equiv \gamma$ for all linear $\gamma \in I^*$. In the sense of Equation (8.61), we can separate $\mathrm{pr}_I(\Lambda)$ and $\mathrm{pr}_I(A_\Omega^\eta)$ by the

hyperplane with normal vector $A_I^{\eta|I} \in \mathrm{span}\{T_b \,|\, b \in I\}$:

$$\mathrm{Tr}\Big(A_I^{\eta|I}\,\mathrm{pr}_I(A_\Omega^\eta)\Big) = \mathrm{Tr}\Big(A_I^{\eta|I} A_I^{\eta|I}\Big) \tag{8.63}$$

$$= \frac{1}{d^{2n}} \sum_{u\in I} \omega^{\eta(u)-\eta(u)}\,\mathrm{Tr}(T_u T_{-u}) \tag{8.64}$$

$$= \frac{|I|}{d^n} > \mathrm{Tr}\Big(A_I^{\eta|I}\Pi_I^\gamma\Big), \tag{8.65}$$

where the strict inequality holds for all $\gamma : I \to \mathbb{Z}_d$ with $\gamma \neq \eta_{|I}$, so in particular for all $\gamma \in I^*$.

(ii) Now suppose that $\eta_{|I} \in I^*$ for any isotropic subspace I contained in Ω. For (ii) assume that there are $a, b \in \Omega$ with $[a, b] = 0$ such that $a + b \notin \Omega$. We may assume that $b \neq -a$ since $0 \in \Omega$. Set $I = \langle a, b\rangle$ (note that $I \nsubseteq \Omega$) and

$$Y = \omega^{\eta(a)} T_a + \omega^{\eta(b)} T_b + \omega^{-\eta(a)} T_{-a} + \omega^{-\eta(b)} T_{-b} + \omega^{\eta(a)+\eta(b)+\lfloor d/2\rfloor} T_{a+b} + \omega^{-(\eta(a)+\eta(b)+\lfloor d/2\rfloor)} T_{-(a+b)} \tag{8.66}$$

$$\in \mathrm{span}\{T_x \,|\, x \in I\}. \tag{8.67}$$

Then

$$\mathrm{Tr}\big(\mathrm{pr}_I(A_\Omega^\gamma) Y\big) = 4 \tag{8.68}$$

but for any Π_I^γ with $\gamma \in I^*$

$$\mathrm{Tr}\big(\Pi_I^\gamma Y\big) = (\omega^{\eta(a)-\gamma(a)} + \omega^{-\eta(a)+\gamma(a)}) + (\omega^{\eta(b)-\gamma(b)} + \omega^{-\eta(b)+\gamma(b)}) \tag{8.69}$$

$$+ (\omega^{\eta(a)+\eta(b)-\gamma(a)-\gamma(b)+\lfloor d/2\rfloor} + \omega^{-(\eta(a)+\eta(b)-\gamma(a)-\gamma(b)+\lfloor d/2\rfloor)}) \tag{8.70}$$

$$= 2\cos\left(\underbrace{\frac{2\pi(\eta(a)-\gamma(a))}{d}}_{=:x}\right) + 2\cos\left(\underbrace{\frac{2\pi(\eta(b)-\gamma(b))}{d}}_{=:y}\right) \tag{8.71}$$

$$+ 2\cos\left(\frac{2\pi(\eta(a)-\gamma(a)+\eta(b)-\gamma(b)+\lfloor d/2\rfloor)}{d}\right) \tag{8.72}$$

$$= 2\left(\cos(x) + \cos(y) + \cos\left(x + y + 2\pi\frac{\lfloor d/2\rfloor}{d}\right)\right) \tag{8.73}$$

$$< 4. \tag{8.74}$$

This proves the theorem. □

8.6 Conclusion

In this work, we have presented the generalization of the CNC construction of Ref. [1] to the setting of odd-prime-dimensional qudits. We provide a characterization of the CNC phase space in this setting, and we describe its relation to other models like the Wigner function [19, 29–31] and the Λ polytope models [2, 7]. The phase space of this model contains the phase space of the Wigner function, but it also includes new phase space points which

cannot be described as convex mixtures of Wigner function phase space points. We show that all vertices of the Λ polytopes with coefficients of absolute value equal to one when expanded in the Pauli basis are CNC-type phase space point operators.

We also introduce a classical simulation algorithm for quantum computation with magic states based on sampling from probability distributions over the CNC phase space. Since the CNC construction outperforms the Wigner function and stabilizer methods in terms of the volume of states that can be positively represented, this new method allows a broader class of magic state quantum circuits to be efficiently classically simulated.

To conclude, we make two additional remarks commenting on possible avenues of future work.

1. The CNC construction applies to qudits of any Hilbert space dimension, but the proof of efficiency of the update rules under Clifford gates and Pauli measurements relies on the characterization of the CNC phase space given in Ref. [1, §IV] for qubits and in Section 8.3 for odd-prime-dimensional qudits. A similar characterization may be possible for qudits with composite Hilbert space dimension, but is likely much more complicated, in part because the set of labels $E := \mathbb{Z}_d^{2n}$ of the Pauli operators fails to be a vector space for composite Hilbert space dimensions d [29].

2. For each fixed number n of qudits, the volume of states that are positively represented by the CNC construction is strictly larger than the corresponding volume for the Wigner function. That said, talking about scaling of simulation cost with the number of qudits is more difficult here. The reason is the following: the phase space point operators of the Wigner function are closed under tensor products, and in fact, every phase space point operator can be constructed as a tensor product of single-qudit phase space point operators. As a result, the Wigner function is multiplicative: $W_{\rho\otimes\sigma}(u \otimes v) = W_\rho(u)W_\sigma(v)$ (see e.g. Ref. [19]). When the representation goes negative for some state, the amount of negative as measured by the 1-norm of the Wigner function is multiplicative in the number of copies of that state. For the case of stabilizer quasimixtures [21], a tensor product of stabilizer states is again a stabilizer state, but there are other entangled stabilizer states that cannot be constructed in this way. As a result, the corresponding measure of negativity, the robustness of magic, is submultiplicative. On the other hand, for the CNC construction, a tensor product of CNC phase space point operators is generally not CNC. Thus, the equivalent of the robustness of magic for the CNC construction, another measure of negativity, is supermultiplicative. For example, two copies of the magic state $|H\rangle$ (see Figure 2.3) are nonnegative, but three copies are not [1]. This means we cannot obtain upper bounds on the negativity of product states by computing the negativity in the representation for few qudits. There are some tricks for partially addressing this, for example see Ref. [1, Appendix D] and [3], but it remains more difficult to talk about scaling of quantum computational resources with the number of qudits for the CNC construction than for previous methods. We regard this as an open problem.

The second issue above is also an present in other models derived from the Λ polytopes, e.g. [6]. While the Λ polytopes are beginning to be explored for multiqubit systems [1, 3, 72, 200, 201], the extension of the Λ polytope models to higher dimensional qudits is relatively

new [7]. Here we provide some initial structural insights, by characterizing the vertices with coefficients with absolute value equal to one in the Pauli basis, but we believe there is much more to be learned here.

In other settings, the odd-prime-dimensional qudit setting is simpler than the even-dimensional case, and was explored first. This is true for example for the link between Wigner functions and quantum computation [19, 27, 28, 31]. Following this pattern, we may be able to continue this work and derive insights from the Λ polytopes for the simpler case of odd-prime-dimensional qudits which can later be imported to the case of qubits.

Chapter 9

Simulation of quantum computation with magic states via Jordan-Wigner transformations

Negativity in certain quasiprobability representations is a necessary condition for a quantum computational advantage. Here we define a new quasiprobability representation exhibiting this property with respect to quantum computations in the magic state model. It is based on generalized Jordan-Wigner transformations and it has a close connection to the probability representation of universal quantum computation based on the Λ polytopes. For each number of qubits it defines a polytope contained in the Λ polytope with some shared vertices. It leads to an efficient classical simulation algorithm for magic state quantum circuits for which the input state is positively represented, and it outperforms previous representations in terms of the states that can be positively represented.

This chapter has previously appeared in Ref. [6].

9.1 Introduction

Quasiprobability representations have long played a crucial role in physics bridging the gap between classical and quantum [43]. Often with such representations, negativity serves as an indicator of genuinely quantum behaviour, with the fragment of quantum theory that is positively represented behaving more classically [45].

Gross' discrete Wigner function [29, 31]—a quasiprobability representation for systems of odd-dimensional qudits—has been particularly useful in describing quantum computations in the magic state model. In fact, Veitch et al. [19] showed that negativity in this representation is a necessary condition for a quantum computational advantage over classical computation. This result is obtained by defining an efficient classical simulation algorithm for magic state quantum computations that applies whenever the input state of the quantum circuit has a nonnegative representation. In the last decade, many other quasiprobability representations have been defined which also exhibit this property [1, 21, 49, 57, 60, 147, 196].

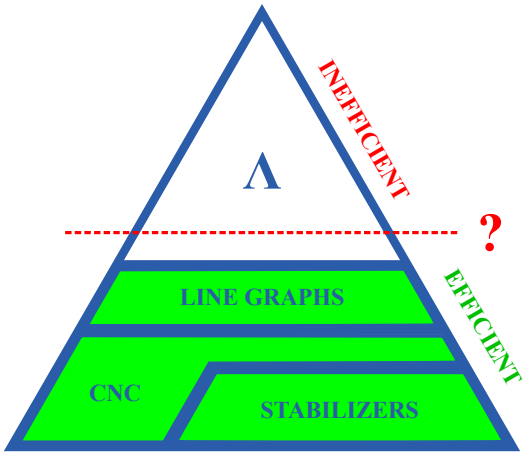


Figure 9.1: The big question about the Λ polytopes [2, 7] is where the line between efficient and inefficient classical simulation of quantum computation lies. Here, we increase the size of the known efficiently simulable region. Our construction is based on Jordan-Wigner transformations, and its range of applicability includes the earlier CNC construction [1], which in turn includes the stabilizer formalism [21, 40, 41].

This viewpoint relating negativity and quantum computational advantage was disrupted in Refs. [2, 7] where a fully probabilistic model describing universal quantum computation was defined. In this model, all quantum states are represented by a probability distribution over a finite set of hidden states and all computational dynamics are represented by stochastic state update rules—no negative probabilities are required. However, this circumvention of negativity comes at a cost: it can no longer be guaranteed that the update rules are efficiently computable classically. Thus, although this model yields a classical simulation algorithm for universal quantum computation, the simulation is likely inefficient in general (as would be expected for any classical simulation of universal quantum computation).

The ultimate question for these classical simulation methods based on quasiprobability representations is: *Where runs the line between efficient and inefficient classical simulation, and which physical or mathematical property determines it?* For qubits, so far it is known that this dividing line lies somewhere between the quantum computations covered by the so-called CNC (closed and noncontextual) construction [1], and the universal quantum computations described by the Λ polytopes [2] (see Fig. 9.1).

In this work, we enlarge the known region of efficient classical simulability inside the Λ-polytopes. Specifically, we define a new quasiprobability representation intermediate between those of the CNC construction and the Λ-polytopes. This model has efficiently computable state update rules under the dynamics of quantum computation, and it yields an efficient classical simulation algorithm for quantum comptutations on the subset of input states that are positively represented. It subsumes the CNC construction, and previous methods such as those based on quasimixtures of stabilizer states [21]. Thus, we effectively push out the boundary of quantum computations which can be efficiently simulated classically.

The model is based on generalized Jordan-Wigner transformations [73], and its conception was influenced by the surprising connection to Majorana fermions first realized in the CNC construction [1]. It was also partially inspired by the techniques of mapping to free fermions in the simulation of other computational models such as matchgate circuits [74], which have recently received renewed interest [75–78].

Outline

The remainder of this chapter is structured as follows. We begin in Section 9.2 by introducing some notation and definitions. In Section 9.3 we define the new quasiprobability representation of quantum computation with magic states. In Section 9.4 we describe the behaviour of the generalized phase space over which this representation is defined with respect to the dynamics of quantum computation with magic states, namely, Clifford gates and Pauli measurements, and we define a classical simulation algorithm for quantum computation with magic states that applies whenever the input state of the quantum circuit is positively represented. We also introduce a monotone for the resource theory of stabilizer quantum computation [42] for the case where the state is not positively represented. In Section 9.5 we elucidate the relationship between this model and the Λ polytopes. Finally, we conclude with a discussion of these results in Section 9.6.

9.2 Preliminaries

The setting of this chapter is quantum computation with magic states (QCM) on systems of qubits [38]. In this model, computation proceeds through the application of a sequence of Clifford gates and Pauli measurements on an initially prepared "magic" input state. In general, the input can be any nonstabilizer quantum state, but for universal quantum computation it suffices to consider input states formed from multiple copies of a single-qubit nonstabilizer state [105]. The output of the computation is derived from the outcomes of the measurements.

In the following we use the notation $\mathcal{O}$ for a general subset of E (labels for a subset of Pauli operators), and $\mathcal{O}^*$ for the set $\mathcal{O} \setminus \{0\}$ (the nonidentity Pauli operators in $\mathcal{O}$). We denote by $\mathrm{Herm}(\mathcal{H})$ the space of Hermitian operators on Hilbert space $\mathcal{H}$, and unless otherwise specified, $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ is the n-qubit Hilbert space. $\mathrm{Herm}_1(\mathcal{H})$ is the affine subspace of $\mathrm{Herm}(\mathcal{H})$ consisting of operators with unit trace and $\mathrm{Herm}_1^{\succ 0}(\mathcal{H})$ is the subset of $\mathrm{Herm}_1(\mathcal{H})$ consisting of positive semidefinite operators. $\mathrm{Herm}_1^{\succ 0}(\mathcal{H})$ contains density operators representing physical quantum states.

9.2.1 Previous quasiprobability representations

The CNC construction

Recently, a quasiprobability representation for QCM was defined based on noncontextual sets of Pauli observables [1], this is the so-called CNC construction. Therein, phase space points are identified with pairs (Ω, γ), where $\Omega \subset E$ is a subset of Pauli operators and

$\gamma : \Omega \to \mathbb{Z}_2$ is a function on Ω satisfying two conditions:[1]

1. *Closure under inference.* For all $a, b \in \Omega$,

$$[a, b] = 0 \Longrightarrow a + b \in \Omega.$$

2. *Noncontextuality.*[2] γ is a noncontextual value assignment on Ω. I.e. a function $\gamma : \Omega \to \mathbb{Z}_2$ such that $\forall a, b \in \Omega$ with $[a, b] = 0$, we have

$$\gamma(a) + \gamma(b) + \gamma(a + b) \equiv \beta(a, b) \mod 2. \tag{9.1}$$

Then the phase space point operator associated to phase space point (Ω, γ) is defined as

$$A^{\gamma}_{\Omega} = \frac{1}{2^n} \sum_{b \in \Omega} (-1)^{\gamma(b)} T_b. \tag{9.2}$$

The terms phase space and phase space point operators are used in analogy with the phase space and phase space point operators of the discrete Wigner function [29, 31].

These operators span the space of Hermitian operators on Hilbert space $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ and so any n-qubit quantum state ρ can be expanded in these operators as

$$\rho = \sum_{(\Omega, \gamma)} W_{\rho}(\Omega, \gamma) A^{\gamma}_{\Omega}. \tag{9.3}$$

It was shown in Ref. [1, §IV] that the admissible pairs (Ω, γ) could be characterized as follows. Let $I \subset E$ be an isotropic subspace and let $\tilde{\Omega} \subset E \setminus I$ be a subset of Pauli operators such that $[a, b] = 1$ for all $a, b \in \tilde{\Omega}^*$. Then the sets Ω can be expressed as $\Omega = \bigcup_{a \in \tilde{\Omega}} \langle a, I \rangle$. The signs γ can be chosen freely on $\tilde{\Omega}^*$ and on a basis of I, and then the signs on the rest of Ω are determined by Eq. (9.1).

In simpler terms, the phase space point operators have the form

$$A^{\gamma}_{\Omega} = g(A^{\tilde{\gamma}}_{\tilde{\Omega}} \otimes |\sigma\rangle \langle\sigma|) g^{\dagger} \tag{9.4}$$

where $g \in \mathcal{C}\ell$ is a Clifford group element, $|\sigma\rangle$ is a stabilizer state, all elements of $\tilde{\Omega}^*$ pair-wise anticommute, and the signs $\tilde{\gamma}$ on $\tilde{\Omega}^*$ can be chosen freely. Interestingly, a set of pair-wise anticommuting Pauli operators has the same algebraic structure as a set of Majorana fermion operators and so, modulo the stabilizer state tensor factor in Eq. (9.4), we can view the operator $A^{\tilde{\gamma}}_{\Omega}$ as linear combinations of Majorana operators.

In the case that a state ρ has a representation in Eq. (9.3) with $W(\Omega, \gamma) \geq 0$, $\forall (\Omega, \gamma)$, the expansion coefficients $W(\Omega, \gamma)$ define a probability distribution over the phase space. Phase space point operators of the form Eq. (9.2) map deterministically to other phase space point operators under any Clifford gate, and they map to convex mixtures of these operators under Pauli measurements (see Ref. [1, §V] for details). Further, these state updates can

[1] "CNC" is for Closed and NonContextual after these conditions.

[2] Implicit in this condition on γ there is a constraint on the set Ω, namely that such a noncontextual value assignment exists. Because of Mermin square-style proofs of contextuality [62], this is not guaranteed for general subsets of E.

be computed efficiently on a classical computer. Thus, for QCM computations in which the input state of the circuit has a nonnegative representation, all aspects of the computation are represented probabilistically by the model. If samples can be obtained efficiently from the distribution $W(\Omega, \gamma)$, then the model yields an efficient classical simulation of the circuit [1, §VI].

The Λ polytopes

In Ref. [2], a probability representation (or a hidden variable model) of quantum computation with magic states was defined. It has the same structure as a quasiprobability representation, except for the fact that no negativity is needed. This representation can be defined for qudits of any Hilbert space dimension [7], but we state only the multiqubit version here. The state space of the model is based on the set

$$\Lambda = \{X \in \mathrm{Herm}_1(\mathcal{H}) \mid \mathrm{Tr}(|\sigma\rangle\langle\sigma| X) \geq 0 \; \forall \, |\sigma\rangle \in \mathcal{S}\} \tag{9.5}$$

where $\mathcal{S}$ denotes the set of pure n-qubit stabilizer states. For each number n of qubits, Λ is a bounded polytope with a finite number of vertices [7]. We denote the vertices of Λ by $\{A_\alpha \mid \alpha \in \mathcal{V}\}$ where $\mathcal{V}$ is an index set. Then the model is defined by the following theorem.

Theorem 25 (Ref. [2]; Theorem 1)**.** *For any number of qubits n,*

1. *Any n-qubit quantum state $\rho \in Herm_1^{\succeq 0}(\mathcal{H})$ can be decomposed as*

$$\rho = \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) A_\alpha, \tag{9.6}$$

   with $p_\rho(\alpha) \geq 0$ for all $\alpha \in \mathcal{V}$, and $\sum_\alpha p_\rho(\alpha) = 1$.

2. *For any A_α, $\alpha \in \mathcal{V}$, and any Clifford gate $g \in \mathcal{C}\ell$, $gA_\alpha g^\dagger$ is a vertex of Λ. This defines an action of the Clifford group on $\mathcal{V}$ as $gA_\alpha g^\dagger =: A_{g\cdot\alpha}$ where $g \cdot \alpha \in \mathcal{V}$.*

3. *For any A_α, $\alpha \in \mathcal{V}$ and any Pauli projector Π_a^s we have*

$$\Pi_a^s A_\alpha \Pi_a^s = \sum_{\beta \in \mathcal{V}} q_{\alpha,a}(\beta, s) A_\beta. \tag{9.7}$$

   with $q_{\alpha,a}(a, s) \geq 0$ for all $\beta \in \mathcal{V}$ and $s \in \mathbb{Z}_2$, and $\sum_{\beta,s} q_{\alpha,a}(\beta, s) = 1$.

This theorem describes a hidden variable model for QCM in which (i) states are represented by probability distributions p_ρ over $\mathcal{V}$, and (ii) Clifford gates and Pauli measurements are represented by stochastic maps $g \cdot \alpha$ and $q_{\alpha,a}$. In this model, no negative values are needed in the representation of states, gates, or measurements—it is a true probability representation, but we can no longer guarantee that the updates under Clifford gates and Pauli measurements are efficiently computable classically. Thus, although this representation does give a classical simulation algorithm for any magic state quantum circuit, the simulation is inefficient in general. Analyzing the efficiency of simulation using the Λ polytope requires a characterization of its vertices.

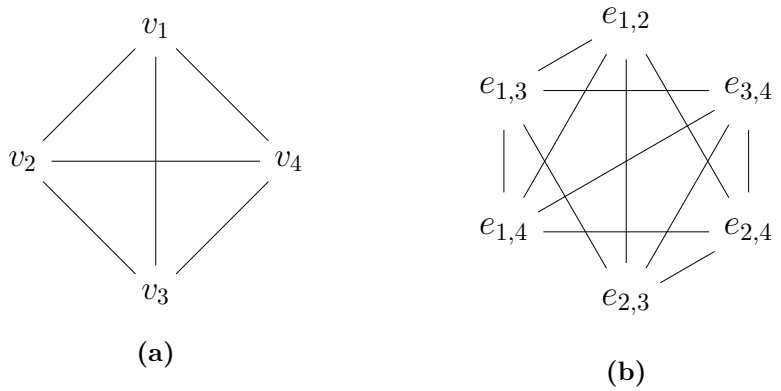


Figure 9.2: (a) The complete graph on four vertices K_4, and (b) the line graph of K_4 denoted $L(K_4)$. The edge between vertices v_i and v_j in K_4 maps to the vertex $e_{i,j}$ in $L(K_4)$. Two vertices in $L(K_4)$ share an edge iff the corresponding edges of K_4 share a vertex.

To date, only the Λ polytopes on one and two qubits have been fully characterized. In addition, some vertices of Λ are known for every qubit number. For example, it is known that the CNC-type phase space point operators associated to maximal CNC sets are vertices of the Λ polytopes [2, 72].

9.3 Multiqubit phase space from Jordan-Wigner transformations

Below we define a new quasiprobability representation intermediate between the CNC construction and the full Λ polytope model based on newly characterized vertices of Λ for every number of qubits. This new model can positively represent more states than the CNC construction (though not all quantum states as in the case of Λ), and it maintains the property of efficiently computable state update rules for the dynamics of QCM.

9.3.1 Line graphs and Jordan-Wigner transformations

The generalized phase space we define will consist of operators which can be described by quadratic polynomials of Majorana fermion operators. We motivate this approach by first considering the earlier CNC construction. As alluded to above, on n-qubits a maximal CNC set (without stabilizer state tensor factors) is given by $2n+1$ pair-wise anticommuting Pauli operators [1, §IV]. Such a set of Pauli operators satisfies the anticommutation relations

$$\{T_a, T_b\} := T_a T_b + T_b T_a = 2\delta_{a,b} \tag{9.8}$$

which are exactly the relations satisfied by a set of Majorana operators. Thus, under a Jordan-Wigner-like transformation, the maximal CNC operators are given by linear combinations of Majorana operators. Furthermore, there is a large body of work studying which operators can be described via quadratic polynomials of Majorana operators. In particular,

recently it has been shown that the operators describable in such a way can be identified by the structure of the graphs describing their anticommutation relations [73].

Given a graph $R = (\mathfrak{V}, \mathfrak{E})$, the line graph of R, $L(R) = (\mathfrak{E}, \mathfrak{E}')$, is the graph whose vertex set is the edge set of R, and whose edge set is $\mathfrak{E}' = \{(e_1, e_2) \in \mathfrak{E} \times \mathfrak{E} | e_1 \cap e_2 \neq \emptyset\}$; that is two vertices in $L(R)$ are neighbours if and only if the corresponding edges in R share a vertex. Given a line graph $G \cong L(R)$, we refer to the graph R as the root graph of $L(R)$. See Figure 9.2 for an example. We will also use the notion of twin vertices. Two vertices $u, v \in \mathfrak{V}$ are twin vertices if for every vertex $w \in \mathfrak{V}$, $(u, w) \in \mathfrak{E}$ if and only if $(v, w) \in \mathfrak{E}$.

Let $\mathcal{O} \subset E$ be a subset of the Pauli operators. We define the frustration graph of $\mathcal{O}$, $F(\mathcal{O}) = (\mathcal{O}, \mathfrak{E})$ as the graph whose vertices are identified with elements of $\mathcal{O}$, and with edges drawn between $a, b \in \mathcal{O}$ if and only if $[a, b] = 1$.

We now show that every line graph, $L(R)$, can be realized as the frustration graph of some set of Pauli operators $\mathcal{O}$ [73] (see Fig. 9.3 for an example). To do this, we first construct a set of $|R|$ pair-wise anticommuting Pauli operators. For even $|R|$, we can find such a set by taking the standard Jordan-Wigner transformation of Majorana fermion operators [202]

$$C_{2j-1} = \prod_{k=1}^{j-1} Z_k \otimes X_j \tag{9.9}$$

$$C_{2j} = \prod_{k=1}^{j-1} Z_k \otimes Y_j \tag{9.10}$$

for $j = 1, \ldots, |R|/2$. For odd $|R|$, we use the Jordan-Wigner transformation on $(|R| - 1)/2$ qubits and also include the operator

$$C_{|R|} = \prod_{k=1}^{(|R|-1)/2} Z_k. \tag{9.11}$$

Now we identify each vertex r in R with an operator C_r. We can then identify each edge (r, s) in R, and consequently each vertex in $L(R)$, with the Pauli operator $\pm iC_rC_s$. Two Pauli operators $\pm iC_rC_s$ and $\pm iC_tC_u$ anticommute if and only if they share exactly one Majorana operator, but this is equivalent to the associated edges in R sharing a vertex. Thus, the frustration graph of this set of quadratic Majorana operators is exactly the line graph $L(R)$, as desired. Note that this is not necessarily the most efficient way to realize each line graph as a frustration graph, though it does work for any line graph.

9.3.2 Definition of the generalized phase space

In this work we are interested in operators with the following structure:

Definition 13. *A Hermitian operator expanded in the Pauli basis as*

$$A^c_{\mathcal{O}} = \frac{1}{2^n}\left(1 + \sum_{b \in \mathcal{O}^*} c_b T_b\right) \tag{9.12}$$

is a line graph operator if the frustration graph of $\mathcal{O}^*$ is a line graph.

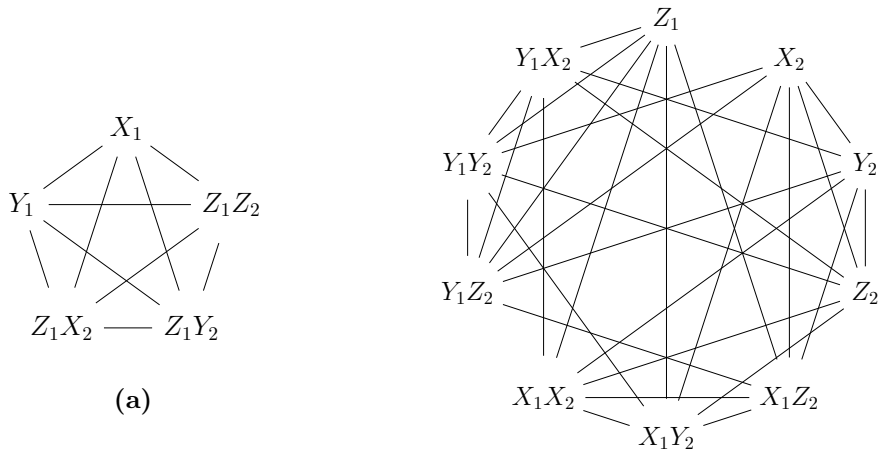


Figure 9.3: An example of a realization of the line graph $L(K_5)$ as the frustration graph of a set of Pauli operators. (a) First we find a set of five pair-wise anticommuting Pauli operators, in this case $\{X_1, Y_1, Z_1X_2, Z_1Y_2, Z_1Z_2\}$. The frustration graph of these operators is K_5. (b) The set of Pauli operators obtained by taking all products of pairs of operators in this set has frustration graph $L(K_5)$.

We actually wish to be a bit more permissive than requiring the frustration graphs to be exactly line graphs. In particular, we also want to include operators that have been obtained by tensoring on stabilizer states and conjugating by a Clifford gate, as in Eq. (9.4). In this case, the frustration graph will only be a line graph up to twin vertices [73]. Suppose we take an operator which is the tensor product of a projector onto an m-qubit stabilizer state, $|\sigma\rangle\langle\sigma|$, and an $n-m$-qubit operator, $A^c_{\mathcal{O}}$, whose frustration graph is a line graph. Then for each nontrivial Pauli operator in the expansion of $A^c_{\mathcal{O}}$ there are 2^m associated vertices in the frustration graph of $A^c_{\mathcal{O}} \otimes |\sigma\rangle\langle\sigma|$. These vertices will all be pair-wise twin vertices. Thus, if we can partition a graph into sets of equally sized pair-wise twin vertices then the associated operator will have a stabilizer tensor factor. It is thus possible to include stabilizer states in this formalism by collapsing sets of twin vertices.

In order for a model constructed from operators with the above structure to correctly reproduce the quantum mechanical predictions for all sequences of Pauli measurements, they must be in the Λ polytopes [2]. Otherwise, the model could predict negative quasiprobabilities for some sequences of Pauli measurements, and the model could fail to be closed under Pauli measurement. Without any more constraints, this defines an infinite set of operators for each number of qubits since the coefficients in Eq. (9.12) can vary continuously.

For the purpose of defining the generalized phase space of our model, we focus our attention on a finite subset of these operators. We choose this set in the following way. Let $\mathcal{O}$ be a set of Pauli operators such that the frustration graph of $\mathcal{O}^*$ is a line graph. The operators of the form Eq. (9.12) with support $\mathcal{O}$ form an affine subspace of $\mathrm{Herm}_1(\mathcal{H})$. Projecting the polytope Λ onto this affine subspace defines a new polytope with a finite

number of vertices, and when interpreted as a polytope embedded in the same $4^n - 1$-dimensional space as Λ with the Pauli coefficients of $E \setminus \mathcal{O}$ set to zero, it is contained in Λ. This projected polytope also has a finite number of vertices. Let $\mathcal{V}_{\mathcal{O}}$ be the index set for these vertices. We define an initial generalized phase space to be the set $\mathcal{V}_{LG} := \bigcup_{\mathcal{O}} \mathcal{V}_{\mathcal{O}}$, where we allow the support $\mathcal{O}$ to vary over all inclusion maximal sets of Pauli operators with line graphs as frustration graphs.

Finally, we define define our generalized phase space $\mathcal{V}$ to be the phase space $\mathcal{V}_{LG}$, augmented with stabilizer state tensor factors and conjugation by Clifford operations, analogous to Eq. (9.4). This augmentation of the line graph operators is natural for our purposes since tensoring on stabilizer states and conjugating by Clifford operators does not increase the cost of classical simulation of quantum computation [69], and furthermore, arises naturally since we allow Clifford gates and Pauli measurements in the computational model.

To summarize, the set of phase space point operators corresponding to our generalized phase space $\mathcal{V}$ is the set of all operators that can be constructed by the following procedure:

1. Start by defining a support $\tilde{\mathcal{O}} \subset E$ such that the frustration graph of $\tilde{\mathcal{O}}^*$ is a line graph, and $\tilde{\mathcal{O}}$ is inclusion maximal relative to this property.

2. Project Λ onto the subspace of $\mathrm{Herm}_1(\mathcal{H})$ spanned by the operators in $\tilde{\mathcal{O}}^*$, choose an extreme point $A^{\tilde{c}}_{\tilde{\mathcal{O}}}$ of this projected polytope.

3. Choose a stabilizer state $|\sigma\rangle \in \mathcal{S}$ and a Clifford gate $g \in \mathcal{C}\ell$, return

$$A^c_{\mathcal{O}} := g(A^{\tilde{c}}_{\tilde{\mathcal{O}}} \otimes |\sigma\rangle\langle\sigma|)g^\dagger \tag{9.13}$$

   as the phase space point operator.

Such an operator can be uniquely labeled by its support $\mathcal{O}$ in the Pauli basis and the corresponding Pauli basis coefficients c.

With these conditions defining the generalized phase space $\mathcal{V}$, any state ρ can be decomposed as

$$\rho = \sum_{(\mathcal{O},c)\in\mathcal{V}} W_\rho(\mathcal{O}, c) A^c_{\mathcal{O}}. \tag{9.14}$$

The function $W_\rho(\mathcal{O}, c)$ constitutes the representation of states in the model. Note that since the phase space point operators are over-complete, this representation is not unique. It is generally preferable to choose a representation that minimizes the amount of negativity as measured by the 1-norm of the coefficients $W_\rho(\mathcal{O}, c)$ [49]. This can be obtained through linear programming.

In the second step of the procedure above defining the generalized phase space $\mathcal{V}$, we take extreme points from a projection of the relevant Λ polytope. Including only extreme points in this way is convenient for the purpose of representing quantum states in Eq. (9.14) since it means $\mathcal{V}$ is finite for each number of qubits. However, this extremal property is not necessary for the classical simulation algorithm in the next section. We introduce the notation $\bar{\mathcal{V}}$ to be the index set corresponding to the set of all operators of the form Eq. (9.13), which are in Λ and for which the support $\tilde{\mathcal{O}}$ has a line graph structure.[3]

[3]Equivalently, for a given support $\mathcal{O}$ with a line graph for a frustration graph, $\bar{\mathcal{V}}_{\mathcal{O}}$ is an index set for the

9.4 Extended classical simulation

The goal of this section is to show that the generalized phase space constructed in Section 9.3 is closed under the dynamics of quantum computation with magic states—Clifford gates and Pauli measurements. This allows us to define a classical simulation algorithm for quantum computation with magic states in Section 9.4.3.

9.4.1 Closure under Clifford gates

First we consider the action of Clifford group on the generalized phase space $\mathcal{V}_{LG}$. We have the following Lemma.

Lemma 36. *For any line graph operator $A^c_{\mathcal{O}}$, $(\mathcal{O}, c) \in \mathcal{V}_{LG}$, and any Clifford group element $g \in \mathcal{C}\ell$, the operator $gA^c_{\mathcal{O}}g^\dagger$ is a line graph operator. That is, the phase space $\mathcal{V}_{LG}$ is closed under the action of the Clifford group.*

Proof of Lemma 36. Let $A^c_{\mathcal{O}}$ be an operator of the form Eq. (9.12) with $(\mathcal{O}, c) \in \mathcal{V}_{LG}$. I.e., the frustration graph of the Pauli operators in $\mathcal{O}^*$ is a line graph, and the coefficients c are chosen so that $A^c_{\mathcal{O}}$ is in Λ.

For any Clifford operation $g \in \mathcal{C}\ell$, we have

$$\begin{aligned} gA^c_{\mathcal{O}}g^\dagger =& \frac{1}{2^n}\sum_{b\in\mathcal{O}^*} c_b g T_b g^\dagger \\ =& \frac{1}{2^n}\sum_{b\in\mathcal{O}^*} c_b (-1)^{\Phi_g(b)} T_{S_g b} =: A^{g\cdot c}_{g\cdot\mathcal{O}}. \end{aligned}$$

This defines an action of g on the support $\mathcal{O}$ as

$$g \cdot \mathcal{O} = \{S_g b \mid b \in \mathcal{O}\}, \tag{9.15}$$

and on the coefficients c as

$$(g \cdot c)_{S_g b} = c_b (-1)^{\Phi_g(b)}. \tag{9.16}$$

Since Λ is closed under Clifford operations [2], if $A^c_{\mathcal{O}}$ is in Λ then $A^{g\cdot c}_{g\cdot\mathcal{O}}$ is also in Λ. Since Clifford operations preserve commutation relations of Pauli operators, the frustration graph of $g \cdot \mathcal{O}^*$ is isomorphic to that of $\mathcal{O}^*$. Thus, if the frustration graph of $\mathcal{O}^*$ is a line graph then so is the frustration graph of $g \cdot \mathcal{O}^*$. □

Corollary 9. *The phase spaces $\mathcal{V}$ and $\bar{\mathcal{V}}$ are each closed under Clifford group operations.*

This corollary follows immediately from Lemma 36 and from the definitions of $\mathcal{V}$ and $\bar{\mathcal{V}}$. In fact, this corollary follows immediately from the definitions even without Lemma 36. The point of introducing Lemma 36 is to note that there is some freedom how we choose to represent phase space point operators, by absorbing Clifford group elements into the coefficients of the line graph operator part via Eqs. (9.15) and (9.16). This freedom will

set of all operators in Λ projected onto $\mathcal{O}$. Then $\bar{\mathcal{V}}_{LG}$ and $\bar{\mathcal{V}}$ are defined relative to $\{\bar{\mathcal{V}}_{\mathcal{O}}\}_{\mathcal{O}}$ the same way $\mathcal{V}_{LG}$ and $\mathcal{V}$ are defined relative to $\{\mathcal{V}_{\mathcal{O}}\}_{\mathcal{O}}$.

become useful in describing the update of the phase space under Pauli measurements in the next section, and in the classical simulation algorithm in Section 9.4.3.

Corollary 10. *The set of states positively representable by $\mathcal{V}$ in Eq.* (9.14) *is closed under Clifford group operations.*

This follows immediately from Corollary 9 by considering the action of a Clifford group element $g \in \mathcal{C}\ell$ on a nonnegative decomposition of a state ρ of the form $\rho = \sum_{\alpha \in \mathcal{V}} W_\rho(\alpha) A_\alpha$. Conjugating by g on the left hand side gives the state $g\rho g^\dagger$, and on the right hand side we obtain a nonnegative decomposition of $g\rho g^\dagger$ in terms of phase space point operators of the phase space $\mathcal{V}$.

9.4.2 Closure under Pauli measurements

Now we consider the effect of Pauli measurements with respect to the generalized phase space $\mathcal{V}$ (or $\bar{\mathcal{V}}$). First, consider the action of a Pauli projector Π_a^s, corresponding to a Pauli measurement T_a, $a \in E$, yielding measurement outcome $s \in \mathbb{Z}_2$, on a phase space point operator $A_{\mathcal{O}}^c$, $(\mathcal{O}, c) \in \mathcal{V}_{LG}$ (or $\bar{\mathcal{V}}_{LG}$). This is the content of the following lemma.

Lemma 37. *For any line graph operator $A_{\mathcal{O}}^c$, $(\mathcal{O}, c) \in \mathcal{V}_{LG}$ (or $\bar{\mathcal{V}}_{LG}$), and any Pauli projector Π_a^s, we have*

1. $\mathrm{Tr}(\Pi_a^s A_{\mathcal{O}}^c) = \begin{cases} (1 + (-1)^s c_a)/2, & \text{if } a \in \mathcal{O}, \\ 1/2, & \text{otherwise.} \end{cases}$
2. *If $\mathrm{Tr}(\Pi_a^s A_{\mathcal{O}}^c) > 0$ then, up to normalization, $\Pi_a^s A_{\mathcal{O}}^c \Pi_a^s$ is a phase space point operator of $\bar{\mathcal{V}}$, and if $\mathrm{Tr}(\Pi_a^s A_{\mathcal{O}}^c) = 0$ then $\Pi_a^s A_{\mathcal{O}}^c \Pi_a^s = 0$*

Proof of Lemma 37. The first statement of the lemma follows from a direct calculation.

$$\begin{aligned} \mathrm{Tr}(A_{\mathcal{O}}^c \Pi_a^s) =& \frac{1}{2^{n+1}} \sum_{b \in \mathcal{O}} c_b \left[\mathrm{Tr}(T_b) + (-1)^s \mathrm{Tr}(T_a T_b)\right] \\ =& \frac{1}{2} \sum_{b \in \mathcal{O}} c_b \left[\delta_{b,0} + (-1)^s \delta_{b,a}\right] \\ =& \begin{cases} (1 + (-1)^s c_a)/2, & \text{if } a \in \mathcal{O}, \\ 1/2, & \text{otherwise.} \end{cases} \end{aligned}$$

Here in the second line we use the orthogonality relations of the Pauli operators $\mathrm{Tr}(T_a T_b) = 2^n \delta_{a,b}$.

For the second statement of the lemma, we take each case separately. First, assume $\mathrm{Tr}(\Pi_a^s A_{\mathcal{O}}^c) > 0$. Then conjugating $A_{\mathcal{O}}^c$ by the projector Π_a^s we have

$$\begin{aligned} \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s =& \frac{1}{2^{n+2}} \sum_{b \in \mathcal{O}} c_b [T_b + T_a T_b T_a + (-1)^s (T_a T_b + T_b T_a)] \\ =& \frac{1}{2^{n+1}} \sum_{b \in \mathcal{O} \cap a^\perp} c_b [T_b + (-1)^s T_a T_b] \\ =& \Pi_a^s A_{\mathcal{O} \cap a^\perp}^{c|_{\mathcal{O} \cap a^\perp}}. \end{aligned}$$

Using Witt's lemma [150], we can find a Clifford group element $g_1 \in \mathcal{C}\ell$ (depending on a and s) such that $g_1^\dagger \Pi_a^s g_1 = \Pi_{z_n}^0$ and $g_1^\dagger A_{\mathcal{O}\cap a^\perp}^{c|_{\mathcal{O}\cap a^\perp}} g_1 = A_{\tilde{\mathcal{O}}}^{\tilde{c}} \otimes 1$. Here $A_{\tilde{\mathcal{O}}}^{\tilde{c}}$ is an operator on $n-1$ qubits. Then we have

$$\Pi_a^s A_{\mathcal{O}}^c \Pi_a^s = \Pi_a^s A_{\mathcal{O}\cap a^\perp}^{c|_{\mathcal{O}\cap a^\perp}} = g_1 (A_{\tilde{\mathcal{O}}}^{\tilde{c}} \otimes \Pi_z^0) g_1^\dagger.$$

In this last expression, Π_z^0 is a single-qubit Pauli projector. The frustration graph of $\tilde{\mathcal{O}}^*$ is isomorphic to that of $(\mathcal{O} \cap a^\perp)^*$ which is a subgraph of the frustration graph of $\mathcal{O}^*$. Line graphs are closed under induced subgraphs [203], and so since the frustration graph of $\mathcal{O}^*$ is a line graph, the frustration graph of $\tilde{\mathcal{O}}^*$ is also a line graph. Since Λ is closed under Pauli measurements, up to normalization $\Pi_a^s A_{\mathcal{O}}^c \Pi_a^s$ is in Λ. Thus, $(\tilde{O}, \tilde{c}) \in \bar{\mathcal{V}_{LG}}$, and the operator $\Pi_a^s A_{\mathcal{O}}^c \Pi_a^s$ is a phase space point operator of $\bar{\mathcal{V}}$.

For the second case, assume that $\mathrm{Tr}(\Pi_a^s A_{\mathcal{O}}^c) = 0$, that is, $c_a = (-1)^{s+1}$. Following the same calculation as in the previous case,

$$\Pi_a^s A_{\mathcal{O}}^c \Pi_a^s = \frac{1}{2^{n+1}} \sum_{b \in \mathcal{O}\cap a^\perp} c_b \left[T_b + (-1)^s T_a T_b \right]. \tag{9.17}$$

I.e., the support of $\Pi_a^s A_{\mathcal{O}}^c \Pi_a^s$ is contained in the union of $\mathcal{O} \cap a^\perp$ and $a + (\mathcal{O} \cap a^\perp)$.

For any $d \in \mathcal{O} \cap a^\perp$ and any $t \in \mathbb{Z}_2$, by direct calculation we obtain

$$\begin{aligned}
&\mathrm{Tr}\big(\Pi_d^t \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s\big) = \mathrm{Tr}\big(\Pi_d^t \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s \Pi_d^t\big) \\
&= \frac{1}{2^{n+1}} \sum_{b \in \mathcal{O}\cap a^\perp} c_b \left[\mathrm{Tr}\big(\Pi_d^t T_b \Pi_d^t\big) + (-1)^s \,\mathrm{Tr}\big(\Pi_d^t T_a T_b \Pi_d^t\big) \right] \\
&= \frac{1}{2^{n+2}} \sum_{b \in \mathcal{O}\cap \{a,d\}^\perp} c_b [\mathrm{Tr}(T_b) + (-1)^s \,\mathrm{Tr}(T_a T_b) \\
&\qquad\qquad + (-1)^t \,\mathrm{Tr}(T_d T_b) + (-1)^{s+t} \,\mathrm{Tr}(T_a T_d T_b)] \\
&= \frac{1}{4} \left[1 + (-1)^s c_a + (-1)^t c_d + (-1)^{s+t+\beta(a,d)} c_{a+d} \right] \\
&= \frac{(-1)^t}{4} (c_d + (-1)^{s+\beta(a,d)} c_{a+d}).
\end{aligned}$$

Here in the first line we use idempotence of the projector Π_d^t and the cyclic property of the trace, in the second line we use Eq. (9.17), in the third line we expand the projectors and simplify, in the fourth line we use the orthogonality relations of the Pauli operators, and in the last line we use the assumption that $c_a = (-1)^{s+1}$.

That is, $\mathrm{Tr}\big(\Pi_d^1 \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s\big) = -\,\mathrm{Tr}\big(\Pi_d^0 \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s\big)$. In order to not contradict the assumption $A_{\mathcal{O}}^c \in \Lambda$, we must have $\mathrm{Tr}\big(\Pi_d^0 \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s\big) = \mathrm{Tr}\big(\Pi_d^1 \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s\big) = 0$. Then

$$\mathrm{Tr}(T_d \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s) = \mathrm{Tr}\big(\Pi_d^0 \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s\big) - \mathrm{Tr}\big(\Pi_d^1 \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s\big) = 0.$$

This holds for all $d \in \mathcal{O} \cap a^\perp$.

A similar argument shows that $\mathrm{Tr}(T_d \Pi_a^s A_{\mathcal{O}}^c \Pi_a^s) = 0$ for all $d \in a + (\mathcal{O} \cap a^\perp)$. Therefore, the support of $\Pi_a^s A_{\mathcal{O}}^c \Pi_a^s$ is empty and so $\Pi_a^s A_{\mathcal{O}}^c \Pi_a^s = 0$. □

Corollary 11. *The phase space $\bar{\mathcal{V}}$ is closed under Pauli measurements.*

Proof of Corollary 11. This corollary follows from a combination of Lemma 37 along with the techniques used in the proof of Ref. [3, Theorem 3]. Ref. [3, Theorem 3] reduces a Pauli measurement on an operator of the form $g(A \otimes |\sigma\rangle\langle\sigma|)g^\dagger$ to measurement only on A, along with Clifford corrections. The statement of this theorem is for the case where the operator A is a vertex of Λ, but it applies to our case as well and the proof is identical. This idea is closely related to the model of Pauli-based quantum computation [111, 112].

Consider a phase space point $(\mathcal{O}, c) \in \bar{\mathcal{V}}$. By definition, the corresponding phase space point operator has the structure

$$A^c_{\mathcal{O}} = g\left(A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma|\right) g^\dagger$$

for some Clifford operator $g \in \mathcal{C}\ell$ and stabilizer state $|\sigma\rangle$, where $(\mathcal{O}', c') \in \bar{\mathcal{V}}_{LG}$. Applying the projection Π^s_a,

$$\begin{aligned}\Pi^s_a A^c_{\mathcal{O}} \Pi^s_a =& \Pi^s_a g\left(A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma|\right) g^\dagger \Pi^s_a \\ =& g\Pi^{s'}_{a'}\left(A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma|\right)\Pi^{s'}_{a'} g^\dagger.\end{aligned}$$

Here in the second line we use the defining property of Clifford operations—that they map Pauli operators to Pauli operators under conjugation—to map g past Π^s_a.

The Pauli observable $T_{a'}$ can be decomposed as $T_{a'} \propto T_{a_1} \otimes T_{a_2}$ where T_{a_1} acts only on the subspace of the Hilbert space supporting $A^{c'}_{\mathcal{O}'}$ and T_{a_2} acts only on the subspace supporting $|\sigma\rangle$. Here we have two cases.

Case 1: If $\pm T_{a_2}$ is in the stabilizer group of $|\sigma\rangle$, the part of $T_{a'}$ acting on the second subspace can be replaced by the eigenvalue of T_{a_2}. In more detail, suppose $(-1)^\nu T_{a_2}$ is in the stabilizer of the stabilizer state $|\sigma\rangle$. Then,

$$\begin{aligned}&\Pi^{s'}_{a'}\left(A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma|\right)\Pi^{s'}_{a'} \\ &=A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma| + (-1)^{s'}\left[\left(T_{a_1}A^{c'}_{\mathcal{O}'} \otimes T_{a_2}|\sigma\rangle\langle\sigma|\right)\right. \\ &\quad\left. + \left(A^{\vec{c'}}_{\mathcal{O}'}T_{a_1} \otimes |\sigma\rangle\langle\sigma| T_{a_2}\right)\right] + T_{a_1}A^{c'}_{\mathcal{O}'}T_{a_1} \otimes T_{a_2}|\sigma\rangle\langle\sigma| T_{a_2} \\ &=A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma| + (-1)^{s'+\nu}\left[\left(T_{a_1}A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma|\right)\right. \\ &\quad\left. + \left(A^{\vec{c'}}_{\mathcal{O}'}T_{a_1} \otimes |\sigma\rangle\langle\sigma|\right)\right] + T_{a_1}A^{c'}_{\mathcal{O}'}T_{a_1} \otimes |\sigma\rangle\langle\sigma| \\ &=\Pi^{s'+\nu}_{a_1}A^{c'}_{\mathcal{O}'}\Pi^{s+\nu}_{a_1} \otimes |\sigma\rangle\langle\sigma|.\end{aligned}$$

Now we can apply Lemma 37.

Case 2: If $\pm T_{a_2}$ is not in the stabilizer of $|\sigma\rangle$, then T_{a_2} anticommutes with some Pauli operator in the stabilizer of $|\sigma\rangle$ and the outcome s' is uniformly random. Consequently, there exists a Clifford element $g_2 \in \mathcal{C}\ell$ (depending on a' and σ) such that $g_2 T_{a'} g_2^\dagger = Z_{n-k+1}$

and $g_2^\dagger \left(A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma| \right) g_2 = A^{\tilde{c}'}_{\tilde{\mathcal{O}}'} \otimes \Pi^0_x \otimes |\sigma'\rangle\langle\sigma'|$. Then we have

$$\begin{aligned}
&\Pi^{s'}_{a'} \left(A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma| \right) \Pi^{s'}_{a'} \\
&\quad = g_2^\dagger \Pi^{s'}_{z_{n-k+1}} \left(A^{\tilde{c}'}_{\tilde{\mathcal{O}}'} \otimes \Pi^0_x \otimes |\sigma'\rangle\langle\sigma'| \right) \Pi^{s'}_{z_{n-k+1}} g_2.
\end{aligned}$$

Up to normalization this last expression is equivalent to

$$\begin{aligned}
&g_2^\dagger \left(H X^{s'} \right)_{n-k+1} \left(A^{\tilde{c}'}_{\tilde{\mathcal{O}}'} \otimes \Pi^0_x \otimes |\sigma'\rangle\langle\sigma'| \right) \left(X^{s'} H \right)_{n-k+1} g_2 \\
&= g_2^\dagger \left(H X^{s'} \right)_{n-k+1} g_2 \left(A^{c'}_{\mathcal{O}'} \otimes |\sigma\rangle\langle\sigma| \right) g_2^\dagger \left(X^{s'} H \right)_{n-k+1} g_2
\end{aligned}$$

where H is the Hadamard gate. This operator has the structure of Eq. (9.12), and thus represents an operator of the generalized phase space $\bar{\mathcal{V}}$. □

Corollary 12. *The set of states positively representable by $\bar{\mathcal{V}}$ in Eq.* (9.14) *is closed under Pauli measurements.*

Proof of Corollary 12. Suppose a state ρ has a nonnegative decomposition in Eq. (9.14). If a Pauli measurement T_a is performed on ρ giving measurement outcome $s \in \mathbb{Z}_2$ the post-measurement state is

$$\begin{aligned}
\frac{\Pi^s_a \rho \Pi^s_a}{\mathrm{Tr}(\Pi^s_a \rho)} &= \frac{1}{\mathrm{Tr}(\Pi^s_a \rho)} \sum_{(\mathcal{O},c)\in\mathcal{V}} W_\rho(\mathcal{O},c) \Pi^s_a A^c_{\mathcal{O}} \Pi^s_a \\
&= \sum_{\substack{(\mathcal{O},c)\in\mathcal{V} \\ \mathrm{Tr}\left(\Pi^s_a A^c_{\mathcal{O}}\right)\neq 0}} \frac{\mathrm{Tr}(\Pi^s_a A^c_{\mathcal{O}})}{\mathrm{Tr}(\Pi^s_a \rho)} W_\rho(\mathcal{O},c) \cdot \frac{\Pi^s_a A^c_{\mathcal{O}} \Pi^s_a}{\mathrm{Tr}\left(\Pi^s_a A^c_{\mathcal{O}}\right)}.
\end{aligned}$$

Which gives a positive decomposition in terms of phase space point operators in $\bar{\mathcal{V}}$ whenever $W_\rho(\mathcal{O},c)$ is positive. □

9.4.3 Classical simulation algorithm

As shown in Sections 9.4.1 and 9.4.2, the generalized phase space over which the quasiprobability representation is defined is closed under Clifford gates and Pauli measurements. This fact allows us to define a classical simulation algorithm for quantum computation with magic states that applies whenever the representation of the input state is nonnegative, Algorithm 6. The proof of correctness for this algorithm is analogous to the proof of Theorem 5 of Ref. [1] and the proof of Theorem 2 of Ref. [2]. With certain additional assumptions, this algorithm is also efficient, this is the result of the following theorem.

Theorem 26. *The classical simulation given in Algorithm 6 is efficient on all input states ρ for which there exists a decomposition Eq.* (9.14) *with nonzero coefficients $W_\rho(\mathcal{O},c)$, and for which samples can be efficiently obtained from the distribution W_ρ.*

Input: p_ρ, sequence of gates and measurements $\mathcal{C}$
Output: sequence of simulated measurement outcomes $\mathcal{M}$

sample a point $(g, |\sigma\rangle, A^c_{\mathcal{O}}) \in \mathcal{V}$ according to the probability distribution p_ρ
while end of circuit has not been reached **do**
  if a Clifford gate $h \in \mathcal{C}\ell$ is encountered **then**
    update $(g, |\sigma\rangle, A^c_{\mathcal{O}}) \leftarrow (hg, |\sigma\rangle, A^c_{\mathcal{O}})$
  end if
  if a Pauli measurement $T_a,\ a \in E$ is encountered **then**
    define $T_{a_1} \otimes T_{a_2} := T_{S_g a}$
    if $T_{a_2} |\sigma\rangle = (-1)^\nu |\sigma\rangle$ **then**
      choose $s = 0$ or $s = 1$ with probability $\frac{1+(-1)^s c_{a_1}}{2}$
      add $(-1)^{s+\nu+\Phi_g(a)}$ to $\mathcal{M}$
      update $g \leftarrow g \cdot g_1$
      update $|\sigma\rangle \leftarrow |s\rangle \otimes |\sigma\rangle$
      update $A^c_{\mathcal{O}} \leftarrow g_1^\dagger \cdot \left(A^{\mathcal{O} \cap a_1^\perp}_{c|_{\mathcal{O} \cap a_1^\perp}} \right) \cdot g_1$
    else
      choose $s = 0$ or $s = 1$ with equal probability
      add $(-1)^{s+\Phi_g(a)}$ to $\mathcal{M}$
      update $g \leftarrow g \cdot g_2^\dagger \cdot (HX^s)_{n-k-1} \cdot g_2$
    end if
  end if
end while

Algorithm 6: One run of the classical simulation of quantum computation with magic states based on the quasiprobability representation defined in Sections 9.3 and the update rules described in Sections 9.4.1 and 9.4.2. We describe points in $\mathcal{V}$ (and $\bar{\mathcal{V}}$) by a triple in accordance with Eq. (9.13) and g_1 and g_2 are defined as in Lemma 37 and Corollary 11 respectively. The algorithm provides samples from the joint probability distribution of the Pauli measurements in a quantum circuit consisting of Clifford gates and Pauli measurements applied to an input state ρ such that $W_\rho \geq 0$.

Proof of Theorem 26. In order to prove efficiency of the classical simulation algorithm, we must prove (I) existence of an efficient representation of each component (i.e. phase space points, Clifford operations, and Pauli measurements), and (II) efficiently computable updates of the phase space points under Clifford operations and Pauli measurements.

(I) Pauli observables are represented by $2n$ bit strings (possibly with an extra bit indicating a sign) as in Eq. (2.3). Clifford operations can be specified by a $2n \times 2n$ bit matrix indicating the action of the Clifford on a basis of Pauli operators, together with an extra $2n$ bits indicating the signs of the Pauli operators.[4] Algorithms for converting between unitary Pauli and Clifford operators and this representation exist [204]. In the simulation,

[4]This description of a Clifford operation contains some redundant information as the constraint that the Clifford operation must preserve the commutation relations of the Pauli operators constrains the binary matrix to be a symplectic map on E.

we represent phase space points as in Eq. (9.13), by a triple $(g, |\sigma\rangle, A^c_{\mathcal{O}})$, where $g \in \mathcal{C}\ell$ is a Clifford group element, $|\sigma\rangle$ is a stabilizer state, and $A^c_{\mathcal{O}}$ is a line graph operator with $(\mathcal{O}, c) \in \mathcal{V}_{LG}$ (or $\bar{\mathcal{V}}_{LG}$). As above, specifying the Clifford operator requires a $2n \times 2n$ binary matrix together with an additional $2n$ bits. The stabilizer state $|\sigma\rangle$ can be specified by the generators of its stabilizer group, which requires $2n^2 + n$ bits, namely, $2n^2$ bits to specify the n Pauli operators that generate the group, and another n bits for their signs. Finally, specifying $A^c_{\mathcal{O}}$ requires specifying its support $\mathcal{O}$ and the coefficients c. Every line graph operator has at most $O(n^2)$ nonzero coefficients when expanded in the Pauli basis [205]. Therefore, specifying the $\mathcal{O}$ requires $O(n^3)$ bits. The coefficients c are $O(n^2)$ real numbers, approximation of the coefficients by bits is a matter of numeric precision.

(II) Proving efficiency of the the updates of phase space points under Clifford gates and Pauli measurements relies on the structure of the phase space points in Eq. (9.13). These operators have a stabilizer-like part consisting of the Clifford operation g and the stabilizer state $|\sigma\rangle$, and a line graph operator $A^{\tilde{c}}_{\tilde{\mathcal{O}}}$. These two parts can be tracked separately in the simulation. Under Clifford gates, the line graph operator part is unaffected, and update of the stabilizer part follows from standard stabilizer techniques. Under Pauli measurements, using now standard stabilizer techniques [3, 111, 112] (described in the previous section), the measurement can be efficiently reduced to a measurement only on the line graph part. Update of this part follows from the fact that the line graph property restricts the number of terms that can appear when the operator is expanded in the Pauli basis, as alluded to above. Some additional details for some of the steps are provided below.

We consider a circuit on an initial n-qubit state consisting of m Clifford gates and m' Pauli measurements. Whenever a Clifford gate is encountered, for the update in line 4 of Algorithm 6, we must compute the composition of two n-qubit Clifford gates. This requires two steps: matrix multiplication in $\mathbb{Z}_2^{2n}$ and $O(n^2)$ evaluations of the β function in Eq. (2.6) to determine the updated signs of the Clifford operation. The time complexity of each of these steps is polynomial in n.

If a Pauli measurement $a \in E$ is encountered we proceed in several steps outlined in lines 6–19 of Algorithm 6. First, we compute $S_g a$, and find the subvectors a_1 and a_2. This involves matrix-vector multiplication in $\mathbb{Z}_2^{2n}$, which has time complexity $O(n^2)$, and a partitioning of the vector into two subvectors depending on the number of qubits supporting the stabilizer state $|\sigma\rangle$. Determining whether $\pm T_{a_2}$ is in the stabilizer group of $|\sigma\rangle$ is a matter of determining whether a_2 is in the span of the Pauli generators of the stabilizer group. This is again a linear algebra problem in $\mathbb{Z}_2^{2n}$ with that can be accomplished using Gaussian elimination (i.e. time complexity $O(n^3)$. In the case where $\pm T_{a_2}$ is in the group, ν can be determined by recursive use of the relation Eq. (9.1) on the signs of the generators of the stabilizer group.

The remaining steps can be easily filled in. They involve linear algebra subroutines in $\mathbb{Z}_2^{2n}$, each with time complexity polynomial in n. The number of the linear algebra subroutines that need to be performed is polynomial in m, m', and n. Therefore, assuming samples can be efficiently obtained from the distribution representing the input state of the circuit, the simulation algorithm is efficient. □

In general a state ρ will not admit a decomposition such that $W_\rho(\mathcal{O}, c) \geq 0$ for all $(\mathcal{O}, c)$.

Consequently we define the robustness of a state as

$$\mathfrak{R}(\rho) := \min_{W} \left\{ ||W||_1 \;\middle|\; \rho = \sum_{(\mathcal{O},c)\in\mathcal{V}} W(\mathcal{O},c) A_{\mathcal{O}}^{c} \right\}.$$

Since line graph operators are preserved under Clifford gates and Pauli measurement the robustness is a monotone with respect to the resource theory of stabilizer quantum computation [42]. The complexity of the simulation algorithm in the presence of negativity can then be related to the robustness [49]. This generalized robustness is bounded above by the phase-space robustness of the CNC construction [1] and the robustness of magic [21]. See Appendix A.5.3 for details.

9.5 New vertices of the Λ polytopes

The generalized phase space of the present model is defined in part by looking at projections of the Λ polytopes. Depending on the subspace on which we project, the projected polytope may share vertices with Λ. In this section we show that this is indeed the case. In particular, we show that, for any number n of qubits, projecting Λ onto the space of operators with support $\mathcal{O}$ for which the frustration graph of $\mathcal{O}^*$ is $L(K_{2n+1})$—the line graph of the complete graph on $2n+1$ vertices—gives a polytope which shares vertices with Λ. This is the content of Theorem 27 below.

In this case we can determine the coefficients of the vertices in the Pauli basis as well, thus we obtain a complete characterization of new families of vertices of the Λ polytopes for every number of qubits. By polar duality [7, 72], we also obtain a complete characterization of new families of facets of the stabilizer polytope for every number of qubits.

Theorem 27. *Define the operator*

$$A_{\mathcal{O}}^{\eta} = \frac{1}{2^n}\left(1 + \frac{1}{n}\sum_{b\in\mathcal{O}^*}(-1)^{\eta(b)} T_b\right) \tag{9.18}$$

where $\mathcal{O}^*$ is any set of Pauli operators whose frustration graph is $L(K_{2n+1})$. Then there exist choices for the signs $\eta : \mathcal{O}^* \to \mathbb{Z}_2$ such that the operators $A_{\mathcal{O}}^{\eta}$ of the form Eq. (9.18) are vertices of Λ.

The rest of this section is devoted to the proof of this Theorem.

Proof of Theorem 27. By Theorem 18.1 of Ref. [160], to prove that an operator $A_{\mathcal{O}}^{\eta}$ is a vertex of Λ, it suffices to show (1) that $A_{\mathcal{O}}^{\eta}$ is in Λ, and (2) that the set of projectors onto stabilizer states which are orthogonal to $A_{\mathcal{O}}^{\eta}$ with respect to the Hilbert-Schmidt inner product has rank $4^n - 1$ when viewed as vectors of Pauli basis coefficients.

To start, we can directly compute the inner product $\mathrm{Tr}\big(\Pi_I^r A_{\mathcal{O}}^{\eta}\big)$ for any choice of signs

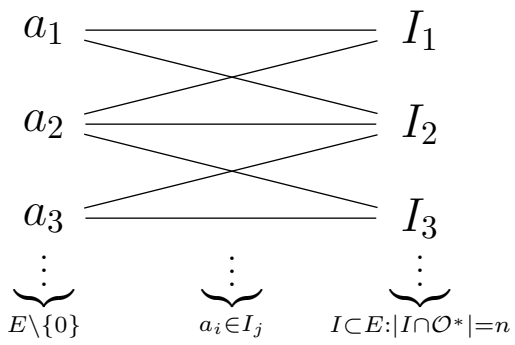


Figure 9.4: A sketch of a bipartite graph describing the inclusion relations of elements $a_i \in E \setminus \{0\}$ in a set of isotropic subspaces $I_j \subset E$ each with the property that $|I_j \cap \mathcal{O}^*| = n$. An edge connects an observable a_i on the left to an isotropic subspace I_j on the right if and only if $a_i \in I_j$.

η and any projector onto a stabilizer state Π_I^r:

$$\begin{aligned}\operatorname{Tr}\left(\Pi_I^r A_{\mathcal{O}}^{\eta}\right) =& \frac{1}{2^n} + \frac{1}{n2^{2n}} \sum_{a\in I} \sum_{b\in\mathcal{O}^*} (-1)^{r(a)+\eta(b)} \operatorname{Tr}(T_a T_b) \\ =& \frac{1}{2^n} + \frac{1}{n2^n} \sum_{a\in I\cap\mathcal{O}^*} (-1)^{r(a)+\eta(a)}.\end{aligned}$$

Since the largest independent set in $L(K_{2n+1})$ has size n, this inner product is always nonnegative and so $A_{\mathcal{O}}^{\eta} \in \Lambda$. Also, the inner product is zero if and only if $|I \cap \mathcal{O}^*| = n$ and $r(a) \neq \eta(a)$ for all $a \in I \cap \mathcal{O}^*$. All that remains is to show that the signs η can be chosen so that the set of stabilizer states for which this inner product is zero has rank $4^n - 1$.

Consider the bipartite graph G constructed as follows: G has a vertex for each Pauli observable $a \in E \setminus \{0\}$, a vertex for each isotropic subspace $I \subset E$ such that $|I \cap \mathcal{O}^*| = n$, and an edge connecting a vertex $a \in E$ to an isotropic subspace $I \subset E$ if and only if $a \in I$. A sketch of this graph is shown in Fig. 9.4. We need to show that there exists a choice of signs for the edges of this graph such that each off-diagonal block of the corresponding signed adjacency matrix has rank $4^n - 1$.

Using a graph theoretic result proven in Appendix A.5.1 (see in particular Corollary 15), to establish this it suffices to show that the graph G has a matching of size $4^n - 1$. By Kőnig's theorem, this is true if and only if the minimum vertex cover of G has order $4^n - 1$. One vertex cover of order $4^n - 1$ is obtained by taking all vertices on the Pauli operator side of the bipartition. Now we must show that a smaller vertex cover cannot be obtained by removing vertices from the Pauli side and adding fewer vertices on the isotropic subspace side. To do this we compute the degree of each vertex in the graph. If we can show that the degree of each vertex on the left hand side of the graph in Fig. 9.4 is larger than the degree of every vertex on the right hand side of the graph, then the minimum vertex degree has size $4^n - 1$ and the result is proved.

Each isotropic subspace $I \subset E$ has order 2^n, one of these elements is 0, and so the degree of each vertex on the right hand side is $2^n - 1$. Now we need to count the number of isotropic subspaces $I \subset E$ containing each observable $a \in E \setminus \{0\}$. Here we have many cases. Up to an overall phase, each Pauli observable T_a can be written as a product of some subset of the $2n+1$ pair-wise anticommuting observables $C_1, C_2, \ldots, C_{2n+1}$ which generate $\mathcal{O}$ by taking pair-wise products. Since the product of all $2n+1$ of these Pauli operators is proportional to the identity, this representation is not unique. Each Pauli observable a will have exactly two factorizations of the form $T_a \propto C_{\mu_1} C_{\mu_2} \cdots C_{\mu_k}$, one where k is odd and one where k is even. Therefore, without loss of generality we can restrict our attention to the factorizations where k is even.

Suppose $T_a \propto C_{\mu_1} C_{\mu_2} \cdots C_{\mu_{2m}}$. In order for $I \subset E$ to contain a, I must contain a set of m generators, each of which is proportional to a product of a pair of the operators $C_{\mu_1}, C_{\mu_2}, \ldots$. Once these m generators are specified, the remaining $n - m$ generators of I must be chosen from products of pairs of operators from the remaining operators $\{C_1, C_2, \ldots\} \setminus \{C_{\mu_1}, C_{\mu_2}, ...\}$. Therefore, the number of isotropic subspaces I containing a is

$$\begin{aligned} f(n,m) = &\left[\frac{1}{m!}\prod_{j=0}^{m-1}\binom{2m-2j}{2}\right] \cdot \frac{1}{(n-m)!}\prod_{k=0}^{n-m-1}\binom{2n+1-2m-2k}{2} \qquad (9.19) \\ =&\frac{1}{m!(n-m)!2^n}\left[\prod_{j=1}^{2m} j\right] \cdot \left[\prod_{j=1}^{2n+1-2m} j\right] \\ =&\frac{(2m)!(2n-2m)!}{m!(n-m)!2^n} \cdot (2n+1-2m). \end{aligned}$$

For this function, we can establish the bound $f(n,m) \geq 2^n - 1$ for all n, m. This proof of this bound is given in Lemma 42 in Appendix A.5.2. Therefore, each vertex on the left hand side of the graph in Figure 9.4 has higher degree than every vertex on the right hand side, and so the minimum vertex cover has size $4^n - 1$. This completes the proof. □

9.6 Discussion

In this work, we presented a new quasiprobability representation for quantum computation with magic states based on generalized Jordan-Wigner transformations. We demonstrated that this representation has efficiently computable update rules with respect to Clifford gates and Pauli measurements. Moreover, it extends previous representations including those based on quasiprobabilistic decompositions in projectors onto stabilizer states [21], and the CNC construction [1]. By leveraging this new construction we can efficiently simulate magic state quantum circuits on a larger class of input states than was previously known, thus pushing back the boundary between efficiently classically simulable, and potentially advantageous quantum circuits.

This model has a close connection to the probabilistic model of quantum computation based on the Λ polytopes. Namely, for each number n of qubits, it defines a new polytope contained inside of the Λ polytope which shares some vertices with the Λ polytope. These

vertices include the previously known CNC vertices [1, 2, 72], as well as some new infinite families of vertices as shown by Theorem 27.

Outlook. Our results provide several avenues for future research. One possible direction is to look for more families of vertices of the Λ polytopes which also have the line graph structure like those described in Theorem 27.

A more speculative idea is to look at the potential application of these results for magic state distillation protocols. One of our results is a characterization of a new family of vertices of the Λ polytope. By polar duality [7], this also gives a full characterization of new families of facets of the stabilizer polytopes for every number of qubits. Facets of the two-qubit stabilizer polytope have in the past been linked to magic state distillation [105]. The CNC type facets [1, 72], as well as the new facets defined by Theorem 27 could be used in a similar way for any number of qubits.

Chapter 10

Simulating quantum computation: how many "bits" for "it"?

A recently introduced classical simulation method for universal quantum computation with magic states operates by repeated sampling from probability functions [M. Zurel et al. PRL 260404 (2020)]. This method is closely related to sampling algorithms based on Wigner functions, with the important distinction that Wigner functions can take negative values obstructing the sampling. Indeed, negativity in Wigner functions has been identified as a precondition for a quantum speed-up. However, in the present method of classical simulation, negativity of quasiprobability functions never arises. This model remains probabilistic for all quantum computations. In this chapter, we analyze the amount of classical data that the simulation procedure must track. We find that this amount is small. Specifically, for any number n of magic states, the number of bits that describe the quantum system at any given time is $2n^2 + O(n)$.

This chapter has previously appeared in Ref. [5].

10.1 Introduction

In an article of 1989 [100], John Archibald Wheeler argued that quantum physics required a new perspective on reality based on information theoretic concepts. He wrote: "No element in the description of physics shows itself as closer to primordial than the elementary quantum phenomenon, that is, the elementary device-intermediated act of posing a yes-no physical question and eliciting an answer or, in brief, the elementary act of observer-participancy. Otherwise stated, every physical quantity, every it, derives its ultimate significance from bits, binary yes-or-no indications, a conclusion we epitomize in the phrase, *it from bit*."

A prototypical realization of this view on physics has been provided in the description of quantum computation with magic states (QCM) through the Λ polytopes [2], in which the quantum phenomena at hand are reproduced, without any deviation or approximation, by repeated sampling processes dependent on and producing bit strings. The "*It*" in this case is universal quantum computation, and hence all nonrelativistic quantum mechanics in finite-dimensional Hilbert spaces. The "*Bits*" represent the binary outcomes of Pauli

measurements and the labels of the vertices of the Λ polytopes in the repeated sampling. There are finitely many such vertices for any number n of magic states.

This description of quantum computation with magic states grew out of the analysis of Wigner function negativity as a precondition for a quantum computational speedup, a research programme that started with Refs. [27] and [19]. Specifically, in Ref. [19] it was established that negativity in Gross' Wigner function [29, 31] is required for a quantum speedup, under the condition that the Hilbert space dimension is odd. Analogous results were subsequently established in even dimension, specifically for rebits [57] and qubits [1]. However, in the end it turned out that once sufficiently general (quasi)probability functions are admitted, there is no need for any negativity at all [2, 7]. Universal quantum computation in the magic state model can be described by repeated sampling from a generalized phase space whose points are labeled by the vertices of the Λ polytopes. This process essentially resembles a random walk, with the complication that the transition function changes from one time step to the next and can depend on the prior sampling history. A summary of this sampling process is given in Section 10.3; see in particular Theorem 28 and Algorithm 7.

The question of interest for the present work is how much classical information the simulation of quantum computations must track, i.e., *What is the length of the bit strings produced in the simulation?* For example, if it turned out that those bit strings were very long, say exponentially long in the number n of magic states, this would provide a convenient explanation for the hardness of classical simulation of universal quantum computation using Λ polytopes. If the information storage itself is inefficient, so is the processing. However, this is not what we find. We find that the bit strings are short. Specifically, they are of length $O(n^2)$. Thus, simulation of universal quantum computation based on Λ polytopes is a *small data problem*. The presumed hardness of this simulation must come from the computational hardness of the sampling processes involved, not from moving around large amounts of data.[1]

Naively, the length of the bit strings labeling the vertices of Λ_n are upper-bounded by $\log_2(|\mathcal{V}_n|)$, with $|\mathcal{V}_n|$ the size of the generalized phase space, i.e., the number of vertices of Λ_n. To date, an upper bound and a lower bound are known for this quantity, namely

$$\frac{n^2}{2} \leq \log_2(|\mathcal{V}_n|) \leq 4^n n^2.$$

The lower bound is by Karanjai, Wallman, and Bartlett [71]. The upper bound comes from the upper bound theorem of polytope theory [207] (see Appendix A.6.1 for details).

The gap between the bounds is extremely wide, consistent with both efficient and inefficient storage of the bit strings. Numerical studies suggest that the number of phase space points is growing very rapidly with n; for $n = 1$ the number of phase space points is 8, for $n = 2$ it is 22320, and for $n > 2$ we don't know the precise numbers, but the estimate for $n = 3$ is already huge.

The following simple insight is crucial for establishing our main result: in the QCM model, for any fixed value n, all quantum computations start in the *same* magic state $|M\rangle^{\otimes n}$.

[1] Elsewhere in the quantum domain there are sampling problems which are algorithmically hard classically but for which samples can be obtained quantum mechanically, for example, boson sampling [84], random circuit sampling [83], and certain instantaneous quantum circuits [206].

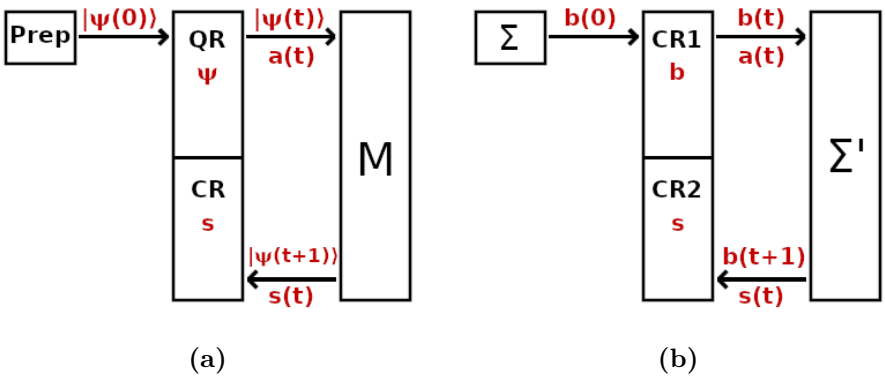


Figure 10.1: Quantum computation with magic states (a), and its simulation based on Λ polytopes (b).

Therefore, the question of interest for classical simulation of QCM using Λ polytopes is not "What is the size of the phase space $\mathcal{V}_n$?", but rather "What is the size of the region of $\mathcal{V}_n$ that can be reached from the initial magic state $|M\rangle^{\otimes n}$?". This size can be computed, and it turns out to be small. It implies a new upper bound for the length of the bit strings that is within a factor of four of the matching lower bound [71]. This is the content of our main result, Theorem 29, presented in Section 10.4 below.

10.2 Quantum state as a string of bits

A crucial feature of the Λ polytope formalism is that the quantum state $|\Psi(t)\rangle$ of the system is replaced by a bit string $b(t)$ of bounded length. *That bit string $b(t)$ is a valid and accurate representation of the quantum system.*

The role of the bit string $b(t)$ is illustrated in Fig. 10.1. QCM consists of preparing a quantum register in a magic state $|M\rangle^{\otimes n}$, followed by a sequence of Pauli measurements; see Fig. 10.1a. This requires a device Prep to deliver the magic states to the quantum register QR, and a classical register CR to store the previous measurement record $\mathbf{s}$, a classical side computation to identify the label $a(t)$ of the Pauli observable measured in step t, and a measurement device M to perform the measurements and to output the corresponding results $s(t)$. The overall structure of the classical simulation is the same, but with the components modified; see Fig. 10.1b. Prep is replaced by a first sampler Σ that samples from the phase space distribution of the initial state $|M\rangle^{\otimes n}$. There are two classical registers, CR1 and CR2. The former stores the phase space samples $b(t)$, and the latter the prior measurement record, as in (a). The measurement device M is replaced by a second sampler Σ' that takes as input a phase space point $b(t)$ and a Pauli label $a(t)$, and outputs a new phase space point $b(t+1)$ as well as a measurement outcome $s(t)$. The same information that in the standard quantum mechanical description is carried by the quantum state $|\Psi(t)\rangle$ is in the Λ polytope description carried by the bit string $b(t)$.

For universal quantum computation in the magic state model, the classical simulation needs to reproduce the quantum mechanical prediction for the joint distribution of outcomes for any sequence of Pauli measurements. The Λ-polytope method does that, without any

approximation. For the predictions, the statistical distribution of the bit strings $\{b(t), \forall t\}$ matters, not individual values $b(t)$. This is the same for the quantum mechanical states $|\Psi(t)\rangle$. They too are conditioned on prior measurement outcomes, hence probabilistic.

10.3 Preliminaries

10.3.1 The Λ polytope model

In Ref. [2], a hidden variable model (HVM) is defined for quantum computation with magic states (QCM) [38]—a universal model of quantum computation in which computation proceeds through a sequence of Clifford gates and Pauli measurements on an initially prepared "magic state".

We denote by $\mathrm{Herm}(\mathcal{H})$ the space of Hermitian operators on Hilbert space $\mathcal{H}$, and unless otherwise specified, $\mathcal{H}$ (or $\mathcal{H}_n$) is the n-qubit Hilbert space $(\mathbb{C}^2)^{\otimes n}$. $\mathrm{Herm}_1(\mathcal{H})$ is the affine subspace of $\mathrm{Herm}(\mathcal{H})$ consisting of operators with unit trace and $\mathrm{Herm}_1^{\succeq 0}(\mathcal{H})$ is the subset of $\mathrm{Herm}_1(\mathcal{H})$ consisting of positive semidefinite operators. $\mathrm{Herm}_1^{\succeq 0}(\mathcal{H})$ contains the density operators representing physical quantum states.

The state space of the hidden variable model of Ref. [2] is based on the Λ polytopes. Denoting the set of pure n-qubit stabilizer states by $\mathcal{S}_n$, the Λ polytope for n qubits is defined as

$$\Lambda_n = \{X \in \mathrm{Herm}_1(\mathcal{H}_n) \mid \mathrm{Tr}(|\sigma\rangle\langle\sigma| X) \geq 0 \; \forall \, |\sigma\rangle \in \mathcal{S}_n\}\,. \tag{10.1}$$

For any fixed number $n \in \mathbb{N}$ of qubits, Λ_n is a bounded polytope with a finite number of vertices [7]. We denote the vertices of Λ_n by $\{A_\alpha \mid \alpha \in \mathcal{V}_n\}$ where $\mathcal{V}_n$ is an index set [2]. The hidden variable model is defined by the following theorem.

Theorem 28 (Ref. [2], Theorem 1 & Ref. [7], Theorem 1)**.** *For any number of qubits $n \in \mathbb{N}$,*

1. *Any n-qubit quantum state $\rho \in Herm_1^{\succeq 0}(\mathcal{H}_n)$ can be decomposed as*

$$\rho = \sum_{\alpha \in \mathcal{V}_n} p_\rho(\alpha) A_\alpha, \tag{10.2}$$

   with $p_\rho(\alpha) \geq 0$ for all $\alpha \in \mathcal{V}_n$, and $\sum_\alpha p_\rho(\alpha) = 1$. I.e. any n-qubit quantum state ρ can be represented by a probability distribution p_ρ over $\mathcal{V}_n$.

2. *For any A_α, $\alpha \in \mathcal{V}_n$, and any Clifford gate $g \in \mathcal{C}\ell_n$, $gA_\alpha g^\dagger$ is a vertex of Λ_n. This defines an action of the Clifford group on $\mathcal{V}_n$ as $gA_\alpha g^\dagger =: A_{g\cdot\alpha}$ where $g \cdot \alpha \in \mathcal{V}_n$.*

3. *For any A_α, $\alpha \in \mathcal{V}_n$, and any Pauli projector Π_a^s, we have*

$$\Pi_a^s A_\alpha \Pi_a^s = \sum_{\beta \in \mathcal{V}_n} q_{\alpha,a}(\beta, s) A_\beta, \tag{10.3}$$

[2] Λ_n refers to a specific mathematical structure, a polytope, for a fixed number of qubits n. We refer to these polytopes collectively as the Λ polytopes, or Λ (no subscript). The set of vertices of Λ_n is $\{A_\alpha \mid \alpha \in \mathcal{V}_n\}$. That is, A_α is a vertex of Λ_n and it is labeled by $\alpha \in \mathcal{V}_n$. $\mathcal{V}_n$ is the set of all labels of vertices of Λ_n. There is no n subscript on A_α (or α) even though it refers to an object associated to a fixed number of qubits n, because this fact is implicit in the statement $\alpha \in \mathcal{V}_n$.

with $q_{\alpha,a}(\beta,s) \geq 0$ for all $\beta \in \mathcal{V}_n$ and $s \in \mathbb{Z}_2$, and $\sum_{\beta,s} q_{\alpha,a}(\beta,s) = 1$. I.e. Pauli measurements are represented by a stochastic map from (phase-space-point, measurement) pairs $(\alpha, a) \in \mathcal{V}_n \times E_n$ to (phase-space-point, measurement outcome) pairs $(\beta, s) \in \mathcal{V}_n \times \mathbb{Z}_2$.

A classical simulation algorithm for QCM based on sampling from the defining probability distributions of this HVM is given in Algorithm 7. The algorithm returns samples from the distribution of measurement outcomes for the quantum circuit being simulated which agree with the predictions of quantum theory [2].

```
sample α ∈ V_n according to p_ρ : V_n → R_{≥0}
propagate α through the circuit
while the end of the circuit has not been reached do
   if a Clifford gate g ∈ Cℓ_n is encountered then
      update the phase space point according to α ↼ g · α
   end if
   if a Pauli measurement a ∈ E_n is encountered then
      sample (β, s) ∈ V_n × Z_2 according to q_{α,a}
      return s ∈ Z_2 as the outcome of the measurement
      update the phase space point according to α ↼ β
   end if
end while
```

Algorithm 7: Classical simulation of a single run of a magic state quantum circuit with input state ρ

10.4 Main result

In this section we present our main result—the amount of classical data that the simulation procedure Algorithm 7 must track is small. First, in section 10.4.1 we give the result for a simplified version of the computational model wherein we only allow Pauli measurements, no Clifford gates, and in particular, we allow only sequences of independent and commuting Pauli measurements. This simplified model is still universal for quantum computation [111, 112]. Then, in Section 10.4.2 we give a more general statement of the main result where we allow computations consisting of any sequence of Clifford gates and Pauli measurements.

10.4.1 Simplified case

Although the most general quantum computation in QCM could consist of any sequence of Clifford gates and Pauli measurements performed on an arbitrary input state, we can make several assumptions simplifying the computational model while preserving the property of quantum computational universality.

First, in QCM, we usually assume that every computation starts from a fixed magic input state. For example, using the standard magic state circuit gadget [12, Figure 10.25], any Clifford+T circuit with n T-gates acting on m qubits can be turned into a QCM circuit (Clifford gates and Pauli measurements only) on $n+m$ qubits acting on a state of the form

$|0\rangle^{\otimes m} \otimes |H\rangle^{\otimes n}$ where $|H\rangle = (|0\rangle + \exp(i\pi/4)\,|1\rangle)/\sqrt{2}$. In general, it suffices to consider input states of the form $|0\rangle^{\otimes m} \otimes |M\rangle^{\otimes n}$ where $|M\rangle$ is a fixed single-qubit magic state [105].

Second, we can do away with the Clifford gates altogether [57, 111]. To see this, note that the Clifford gates can always be propagated forward in time through the circuit, conjugating the Pauli measurements into other Pauli measurements. Once they are propagated past the final measurement in the circuit they can be removed since they no longer affect the statistics of the measurement outcomes.

We can also do away with the stabilizer part of the input. The Pauli circuit on the $(m+n)$-qubit state $|0\rangle^{\otimes m} \otimes |M\rangle^{\otimes n}$ can be simulated by an adaptive sequence of Pauli measurements acting only on the n-qubit magic part of the input [3, 112]. The stabilizer part is handled by extra classical side processing.

Finally, it suffices to consider only sequences of commuting Pauli measurements up to length n [112]. This is because when a Pauli measurement is encountered which anticommutes with a measurement that was previously performed, the outcome of the measurement will be uniformly random, and the update of the state after the measurement can be implemented by a Clifford gate. Therefore, the anticommuting measurement can always be replaced by a coin flip to determine the measurement outcome and a Clifford gate to implement the state update which can then be propagated past the future measurements. The longest sequences of independent pair-wise commuting Pauli measurements on n qubits have length n.

To summarize, for universal quantum computation, it suffices to consider adaptive sequences of pair-wise commuting Pauli measurements of length n acting on a fixed magic state of the form $|M\rangle^{\otimes n}$. A simplified version of the classical simulation algorithm above for circuits of this form is given by Algorithm 8. For a complete description of the above circuit simplifications, along with a method for compiling a given QCM circuit into an adaptive Pauli circuit, see Ref. [112].

1: sample $\alpha_0 \in \mathcal{V}_n$ according to $p_\rho : \mathcal{V}_n \to \mathbb{R}_{\geq 0}$
2: **for all** $a_t, t \in \{1, 2, \ldots, n\}$ **do**
3: sample $(\alpha_t, s_t) \in \mathcal{V}_n \times \mathbb{Z}_2$ according to q_{α_{t-1}, a_t}
4: **return** $s_t \in \mathbb{Z}_2$ as the outcome of measurement a_t
5: **end for**

Algorithm 8: Classical simulation of a single run of a Pauli-based quantum circuit with input state ρ

We can now state our main result.

Theorem 29 (Main result)**.** *Any quantum computation consisting of a sequence of n independent, pair-wise commuting Pauli measurements on a fixed magic state $|M\rangle^{\otimes n}$ can be simulated by Algorithm 8 using a memory of $2n^2 + 3n$ bits to specify the phase space points reached.*

Note that the precise form of the magic input state used is not important for the proof of this theorem, so long as the state is fixed for each number of qubits. For universal quantum computation, it suffices to consider n copies of a fixed single-qubit magic state such as

$|H\rangle^{\otimes n}$. The proof of this theorem relies on the following result from convex geometry (see e.g. Ref. [161, §1.6]).

Theorem (Carathéodory's theorem)**.** *If a point x of $\mathbb{R}^D$ lies in the convex hull of a set V, then x can be written as the convex combination of at most $D+1$ points in V.*

Proof of Theorem 29. Since the generalized phase space point operators $\{A_\alpha \,|\, \alpha \in \mathcal{V}_n\}$ are not a basis for $\mathrm{Herm}_1(\mathcal{H}_n)$, they are overcomplete, the distributions p_ρ that represent states and the distributions $q_{\alpha,a}$ that represent Pauli measurements in the model of Theorem 28 are not unique. Since Λ_n lives in $\mathrm{Herm}_1(\mathcal{H}_n)$, a real affine space of dimension $4^n - 1$, by Carathéodory's theorem there exist choices for the distribution p_ρ such that $|\mathrm{supp}(p_\rho)| \leq 4^n$. Similarly, for each $s \in \mathbb{Z}_2$, there exist choices for $q_{\alpha,a}(-, s)$ such that $|\mathrm{supp}(q_{\alpha,a}(-, s))| \leq 4^n$. To start, we fix a canonical choice of the probability distributions p_ρ and $q_{\alpha,a}$ satisfying these properties. With the distribution p_ρ representing the input state of the circuit fixed, we can label the elements of the support by the numbers $1, 2, \ldots, 4^n$, and so specifying a sample from this distribution (i.e., specifying an element of the support of p_ρ), amounts to specifying an integer between 1 and 4^n, which takes no more than $\log_2(4^n) = 2n$ bits.

There are $4^n - 1$ nontrivial n-qubit Pauli measurements, therefore, specifying each measurement requires no more than $2n$ bits. For the t^{th} measurement a_t, the distribution q_{α_{t-1},a_t} is uniquely specified by the sampling record consisting of states $\alpha_0, \alpha_1, \ldots, \alpha_{t-1}$, measurements $a_1, a_2, \ldots, a_{t-1}$, and measurement outcomes $s_1, s_2, \ldots, s_{t-1}$. Once this distribution is fixed, with the canonical choice above, specifying a sample from this distribution requires no more than $2n + 1$ bits (1 bit for s_t, and as before $2n$ bits for α_t).

Since the length of the measurement sequence is no more than n, the number of classical bits required to specify the complete sampling record is no more than

$$\underbrace{2n}_{\alpha_0} + \sum_{t=1}^{n} \left[\underbrace{2n}_{a_t} + \underbrace{1}_{s_t} + \underbrace{2n}_{\alpha_t} \right] = 4n^2 + 3n.$$

See Figure 10.2 for a summary of the accounting used to get this bound.

This initial bound can be improved in a number of ways. First, for the purpose of simulation, we don't need to store α_n since there are no more measurements. This immediately removes $2n$ bits. Second, after the t^{th} measurement, the value of any measurement in the span of $a_1, a_2, \ldots, a_t$ is already determined. Therefore, if the $t + 1^{th}$ measurement a_{t+1} is to be independent and commute with the previous measurements, it is chosen from $\{a_1, a_2, \ldots, a_t\}^\perp / \mathrm{span}(a_1, a_2, \ldots, a_t) \cong E_{n-t}$. Specifying a measurement chosen from this set requires only $2(n - t)$ bits, not the full $2n$ bits.

Finally, we can perform another simplification which reduces the number of qubits by 1 after each measurement. After a measurement of $a \in E_n$ giving outcome $s \in \mathbb{Z}_2$, the relevant state space is projected to $\Pi_a^s \Lambda_n \Pi_a^s$. This is contained in a $4^{n-1} - 1$ dimensional subspace of $\mathrm{Herm}_1(\mathcal{H}_n)$. There exists a Clifford gate $g \in \mathcal{C}\ell_n$ such that $g\Pi_{z_n}^0 g^\dagger = \Pi_a^s$, and $\Pi_a^s \Lambda_n \Pi_a^s = g\Pi_{z_n}^0 \Lambda_n \Pi_{z_n}^0 g^\dagger = g(\Lambda_{n-1} \otimes |0\rangle\langle 0|) g^\dagger$. Therefore, after the measurement the Clifford gate g can be propagated out yielding a Pauli measurement circuit on an input state with a stabilizer state tensor factor. Then the above mentioned circuit simplifications [3, 112] can be used to remove the stabilizer part of the input. This reduction can be performed

$$\rho \xrightarrow{a_1} s_1 \xrightarrow{a_2} s_2 \xrightarrow{a_3} s_3 \cdots$$
(a)

$$p_\rho \xrightarrow{q_{\alpha_0,a_1}} (\alpha_1, s_1) \xrightarrow{a_{\alpha_1,a_2}} (\alpha_2, s_2) \xrightarrow{q_{\alpha_2,a_3}} (\alpha_3, s_3) \cdots$$
(b)

$$\alpha_0 \xrightarrow{a_1} (\alpha_1, s_1) \xrightarrow{a_2} (\alpha_2, s_2) \xrightarrow{a_3} (\alpha_3, s_3) \cdots$$
(c)

$$2n \quad +2n \quad +2n+1 \quad +2n \quad +2n+1 \quad +2n \quad +2n+1 \quad \cdots$$
(d)

Figure 10.2: Three descriptions of a quantum computation. (a) A circuit level description. Pauli measurements $a_1, a_2, \ldots$ are performed on the input state ρ yielding measurement outcomes $s_1, s_2, \ldots$. (b) The representation of this computation in the model of Theorem 28. The input state is represented by the probability distribution p_ρ, each measurement a_t is represented by a probability distribution q_{α_{t-1},a_t}. (c) A single run of the simulation algorithm based on the probabilistic representation of the computation. We start by sampling from p_ρ to obtain α_0. For each measurement a_t, we sample from q_{α_{t-1},a_t} to obtain measurement outcome s_t and updated state α_t. (d) An upper bound on the number of classical bits required to store each piece of the sampling record of (c) according to Theorem 29.

after each measurement. This is similar to the idea behind the reduced classical simulation of Theorem 3 of Ref. [3]. With this dimension reduction after each measurement, to specify the sample (α_t, s_t) requires no more than $2(n-t)+1$ bits. This dimension reduction implicitly implements both of the reductions of the previous paragraph. Note that this explicit dimension reduction is not required to get the memory savings since, according to Ref. [3, Theorem 2], $\Lambda_{n-t} \otimes |0\rangle\langle 0|^{\otimes t}$ is a subpolytope of Λ_n contained in a $4^{n-t}-1$-dimensional affine subspace of $\mathrm{Herm}_1(\mathcal{H}_n)$, and so by Carathéodory's theorem, an operator of the form $X \otimes |0\rangle\langle 0|^{\otimes t}$ can be decomposed in vertices of this subpolytope with support of size no more than 4^{n-t}.

With these reductions, the complete measurement record can be specified with no more than

$$\underbrace{2n}_{\alpha_0} + \sum_{t=1}^{n} \left[\underbrace{2(n-t+1)}_{a_t} + \underbrace{1}_{s_t} + \underbrace{2(n-t)}_{\alpha_t} \right] = 2n^2 + 3n$$

classical bits, which is the claimed bound. □

A similar idea as the one used in the proof of Theorem 29 of explicitly tracking the measurement record of a quantum circuit has previously been used to define a contextual hidden variable model for the stabilizer subtheory [208].

10.4.2 General case

Theorem 29 used a simplified version of QCM, where, without loss in computational power, all Clifford gates have been eliminated and the measurement sequence is shrunk to at most n commuting Pauli measurements. In the remainder of this section we demonstrate that the bound on the size of the reachable phase space region does not increase much if we do not make these simplifications, i.e., if we admit arbitrarily long sequences of (potentially noncommuting) Pauli measurements, and Clifford gates between them.

Specifically, we have the following additional result.

Corollary 13. *Any quantum computation consisting of an arbitrarily long sequence of Pauli measurements and Clifford unitaries, applied to a fixed magic state $|M\rangle^{\otimes n}$, can be simulated using a memory of $4n^2 + 6n$ bits to specify the reachable phase space points.*

Comparing with Theorem 29 we find that the memory requirement merely doubles, in particular, the quadratic scaling with the number n of magic states remains unchanged.

Proof of Corollary 13. Every vertex of Λ_n is in exactly one orbit of vertices with respect to the action of the Clifford group. Therefore, the number of Clifford orbits travelled by Algorithm 8 is smaller than or at most equal to the number of vertices travelled. A consequence of Theorem 29 thus is that the number of equivalence classes travelled by Algorithm 8 is no more than 2^{2n^2+3n}.

At every given point in the classical simulation, the single vertex A_ν under consideration by the simulation algorithm has a product structure, $A_\nu = g(\tilde{A}_\nu \otimes |S\rangle\langle S|)g^\dagger$, where $|S\rangle\langle S|$ is the projector onto an m-qubit stabilizer state with stabilizer group S, $\tilde{A}_\nu$ is a vertex of Λ_{n-m}, and g is a Clifford unitary [3].

When switching from Algorithm 8 specialized to the simplified computational model back to the general Algorithm 7, the following additional situations may occur at any given step: (i) the next operation is a Clifford unitary, (ii) the next operation is the measurement of a Pauli observable T_a, $a \in E_n$, such that $\pm T_a \in gSg^\dagger$, and (iii) the next operation is the measurement of a Pauli observable T_a that does not commute with all elements of $gSg^\dagger$.

In case (i), the Clifford orbit $[A_\nu]$ of A_ν doesn't change. In case (ii), the measurement outcome deterministically follows from the stabilizer, and no update of A_ν occurs at all. Hence, no update of $[A_\nu]$ occurs either. In case (iii), the measurement outcome is uniformly random, and for either outcome the ensuing transformation can be replaced by a Clifford unitary. Thus we are back to case (i)—no update of $[A_\nu]$ occurs. The conclusion is that the number of Clifford orbits of vertices reachable by the more general Algorithm 7 equals the number of orbits reachable by the specialized Algorithm 8 applying to the canonical form of QCMs.

Now, the size of every Clifford orbit of vertices of Λ_n is upper-bounded by the size $|\mathcal{C}\ell_n|$ of the n-qubit Clifford group. Therefore, the number of vertices reachable by Algorithm 7 is bounded by $2^{2n^2+3n} \times |\mathcal{C}\ell_n|$. The number of bits required to specify an n-qubit Clifford gate is no more than $2n^2 + 3n$. Namely, the conjugation action of a Clifford unitary on a Pauli operator is described by a binary-valued $2n \times 2n$ matrix plus a $2n$-component binary vector for the signs of Pauli operators [12], requiring $2n \times (2n + 1)$ bits. From this, the $2n^2 - n$ constraints imposed by preservation of the Pauli commutation relations need to be subtracted. Hence, to specify a vertex reachable in the simulation using Algorithm 7, we

specify a Clifford orbit $[A_\nu]$ and a Clifford operation g, which takes no more than $4n^2 + 6n$ bits, as claimed. □

We may generalize beyond the scenario of Corollary 13 by supplementing the initial magic state $|H\rangle^{\otimes n}$ with an m-qubit ancilla in a stabilizer state, where the stabilizer part can be kept separate and tracked independently in the simulation. As in the previous generalization, the computational power does not increase, but a broader physical setting is represented.

Then the number of bits needed depends on both the number of qubits m supporting the stabilizer part of the input, and the number of magic states n, namely,

$$[2m^2 + m] + [4n^2 + 6n] + [2(n + m)^2 + 3(n + m)]. \tag{10.4}$$

Here, the first term accounts for the stabilizer part of the state, the second term for the the magic part (this is the first version of the bound given in the proof of Theorem 29 shown in Fig. 10.2), and the third term is for the $m + n$ qubit Clifford group element that needs to be tracked, as in the proof of Corollary 13.

In deriving Eq. (10.4), we make use of the fact that a tensor product of a projector onto an m-qubit stabilizer state with a vertex of Λ_n is a vertex of Λ_{m+n} [3, Theorem 2], and we use the stabilizer formalism techniques from Ref. [3, §4] to isolate update of the stabilizer part from the magic part of the state. The number of bits tracked remains quadratic in n and m.

10.5 Discussion

To summarize, in this work we have shown that the classical simulation of universal quantum computation using the Λ-polytopes [2] is a small data problem. Specifically, it is shown that, with respect to the model of quantum computation with magic states, the number of bits that represent the quantum system at any stage of the simulation is quadratic in the number n of the magic states. Classical simulation of quantum computation is (presumably) still hard. In the present case this hardness does not stem from shuffling around lots of data, but instead from complicated operations on little data.

For illustration, we compare the above classical simulation method to two others, one very different, one rather similar. The first method is the straightforward simulation of a quantum system obtained by choosing a specific basis of the Hilbert space at hand, mapping operators to matrices and states to vectors. The state of the system at any moment in time is now described by exponentially many complex-valued amplitudes. The bit equivalent of a complex number is a matter of numeric precision, but in any case the amount of data to be processed is large. The state update is conceptually simple, and the computational hardness derives from the size of the objects involved. This simulation method and that of Ref. [2] thus represent opposite ends of the spectrum.

The second simulation method we compare to is that of sampling from Wigner functions [19], applicable to odd Hilbert space dimension. Here, the overall structure of the simulation is the same as in Ref. [2], i.e., repeated sampling from a phase space. There are two important differences, however. In Ref. [19], (i) the sampling is computationally efficient whenever it applies, but (ii) the sampling procedure does not apply to all initial magic

states. Specifically, it only applies when the Wigner function [29, 31] of the initial state is positive. Indeed, this is why negativity of the Wigner function is a precondition for computational speedup. In Ref. [2], (i) the sampling is not guaranteed to be computationally efficient, but (ii) it applies to all possible initial states.

We highlight two further aspects of the simulation method [2]:

- In the classical simulation of quantum computation using Λ-polytopes, the description of the system's state by $2n^2 + 3n$ bits does not invoke any approximation. The distributions of measurement outcomes sampled from are the exact quantum-mechanical ones. Thus, the data representing the system is genuinely discrete [3], and for this reason we regard Ref. [2] as a realization of Wheeler's "*it from bit*" proposal.

- Theorem 2 can be generalized to the odd-prime-dimensional qudit case [7]. In this case, the precompilation step reducing the computation to sequences of commuting Pauli measurements is also possible [113], but again not necessary.

The Λ-polytopes are only beginning to be explored, and they may hold many more surprises. From the perspective of quantum computation, an important question is the following. Two known efficient classical simulation methods, the stabilizer formalism [101] and sampling from nonnegative Wigner functions [19] are reproduced and extended by the Λ-polytope method, at the level of the simplest vertices of the polytopes. Where does efficient classical simulation end?

[3] Any continuous parameter dependence is relegated to the sampling probabilities in the first sampling step.

Chapter 11

Conclusion

At this point, it could be helpful to take a step back and summarize the story up to now. Motivation for a lot of the work presented in this thesis comes from previous work demonstrating the link between discrete Wigner functions and quantum computational advantage [19, 27–32]. This in turn is motivated by the traditional use of quasiprobability functions in physics [43, 116–118], where negativity in these representations serves as an indicator of genuinely quantum behaviour of physical systems [44, 45].

For odd-dimensional systems, everything is very simple [19, 30] (and particularly so for odd-prime dimensional qudits [30]). In this case, there is a unique well-behaved Wigner function which looks like a direct generalization of the original Wigner function to finite-dimensional quantum mechanics [4, 29, 31, 52]. By this we mean, a unique Wigner function satisfying the finite-dimensional generalization of the Stratonovich-Weyl criteria [48]. The fact that it is also happens to have properties conducive to describing quantum computations in the magic state model, namely Clifford covariance and preservation of positivity under Pauli measurements, is a nice added bonus.

But the fact that this simple picture falls apart when we go to even dimensions is somewhat counter-intuitive. There are many possible explanations for the nonexistence of "nice" Wigner functions in even dimensions, relating to, for example, quantum contextuality [3, 52], algebraic and representation theoretic properties of symplectic groups and Heisenberg groups [], and number theoretic properties of modular arithmetic (see Section 2.7.1).

Many have tried to reestablish the link between quasiprobability representations and quantum computation in even dimensions even dimensions with varying degrees of success [1, 60, 137, 141], but these typically require violating some of Stratovich-Weyl criteria that define Wigner functions. In some special cases [58, 63], it had been shown that there was no proper Wigner function which had the properties required for describing quantum computation. In Chapter 5, we prove this for the general case. That is, we show that for a large family of Wigner functions satisfying a suitable finite-dimensional verison of the Stratonovich-Weyl criteria, no Wigner function exists for even-dimensional qudits which is Clifford covariant (Theorem 11) or which preserves positivity under Pauli measurements (Theorem 14).

This means, in order to get a quasiprobability representation that can describe quantum computations in the magic state model, some of the Stratonovich-Weyl criteria must be abandoned. A theorem of Karanjai, Wallman, and Bartlett [71] gives a hint as to which

modifications need to be made. They give a memory lower bound for simulating the kind of quantum contextuality that comes up in magic state quantum circuits on even-dimensional qudits, i.e., proofs of contextuality involving the Pauli observables. A consequence of this is that if a quasiprobability representation can be used to simulate quantum computations in this model, the phase space point operators that define it cannot define an operator basis for the space of Hermitian operators on n-qudit Hilbert space. There must more phase space point operators than one would expect for a Wigner function, resulting in an overcomplete set.

A number of alternative quasiprobability representations have been proposed for quantum theory in even-dimensions [21, 60, 137], in particular, the so-called CNC construction [1] (presented in Chapter 3) has been proposed as the generalization of the results on odd-dimensional Wigner functions in quantum computation to systems of any dimension. This construction reproduces some features of the odd-dimensional Wigner function, including those needed for describing quantum computations like covariance with respect to the Clifford group, but at the cost of requiring a much larger phase space than one would expect for a Wigner function describing qubits.

In the original paper on the subject [1], it was claimed that, when applied to odd-dimensions, the CNC constructions reduced to the previous odd-dimensional Wigner function. This claim turned out to be incorrect. In Chapter 8, we presented the generalization of the CNC construction to qubits of odd-prime dimension. There we found that the CNC phase space includes all the points of the Wigner function phase space, but it includes new cnc-type phase space points as well. The resulting quasiprobability representation can represent more states positively, and thus can simulate more quantum computations that the Wigner function.

In pursuing useful quasiprobability representations for quantum computation, we arrived at the idea of removing all of the unnecessary assumptions that define Wigner functions, and focus on the ones that are required for simulating quantum computations. We still want a real, normalized quasiprobability function, but for example we don't need to assume the quasiprobability function comes from the expansion coefficients of a fixed operator basis. In addition, we explicitly enforce Clifford covariance and preservation of positivity for Pauli measurements. The way to construct this representation is to look for the largest subset of the space of unit-trace Hermitian operators which is closed under Clifford unitaries and Pauli measurement. This is how we get the Λ polytope model presented in Chapter 4 for qubits, and Chapter 7 for qudits of any Hilbert space dimension (an explanation of how the Λ polytopes fall out of this simple idea is given in Section A.3.1).

The model resulting from the Λ polytope construction, looks a lot like some of the previous quasiprobability representations for quantum computation on even dimensional-qudits. It has a deterministic update rule for Clifford group operations, and a probabilistic update rule for Pauli measurements, as well as a real-valued, normalized function representing states. The main feature setting the Λ polytope model apart is the fact that all states have a nonnegative representation.

Initally the Λ was defined only for systems of multiple qubits [2], but it has since been extended to systems of any dimension [7]. The qudit Λ polytopes share many properties with the qubit version. For example, the polytopes are closed under tensor products with projectors onto stabilizer states (Theorem 21), and certain CNC operators are vertices of

Λ. In the case of odd qudit Hilbert space dimensions, the phase space point operators of Gross' Wigner function are vertices of Λ (Theorem 20).

The main task in the rest of the chapters in this thesis is the further elucidate the hidden variable model defined in Chapters 4 and 7, and to explore the scope and the limits of the classical simulation algorithm that it spawns. First, in Chapter 6 we provide some basic structural properties of the Λ polytopes, for example, the model is closed under the operation of taking a tensor product with a projector onto a stabilizer state. In terms of the simulation algorithm, this means the cost of classical simulation does not increase under this operation, this is the result of Theorem 16 in Section 6.4. We also provide a first foray into territory beyond the CNC framework by characterization a new class of vertices of the polytope Λ_2.

In Chapter 9, we continue this venture by introducing a new quasiprobability represenation intermediate between the CNC framework and the full Λ polytope model. This new representation takes the motivation given by the structure of the multiqubit CNC phase space (Section 3.3.3) in which Jordan-Wigner transformations of Majorana fermion operators come up.

In the last research chapter of this thesis, Chapter 10, we present two new contributions regarding the Λ polytope model—one technical and one a matter of interpretation. The technical contribution is about the amount of classical data that needs to be tracked in the classical simulation of quantum computations using the Λ polytope models.

Since this model can simulate any quantum computation, the simulation must be inefficient in general. An early suggestion for the source of this inefficiency was that it must come from a blow-up in the size of the phase space. We show that, in fact, this is not the case. Since in the magic state model, for quantum computational universality it suffices to consider circuits starting from a fixed magic input state for each number of qubits, e.g., an n-fold tensor product of the single-qubit magic state $|H\rangle$. In this setting, the relevant quantity for determining the size of the phase space is not the total number of vertices of the Λ polytopes, but rather the number of vertices reachable starting from a fixed state. This we can count, and it turns out to be pretty small (Theorem 29 and Corollary 13).

Thus, the explanation for the hardness of the classical simulation of quantum computation is not the size of the phase space. The simulation algorithms, Algorithm 3 and Algorithm 4, are fairly simple. They consist simply of sampling from a sequence of probability distributions. Therefore, the only place left for the hardness to be hiding is in obtaining these samples.

Future work

Although progress has been made, the Λ polytopes, along with their hidden variable models and classical simulation algorithms, are relatively new, and there remains much to be discovered about them. The observations provided by the existing Λ subpolytope models, e.g. the CNC model of Chapters 3 and 8, and the line graph model of Chapter 9, leave us with a clear path forward. It proceeds in stages.

1. First, some sub-polytope models inside Λ provide efficient classical simulation of quantum computations, but if a quantum computational advantage is possible as is commonly believed then this efficiency must breakdown somewhere in between the CNC

sub-polytope model and the full Λ polytope. Now we can probe this phenomenon by characterizing more vertices of Λ. The more vertices we can efficiently characterize, the more quantum computations we can efficiently simulate classically, thus effectively pushing back the boundary between classical and quantum. Some progress has been made in this regard. In Ref. [3], some new vertices were characterized, and in Ref. [6] we found new infinite families of vertices for all numbers of qubits. That said, the vertices characterized so far remain a small fraction of the total number of vertices for $n \geq 3$ qubits, and so there is still much work to be done here and much to be gained from it.

2. Second, since negativity in a quasiprobability function is no longer the source of inefficiency of classical simulation, inefficiency must come another source. One early suggestion was that it came from a blowup in the size of the generalized phase space of the model. Although seemingly a reasonable suggestion, we recently ruled this out as the source of inefficiency as well [5]. The only possibility remaining is that it lies somewhere in the sampling from the probability distributions that represent states and measurements in the model. We would like to be able to better describe this source of the inefficiency in the classical simulation algorithm.

3. Finally, in the past sources of inefficiency in the classical simulation of quantum computation were linked to traditional indicators of nonclassicality in quantum theory, such as contextuality [20, 51]. We would like to determine how the inefficiency in the classical simulation based on the Λ polytopes relates to other notions of nonclassicality like generalized contextuality. Alternatively, here we could reverse this process, and derive an indicator of nonclassicality from the source of inefficiency found in the first two steps above. This indicator could then be applied to other quantum information processing tasks, and possibly to the foundations of quantum computation and quantum theory.

Although significant progress has been made, there is still much work to be done and much to be gained from pursuing this line of research further. In addition, the Λ polytopes are still relatively new objects [2, 7]. As a result, some information is known about them [3, 5–7], but there is much still to be discovered. In particular, they have so far only been used for the problem of characterizing quantum computational advantages in the magic state model. Some progress has been made in extending their applicability beyond this model [5, 200], but we believe there could be more applications yet to be discovered.

Bibliography

[1] Robert Raussendorf, Juani Bermejo-Vega, Emily Tyhurst, Cihan Okay, and Michael Zurel. "Phase-space-simulation method for quantum computation with magic states on qubits". In: *Physical Review A* 101.1 (2020), p. 012350. DOI: 10.1103/PhysRevA.101.012350. arXiv: 1905.05374.

[2] Michael Zurel, Cihan Okay, and Robert Raussendorf. "Hidden Variable Model for Universal Quantum Computation with Magic States on Qubits". In: *Physical Review Letters* 125.26 (2020), p. 260404. DOI: 10.1103/PhysRevLett.125.260404. arXiv: 2004.01992.

[3] Cihan Okay, Michael Zurel, and Robert Raussendorf. "On the extremal points of the Λ-polytopes and classical simulation of quantum computation with magic states". In: *Quantum Information & Computation* 21.13&14 (2021), pp. 1091–1110. DOI: 10.26421/QIC21.13-14-2. arXiv: 2104.05822.

[4] Robert Raussendorf, Cihan Okay, Michael Zurel, and Polina Feldmann. "The role of cohomology in quantum computation with magic states". In: *Quantum* 7 (2023), p. 979. DOI: 10.22331/q-2023-04-13-979. arXiv: 2110.11631.

[5] Michael Zurel, Cihan Okay, and Robert Raussendorf. "Simulating quantum computation with magic states: how many "bits" for "it"?" [to be published in PRX Quantum]. 2023. arXiv: 2305.17287.

[6] Michael Zurel, Lawrence Z. Cohen, and Robert Raussendorf. "Simulation of quantum computation with magic states via Jordan-Wigner transformations". 2023. arXiv: 2307.16034.

[7] Michael Zurel, Cihan Okay, Robert Raussendorf, and Arne Heimendahl. "Hidden variable model for quantum computation with magic states on qudits of any dimension". In: *Quantum* 8 (2024), p. 1323. DOI: 10.22331/q-2024-04-30-1323. arXiv: 2110.12318.

[8] Michael Zurel and Arne Heimendahl. "Efficient classical simulation of quantum computation beyond Wigner positivity". 2024. arXiv: 2407.10349.

[9] Michael Zurel. "Hidden variable models and classical simulation algorithms for quantum computation with magic states on qubits". Master's Thesis. University of British Columbia, 2020. DOI: 10.14288/1.0394790.

[10] Robert Raussendorf, Juani Bermejo-Vega, Emily Tyhurst, Cihan Okay, and Michael Zurel. "Erratum: Phase-space-simulation method for quantum computation with magic states on qubits [Phys. Rev. A 101, 012350 (2020)]". In: *Physical Review A* 105.3 (2022), p. 039902. DOI: 10.1103/PhysRevA.105.039902.

[11] Cihan Okay, Michael Zurel, and Robert Raussendorf. "Erratum: On the extremal points of the Lambda-polytopes and classical simulation of quantum computation with magic states". In: *Quantum Information & Computation* 22.7&8 (2022), pp. 627–628. DOI: 10.26421/QIC22.7-8-4.

[12] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. 10th Anniversary edition. Cambridge University Press, 2010. DOI: 10.1017/CBO9780511976667.

[13] Yuri Ivanovich Manin. *Vychislimoe i nevychislimoe*. Soviet Radio, 1980.

[14] Richard P. Feynman. "Simulating physics with computers". In: *International Journal of Theoretical Physics* 21 (1982), pp. 467–488. DOI: 10.1007/BF02650179.

[15] Peter W. Shor. "Algorithms for quantum computation: discrete logarithms and factoring". In: *Proceedings, 35th Annual Symposium on Foundations of Computer Science*. 1994, pp. 124–134.

[16] Lov K. Grover. "A fast quantum mechanical algorithm for database search". In: *Proceedings, 28th Annual ACM Symposium on the Theory of Computing (STOC)*. 1996, pp. 212–219.

[17] Aram W. Harrow, Avinatan Hassidim, and Seth Lloyd. "Quantum Algorithm for Linear Systems of Equations". In: *Physical Review Letters* 103.15 (2009), p. 150502. DOI: 10.1103/PhysRevLett.103.150502. arXiv: 0811.3171.

[18] Guifre Vidal. "Efficient classical simulation of slightly entangled quantum computations". In: *Physical Review Letters* 91.14 (2003), p. 147902. DOI: 10.1103/PhysRevLett.91.147902. arXiv: quant-ph/0301063.

[19] Victor Veitch, Christopher Ferrie, David Gross, and Joseph Emerson. "Negative quasi-probability as a resource for quantum computation". In: *New Journal of Physics* 14.11 (2012), p. 113011. DOI: 10.1088/1367-2630/14/11/113011. arXiv: 1201.1256.

[20] Mark Howard, Joel J. Wallman, Victor Veitch, and Joseph Emerson. "Contextuality supplies the 'magic' for quantum computation". In: *Nature* 510.7505 (2014), pp. 351–355. DOI: 10.1038/nature13460. arXiv: 1401.4174.

[21] Mark Howard and Earl T. Campbell. "Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing". In: *Physical Review Letters* 118.9 (2017). DOI: 10.1103/physrevlett.118.090501. arXiv: 1609.07488.

[22] Robert Raussendorf and Hans J. Briegel. "A One-Way Quantum Computer". In: *Physical Review Letters* 86.22 (2001), pp. 5188–5191. DOI: 10.1103/PhysRevLett.86.5188.

[23] Yaoyun Shi. "Both Toffoli and Controlled-NOT need little help to do universal quantum computation". In: *Quantum Information & Computation* 3.1 (2003), pp. 84–92. DOI: 10.26421/QIC3.1-7. arXiv: quant-ph/0205115.

[24] Carlton M. Caves, Ivan Deutsch, and Robin Blume-Kohout. "Physical-resource demands for scalable quantum computation". In: *Fluctuations and Noise in Photonics and Quantum Optics*. Ed. by Derek Abbott, Jeffrey H. Shapiro, and Yoshihisa Yamamoto. Vol. 5111. International Society for Optics and Photonics. SPIE, 2003, pp. 425–433. DOI: 10.1117/12.504760. arXiv: quant-ph/0304083.

[25] Janet Anders and Dan E. Browne. "Computational Power of Correlations". In: *Physical Review Letters* 102.5 (2009), p. 050502. DOI: 10.1103/PhysRevLett.102.050502. arXiv: 0805.1002.

[26] Robert Raussendorf. "Contextuality in measurement-based quantum computation". In: *Physical Review A* 88.2 (2013), p. 022322. DOI: 10.1103/physreva.88.022322. arXiv: 0907.5449.

[27] Ernesto F. Galvão. "Discrete Wigner functions and quantum computational speedup". In: *Physical Review A* 71.4 (2005), p. 042302. DOI: 10.1103/PhysRevA.71.042302. arXiv: quant-ph/0405070.

[28] Cecilia Cormick, Ernesto F. Galvão, Daniel Gottesman, Juan Pablo Paz, and Arthur O. Pittenger. "Classicality in discrete Wigner functions". In: *Physical Review A* 73.1 (2006), p. 012301. DOI: 10.1103/PhysRevA.73.012301. arXiv: quant-ph/0506222.

[29] David Gross. "Hudson's theorem for finite-dimensional quantum systems". In: *Journal of Mathematical Physics* 47.12 (2006), p. 122107. DOI: 10.1063/1.2393152. arXiv: quant-ph/0602001.

[30] David Gross. "Non-negative Wigner functions in prime dimensions". In: *Applied Physics B* 86.3 (2006), pp. 367–370. DOI: 10.1007/s00340-006-2510-9. arXiv: quant-ph/0702004.

[31] David Gross. "Computational power of quantum many-body states and some results on discrete phase spaces". PhD Thesis. Institute for Mathematical Sciences, Imperial College London, 2008. URL: https://www.thp.uni-koeln.de/gross/files/diss.pdf.

[32] Andrea Mari and Jens Eisert. "Positive Wigner Functions Render Classical Simulation of Quantum Computation Efficient". In: *Physical Review Letters* 109.23 (2012), p. 230503. DOI: 10.1103/PhysRevLett.109.230503. arXiv: 1208.3660.

[33] Eric Chitambar and Gilad Gour. "Quantum resource theories". In: *Reviews of Modern Physics* 91.2 (2019), p. 025001. DOI: 10.1103/RevModPhys.91.025001. arXiv: 1806.06107.

[34] Eric Chitambar, Debbie Leung, Laura Mančinska, Maris Ozols, and Andreas Winter. "Everything You Always Wanted to Know About LOCC (But Were Afraid to Ask)". In: *Communications in Mathematical Physics* 328.1 (2014), pp. 303–326. DOI: 10.1007/s00220-014-1953-9. arXiv: 1210.4583.

[35] Francesco Buscemi. "All Entangled Quantum States Are Nonlocal". In: *Physical Review Letters* 108.20 (2012), p. 200401. DOI: 10.1103/PhysRevLett.108.200401. arXiv: 1106.6095.

[36] David Schmid, Denis Rosset, and Francesco Buscemi. "The type-independent resource theory of local operations and shared randomness". In: *Quantum* 4 (2020), p. 262. DOI: 10.22331/q-2020-04-30-262. arXiv: 1909.04065.

[37] Daniel Gottesman, Alexei Kitaev, and John Preskill. "Encoding a qubit in an oscillator". In: *Physical Review A* 64.1 (2001), p. 012310. DOI: 10.1103/PhysRevA.64.012310. arXiv: quant-ph/0008040.

[38] Sergey Bravyi and Alexei Kitaev. "Universal quantum computation with ideal Clifford gates and noisy ancillas". In: *Physical Review A* 71.2 (2005), p. 022316. DOI: 10.1103/PhysRevA.71.022316. arXiv: quant-ph/0403025.

[39] Earl T. Campbell, Barbara M. Terhal, and Christophe Vuillot. "Roads towards fault-tolerant universal quantum computation". In: *Nature* 549.7671 (2017), pp. 172–179. DOI: 10.1038/nature23460. arXiv: 1612.07330.

[40] Daniel Gottesman. "The Heisenberg representation of quantum computers". In: *Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*. Ed. by S.P. Corney, R. Delbourgo, and P.D. Jarvis. 1999, pp. 32–43. arXiv: quant-ph/9807006.

[41] Scott Aaronson and Daniel Gottesman. "Improved simulation of stabilizer circuits". In: *Physical Review A* 70.5 (2004), p. 052328. DOI: 10.1103/PhysRevA.70.052328. arXiv: quant-ph/0406196.

[42] Victor Veitch, S.A. Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. "The resource theory of stabilizer quantum computation". In: *New Journal of Physics* 16.1 (2014), p. 013009. DOI: 10.1088/1367-2630/16/1/013009. arXiv: 1307.7171.

[43] Eugene Wigner. "On the Quantum Correction For Thermodynamic Equilibrium". In: *Physical Review* 40.5 (1932), pp. 749–759. DOI: 10.1103/PhysRev.40.749.

[44] R.L. Hudson. "When is the wigner quasi-probability density non-negative?" In: *Reports on Mathematical Physics* 6.2 (1974), pp. 249–252. DOI: 10.1016/0034-4877(74)90007-X.

[45] Anatole Kenfack and Karol Życzkowski. "Negativity of the Wigner function as an indicator of non-classicality". In: *Journal of Optics B: Quantum and Semiclassical Optics* 6.10 (2004), pp. 396–404. DOI: 10.1088/1464-4266/6/10/003. arXiv: quant-ph/0406015.

[46] William K. Wootters. "A Wigner-function formulation of finite-state quantum mechanics". In: *Annals of Physics* 176.1 (1987), pp. 1–21. DOI: 10.1016/0003-4916(87)90176-X.

[47] Kathleen S. Gibbons, Matthew J. Hoffman, and William K. Wootters. "Discrete phase space based on finite fields". In: *Physical Review A* 70.6 (2004), p. 062101. DOI: 10.1103/PhysRevA.70.062101. arXiv: quant-ph/0401155.

[48] Ruslan Leont'evich Stratonovich. "On Distributions in Representation Space". In: *Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki* 31.6 (1956). English translation: *Soviet Physics JETP*, 4(6):891–898, 1957., pp. 1012–1020. URL: http://jetp.ras.ru/cgi-bin/dn/e_004_06_0891.pdf.

[49] Hakop Pashayan, Joel J. Wallman, and Stephen D. Bartlett. "Estimating Outcome Probabilities of Quantum Circuits Using Quasiprobabilities". In: *Physical Review Letters* 115.7 (2015), p. 070501. DOI: 10.1103/physrevlett.115.070501. arXiv: 1503.07525.

[50] Robert W. Spekkens. "Negativity and Contextuality are Equivalent Notions of Nonclassicality". In: *Physical Review Letters* 101.2 (2008), p. 020401. DOI: 10.1103/PhysRevLett.101.020401. arXiv: 0710.5549.

[51] Nicolas Delfosse, Cihan Okay, Juan Bermejo-Vega, Dan E. Browne, and Robert Raussendorf. "Equivalence between contextuality and negativity of the Wigner function for qudits". In: *New Journal of Physics* 19.12 (2017), p. 123024. DOI: 10.1088/1367-2630/aa8fe3. arXiv: 1610.07093.

[52] David Schmid, Haoxing Du, John H. Selby, and Matthew F. Pusey. "Uniqueness of Noncontextual Models for Stabilizer Subtheories". In: *Physical Review Letters* 129.12 (2022), p. 120403. DOI: 10.1103/PhysRevLett.129.120403. arXiv: 2101.06263.

[53] Samson Abramsky, Shane Mansfield, and Rui Soares Barbosa. "The Cohomology of Non-Locality and Contextuality". In: *Electronic Proceedings in Theoretical Computer Science* 95 (2012), pp. 1–14. DOI: 10.4204/eptcs.95.1. arXiv: 1111.3620.

[54] Samson Abramsky, Rui Soares Barbosa, and Shane Mansfield. "Contextual Fraction as a Measure of Contextuality". In: *Physical Review Letters* 119.5 (2017), p. 050504. DOI: 10.1103/PhysRevLett.119.050504. arXiv: 1705.07918.

[55] Farid Shahandeh. "Quantum computational advantage implies contextuality". 2021. arXiv: 2112.00024.

[56] Pierre-Emmanuel Emeriau, Mark Howard, and Shane Mansfield. "Quantum Advantage in Information Retrieval". In: *PRX Quantum* 3.2 (2022), p. 020307. DOI: 10.1103/PRXQuantum.3.020307. arXiv: 2007.15643.

[57] Nicolas Delfosse, Philippe Allard Guerin, Jacob Bian, and Robert Raussendorf. "Wigner Function Negativity and Contextuality in Quantum Computation on Rebits". In: *Physical Review X* 5.2 (2015), p. 021003. DOI: 10.1103/PhysRevX.5.021003. arXiv: 1409.5170.

[58] Huangjun Zhu. "Permutation Symmetry Determines the Discrete Wigner Function". In: *Physical Review Letters* 116.4 (2016), p. 040501. DOI: 10.1103/PhysRevLett.116.040501. arXiv: 1504.03773.

[59] Juan Bermejo-Vega, Nicolas Delfosse, Dan E. Browne, Cihan Okay, and Robert Raussendorf. "Contextuality as a Resource for Models of Quantum Computation with Qubits". In: *Physical Review Letters* 119.12 (2017). DOI: 10.1103/physrevlett.119.120505. arXiv: 1610.08529.

[60] Lucas Kocia and Peter Love. "Discrete Wigner formalism for qubits and noncontextuality of Clifford gates on qubit stabilizer states". In: *Physical Review A* 96.6 (2017), p. 062134. DOI: 10.1103/PhysRevA.96.062134. arXiv: 1705.08869.

[61] Robert Raussendorf, Dan E. Browne, Nicolas Delfosse, Cihan Okay, and Juan Bermejo-Vega. "Contextuality and Wigner-function negativity in qubit quantum computation". In: *Physical Review A* 95.5 (2017), p. 052334. DOI: 10.1103/physreva.95.052334. arXiv: 1511.08506.

[62] N. David Mermin. "Hidden variables and the two theorems of John Bell". In: *Reviews of Modern Physics* 65.3 (1993), pp. 803–815. DOI: 10.1103/RevModPhys.65.803. arXiv: 1802.10119.

[63] Zak Webb. "The Clifford group forms a unitary 3-design". In: *Quantum Information and Computation* 16.15&16 (2016), pp. 1379–1400. DOI: 10.26421/QIC16.15-16-8. arXiv: 1510.02769.

[64] Mikio Nakahara. *Geometry, topology and physics*. Taylor & Francis, 1990. DOI: 10.1201/9781315275826.

[65] Alexei Yu. Kitaev. "Fault-tolerant quantum computation by anyons". In: *Annals of Physics* 303.1 (2003), pp. 2–30. DOI: 10.1016/S0003-4916(02)00018-0. arXiv: quant-ph/9707021.

[66] Eric Dennis, Alexei Kitaev, Andrew Landahl, and John Preskill. "Topological quantum memory". In: *Journal of Mathethematical Physics* 43 (2002), pp. 4452–4505. DOI: 10.1063/1.1499754. arXiv: quant-ph/0110143.

[67] Robert Raussendorf, Jim Harrington, and Kovid Goyal. "A fault-tolerant one-way quantum computer". In: *Annals of Physics* 321.9 (2006), pp. 2242–2270. DOI: 10.1016/j.aop.2006.01.012. arXiv: quant-ph/0510135.

[68] Samson Abramsky and Adam Brandenburger. "The sheaf-theoretic structure of non-locality and contextuality". In: *New Journal of Physics* 13.11 (2011), p. 113036. DOI: 10.1088/1367-2630/13/11/113036. arXiv: 1102.0264.

[69] Cihan Okay, Sam Roberts, Stephen D. Bartlett, and Robert Raussendorf. "Topological proofs of contextuality in quantum mechanics". In: *Quantum Information and Computation* 17.13&14 (2017), pp. 1135–1166. DOI: 10.26421/QIC17.13-14-5. arXiv: 1701.01888.

[70] Robert Raussendorf. "Cohomological framework for contextual quantum computations". In: *Quantum Information and Computation* 19.13&14 (2019), pp. 1141–1170. DOI: 10.26421/QIC19.13-14-4. arXiv: 1602.04155.

[71] Angela Karanjai, Joel J. Wallman, and Stephen D. Bartlett. "Contextuality bounds the efficiency of classical simulation of quantum processes". 2018. arXiv: 1802.07744.

[72] Arne Heimendahl. "The Stabilizer Polytope and Contextuality for Qubit Systems". Master's Thesis. University of Cologne, 2019. URL: http://www.mi.uni-koeln.de/opt/wp-content/uploads/2020/07/MT_Arne_Heimendahl.pdf.

[73] Adrian Chapman and Steven T. Flammia. "Characterization of solvable spin models via graph invariants". In: *Quantum* 4 (2020), p. 278. DOI: 10.22331/q-2020-06-04-278. arXiv: 2003.05465.

[74] Richard Jozsa and Akimasa Miyake. "Matchgates and classical simulation of quantum circuits". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 464.2100 (2008), pp. 3089–3106. DOI: 10.1098/rspa.2008.0189. arXiv: 0804.4050.

[75] Avinash Mocherla, Lingling Lao, and Dan E. Browne. "Extending Matchgate Simulation Methods to Universal Quantum Circuits". 2023. arXiv: 2302.02654.

[76] Joshua Cudby and Sergii Strelchuk. "Gaussian decomposition of magic states for matchgate computations". 2023. arXiv: 2307.12654.

[77] Beatriz Dias and Robert Koenig. "Classical simulation of non-Gaussian fermionic circuits". In: *Quantum* 8 (2024), p. 1350. DOI: 10.22331/q-2024-05-21-1350. arXiv: 2307.12912.

[78] Oliver Reardon-Smith, Michał Oszmaniec, and Kamil Korzekwa. "Improved simulation of quantum circuits dominated by free fermionic operations". 2023. arXiv: arXiv:2307.12702.

[79] Daniel R. Simon. "On the Power of Quantum Computation". In: *SIAM Journal on Computing* 26.5 (1997), pp. 1474–1483. DOI: 10.1137/S0097539796298637.

[80] Deutsch David and Jozsa Richard. "Rapid solution of problems by quantum computation". In: *Proc. R. Soc. Lond. A* 439 (1992), pp. 553–558. DOI: 10.1098/rspa.1992.0167.

[81] Sergey Bravyi, David Gosset, and Robert König. "Quantum advantage with shallow circuits". In: *Science* 362.6412 (2018), pp. 308–311. DOI: 10.1126/science.aar3106. arXiv: 1704.00690.

[82] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, et al. "Quantum supremacy using a programmable superconducting processor". In: *Nature* 574.7779 (2019), pp. 505–510. DOI: 10.1038/s41586-019-1666-5. arXiv: 1910.11333.

[83] Adam Bouland, Bill Fefferman, Chinmay Nirkhe, and Umesh Vazirani. "On the complexity and verification of quantum random circuit sampling". In: *Nature Physics* 15.2 (2019), pp. 159–163. DOI: 10.1038/s41567-018-0318-2. arXiv: 1803.04402.

[84] Scott Aaronson and Alex Arkhipov. "The computational complexity of linear optics". In: *Proceedings of the Forty-Third Annual ACM Symposium on Theory of Computing*. 2011, pp. 333–342. DOI: 10.1145/1993636.1993682. arXiv: 1011.3245.

[85] Tzu-Chieh Wei, Ian Affleck, and Robert Raussendorf. "Affleck-Kennedy-Lieb-Tasaki State on a Honeycomb Lattice is a Universal Quantum Computational Resource". In: *Physical Review Letters* 106.7 (2011), p. 070501. DOI: 10.1103/PhysRevLett.106.070501. arXiv: 1102.5064.

[86] Tzu-Chieh Wei, Ian Affleck, and Robert Raussendorf. "Two-dimensional Affleck-Kennedy-Lieb-Tasaki state on the honeycomb lattice is a universal resource for quantum computation". In: *Physical Review A* 86.3 (2012), p. 032328. DOI: 10.1103/PhysRevA.86.032328. arXiv: 1009.2840.

[87] Robert Raussendorf, Dong-Sheng Wang, Abhishodh Prakash, Tzu-Chieh Wei, and David T. Stephen. "Symmetry-protected topological phases with uniform computational power in one dimension". In: *Physical Review A* 96.1 (2017), p. 012302. DOI: 10.1103/PhysRevA.96.012302. arXiv: 1609.07549.

[88] David T. Stephen, Dong-Sheng Wang, Abhishodh Prakash, Tzu-Chieh Wei, and Robert Raussendorf. "Computational Power of Symmetry-Protected Topological Phases". In: *Physical Review Letters* 119.1 (2017), p. 010504. DOI: 10.1103/PhysRevLett.119.010504. arXiv: 1611.08053.

[89] Robert Raussendorf, Cihan Okay, Dong-Sheng Wang, David T. Stephen, and Hendrik Poulsen Nautrup. "Computationally Universal Phase of Quantum Matter". In: *Physical Review Letters* 122.9 (2019), p. 090501. DOI: 10.1103/PhysRevLett.122.090501. arXiv: 1803.00095.

[90] David T. Stephen, Hendrik Poulsen Nautrup, Juani Bermejo-Vega, Jens Eisert, and Robert Raussendorf. "Subsystem symmetries, quantum cellular automata, and computational phases of quantum matter". In: *Quantum* 3 (2019), p. 142. DOI: 10.22331/q-2019-05-20-142. arXiv: 1806.08780.

[91] Paul Herringer and Robert Raussendorf. "Classification of measurement-based quantum wire in stabilizer PEPS". In: *Quantum* 7 (2023), p. 1041. DOI: 10.22331/q-2023-06-12-1041. arXiv: 2207.00616.

[92] Robert Raussendorf, Wang Yang, and Arnab Adhikary. "Measurement-based quantum computation in finite one-dimensional systems: string order implies computational power". In: *Quantum* 7 (2023), p. 1215. DOI: 10.22331/q-2023-12-28-1215. arXiv: 2210.05089.

[93] Robert Raussendorf, Daniel E. Browne, and Hans J. Briegel. "Measurement-based quantum computation on cluster states". In: *Physical Review A* 68.2 (2003), p. 022312. DOI: 10.1103/PhysRevA.68.022312. arXiv: quant-ph/0301052.

[94] Erwin Schrödinger. "Die gegenwärtige Situation in der Quantenmechanik". In: *Naturwissenschaften* 23 (1935), pp. 807–812. DOI: https://doi.org/10.1007/BF01491891.

[95] Eugene P. Wigner. "Remarks on the Mind-Body Question". In: *The Scientist Speculates: An Anthology of Partly-Baked Ideas*. Ed. by Irving John Good. Heinemann, London, 1962. DOI: 10.1007/978-3-642-78374-6_20.

[96] Albert Einstein, Boris Podolsky, and Nathan Rosen. "Can Quantum-Mechanical Description of Physical Reality Be Considered Complete?" In: *Physical Review* 47.10 (1935), pp. 777–780. DOI: 10.1103/PhysRev.47.777.

[97] Louis de Broglie. "La nouvelle dynamique des quanta". In: *Électrons et photons. Rapports et discussions du cinquième Conseil de physique tenu à Bruxelles du 24 au 29 octobre 1927 sous les auspices de l'institut international de physique Solvay.* Ed. by Hendrik Antoon Lorentz. Gauthier-Villars et cie, 1928, pp. 105–132.

[98] David Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables. I". In: *Physical Review* 85.2 (1952), pp. 166–179. DOI: 10.1103/PhysRev.85.166.

[99] David Bohm. "A Suggested Interpretation of the Quantum Theory in Terms of 'Hidden' Variables. II". In: *Physical Review* 85.2 (1952), pp. 180–193. DOI: 10.1103/PhysRev.85.180.

[100] John Archibald Wheeler. "Information, Physics, Quantum: The Search for Links". In: *Proceedings of the 3rd International Symposium on Foundations of Quantum Mechanics in the Light of New Technology*. Ed. by Hiroshi Ezawa, Shun Ichi Kobayashi, and Yoshimasa Murayama. (Also published in Ref. [209]). Physical Society of Japan, 1989, pp. 354–368.

[101] Daniel Gottesman. "Fault-Tolerant Quantum Computation with Higher-Dimensional Systems". In: *Chaos, Solitons & Fractals* 10.10 (1999), pp. 1749–1758. DOI: 10.1016/s0960-0779(98)00218-5. arXiv: quant-ph/9802007.

[102] Asher Peres. *Quantum Theory: Concepts and Methods*. Kluwer Academic Publishers, 1993.

[103] Alexei Yu. Kitaev, Alexander H. Shen, and Mikhail N. Vyalyi. *Classical and Quantum Computation*. Vol. 47. Graduate studies in mathematics. Translated from Russian by Lester J. Senechal. American Mathematical Society, 2002.

[104] Sean Clark. "Valence bond solid formalism ford-level one-way quantum computation". In: *Journal of Physics A: Mathematical and General* 39.11 (2006), pp. 2701–2721. DOI: 10.1088/0305-4470/39/11/010. arXiv: quant-ph/0512155.

[105] Ben W. Reichardt. "Quantum universality by state distillation". In: *Quantum Information and Computation* 9.11&12 (2009), pp. 1030–1052. DOI: 10.26421/QIC9.11-12-7. arXiv: quant-ph/0608085.

[106] A. Chi-Chih Yao. "Quantum circuit complexity". In: *Proceedings of 1993 IEEE 34th Annual Foundations of Computer Science*. 1993, pp. 352–361. DOI: 10.1109/SFCS.1993.366852.

[107] Hussain Anwar, Earl T. Campbell, and Dan E. Browne. "Qutrit magic state distillation". In: *New Journal of Physics* 14.6 (2012), p. 063006. DOI: 10.1088/1367-2630/14/6/063006. arXiv: 1202.2326.

[108] Earl T. Campbell, Hussain Anwar, and Dan E. Browne. "Magic-State Distillation in All Prime Dimensions Using Quantum Reed-Muller Codes". In: *Physical Review X* 2.4 (2012), p. 041021. DOI: 10.1103/PhysRevX.2.041021. arXiv: 1205.3104.

[109] Mark Howard and Jiri Vala. "Qudit versions of the qubit $\pi/8$ gate". In: *Physical Review A* 86.2 (2012), p. 022316. DOI: 10.1103/PhysRevA.86.022316. arXiv: 1206.1598.

[110] Nadish de Silva. "Efficient quantum gate teleportation in higher dimensions". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 477.2251 (2021), p. 20200865. DOI: 10.1098/rspa.2020.0865. arXiv: 2011.00127.

[111] Sergey Bravyi, Graeme Smith, and John A Smolin. "Trading classical and quantum computational resources". In: *Physical Review Letters* 6.2 (2016), p. 021043. DOI: 10.1103/PhysRevX.6.021043. arXiv: 1506.01396.

[112] Filipa C.R. Peres and Ernesto F. Galvão. "Quantum circuit compilation and hybrid computation using Pauli-based computation". In: *Quantum* 7 (2023), p. 1126. DOI: 10.22331/q-2023-10-03-1126. arXiv: 2203.01789.

[113] Filipa C. R. Peres. "Pauli-based model of quantum computation with higher-dimensional systems". In: *Physical Review A* 108.3 (2023), p. 032606. DOI: 10.1103/PhysRevA.108.032606. arXiv: 2302.13702.

[114] Daniel Gottesman. "Stabilizer Codes and Quantum Error Correction". PhD Thesis. California Institute of Technology, 1997. arXiv: quant-ph/9705052.

[115] Vlad Gheorghiu. "Standard form of qudit stabilizer groups". In: *Physics Letters A* 378.5 (2014), pp. 505–509. DOI: 10.1016/j.physleta.2013.12.009. arXiv: 1101.1519.

[116] E.C.G. Sudarshan. "Equivalence of Semiclassical and Quantum Mechanical Descriptions of Statistical Light Beams". In: *Physical Review Letters* 10.7 (1963), pp. 277–279. DOI: 10.1103/PhysRevLett.10.277.

[117] Roy J. Glauber. "Coherent and Incoherent States of the Radiation Field". In: *Physical Review* 131.6 (1963), pp. 2766–2788. DOI: 10.1103/PhysRev.131.2766.

[118] Kôdi Husimi. "Some Formal Properties of the Density Matrix". In: *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series* 22.4 (1940), pp. 264–314. DOI: 10.11429/ppmsj1919.22.4_264.

[119] C. Brif and A. Mann. "A general theory of phase-space quasiprobability distributions". In: *Journal of Physics A: Mathematical and General* 31.1 (1998), pp. L9–L17. DOI: 10.1088/0305-4470/31/1/002. arXiv: quant-ph/9707010.

[120] Richard P. Feynman. "Negative probability". In: *Quantum Implications: Essays in Honour of David Bohm*. Ed. by F. David Peat and Basil Hiley. Routledge & Keegan Paul Ltd., 1987, pp. 235–248. URL: https://cds.cern.ch/record/154856/files/pre-27827.pdf.

[121] Christopher Ferrie. "Quasi-probability representations of quantum theory with applications to quantum information science". In: *Reports on Progress in Physics* 74.11 (2011), p. 116001. DOI: 10.1088/0034-4885/74/11/116001. arXiv: 1010.2701.

[122] Hakop Pashayan, Stephen D. Bartlett, and David Gross. "From estimation of quantum probabilities to simulation of quantum circuits". In: *Quantum* 4 (2020), p. 223. DOI: 10.22331/q-2020-01-13-223. arXiv: 1712.02806.

[123] John S. Bell. "On the Einstein Podolsky Rosen paradox". In: *Physics Physique Fizika* 1.3 (1964), pp. 195–200. DOI: 10.1103/PhysicsPhysiqueFizika.1.195.

[124] John von Neumann. *Mathematische Grundlagen der Quantenmechanik*. English translation: *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955. Springer-Verlag, 1932. DOI: 10.1007/978-3-642-61409-5.

[125] Grete Hermann. "Die naturphilosophischen Grundlagen der Quantenmechanik". In: *Naturwissenschaften* 23 (1935), pp. 718–721. DOI: 10.1007/BF01491142.

[126] John S. Bell. "On the Problem of Hidden Variables in Quantum Mechanics". In: *Reviews of Modern Physics* 38.3 (1966), pp. 447–452. DOI: 10.1103/RevModPhys.38.447.

[127] Simon Kochen and E.P. Specker. "The Problem of Hidden Variables in Quantum Mechanics". In: *Journal of Mathematics and Mechanics* 17.1 (1967), pp. 59–87. DOI: 10.1512/iumj.1968.17.17004.

[128] Robert W. Spekkens. "Contextuality for preparations, transformations, and unsharp measurements". In: *Physical Review A* 71.5 (2005), p. 052108. DOI: 10.1103/PhysRevA.71.052108. arXiv: quant-ph/0406166.

[129] Adán Cabello, Simone Severini, and Andreas Winter. "Graph-Theoretic Approach to Quantum Correlations". In: *Physical Review Letters* 112.4 (2014), p. 040401. DOI: 10.1103/PhysRevLett.112.040401. arXiv: 1401.7081.

[130] Shane Mansfield and Elham Kashefi. "Quantum Advantage from Sequential-Transformation Contextuality". In: *Physical Review Letters* 121.23 (2018), p. 230401. DOI: 10.1103/PhysRevLett.121.230401. arXiv: 1801.08150.

[131] C. Brif and A. Mann. "Phase-space formulation of quantum mechanics and quantum-state reconstruction for physical systems with Lie-group symmetries". In: *Physical Review A* 59.2 (1999), pp. 971–987. DOI: 10.1103/PhysRevA.59.971. arXiv: quant-ph/9809052.

[132] Christopher Ferrie and Joseph Emerson. "Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations". In: *Journal of Physics A: Mathematical and Theoretical* 41.35 (2008), p. 352001. DOI: 10.1088/1751-8113/41/35/352001. arXiv: 0711.2658.

[133] Christopher Ferrie and Joseph Emerson. "Framed Hilbert space: hanging the quasi-probability pictures of quantum theory". In: *New Journal of Physics* 11.6 (2009), p. 063040. DOI: 10.1088/1367-2630/11/6/063040. arXiv: 0903.4843.

[134] Jack Davis, Meenu Kumari, Robert B. Mann, and Shohini Ghose. "Wigner negativity in spin-j systems". In: *Physical Review Research* 3 (2021), p. 033134. DOI: 10.1103/PhysRevResearch.3.033134. arXiv: 2008.10167.

[135] David Schmid, John H. Selby, Matthew F. Pusey, and Robert W. Spekkens. "A structure theorem for generalized-noncontextual ontological models". In: *Quantum* 8 (2024), p. 1283. DOI: 10.22331/q-2024-03-14-1283. arXiv: 2005.07161.

[136] Joel J. Wallman and Stephen D. Bartlett. "Non-negative subtheories and quasiprobability representations of qubits". In: *Physiew Review A* 85.6 (2012), p. 062121. DOI: 10.1103/PhysRevA.85.062121. arXiv: 1203.2652.

[137] John B. DeBrota and Blake C. Stacey. "Discrete Wigner functions from informationally complete quantum measurements". In: *Physical Review A* 102.3 (2020), p. 032221. DOI: 10.1103/PhysRevA.102.032221. arXiv: 1912.07554.

[138] Paul Ehrenfest. "Bemerkung über die angenäherte Gültigkeit der klassichen Mechanik innerhalb der Quanatenmechanik". In: *Zeitschrift für Physik* 45.7 (1927), pp. 455–457. DOI: 10.1007/BF01329203.

[139] Deutsch David. "Quantum theory, the Church–Turing principle and the universal quantum computer". In: *Proceedings of the Royal Society A* 400 (1985), pp. 97–117. DOI: 10.1098/rspa.1985.0070.

[140] Ethan Bernstein and Umesh Vazirani. "Quantum Complexity Theory". In: *SIAM Journal on Computing* 26.5 (1997), pp. 1411–1473. DOI: 10.1137/S0097539796300921.

[141] Lucas Kocia and Peter Love. "The non-disjoint ontic states of the Grassmann ontological model, transformation contextuality, and the single qubit stabilizer subtheory". In: *Journal of Physics A: Mathematical and Theoretical* 52.9 (2019), p. 095303. DOI: 10.1088/1751-8121/aafca7. arXiv: 1805.09514.

[142] Barbara M. Terhal and David P. DiVincenzo. "Adaptive Quantum Computation, Constant Depth Quantum Circuits and Arthur-Merlin Games". In: *Quantum Information and Computation* 4.2 (2004), pp. 134–145. DOI: 10.26421/QIC4.2-5. arXiv: quant-ph/0205133.

[143] Maarten Van den Nest. "Classical simulation of quantum computation, the Gottesman-Knill theorem, and slightly beyond". In: *Quantum Information and Computation* 10.3&4 (2010), pp. 258–271. DOI: 10.26421/QIC10.3-4-6. arXiv: arXiv:0811.0898.

[144] Maarten Van Den Nest. "Simulating quantum computers with probabilistic methods". In: *Quantum Information and Computation* 11.9&10 (), pp. 784–812. DOI: 10.26421/QIC11.9-10-5. arXiv: 0911.1624.

[145] Adán Cabello, Mile Gu, Otfried Gühne, and Zhen-Peng Xu. "Optimal Classical Simulation of State-Independent Quantum Contextuality". In: *Physical Review Letters* 120.13 (2018), p. 130401. DOI: 10.1103/PhysRevLett.120.130401. arXiv: 1709.07372.

[146] Mark Howard, Eoin Brennan, and Jiri Vala. "Quantum contextuality with stabilizer states". In: *Entropy* 15.6 (2013), pp. 2340–2362. DOI: 10.3390/e15062340. arXiv: 1501.04342.

[147] Patrick Rall, Daniel Liang, Jeremy Cook, and William Kretschmer. "Simulation of qubit quantum circuits via Pauli propagation". In: *Physical Review A* 99.6 (2019), p. 062337. DOI: 10.1103/PhysRevA.99.062337. arXiv: 1901.09070.

[148] Martin Charles Golumbic. "Trivially perfect graphs". In: *Discrete Mathematics* 24.1 (1978), pp. 105–107. DOI: 10.1016/0012-365X(78)90178-4.

[149] William M. Kirby and Peter J. Love. "Contextuality Test of the Nonclassicality of Variational Quantum Eigensolvers". In: *Physical Review Letters* 123.20 (2019), p. 200501. DOI: 10.1103/physrevlett.123.200501. arXiv: 1904.02260.

[150] Michael Aschbacher. *Finite Group Theory*. Cambridge University Press, 1986. DOI: 10.1017/CBO9781139175319.

[151] Piers Lillystone and Joseph Emerson. "A Contextual ψ-Epistemic Model of the n-Qubit Stabilizer Formalism". 2019. arXiv: 1904.04268.

[152] Fernando G.S.L. Brandão and Gilad Gour. "Reversible Framework for Quantum Resource Theories". In: *Physical Review Letters* 115.7 (2015), p. 070503. DOI: 10.1103/PhysRevLett.115.070503. arXiv: 1502.03149.

[153] Steven Diamond and Stephen Boyd. "CVXPY: A Python-embedded modeling language for convex optimization". In: *Journal of Machine Learning Research* 17.83 (2016), pp. 1–5.

[154] Gurobi Optimization, LLC. *Gurobi Optimizer Reference Manual*. 2023. URL: https://www.gurobi.com.

[155] Markus Heinrich and David Gross. "Robustness of Magic and Symmetries of the Stabiliser Polytope". In: *Quantum* 3 (2019), p. 132. DOI: 10.22331/q-2019-04-08-132. arXiv: 1807.10296.

[156] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer New York, NY, 2006. DOI: https://doi.org/10.1007/978-0-387-40065-5.

[157] Richard Feynman, Robert Leighton, and Matthew Sands. "The Feynman Lectures on Physics". In: Addison-Wesley, 1963. Chap. Probability Amplitudes. URL: https://www.feynmanlectures.caltech.edu.

[158] John B. DeBrota, Christopher A. Fuchs, and Blake C. Stacey. "Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems". In: *Physical Review Research* 2.1 (2020), p. 013074. DOI: 10.1103/PhysRevResearch.2.013074. arXiv: 1805.08721.

[159] Matthew F. Pusey, Jonathan Barrett, and Terry Rudolph. "On the reality of the quantum state". In: *Nature Physics* 8.6 (2012), pp. 475–478. DOI: 10.1038/nphys2309. arXiv: 1111.3328.

[160] Vašek Chvátal. *Linear Programming*. W.H. Freeman and Company, 1983.

[161] Günter M. Ziegler. *Lectures on polytopes*. Springer-Verlag, 1995.

[162] E.G. Beltrametti and S. Bugajski. "The Bell phenomenon in classical frameworks". In: *Journal of Physics A: Mathematical and General* 29.2 (1996), p. 247. DOI: 10.1088/0305-4470/29/2/005.

[163] Michael J. Bremner, Richard Jozsa, and Dan J. Shepherd. "Classical simulation of commuting quantum computations implies collapse of the polynomial hierarchy". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 467.2126 (2011), pp. 459–472. DOI: 10.1098/rspa.2010.0301. arXiv: 1005.1407.

[164] Matthew S. Leifer and Owen J.E. Maroney. "Maximally Epistemic Interpretations of the Quantum State and Contextuality". In: *Physical Review Letters* 110.12 (2013), p. 120401. DOI: 10.1103/PhysRevLett.110.120401. arXiv: 1208.5132.

[165] Ravi Kunjwal and Robert W. Spekkens. "From the Kochen-Specker Theorem to Noncontextuality Inequalities without Assuming Determinism". In: *Physical Review Letters* 115.11 (2015), p. 110403. DOI: 10.1103/PhysRevLett.115.110403. arXiv: 1506.04150.

[166] Nicholas Harrigan and Robert W. Spekkens. "Einstein, Incompleteness, and the Epistemic View of Quantum States". In: *Foundations of Physics* 40.2 (2010), pp. 125–157. DOI: 10.1007/s10701-009-9347-0. arXiv: 0706.2661.

[167] Andrew M. Gleason. "Measures on the Closed Subspaces of a Hilbert Space". In: *Indiana University Mathematics Journal* 6.4 (1957), pp. 885–893. DOI: 10.1512/iumj.1957.6.56050.

[168] Jonas Haferkamp and Juani Bermejo-Vega. "Equivalence of contextuality and Wigner function negativity in continuous-variable quantum optics". 2021. arXiv: 2112.14788.

[169] Robert I. Booth, Ulysse Chabaud, and Pierre-Emmanuel Emeriau. "Contextuality and Wigner Negativity Are Equivalent for Continuous-Variable Quantum Measurements". In: *Physical Review Letters* 129.23 (2022), p. 230401. DOI: 10.1103/PhysRevLett.129.230401. arXiv: 2111.13218.

[170] Akimasa Miyake. "Quantum Computation on the Edge of a Symmetry-Protected Topological Order". In: *Physical Review Letters* 105.4 (2010), p. 040501. DOI: 10.1103/PhysRevLett.105.040501. arXiv: 1003.4662.

[171] Dominic V. Else, Ilai Schwarz, Stephen D. Bartlett, and Andrew C. Doherty. "Symmetry-Protected Phases for Measurement-Based Quantum Computation". In: *Physical Review Letters* 108.24 (2012), p. 240505. DOI: 10.1103/PhysRevLett.108.240505. arXiv: 1201.4877.

[172] Jacob Miller and Akimasa Miyake. "Resource Quality of a Symmetry-Protected Topologically Ordered Phase for Quantum Computation". In: *Physical Review Letters* 114.12 (2015), p. 120506. DOI: 10.1103/PhysRevLett.114.120506. arXiv: 1409.6242.

[173] Trithep Devakul and Dominic J. Williamson. "Universal quantum computation using fractal symmetry-protected cluster phases". In: *Physical Review A* 98.2 (2018), p. 022332. DOI: 10.1103/PhysRevA.98.022332. arXiv: 1806.04663.

[174] Austin K. Daniel, Rafael N. Alexander, and Akimasa Miyake. "Computational universality of symmetry-protected topologically ordered cluster phases on 2D Archimedean lattices". In: *Quantum* 4 (2020), p. 228. DOI: 10.22331/q-2020-02-10-228. arXiv: 1907.13279.

[175] Austin K. Daniel and Akimasa Miyake. "Quantum Computational Advantage with String Order Parameters of One-Dimensional Symmetry-Protected Topological Order". In: *Physical Review Letters* 126.9 (2021), p. 090505. DOI: 10.1103/PhysRevLett.126.090505. arXiv: 2007.16160.

[176] Alejandro Adem and R. James Milgram. *Cohomology of finite groups*. Springer Berlin, Heidelberg, 2013. DOI: 10.1007/978-3-662-06280-7.

[177] Steven Roman. *Advanced Linear Algebra*. Springer New York, NY, 2008. DOI: 10.1007/978-0-387-72831-5.

[178] Huangjun Zhu. "Multiqubit Clifford groups are unitary 3-designs". In: *Physical Review A* 96.6 (2017), p. 062336. DOI: 10.1103/PhysRevA.96.062336. arXiv: 1510.02619.

[179] D. Marcus Appleby. "Symmetric informationally complete-positive operator valued measures and the extended Clifford group". In: *Journal of Mathematical Physics* 46.5 (2005), p. 052107. DOI: 10.1063/1.1896384. arXiv: quant-ph/0412001.

[180] Beverley Bolt, T.G. Room, and G.E. Wall. "On the Clifford collineation, transform and similarity groups. II." In: *Journal of the Australian Mathematical Society* 2.1 (1961), pp. 80–96. DOI: 10.1017/S1446788700026380.

[181] Ravi Kunjwal. "Contextuality beyond the Kochen-Specker theorem". PhD Thesis. Homi Bhabha National Institute, 2016. arXiv: 1612.07250.

[182] Richard Cleve, Artur Ekert, Chiara Macchiavello, and Michele Mosca. "Quantum algorithms revisited". In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 454.1969 (1998), pp. 339–354. DOI: 10.1098/rspa.1998.0164. arXiv: quant-ph/9708016.

[183] Maarten Van den Nest, Wolgang Dür, Guiffre Vidal, and Hans J. Briegel. "Classical simulation versus universality in measurement-based quantum computation". In: *Physical Review A* 75.1 (2007), p. 012337. DOI: 10.1103/PhysRevA.75.012337. arXiv: quant-ph/0608060.

[184] Maarten Van den Nest. "Universal Quantum Computation with Little Entanglement". In: *Physical Review Letters* 110.6 (2013), p. 060504. DOI: 10.1103/PhysRevLett.110.060504. arXiv: 1204.3107.

[185] David Gross, Steven T. Flammia, and Jens Eisert. "Most Quantum States Are Too Entangled To Be Useful As Computational Resources". In: *Physical Review Letters* 102.19 (2009), p. 190501. DOI: 10.1103/PhysRevLett.102.190501. arXiv: 0810.4331.

[186] David Poulin, Angie Qarry, Rolando Somma, and Frank Verstraete. "Quantum Simulation of Time-Dependent Hamiltonians and the Convenient Illusion of Hilbert Space". In: *Physical Review Letters* 106.17 (2011), p. 170501. DOI: 10.1103/PhysRevLett.106.170501. arXiv: 1102.1360.

[187] Richard Jozsa and Noah Linden. "On the role of entanglement in quantum-computational speed-up". In: *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences* 459.2036 (2003), pp. 2011–2032. DOI: 10.1098/rspa.2002.1097. arXiv: quant-ph/0201143.

[188] Leslie G. Valiant. "Expressiveness of matchgates". In: *Theoretical Computer Science* 289.1 (2002), pp. 457–471. DOI: 10.1016/S0304-3975(01)00325-5.

[189] Sergey Bravyi and David Gosset. "Improved Classical Simulation of Quantum Circuits Dominated by Clifford Gates". In: *Physical Review Letters* 116.25 (2016), p. 250501. DOI: 10.1103/PhysRevLett.116.250501. arXiv: 1601.07601.

[190] Sergey Bravyi, Dan Browne, Padraic Calpin, Earl Campbell, et al. "Simulation of quantum circuits by low-rank stabilizer decompositions". In: *Quantum* 3 (2019), p. 181. DOI: 10.22331/q-2019-09-02-181. arXiv: 1808.00128.

[191] David Avis. "A Revised Implementation of the Reverse Search Vertex Enumeration Algorithm". In: *Polytopes — Combinatorics and Computation*. Ed. by Gil Kalai and Günter M. Ziegler. Vol. 29. DMV Seminar. Birkhäuser, Basel, 2000, pp. 177–198. DOI: 10.1007/978-3-0348-8438-9_9.

[192] Komei Fukuda and Alain Prodon. "Double description method revisited". In: *Combinatorics and Computer Science*. Ed. by M. Deza, R. Euler, and I. Manoussakis. Lecture Notes in Computer Science, vol 1120. Springer, Berlin, Heidelberg, 1996, pp. 91–111. DOI: 10.1007/3-540-61576-8_77.

[193] Thomas Rehn. "Polyhedral Description Conversion up to Symmetries". Diplona Thesis. Otto von Guericke University Magdeburg, 2010. URL: http://www.math.uni-rostock.de/~rehn/docs/diploma-thesis-ma-rehn.pdf.

[194] Branko Grünbaum. *Convex Polytopes*. Springer New York, NY, 2003. DOI: 10.1007/978-1-4613-0019-9.

[195] Chuangxun Cheng. "A character theory for projective representations of finite groups". In: *Linear Algebra and its Applications* 469 (2015), pp. 230–242. DOI: 10.1016/j.laa.2014.11.027.

[196] James R. Seddon, Bartosz Regula, Hakop Pashayan, Yingkai Ouyang, and Earl T. Campbell. "Quantifying Quantum Speedups: Improved Classical Simulation From Tighter Magic Monotones". In: *PRX Quantum* 2.1 (2021), p. 010345. DOI: 10.1103/PRXQuantum.2.010345. eprint: arXiv:2002.06181.

[197] Boseong Kim and Samson Abramsky. "State-independent all-versus-nothing arguments". 2023. arXiv: 2311.11218.

[198] Angelina Y. M. Chin. "On non-commuting sets in an extraspecial p-group". In: *Journal of Group Theory* 8.2 (2005), pp. 189–194. DOI: 10.1515/jgth.2005.8.2.189.

[199] Valentin Obst, Arne Heimendahl, Tanmay Singal, and David Gross. "Wigner's Theorem for stabilizer states and quantum designs". 2024. arXiv: 2405.17565.

[200] Cihan Okay, Ho Yiu Chung, and Selman Ipek. "Mermin polytopes quantum computation and foundations". In: *Quantum Information and Computation* 23.9&10 (2023), pp. 733–782. DOI: 10.26421/QIC23.9-10-2. arXiv: 2210.10186.

[201] Selman Ipek and Cihan Okay. "The degenerate vertices of the 2-qubit Λ-polytope and their update rules". 2023. arXiv: 2312.10734.

[202] Pascual Jordan and Eugene Wigner. "Über das Paulische Äquivalenzverbot". In: *Zeitschrift für Physik* 47.9 (1928), pp. 631–651. DOI: 10.1007/BF01331938.

[203] Lowell W. Beineke. "Characterizations of derived graphs". In: *Journal of Combinatorial Theory* 9.2 (1970), pp. 129–135. DOI: 10.1016/S0021-9800(70)80019-9.

[204] Nadish de Silva, Wilfred Salmon, and Ming Yin. "Fast algorithms for classical specifications of stabiliser states and Clifford gates". 2023. arXiv: 2311.10357.

[205] Adrian Chapman. Private communication. 2024.

[206] Dan Shepherd and Michael J. Bremner. "Temporally unstructured quantum computation". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 465.2105 (2009), pp. 1413–1439. DOI: 10.1098/rspa.2008.0443. arXiv: 0809.0847.

[207] Peter McMullen. "The maximum numbers of faces of a convex polytope". In: *Mathematika* 17.2 (1970), pp. 179–184. DOI: 10.1112/S0025579300002850.

[208] Angela Karanjai. "Statistical Modelling of Quantum Data". PhD Thesis. School of Physics, University of Sydney, 2019. URL: https://hdl.handle.net/2123/22134.

[209] John Archibald Wheeler. "Information, Physics, Quantum: The Search for Links". In: *Complexity, Entropy and the Physics of Information*. Ed. by Wokciech H. Zurek. 1st Edition. The proceedings of the 1988 Workshop on Complexity, Entropy, and the Physics of Information, held May-June, 1989 in Santa Fe, New Mexico. CRC Press, LLC, 1990, pp. 3–28. DOI: 10.1201/9780429502880.

[210] David S. Dummit and Richard M. Foote. *Abstract Algebra*. John Wiley and Sons, Inc., 2004.

[211] S. Akbari, A. Ghafari, K. Kazemian, and M. Nahvi. "Some criteria for a signed graph to have full rank". In: *Discrete Mathematics* 343.8 (2020), p. 111910. DOI: https://doi.org/10.1016/j.disc.2020.111910. eprint: arXiv:1708.07118.

[212] Tanmay Singal, Che Chiang, Eugene Hsu, Eunsang Kim, et al. "Counting stabiliser codes for arbitrary dimension". In: *Quantum* 7 (2023), p. 1048. DOI: 10.22331/q-2023-07-06-1048. arXiv: 2209.01449.

[213] Stanislaw Szarek and Guillaume Aubrun. *Alice and Bob Meet Banach: The Interface of Asymptotic Geometric Analysis and Quantum Information Theory*. American Mathematical Society, 2017.

[214] Richard A. Brualdi. *Introductory Combinatorics*. Pearson Education International, 2012.

Appendix A

Supplementary Materials

A.1 Supplementary materials for Chapter 2

A.1.1 Functional completeness of NAND gates

Theorem. *The set* {`NAND`} *is universal.*

Sketch of a proof. Any Boolean function f can be written as a polynomial over $\mathbb{F}_2$, the field with two elements, e.g.,

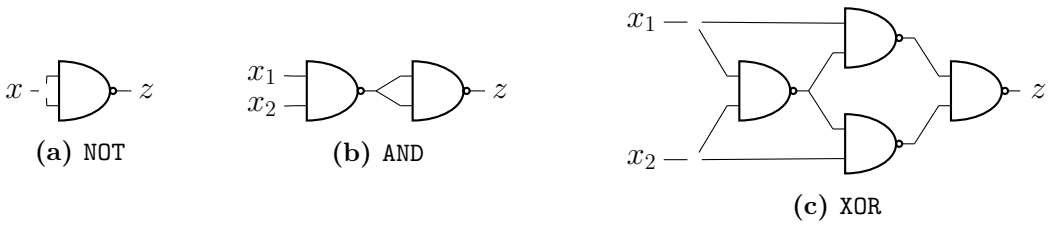
$$f(x_1, x_2, x_3, \dots) = 1 \oplus x_1 \oplus x_2 \oplus (x_1 \wedge x_2) \oplus \cdots , \tag{A.1}$$

where $\wedge$ is multiplication in $\mathbb{F}_2$, and $\oplus$ is addition mod 2. We can construct this polynomial explicitly. Start with the constant polynomial $p(\vec{x}) = 0$ if $f(0, 0, \dots, 0) = 0$, or $p(\vec{x}) = 1$ if $f(0, 0, \dots, 0) = 1$. This constant will agree with f on some inputs and will disagree with f on other inputs. The strategy is to add terms to the polynomial to correct it until it agrees with f on all inputs.

First, run through all inputs with Hamming weight 1, (i.e. all inputs with one bit set to 1 and all other bits set to 0). Add linear terms to $p(\vec{x})$ as needed to get the polynomial to agree with f on these inputs. Next run through all inputs with Hamming weight 2, and add quadratic terms as needed. Continue in this fashion until we reach the degree n term, at which point the polynomial will agree with f on all inputs. This strategy works because the degree d terms added do not change the output of the polynomial on any of the inputs with Hamming weight less than d.

With the function represented as a polynomial over $\mathbb{F}_2$ as in Eq. (A.1), this can immediately be mapped to a circuit consisting of gates `AND`, `NOT`, and `XOR`. Multiplication in $\mathbb{F}_2$ is represented by `AND` gates, addition by `XOR` gates, and a `NOT` gate is added at the end of the circuit if there is a nonzero constant term in the polynomial. Then these three gates can be constructed from compositions of `NAND` gates as shown in Figure A.1.

The case of multiple output bits follows immediately. □



(a) NOT **(b)** AND **(c)** XOR

Figure A.1: Constructions of (a) a NOT gate, (b) an AND gate, and (c) an XOR gate out of circuits of NAND gates.

A.2 Supplementary materials for Chapter 3

A.2.1 Number of CNC phase space points

As shown in Section 2.4, when the local Hilbert space dimension is odd, the phase space of the discrete Wigner function has cardinality $|V| = |\mathbb{Z}_d^{2n}| = d^{2n}$.

For multiple qubits, state-independent contextuality causes the phase space to splinter. The cardinality of the multiqubit phase space is much larger than the 2^{2n} that might be expected for a generalization of the odd-dimensional Wigner function. This is an important feature of the generalized phase space which allows it to simulate contextual scenarios like the Mermin square and that allows the generalized Wigner function to be Clifford covariant [58, 61, 71]. In this section, we explore this feature in more detail.

We consider the size of the generalized phase space that is used for representing states consisting of pairs $(\Omega, \gamma) \in \mathcal{V}$ subject to the constraints of Definition 7 with the additional constraint that Ω is maximal. Using the classification of multiqubit phase space points described in Section 3.3.3 we can count the number of multiqubit phase space points.

Theorem 30. *The cardinality of the n-qubit phase space is*

$$|\mathcal{V}| = \sum_{m=1}^{n} \left\{ \frac{2^{n+m^2+m+1}}{(2m+1)!} \begin{bmatrix} n \\ m \end{bmatrix}_4 \left(\prod_{k=1}^{n-m} 2^k + 1 \right) \left(\prod_{j=1}^{m} 4^j - 1 \right) \right\} \tag{A.2}$$

where $\begin{bmatrix} n \\ m \end{bmatrix}_4 := \prod_{k=1}^{m} \frac{4^{n-k+1}-1}{4^k-1}$ *is the Gaussian binomial coefficient.*

Proof of Theorem 30. Per Theorem 2, maximal cnc sets have the form

$$\Omega = \bigcup_{k=1}^{2m+1} \langle a_k, \tilde{I} \rangle$$

where $1 \leq m \leq n$, $\tilde{I}$ is a $n-m$-dimensional isotropic subspace of E, and the vectors $a_1, a_2, \ldots, a_{2m+1} \in \tilde{I}^{\perp}/\tilde{I}$ pairwise anticommute.

The number of phase space points (Ω, γ) with parameters n and m is the product of three factors: (i) the number of $n-m$-dimensional isotropic subspaces $\tilde{I} \subset E$, (ii) given each subspace $\tilde{I}$, the number of ways of choosing the remaining generators $a_1, a_2, \ldots, a_{2m+1}$ of the set, and (iii) the number of noncontextual value assignments on the set.

First we consider (i), the number of isotropic subspaces $\tilde{I}$ (a similar counting argument to the one used in this part of the proof is given in the proof of proposition 2 in Ref. [41]). The number of $n-m$-dimensional isotropic subspaces is equal to the number of ways of choosing a set of generators $g_1, g_2, \ldots, g_{n-m}$ for $\tilde{I}$ divided by the number of generating sets that give the same isotropic subspace. The first generator, g_1, can be any element of $E \setminus \{0\}$. There are $4^n - 1$ choices. The second generator, g_2, must commute with g_1. There are $\frac{4^n}{2} - 2$ choices. In general, the k^{th} generator must commute with the first $k-1$ generators and not lie in their span so there are $\frac{4^n}{2^{k-1}} - 2^{k-1}$ choices. Therefore, the number of ways of choosing the $n-m$ generators of the isotropic subspace $\tilde{I}$ is

$$N_1 = \prod_{k=1}^{n-m} \frac{4^n}{2^{k-1}} - 2^{k-1} = 2^{(n-m)(n-m-1)/2} \prod_{k=1}^{n-m} 4^{n-k+1} - 1.$$

Similarly, we can count the number of generating sets which give a fixed isotropic subspace $\tilde{I}$. The first generator, g_1, can be any of the $2^{n-m} - 1$ nonzero elements of $\tilde{I}$. The second, g_2, can be any element of $\tilde{I} \setminus \langle g_1 \rangle$. There are $2^{n-m} - 2$ choices. In general, the k^{th} generator, g_k, can be any element of $\tilde{I} \setminus \langle g_1, g_2, \ldots, g_{k-1} \rangle$. There are $2^{n-m} - 2^{k-1}$ choices. Therefore, the number of possible generating sets for each isotropic subspace is

$$N_2 = \prod_{k=1}^{n-m} 2^{n-m} - 2^{k-1} = 2^{(n-m)(n-m-1)/2} \prod_{k=1}^{n-m} (2^k - 1).$$

Therefore, the number of $n-m$-dimensional isotropic subspaces of E is

$$N_{(i)} = \frac{N_1}{N_2} = \prod_{k=1}^{n-m} \frac{4^{n-k+1} - 1}{2^k - 1} = \begin{bmatrix} n \\ m \end{bmatrix}_4 \cdot \prod_{k=1}^{n-m} 2^k + 1$$

where $\begin{bmatrix} n \\ m \end{bmatrix}_4$ is the Gaussian binomial coefficient.[1]

Now consider (ii), given a $n-m$-dimensional isotropic subspace $\tilde{I}$, the number of ways of choosing $a_1, a_2, \ldots, a_{2m+1}$ from $\tilde{I}^\perp / \tilde{I}$. Because of the the isomorphism $\tilde{I}^\perp / \tilde{I} \simeq \mathbb{Z}_2^{2m}$, this is equal to the number of maximal sets of pairwise anticommuting elements in the symplectic vector space $\mathbb{Z}_2^{2m}$ of size $2m+1$. Let $\mathcal{A} \subset \mathcal{P}(\mathbb{Z}_2^{2m})$ be the set of all such sets, and $A = \{a_1, a_2, \ldots, a_{2m+1}\} \in \mathcal{A}$ be one particular set.

To find $|\mathcal{A}|$, consider the action of the symplectic group, $\mathrm{Sp}(\mathbb{Z}_2^{2m})$, on $\mathcal{A}$ induced by the action of $\mathrm{Sp}(\mathbb{Z}_2^{2m})$ on the elements of the sets in $\mathcal{A}$. As shown in the proof of Theorem 2, the symplectic group acts transitively on $\mathcal{A}$. Therefore, by the orbit-stabilizer theorem and

[1] Incidentally, the second factor in the final expression for $N_{(i)}$ is equal to the number of maximal isotropic subspaces in $\mathbb{Z}_2^{2(n-m)}$. Therefore, the Gaussian binomial coefficient $\begin{bmatrix} n \\ m \end{bmatrix}_4$ can be interpreted as the number of inequivalent symplectic maps taking all maximal isotropic subspaces in $\mathbb{Z}_2^{2(n-m)}$ to $n-m$-dimensional isotropic subspaces in $\mathbb{Z}_2^{2n}$.

Lagrange's theorem [210],

$$|\mathcal{A}| = |\mathrm{Sp}(\mathbb{Z}_2^{2m}) \cdot A| = [\mathrm{Sp}(\mathbb{Z}_2^{2m}) \ : \ \mathrm{Sp}(\mathbb{Z}^{2m})_A] = \frac{|\mathrm{Sp}(\mathbb{Z}_2^{2m})|}{|\mathrm{Sp}(\mathbb{Z}_2^{2m})_A|}$$

where $\mathrm{Sp}(\mathbb{Z}_2^{2m})_A = \{M \in \mathrm{Sp}(\mathbb{Z}_2^{2m}) \mid M \cdot A = A\}$ is the stabilizer of A in $\mathrm{Sp}(\mathbb{Z}_2^{2m})$.

Consider a permutation σ of $\{1, 2, \ldots, 2m+1\}$ and a linear map that sends a_k to $a_{\sigma(k)}$ for $1 \leq k \leq 2m$. By Witt's lemma [150, §20], this map extends to a symplectic transformation $f : \mathbb{Z}_2^{2m} \to \mathbb{Z}_2^{2m}$ and since $\langle a_1, a_2, \ldots, a_{2m}\rangle = \mathbb{Z}_2^{2m}$, this extension is unique. Also, since the two systems of equations $[a_k, c] = 1,\ 1 \leq k \leq 2m$ and $[f(a_k), c] = 1,\ 1 \leq k \leq 2m$ each have a unique solution, $f(a_{2m+1}) = a_{\sigma(2m+1)}$. Therefore, any permutation of the elements of the set $A = \{a_1, a_2, \ldots, a_{2m+1}\}$ defines a unique symplectic transformation of $\mathbb{Z}_2^{2m}$. I.e. $|\mathrm{Sp}(\mathbb{Z}_2^{2m})_A| = (2m+1)!$. Also, $|\mathrm{Sp}(\mathbb{Z}_2^{2m})| = 2^{m^2}\prod_{j=1}^{m} 4^j - 1$. Thus, the number of maximal sets of pairwise anticommuting elements in $\mathbb{Z}_2^{2m}$ of size $2m+1$ is

$$N_{(ii)} = \frac{|\mathrm{Sp}(\mathbb{Z}_2^{2m})|}{|\mathrm{Sp}(\mathbb{Z}_2^{2m})_A|} = \frac{1}{(2m+1)!} 2^{m^2} \prod_{j=1}^{m} 4^j - 1.$$

Now consider (iii), the number of noncontextual value assignments. A noncontextual value assignment γ for a set Ω with parameters n and m is uniquely defined by its values on the generators of Ω through the relation Eq. (3.5), and its values on the generators can be chosen freely. For a set Ω with parameters n and m, there are $(n-m)+(2m+1) = n+m+1$ generators and so for a set with these parameters there are $N_{(iii)} = 2^{n+m+1}$ noncontextual value assignments γ.

The product $N_{(i)} \cdot N_{(ii)} \cdot N_{(iii)}$ gives a generic term in the sum Eq. (A.2). Then, summing m from 1 to n gives the result. □

Table A.1 shows the size of the generalized phase space for a few different cases. From this we can see that the size of the full generalized phase space for multiple qubits is very large, much larger than 2^{2n} as would be expected for a naïve generalization of the odd-dimensional phase space, $V = \mathbb{Z}_d^{2n}$. It turns out that for the purpose of representing states, not all of these phase space points are needed simultaneously. First, for any positively representable state ρ, there is a nonnegative Wigner function $W_\rho : \mathcal{V} \to \mathbb{R}_{\geq 0}$ representing ρ with $|\mathrm{Supp}(W_\rho)| \leq 2^{2n}$. That is, no more than 2^{2n} phase space points are required in the optimal representation of any positively representable state. This follows immediately from Caratheodory's theorem [161, §1.6]. A similar result holds for states which are not positively representable, as shown in Section 3.6.5.

A.2.2 Computing W-representations of many copies of magic states

Here we describe how to construct valid quasiprobabilities $W_{\mu^{\otimes n}}$ for n copies of a magic state μ, at bounded computational cost. As with robustness of magic [21], we merge expansions for small numbers of magic states into valid expansions for larger numbers of copies.

Denote by Ω_n^m cnc sets Ω with parameters $n, m \leq n$, and choose the phases ϕ in Eq. (2.3) such that

$$T_{a+b} = T_a \otimes T_b, \quad \forall a \in \Omega_{n_1}^{m_1},\ b \in \Omega_{n_2}^{m_2}. \tag{A.3}$$

n	$\lvert V\rvert$ $(d=3)$	$\lvert V\rvert$ $(d=5)$	$\lvert V\rvert$ $(d=7)$
1	9	25	49
2	81	625	2 401
3	729	15 625	117 649
4	6561	390 625	5 764 801
5	59 049	9 765 625	282 475 249

(a) d odd.

n	$\lvert\mathcal{V}\rvert$ $(m=0)$	$\lvert\mathcal{V}\rvert$ $(m=1)$	$\lvert\mathcal{V}\rvert$ (full)	2^{2n}
1	6	8	8	4
2	60	240	432	16
3	1 080	10 080	71 136	64
4	36 720	734 400	90 494 400	256
5	2 423 520	100 172 160	1 424 082 031 488	1024

(b) $d=2$.

Table A.1: The cardinality of the multiqudit phase space for different numbers n of qudits and different local Hilbert space dimensions d. (a) The cardinality of the phase space when d is odd is d^{2n}. (b) When $d=2$ the cardinality of phase space is larger than 2^{2n}. Here the second column is the number of stabilizer states (phase space points corresponding to cnc sets of the form Eq. (3.10) with $m=0$), the third column is the size of the phase space obtained using only maximal cnc sets with $m=1$, the fourth column is the size of the full multiqubit phase space.

Here we identified a and b as elements of $\mathbb{Z}_d^{2(n_1+n_2)}$ by writing $((a_X,0),(a_Z,0))$ and $((0,b_X),(0,b_Z))$, respectively. We then have the following result.

Lemma 38. *Be $\Omega_{n_1}^{m_1}$ and $\Omega_{n_2}^0$ two cnc sets with parameters n_1, $m_1 \leq n$, and n_2, $m_2=0$, respectively. Then, $A^{\gamma}_{\Omega_{n_1}^{m_1}\oplus\Omega_{n_2}^0} := A^{\gamma_1}_{\Omega_{n_1}^{m_1}} \otimes A^{\gamma_1}_{\Omega_{n_2}^0}$, with the function $\gamma : \Omega_{n_1}^{m_1}\oplus\Omega_{n_2}^0 \longrightarrow \mathbb{Z}_2$ defined by*

$$\gamma(a_1+a_2) := \gamma_1(a_1)+\gamma_2(a_2), \tag{A.4}$$

for all $a_1 \in \Omega_{n_1}^{m_1}$, $a_2 \in \Omega_{n_1}^0$ is a valid phase space point operator on n_1+n_2 qubits.

Proof of Lemma 38. We need to verify the properties of Def. 7, namely (a) that $\Omega_{n_1}^{m_1}\oplus\Omega_{n_2}^0$ is cnc, and (b) that the function γ defined in Lemma 38 satisfies Eq. (3.5), i.e., $d\gamma=\beta$.

Regarding (a), with Eq. (3.10), $\Omega_{n_1}^{m_1} = \bigcup_{k=1}^{m_1}\langle a_k,\tilde{I}\rangle$, and $\Omega_{n_2}^0 = I$, with $\tilde{I}$, I isotropic subspaces. Then,

$$\Omega_{n_1}^{m_1}\oplus\Omega_{n_2}^0 = \bigcup_{k=1}^{m_1}\langle a_k,\tilde{I}\oplus I\rangle,$$

and $\tilde{I}\oplus I$ is also an isotropic subspace. Hence the set $\Omega_{n_1}^{m_1}\oplus\Omega_{n_2}^0$ is cnc by Lemma 10.

Regarding (b), we need to show that for all $a,b\in\Omega_{n_1}^{m_1}\oplus\Omega_{n_2}^0$ with $[a,b]=0$ it holds that

$\gamma(a+b)+\gamma(a)+\gamma(b)=\beta(a,b)$. To this end, for any given commuting pair a, b, we split

$$a=a_1+a_2,\ b=b_1+b_2,$$

with $a_1,b_1\in\Omega^{m_1}_{n_1}$, $a_2,b_2\in\Omega^0_{n_2}$. The decompositions are unique. Since a and b commute, $[a_1,b_1]+[a_2,b_2]=0$. Furthermore, a_2 commutes with b_2, since $\Omega^0_{n_2}$ is an isotropic subspace. Thus,

$$[a_1,b_1]=[a_2,b_2]=0. \tag{A.5}$$

Further, Eq. (A.3) implies that

$$\beta(c,d)=0,\ \forall c\in\Omega^{m_1}_{n_1},\ d\in\Omega^0_{n_2}. \tag{A.6}$$

Now we rewrite

$$\begin{aligned}
\gamma(a+b) =&\gamma((a_1+b_1)+(a_2+b_2))\\
=&\gamma_1(a_1+b_1)+\gamma_2(a_2+b_2)\\
=&\gamma_1(a_1)+\gamma_1(b_1)+\gamma_2(a_2)+\gamma_2(b_2)+\beta(a_1,b_1)+\beta(a_2,b_2)\\
=&\gamma(a_1+a_2)+\gamma(b_1+b_2)+\beta(a_1,b_1)+\beta(a_2,b_2)+\beta(a_1,a_2)+\beta(b_1,b_2)\\
=&\gamma(a)+\gamma(b)+\beta(a_1+a_2,b_1+b_2)+d\beta(a_1,a_2,b_1+b_2)+d\beta(a_2,b_2,b_1)+d\beta(a_1,b_1,a_2+b_2)\\
=&\gamma(a)+\gamma(b)+\beta(a,b).
\end{aligned}$$

Therein, in the second line we have used the definition of γ in Lemma 38. In the third line Eq. (A.5), and in the fourth line the definition of γ again. In the fifth line we have used Eq. (3.6), Eq. (2.8), and Eq. (A.6) on $\beta(a_1+b_1,a_2+b_2)$. In the last line we used Eq. (2.7) ($d\beta=0$). □

Denote by $W^{(0)}_\sigma$ an expansion Eq. (3.3) of the state σ, but only containing phase space point operators with parameter $m=0$, i.e., an expansion into stabilizer states. W_ρ is a valid expansion of ρ, according to Eq. (3.3). Then, it follows from Lemma 38 that

$$W_{\rho\otimes\sigma}:=W_\rho\otimes W^{(0)}_\sigma \tag{A.7}$$

is a valid expansion of $\rho\otimes\sigma$.

Let k be the largest integer for which decompositions $W_{\mu^{\otimes k}}$ and $W^{(0)}_{\mu^{\otimes k}}$ are obtainable. Then, with Eq. (A.7), $W_{\mu^{\otimes n}}=W_{\mu^{\otimes k}}\otimes W^{(0)}_{\mu^{\otimes n-k}}$, and the second factor may be further decomposed as $W^{(0)}_{\mu^{\otimes n-k}}=W^{\otimes(n/k-1)}_{\mu^{\otimes k}}$, if k divides n. Thus, we arrive at an explicit decomposition for $\mu^{\otimes n}$, with 1-norm

$$\left\|W_{\mu^{\otimes n}}\right\|=\mathfrak{R}\left(\mu^{\otimes k}\right)\mathfrak{R}_S\left(\mu^{\otimes k}\right)^{n/k-1}.$$

Thus the reduction to blocks of k magic states familiar from the stabilizer case [21] can be applied in the present setting as well. By Lemma 16, the resulting 1-norm is lower by a constant factor than of the corresponding expansion into stabilizer states.

A.3 Supplementary materials for Chapter 4

A.3.1 Explanation of the origin of the Λ polytopes

The existence of the hidden variable model defined by Theorem 6 is unintuitive. It allows any quantum computation—and therefore by extension all finite-dimensional quantum mechanics—to be described using only classical probabilities. All states are represented by probability distributions over a finite set of elements and the formulation of the Born rule in Eq. (4.4) has the form of a law of total probability. The dynamics described by Eq. (4.3) look like the transition matrix dynamics of a finite Markov process. In this section we will attempt to provide an explanation for where the hidden variable comes from and why the construction works.

This is the most general kind of quantum circuit allowed within the QCM framework consist of arbitrarily long sequences of Clifford gates and Pauli measurements, possibly with gates and measurements conditioned on the outcomes of previous measurements. Let $\mathfrak{C}$ denote the set of all circuits of this form. For a circuit $C \in \mathfrak{C}$ consisting of Pauli measurements $a_1, a_2, \cdots \in E$ interlaced with Clifford gates $g_1, g_2, \ldots$, denote by Φ_C^s the superoperator

$$\Phi_C^s(\rho) = \cdots \Pi_{a_1}^{s_2} g_2 \Pi_{a_1}^{s_1} g_1 \rho g_1^\dagger \Pi_{a_1}^{s_1} g_2 \Pi_{a_2}^{s_2} \cdots$$

corresponding to the circuit C giving measurement outcomes $s = (s_1, s_2, \dots)$. The probability of obtaining the measurement outcomes s is given the circuit C is executed on state ρ is $\mathrm{Tr}(\Phi_C^s(\rho))$.

All of the classical simulation algorithms for quantum computation with magic states discussed in this thesis apply to a set of input "states" which forms a polytope in the space $\mathrm{Herm}(\mathcal{H})$. In the case of Algorithm 3, the polytope is Λ defined in Eq. (4.1)). In the cases of the simulation algorithms based on the discrete Wigner functions, the corresponding polytopes are the convex hulls of the phase space point operators, $\mathrm{ConvHull}\{A_u \mid u \in V\}$ for Algorithm 1 and $\mathrm{ConvHull}\left\{A_\Omega^\gamma \mid (\Omega, \gamma) \in \mathcal{V}\right\}$ for Algorithm 2.

The essential properties that these polytopes share that allow them to spawn classical simulation algorithms for quantum computation are:

1. States in these polytopes give valid probabilities associated with all Pauli measurements. That is, for any $X \in \Lambda$, $a \in E$, and $s \in \mathbb{Z}_d$, $0 \leq \mathrm{Tr}(\Pi_a^s X) \leq 1$.

2. The polytopes map back into themselves under Clifford unitaries and Pauli measurements. That is, for any $X \in \Lambda$, $g \in \mathcal{C}\ell$, $a \in E$, and $s \in \mathbb{Z}_d$, we have $gXg^\dagger \in \Lambda$, and if $\mathrm{Tr}(\Pi_a^s X) > 0$ then $X' := \Pi_a^s X \Pi_a^s / \mathrm{Tr}(\Pi_a^s X) \in \Lambda$.

The first property allows us to extract the expectation values for single Pauli measurements from the Wigner/probability function representation of the state. The second property, by the Krein-Milman theorem, gives the probabilistic update rules under the dynamics of QCM circuits like those in Lemmas 6, 12, and in Theorem 6 part (ii). This allows the process above to be iterated to give expectation values for all Pauli measurements in a QCM circuit.

It can easily be shown that the largest subset of $\mathrm{Herm}_1(\mathcal{H})$ satisfying these two properties is the set

$$\tilde{\Lambda} = \{X \in \mathrm{Herm}_1(\mathcal{H}) \mid \mathrm{Tr}(\Phi_C^s(X)) \geq 0 \; \forall C \in \mathfrak{C}, \forall s\}\,. \tag{A.8}$$

where we range over all circuits in $\mathfrak{C}$ and all physically possible sequences of measurement outcomes s. $\tilde{\Lambda}$ satisfies the two properties above, and for any set $\Omega \subset \mathrm{Herm}_1(\mathcal{H})$ satisfying these two properties, $\Omega \subset \tilde{\Lambda}$. Therefore, any polytope in $\mathrm{Herm}_1(\mathcal{H})$ that could be used to derive a classical simulation algorithm like those given above for QCM must be contained in $\tilde{\Lambda}$. Also, clearly, $\tilde{\Lambda}$ contains all physical states.

The set $\tilde{\Lambda}$ has an obvious connection to quantum computation with magic states. It is simply the set of "states" which give valid (nonnegative) probabilities for all possible outcomes of of QCM circuits. It is not however obvious that $\tilde{\Lambda}$ can be used to define a hidden variable model like that of Theorem 6, since it is defined as the intersection of an infinite number of halfspaces. Therefore, it is not obvious that it has a finite number of extreme points for any number of qubits $n \in \mathbb{N}$. The surprising aspect of the existence of the hidden variable model defined in Theorem 6 is captured by the following theorem.

Theorem 31. $\tilde{\Lambda} = \Lambda$.

The proof of this theorem is given in [9, §4.3].

A.3.2 Stratonovich-Weyl correspondence

In the field of quantum optics, the Stratonovich-Weyl (SW) correspondence is a set of criteria that well-behaved quasiprobability distributions over phase space have to satisfy. Denote by $F_A^{(s)} : X \longrightarrow \mathbb{C}$ the quasiprobability distribution corresponding to the (not necessarily Hermitian) operator A, with X the phase space and s a real parameter in the interval $[-1, 1]$. In the standard formalism for infinite-dimensional Hilbert spaces, $s = -1, 0, 1$ correspond to the Glauber-Sudarshan P, Wigner, and Husimi Q function, respectively. Then, the following set of criteria is imposed on the $F_A^{(s)}$ [48]; also see [119],

(0) Linearity: $A \longrightarrow F_A^{(s)}$ is a one-to-one linear map.

(1) Reality:
$$F_{A^\dagger}^{(s)}(u) = \left(F_A^{(s)}(u)\right)^*, \forall u \in X.$$

(2) Standardization:
$$\int_X d\mu(u) F_A^{(s)}(u) = \mathrm{Tr}\, A.$$

(3) Covariance:
$$F_{g\cdot A}^{(s)}(u) = F_A^{(s)}(g^{-1}u),\ g \in G,$$
with G the dynamical symmetry group.

(4) Traciality:
$$\int_X d\mu(u) F_A^{(s)}(u) F_B^{(-s)}(u) = \mathrm{Tr}\, AB.$$

To investigate the SW criteria in the present setting, we first extend the probability distributions p_ρ defined in Eq. (4.2) for proper density matrices to a quasiprobability function

W defined for all operators A, via

$$A = \sum_{\alpha} W_A(\alpha) A_\alpha. \tag{A.9}$$

We note that W does not come with a parameter s; there is only a single quasiprobability function W. This will affect the formulation of traciality.

Further, the mapping $A \longrightarrow W_A$ is linear, $A + B$ can be represented as $W_A + W_B$. However, the mapping is one-to-many, and the Stratonovich-Weyl criterion (0) is thus not satisfied. In fact, this is a general consequence of Kochen-Specker contextuality, as has been demonstrated in [164].

The remaining SW conditions apply.

(1) Reality. All phase space point operators A_α are Hermitian by definition. Therefore $A^\dagger$ can be represented by the quasiprobability distribution $\alpha \mapsto W_A(\alpha)^*$.

(2) Standardization. By their definition, the phase space point operators satisfy $\mathrm{Tr}\, A_\alpha = 1$, for all $\alpha \in \mathcal{V}_n$. Standardization,

$$\mathrm{Tr} A = \sum_{\alpha} W_A(\alpha), \tag{A.10}$$

follows by taking the trace of Eq. (A.9).

(3) Covariance. Let Cl_n denote the n-qubit Clifford group. We have the following result.

Lemma 39. *For any operator A it holds that*

$$W_{gAg^\dagger}(\alpha) = W_A(g^{-1}\alpha), \;\; \forall g \in Cl_n. \tag{A.11}$$

Proof of Lemma 39. First we show that Λ_n is mapped into itself under the action of the Clifford group. Namely, for all stabilizer sates $|\sigma\rangle \in \mathcal{S}$,

$$\begin{aligned} \mathrm{Tr}(gA_\alpha g^\dagger |\sigma\rangle\langle\sigma|) &= \mathrm{Tr}(A_\alpha\, g^\dagger |\sigma\rangle\langle\sigma| g) \\ &= \mathrm{Tr}(A_\alpha\, |\sigma'\rangle\langle\sigma'|) \\ &\geq 0. \end{aligned}$$

Furthermore, $\mathrm{Tr}(gA_\alpha g^\dagger) = \mathrm{Tr} A_\alpha = 1$. Hence, with the definition of Λ_n, it holds that $gA_\alpha g^\dagger \in \Lambda_n$, for all $\alpha \in \mathcal{V}_n$ and all $g \in \mathrm{Cl}_n$.

Now we show that for every $\alpha \in \mathcal{V}_n$ and every $g \in \mathrm{Cl}_n$ there is a unique $\beta \in \mathcal{V}_n$ such that

$$gA_\alpha g^\dagger = A_\beta. \tag{A.12}$$

Let $\mathcal{S}_\alpha$ be the subset of stabilizer states that specifies A_α, i.e. A_α is the unique solution in Λ_n to the set of constraints $\mathrm{Tr}(X|\sigma\rangle\langle\sigma|) = 0$ for all $|\sigma\rangle \in \mathcal{S}_\alpha$. In fact, we can choose the size of $\mathcal{S}_\alpha$ to be equal to $2^{2n} - 1$ [160, Theorem 18.1]. Let $g \cdot S_\alpha$ denote the set of stabilizers $g|\sigma\rangle\langle\sigma|g^\dagger$ where $|\sigma\rangle \in \mathcal{S}_\alpha$. Then the action of g gives a one-to-one correspondence between the set of solutions to the constraints specified by $\mathcal{S}_\alpha$ and $g^\dagger \cdot \mathcal{S}_\alpha$ since if X is a solution to the former then $gXg^\dagger$ is a solution to the latter and vice versa. Moreover, $gXg^\dagger$ belongs to the polytope Λ_n. Therefore $gA_\alpha g^\dagger$ specifies a vertex. In other words, given $\alpha \in \mathcal{V}_n$ and $g \in \mathrm{Cl}_n$, Eq. (A.12) holds for a suitable $\beta \in \mathcal{V}_n$. We thus define $g\alpha := \beta$, and Eq. (A.12)

becomes

$$gA_\alpha g^\dagger = A_{g\alpha}. \tag{A.13}$$

Therefore,

$$\begin{aligned} \textstyle\sum_\alpha W_{gAg^\dagger}(\alpha)A_\alpha &= gAg^\dagger \\ &= \textstyle\sum_\alpha W_A(\alpha)\, gA_\alpha g^\dagger \\ &= \textstyle\sum_\alpha W_A(\alpha)A_{g\alpha} \\ &= \textstyle\sum_\alpha W_A(g^{-1}\alpha)A_\alpha. \end{aligned}$$

Thus, $W_{gAg^\dagger}(\alpha) = W_A(g^{-1}\alpha)A_\alpha$. □

We remark that, for qubits, only nonunique quasiprobability functions can be Clifford covariant. Namely, if the phase space point operators form an operator basis, i.e., are linearly independent, then the resulting quasiprobability function cannot be Clifford covariant [58].

The covariance property can be used to efficiently simulate the effect of Clifford unitaries in QCM, as an alternative to the method of treating Clifford unitaries discussed in the main text.

(4) Traciality. In the absence of a continuously varying parameter s, we introduce a dual quasiprobability function $\tilde{W}$ in addition to W, to stand in for $F^{(-s)}$. Namely, for all projectors Π_a^s, corresponding to measurements of Pauli observables T_a with outcome s, we define

$$\tilde{W}_{\Pi_a^s}(\alpha) := Q_a(s|\alpha).$$

By linearity, this implies expressions for all $\tilde{W}_{T_a}(\alpha)$. Since the Pauli operators form an operator basis, again by linearity one obtains $\tilde{W}_A$ for any operator A. Then,

$$\mathrm{Tr}AB = \sum_\alpha \tilde{W}_A(\alpha)W_B(\alpha)$$

follows.

We thus satisfy the SW criteria (1) - (4).

To conclude, we emphasize that for the present purpose of classically simulating QCM, a crucial property of W is positivity preservation under Pauli measurement. This property has no counterpart in the Stratonovich-Weyl correspondence.

A.3.3 Complexity parameter of the state polytope Λ_n

A question that arises with Theorem 7 is what determines the value of n labelling the state polytope Λ_n, and hence the complexity of the simulation. In this regard, we make the following observation.

Lemma 40. *Any quantum computation in the magic state model that operates on an initial state $|\mu\rangle_A \otimes |\sigma\rangle_B$, where $|\mu\rangle$ is an n-qubit magic state and $|\sigma\rangle$ is an m-qubit stabilizer state, can with the same efficiency be run on the magic state $|\mu\rangle$ alone.*

Supplementing the nonstabilizer magic state $|\mu\rangle$ with stabilizer states is thus redundant. For example, if the magic states used in a given QCM are all of T-type, then n can be taken to be the number of those states.

Proof of Lemma 40. Without loss of generality, we discuss the version of QCM where the quantum computation consists of a sequence of only Pauli measurements. We give an explicit procedure to replace the sequence τ on $A \otimes B$ by an equivalent sequence $\tilde{\tau}^{(A)}$ of measured observables that act only on the subsystem A. The proof is by induction, and the induction hypothesis is that, at time t, the sequence $\tau_{\leq t}$ of measurements has been replaced by a computationally equivalent sequence $\tilde{\tau}^{(A)}_{\leq t}$ of Pauli measurements on the register A only. This statement is true for $t = 0$, i.e., the empty measurement sequence. We now show that the above statement for time t implies the analogous statement for time $t + 1$.

At time t, the state of the quantum register evolved under the computationally equivalent measurement sequence $\tilde{\tau}^{(A)}_{\leq t}$ is $|\Psi(t)\rangle = |\psi(t)\rangle_A \otimes |\sigma\rangle_B$. We now consider the Pauli observable $T(t+1) \in \tau$ to be measured next, and write $T(t+1) = S_A(t+1) \otimes R_B(t+1)$. There are two cases.

Case I: $T(t+1)$ *commutes with the entire stabilizer* $\mathcal{S}$ *of* $|\sigma\rangle$. Hence, also $R_B(t+1)$ commutes with $\mathcal{S}$. But then, either $R_B(t+1)$ or $-R_B(t+1)$ is in $\mathcal{S}$, and $R_B(t+1)$ may be replaced by its eigenvalue ± 1 in the measurement. Hence, the measurement of $T(t+1)$ is equivalent to the measurement of $\pm S_A(t+1)$.

Case II: $T(t+1)$ *does not commute with the entire stabilizer* $\mathcal{S}$ *of* $|\sigma\rangle$. Then, the measurement outcome s_{t+1} is completely random. Further, there exists a Clifford unitary U such that

$$\begin{aligned} U\mathcal{S}U^\dagger &= \langle X_{B:1}, X_{B:2}, .., X_{B:m} \rangle, \\ UT(t+1)U^\dagger &= Z_{B:1}. \end{aligned}$$

Therefore, the state resulting from the measurement of $T(t+1)$, with outcome s_{t+1} on the state $|\Psi(t)\rangle$ is the same state as the one resulting from the following procedure:

1. Apply the Clifford unitary U to $|\Psi(t)\rangle = |\psi(t)\rangle_A \otimes |\sigma(t)\rangle_B$, leading to

   $$U\,|\Psi(t)\rangle = |\tilde{\psi}(t)\rangle \otimes |\overline{+}\rangle_B,$$

   where $|\overline{+}\rangle_B := \bigotimes_{i \in B} |+\rangle_{B:i}$.

2. Measure $Z_{B:1}$ on $|\tilde{\psi}(t)\rangle \otimes |\overline{+}\rangle_B$, with outcome s_{t+1}.

3. Apply $U^\dagger$.

Now, note that the measurement in Step 2, of the Pauli observable $Z_{B:1}$ is applied to the stabilizer state $|\overline{+}\rangle_B$. The result is $|\tilde{\sigma}(t+1)\rangle = |s_{t+1}\rangle_{B:1} \bigotimes_{j=2}^{m} |+\rangle_{B:j}$. Therefore, after normalization, the effect of the measurement can be replaced by the unitary $(X_{B:1})^{s_{t+1}}\, H_{B:1}$.

Thus, the whole procedure may be replaced by the Clifford unitary $U^\dagger\, (X_{B:1})^{s_{t+1}}\, H_{B:1}\, U$. But Clifford unitaries don't need to be implemented. They are just propagated past the last measurement, thereby affecting the measured observables by conjugation whereby their Pauli-ness is preserved. In result, in Case II, the measurement of $T(t+1)$ doesn't need to be performed at all. It is replaced by classical post-processing of the subsequent measurement sequence.

We conclude that in both the cases I and II, given the induction assumption, the original measurement sequence $\tau_{\leq t+1}$ can be replaced by a computationally equivalent measurement

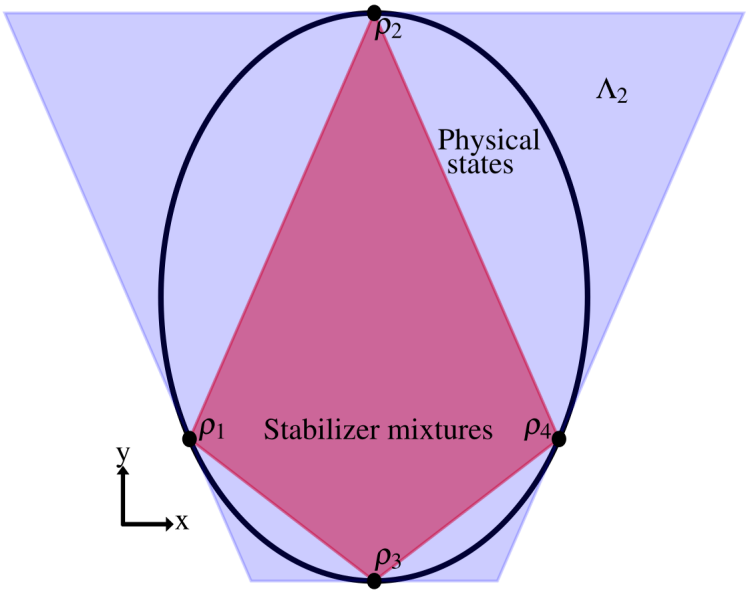


Figure A.2: Cross section of the space $\mathrm{Herm}_1(4)$ parameterized by Eq. (A.14). The two-qubit stabilizer polytope is inscribed in the set of physical states and the set of physical states is inscribed in the polytope Λ_2. The states labelled by ρ_1–ρ_4 are given in Eq. (A.15).

sequence $\tilde{\tau}^{(A)}_{\leq t+1}$ acting on register A only. By induction, the complete measurement sequence τ can be replaced by a computationally equivalent sequence $\tilde{\tau}^{(A)}$ acting on A only.

Since the measurements $\tilde{\tau}^{(A)}$ are applied to an unentangled initial state $|\mu\rangle_A \otimes |\sigma\rangle_B$, the register B can be dropped. Finally, the measurement sequence $\tilde{\tau}^{(A)}$ is of the same length or shorter than τ, and can be efficiently computed from the latter. Hence its implementation is at least as efficient. □

A.3.4 The two-qubit polytope Λ_2

Fig. 1 in the main text shows what the polytope Λ_n looks like for a single qubit, $n = 1$. The polytope Λ_1 is a cube inscribing the Bloch ball—the set of physical quantum states. The situation is similar for multiple qubits. In general, Λ_n is not a hypercube, it is a more general polytope, but it still inscribes the set of physical states.

Fig. A.2 shows a cross section of the space $\mathrm{Herm}_1(4)$, indicating the states which are contained in the two-qubit stabilizer polytope—the set of mixtures of pure two-qubit stabilizer states, the set of physical states, and the polytope Λ_2. The cross section is parameterized by

$$\rho(x, y) = \frac{1}{4} I_{12} + x(Z_1 + Z_2) + y(X_1X_2 + Z_1Z_2 - Y_1Y_2). \tag{A.14}$$

The four states labelled in the figure are

$$\begin{aligned}\rho_1 &= \frac{1}{4}I_{12} - \frac{1}{8}(Z_1 + Z_2),\\ \rho_2 &= \frac{1}{4}I_{12} + \frac{1}{4}(X_1X_2 + Z_1Z_2 - Y_1Y_2),\\ \rho_3 &= \frac{1}{4}I_{12} - \frac{1}{12}(X_1X_2 + Z_1Z_2 - Y_1Y_2),\\ \rho_4 &= \frac{1}{4}I_{12} + \frac{1}{8}(Z_1 + Z_2).\end{aligned} \tag{A.15}$$

A.3.5 Λ_n is bounded

The set E_n has the structure of a vector space over $\mathbb{Z}_2$. For an isotropic subspace $I \subset E_n$ and a value assignment $\lambda : I \to \mathbb{Z}_2$ we define a projector

$$\Pi_{I,\lambda} = \frac{1}{|I|}\sum_{a\in I}(-1)^{\lambda(a)}T_a.$$

Summing over all value assignments gives a resolution of the identity: $\sum_\lambda \Pi_{I,\lambda} = 1$. For each stabilizer state $|\sigma\rangle$ there is a unique pair (I,λ) consisting of a maximal isotropic subspace and a value assignment defined on it such that $\Pi_{I,\lambda} = |\sigma\rangle\langle\sigma|$. Then for $X \in \Lambda_n$ we have

$$\begin{aligned}\mathrm{Tr}(X\Pi_a^s) &= \mathrm{Tr}\big(X\Pi_a^s \textstyle\sum_{\lambda'}\Pi_{I',\lambda'}\big)\\ &= \textstyle\sum_{\lambda'}\mathrm{Tr}\big(X\Pi_a^s\Pi_{I',\lambda'}\big)\\ &= \textstyle\sum_{\lambda|\ \lambda(a)=s}\mathrm{Tr}(X\Pi_{I,\lambda}) \geq 0.\end{aligned}$$

Therefore Λ_n is contained in the hypercube defined by

$$\{X \in \mathrm{Herm}_1(2^n)|\ \mathrm{Tr}(\Pi_a^s X) \geq 0,\ \forall a \in E_n - \{0\},\ s = 0,1\}$$

and thus it is bounded.

A.4 Supplementary materials for Chapter 5

A.4.1 Relation between the 1st and 2nd cohomology group of Q

Here we provide an alternative derivation of the mathematical fact represented by the lower horizontal arrow in the diagram of Eq. (5.32), namely $[\Phi_{\mathrm{cov}}] = 0 \Longrightarrow [\zeta] = 0$, by purely mathematical reasoning that bypasses the discussion of Wigner functions. The result we obtain goes slightly beyond the one presented in the main text in that it applies to any symmetry group $G \subseteq \mathcal{C}\ell_n$, regardless of whether $\mathcal{P}_n \subset G$. The idea is similar to the use of long exact sequences in cohomology that appears in Section 5.6 of [69].

We consider the exact sequence of $\mathbb{Z}_d$-modules

$$0 \to E \to C^1 \to U_{\mathrm{cov}} \to 0, \tag{A.16}$$

where the second map is defined by sending $a \in E$ to the 1-cochain $[a,-] : E \to \mathbb{Z}_d$. There

is an associated long exact sequence in cohomology

$$\cdots H^1(Q,E) \to H^1(Q,C^1) \to H^1(Q,U_{\text{cov}}) \xrightarrow{\sigma} H^2(Q,E) \to \cdots$$

which sends $[\Phi_{\text{cov}}] \in H^1(Q,U_{\text{cov}})$ to the cohomology class $\sigma([\Phi_{\text{cov}}]) = [\zeta]$, where

$$\zeta = d^h(\tilde{\Phi} \circ \theta) \in C^2(Q,C^1), \tag{A.17}$$

see Section 5.2.3. To see that the cocycle ζ belongs, in the sense of the exact sequence in Eq. (A.16), to $C^2(Q,E)$ observe that $d^v\zeta = 0$ since $d^h\Phi_{\text{cov}} = 0$. The symmetry group G, which is the extension of Q by $N \subset E$, is a subgroup of $\tilde{G}$ obtained by extending Q by E using the cocycle ζ. This is a consequence of the way ζ is defined. If $[\Phi_{\text{cov}}] = 0$ then the extension class $[\zeta] = \sigma([\Phi_{\text{cov}}])$ of $\tilde{G}$ vanishes. In this case $\tilde{G}$ splits, which also implies that the subgroup G splits.

A.5 Supplementary materials for Chapter 9

A.5.1 Graph theory results

In the proof of Theorem 27, we use some concepts from graph theory. A signed graph G^σ is a undirected graph G together with a sign function $\sigma : \text{Edge}(G) \to \mathbb{Z}_2$ that assigns to each edge e of G a sign $(-1)^{\sigma(e)}$. We say a graph has full rank if its adjacency matrix has full rank. The perrank of a graph G is the order of the largest subgraph of G which is a disjoint union of copies of K_2 and cycles. If G has order n we say G has full perrank if $perrank(G) = n$.

We use the following Lemma from Ref. [211]

Lemma 41 (Ref. [211], Theorem 2.1)**.** *Let G be a graph. Then there exists a sign function σ for G so that G^σ has full rank if and only if G has full perrank.*

In a bipartite graph, all cycles are even, so edges can be removed from any cycle of a bipartite graph to give copies of K_2 which cover the same vertices. Therefore, for bipartite graphs the existence of a $\{1,2\}$-factor of order n implies the existence of a matching of size n. Thus we get the following corollaries

Corollary 14. *Let G be a bipartite graph with adjacency matrix*

$$A = \begin{bmatrix} 0 & B \\ B^T & 0 \end{bmatrix}. \tag{A.18}$$

where B is a square matrix. Then there exists a choice of signs σ for G so that B has full rank if and only if there exists a perfect matching of G^σ.

Corollary 15. *Let G be a bipartite graph with adjacency matrix of the form Eq.* (A.18) *where B is a (not necessarily square) $m \times n$ matrix with $m \leq n$. Then there exists a choice of signs σ for G so that B has full rank if and only if there exists a matching of order m.*

A.5.2 Lower bound on $f(n,m)$

Here we prove the lower bound on the function $f(n,m)$ defined in Eq. (9.19) and used in the proof of Theorem 27. We restate the bound in the following lemma.

Lemma 42. *For the function $f(n,m)$ defined in Eq. (9.19), we have the bound $f(n,m) \geq 2^n - 1$.*

Proof of Lemma 42. The bound is established if we can prove the bound $\log_2(2^n f(n,m)) \geq 2n$. For this we use the exact form of Stirling's formula

$$\sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n+1}} < n! < \sqrt{2\pi n}\left(\frac{n}{e}\right)^n e^{\frac{1}{12n}}.$$

Applying this to $\log_2(2^n f(n,m))$, we have

$$\begin{aligned}
\log_2(2^n f(n,m)) =& \log_2((2m)!) + \log_2((2n-2m)!) - \log_2(m!) - \log_2((n-m)!) + \log_2(2n+1-2m) \\
\geq& \quad \frac{1}{2}\log_2(2m) + 2m\log_2(2m) - 2m\log_2(e) + \frac{\log_2(e)}{24m+1} \\
&+ \frac{1}{2}\log_2(2(n-m)) + 2(n-m)\log_2(2(n-m)) - 2(n-m)\log_2(e) + \frac{\log_2(e)}{24(n-m)+1} \\
&- \frac{1}{2}\log_2(m) - m\log_2(m) + m\log_2(e) - \frac{\log_2(e)}{12m} \\
&- \frac{1}{2}\log_2(n-m) - (n-m)\log_2(n-m) + (n-m)\log_2(e) - \frac{\log_2(e)}{12(n-m)} \\
&+ \log_2(2n+1-2m) \\
=& 1 + 2n + m\log_2(m) + (n-m)\log_2(n-m) - n\log_2(e) + \log_2(2n+1-2m) \\
&+ \log_2(e)\left[\frac{1}{24m+1} + \frac{1}{24(n-m)+1} - \frac{1}{12m} - \frac{1}{12(n-m)}\right] \\
\geq& 2n + m\log_2(m) + (n-m)\log_2(n-m) - n\log_2(e) + 1 - \frac{\log_2(e)}{6} \\
\geq& 2n + m\log_2(m) + (n-m)\log_2(n-m) - n\log_2(e)
\end{aligned}$$

Here in the second last step we simply throw away some positive terms and use the fact $-\frac{1}{12m} - \frac{1}{12(n-m)} \geq -\frac{1}{6}$. In the last step we use $1 - \log_2(e)/6 > 0$. This expression is minimized at $m = n/2$, and so we have

$$\log_2(2^n f(n,m)) \geq 2n + n(\log_2(n/2) - \log_2(e)).$$

For $n \geq 6$, $\log_2(n/2) > \log_2(e)$, and so we can establish the bound $\log_2(2^n f(n,m)) \geq 2n$. For the remaining cases $n < 6$, we can compute the function and show directly that $f(n,m) \geq 2^n - 1$. These remaining cases are shown in Table A.2. Therefore, $f(n,m) \geq 2^n - 1$ for all n, m.

□

$f(n,m)$ $n\backslash m$	1	2	3	4	5	2^n-1
1	1					1
2	3	3				3
3	15	9	15			7
4	105	45	45	105		15
5	945	315	225	315	945	31

Table A.2: The function $f(n,m)$ of Eq. (9.19) evaluated for all m and for $1 \leq n \leq 5$. This table, together with the calculation below show that $f(n,m) \geq 2^n - 1$ for all n and m.

A.5.3 Upper bound on robustness of Λ vertices

Lemma 43. *For all the vertices described by Theorem 27, on any number of qubits n, the robustness of magic $\mathfrak{R}_S(A_{\mathcal{O}}^{\eta})$, satisfies the inequalities*

$$\mathfrak{R}_S(A_{\mathcal{O}}^{\eta}) \leq n^2 + \frac{n}{2} + 1 \tag{A.19}$$

Proof of Lemma 43. Let $a, b \in \mathbb{R}^d$ and let $c \in \mathbb{R}^d$ collinear to a and b but not in the convex hull of a and b. Then the negativity of c relative to a and b is

$$||c||_1 = \frac{||a-c||_2 + ||b-c||_2}{||a-b||_2}$$

Let c be a vertex of Λ_n as given in Theorem 27 (in Cartesian coordinates). Let a be the point on the stabilizer facet $1 + c \cdot x = 0$ closest to c and let b be the origin, these three points are all collinear. Then we have

$$||a-c||_2 = \frac{\frac{n(2n+1)}{4} + 1}{\left(\frac{n(2n+1)}{4}\right)^{1/2}} = \left(\frac{n(2n+1)}{4}\right)^{1/2} + \left(\frac{n(2n+1)}{4}\right)^{-1/2},$$

$$||b-c||_2 = \left(\frac{n(2n+1)}{4}\right)^{1/2} \quad \text{and} \quad ||a-b||_2 = \left(\frac{n(2n+1)}{4}\right)^{-1/2}.$$

Thus,

$$\mathfrak{R}_S(c) \leq ||c||_1 = \frac{n(2n+1)}{2} + 1$$

□

A.6 Supplementary materials for Chapter 10

A.6.1 Upper bound on the number of vertices of Λ

Lemma 44. *For any number of qubits $n \in \mathbb{N}$, the number of vertices $|\mathcal{V}_n|$ of Λ_n satisfies*

$$\log_2(|\mathcal{V}_n|) \leq 1 + 4^{n-1}\left[n^2 - n + 2\log_2(10e)\right].$$

Proof of Lemma 44. According to the upper bound theorem of polytope theory [207] (for a more detailed exposition on the upper bound theorem see Ref. [161, §8.4]), the number of facets of a D-dimensional polytope with v vertices is bounded by the number of facets of the D-dimensional cyclic polytope with v vertices, denoted $C(v, D)$. By duality, the number of vertices of the Λ polytope on n qubits is bounded by the number of vertices of the polar dual of the $4^n - 1$-dimensional cyclic polytope with $|\mathcal{S}_n|$ vertices.

The number of vertices of the dual of $C(v, D)$ (equivalently the number of facets of $C(v, D)$) is [194, §4.7],

$$f_{d-1}(C(v,D)) = \begin{cases} \frac{v}{v-m}\binom{v-m}{m} & \text{for even } D = 2m, \\ 2\binom{v-m-1}{m} & \text{for odd } D = 2m+1. \end{cases}$$

In the case of Λ_n, $D = 4^n - 1$, i.e. D is odd and $m = 2^{2n-1} - 1$. The number of facets of Λ_n is $|\mathcal{S}_n|$. Therefore,

$$\begin{aligned} |\mathcal{V}_n| \leq & f_{4^n-2}\left(C(|\mathcal{S}_n|, 4^n - 1)\right) \\ = & 2\binom{|\mathcal{S}_n| - 2^{2n-1}}{2^{2n-1} - 1} \\ \leq & 2\binom{|\mathcal{S}_n| - 2^{2n-1}}{2^{2n-1}} \\ \leq & 2\left(\frac{e(|\mathcal{S}_n| - 2^{2n-1})}{2^{2n-1}}\right)^{2^{2n-1}}. \end{aligned}$$

In the last line we use a standard upper bound for the binomial coefficient $\binom{n}{k} \leq \left(\frac{en}{k}\right)^k$. Then

$$\begin{aligned} \log_2(|\mathcal{V}_n|) \leq & 1 + 2^{2n-1}\left[\log_2\left(|\mathcal{S}_n| - 2^{2n-1}\right) - 2n + 1 + \log_2(e)\right] \\ \leq & 1 + 2^{2n-1}\left[\log_2\left(5 \cdot 2^{(n^2+3n)/2}\right) - 2n + 1 + \log_2(e)\right] \\ \leq & 1 + 2^{2n-1}\left[\frac{n^2+3n}{2} - 2n + 1 + \log_2(5e)\right] \\ = & 1 + 4^{n-1}\left[n^2 - n + 2\log_2(10e)\right]. \end{aligned}$$

Here in the second line we use the upper bound on the number of n-qubits stabilizer states,

$|\mathcal{S}_n| \leq 5 \cdot 2^{(n^2+3n)/2}$, from Ref. [212]. This proves the upper bound. □

We also have the following somewhat simpler bound.

Corollary 16. *For any number of qubits* $n \in \mathbb{N}$, $\log_2(|\mathcal{V}_n|) \leq 4^n n^2$.

Proof of Corollary 16. For $n = 1$ and $n = 2$, we can enumerate the vertices of Λ_n and we find that the numbers of vertices are 8 and 22320 respectively. These both satisfy the bound. Further, for $n \geq 2$, $1 + 4^{n-1}[n^2 - n + 2\log_2(10e)] \leq 4^n n^2$ and so the remaining cases follow immediately from Lemma 44. □

Appendix B

Mathematical proofs

B.1 Proofs for Chapter 2

B.1.1 Proof of Lemma 1

Proof of Lemma 1. First note that for the local Pauli operators we have $ZX = \omega XZ$. This can be verified by direct computation as follows:

$$\begin{aligned} ZX =& \sum_{i,j\in\mathbb{Z}_d} \omega^i |i\rangle \langle i|j+1\rangle \langle j| \\ =& \sum_{j\in\mathbb{Z}_d} \omega^{j+1} |j+1\rangle \langle j| \\ =& \omega \sum_{i,j\in\mathbb{Z}_d} \omega^j |i+1\rangle \langle i|j\rangle \langle j| \\ =& \omega XZ. \end{aligned}$$

Now for any $a, b \in E$, we compute

$$\begin{aligned} T_a T_b =& \mu^{\phi(a)+\phi(b)} \bigotimes_{k=1}^{n} Z^{a_z[k]} X^{a_x[k]} Z^{b_z[k]} X^{b_x[k]} \\ =& \mu^{\phi(a)+\phi(b)} \omega^{\langle a_x|b_z\rangle} \bigotimes_{k=1}^{n} Z^{a_z[k]} Z^{b_z[k]} X^{a_x[k]} X^{b_x[k]} \\ =& \mu^{\phi(a)+\phi(b)} \omega^{\langle a_x|b_z\rangle} \bigotimes_{k=1}^{n} Z^{b_z[k]} Z^{a_z[k]} X^{b_x[k]} X^{a_x[k]} \\ =& \mu^{\phi(a)+\phi(b)} \omega^{\langle a_z|b_x\rangle - \langle a_x|b_z\rangle} \bigotimes_{k=1}^{n} Z^{b_z[k]} X^{b_x[k]} Z^{a_z[k]} X^{a_x[k]} \\ =& \omega^{[a,b]} T_b T_a. \end{aligned}$$

□

B.1.2 Proof of Lemma 3

Proof of Lemma 3. The proof is obtained by adapting the proof of Ref. [1, Lemma 1]. Let I, I' be isotropic subgroups of E such that $I \subsetneq I'$. Let $r : I \to \mathbb{Z}_d$ be a noncontextual value assignment for I and $\Gamma_{I,r}^{I'}$ be the set of all noncontextual value assignments on I' satisfying

$$r'|_I = r \quad \forall r' \in \Gamma_{I,r}^{I'}. \tag{B.1}$$

Then

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \Pi_{I'}^{r'} = \sum_{r' \in \Gamma_{I,r}^{I'}} \frac{1}{|I'|} \sum_{a \in I'} \omega^{-r'(a)} T_a \tag{B.2}$$

$$= \frac{1}{|I'|} \sum_{a \in I'} \left[\sum_{r' \in \Gamma_{I,r}^{I'}} \omega^{-r'(a)} \right] T_a. \tag{B.3}$$

We have two cases for the inner sum in the last expression. If $a \in I$ then $r'(a) = r(a)$ for all $r' \in \Gamma_{I,r}^{I'}$. Therefore,

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \omega^{-r'(a)} = \left| \Gamma_{I,r}^{I'} \right| \omega^{-r(a)}. \tag{B.4}$$

In the second case, $a \notin I$. $\Gamma_{I,r}^{I'}$ is the coset of a vector space, the proof of this is analogous to the proof of Ref. [1, Lemma 2] which applies only to qubits. Therefore, by character orthogonality,

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \omega^{-r'(a)} = 0. \tag{B.5}$$

Thus,

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \Pi_{I'}^{r'} = \frac{1}{|I'|} \sum_{a \in I'} \left| \Gamma_{I,r}^{I'} \right| \delta_{a \in I} \omega^{-r(a)} T_a \tag{B.6}$$

$$= \frac{\left| \Gamma_{I,r}^{I'} \right|}{|I'|} \sum_{a \in I} \omega^{-r(a)} T_a. \tag{B.7}$$

We have $|\Gamma_{I,r}^{I'}| = |I'|/|I|$. Therefore,

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \Pi_{I'}^{r'} = \frac{1}{|I|} \sum_{a \in I} \omega^{-r(a)} T_a = \Pi_I^r \tag{B.8}$$

which proves the lemma. □

B.1.3 Proof of Lemma 4

Proof of Lemma 4. The action of the Clifford group on the Pauli operators is defined in Eq. (2.9). With this equation, for a projector onto a stabilizer state $|\sigma\rangle$ corresponding to maximal isotropic subgroup $I \subset E$ and noncontextual value assignment $r : I \to \mathbb{Z}_d$ we have

$$g |\sigma\rangle \langle\sigma| g^\dagger = g\Pi_I^r g^\dagger = \frac{1}{|I|} \sum_{a \in I} \omega^{-r(a)} g T_a g^\dagger \tag{B.9}$$

$$= \frac{1}{|I|} \sum_{a \in I} \omega^{-r(a)+\Phi_g(a)} T_{S_g(a)} \tag{B.10}$$

$$= \frac{1}{|I|} \sum_{a \in g\cdot I} \omega^{-g\cdot r(a)} T_a = \Pi_{g\cdot I}^{g\cdot r} \tag{B.11}$$

where $g \cdot I = \{S_g(a) \mid a \in I\}$ and $g \cdot r$ is defined by the relation

$$g \cdot r(S_g(a)) = r(a) - \Phi_g(a) \quad \forall a \in I. \tag{B.12}$$

In order to show that $g |\sigma\rangle \langle\sigma| g^\dagger$ is a projector onto a stabilizer state it suffices to show that $g \cdot I$ is a maximal isotropic subgroup of E and that $g \cdot r : g \cdot I \to \mathbb{Z}_d$ is a noncontextual value assignment for $g \cdot I$.

$g\cdot I$ is isotropic since I is isotropic and S_g is a symplectic operation. Also, $|g\cdot I| = |I| = d^n$ so by Lemma 2, $g \cdot I$ is a maximal isotropic subgroup.

Since $T_0 \propto \mathbb{1}$

$$g T_0 g^\dagger = T_0 \tag{B.13}$$

so $\Phi_g(0) = 0$ for any $g \in \mathcal{C}\ell$. Therefore,

$$\omega^{-g\cdot r(0)} T_0 = \omega^{-r(0)+\Phi_g(0)} T_0 = \omega^{-r(0)} T_0 = \mathbb{1}. \tag{B.14}$$

For any $a, b \in g \cdot I$, there exist $c, d \in I$ such that $S_g(c) = a$ and $S_g(d) = b$. Computing the product $g T_c T_d g^\dagger$ in two different ways we have

$$g T_c T_d g^\dagger = g T_c g^\dagger g T_d g^\dagger \tag{B.15}$$

$$= \omega^{\Phi_g(c)+\Phi_g(d)} T_{S_g(c)} T_{S_g(d)} \tag{B.16}$$

$$= \omega^{\Phi_g(c)+\Phi_g(d)-\beta(a,b)} T_{a+b} \tag{B.17}$$

and

$$g T_c T_d g^\dagger = \omega^{-\beta(c,d)} g T_{c+d} g^\dagger \tag{B.18}$$

$$= \omega^{-\beta(c,d)+\Phi_g(c+d)} T_{a+b}. \tag{B.19}$$

Therefore, $-\Phi_g(c) - \Phi_g(d) + \Phi_g(c+d) = \beta(c,d) - \beta(a,b)$ and so

$$\begin{aligned} & g \cdot r(a) + g \cdot r(b) - g \cdot r(a+b) && \text{(B.20)} \\ &= r(c) + r(d) - r(c+d) - \Phi_g(c) - \Phi_g(d) + \Phi_g(c+d) && \text{(B.21)} \\ &= -\beta(a,b). && \text{(B.22)} \end{aligned}$$

Thus, $g \cdot r$ satisfies Eq. (2.10) and $\Pi_{g\cdot I}^{g\cdot r}$ is a projector onto a stabilizer state. □

B.1.4 Proof of Lemma 5

Proof of Lemma 5. Let $I, J \subset E$ be isotropic subgroups with noncontextual value assignments $r : I \to \mathbb{Z}_d$ and $s : J \to \mathbb{Z}_d$.

Case 1: $r|_{I\cap J} = s|_{I\cap J}$. Let $r \star s$ denote the unique noncontextual value assignment on the set $I + J \cap I^\perp$ such that $r \star s|_I = r$ and $r \star s|_{J\cap I^\perp} = s|_{J\cap I^\perp}$. We calculate

$$\begin{aligned} \Pi_I^r \Pi_J^s \Pi_I^r =& \frac{1}{|I|^2|J|} \sum_{a,c\in I} \sum_{b\in J} \omega^{-r(a)-r(c)-s(b)} T_a T_b T_c && \text{(B.23)} \\ =& \frac{1}{|I|^2|J|} \sum_{a,c\in I} \sum_{b\in J} \omega^{-r(a)-r(c)-s(b)-\beta(a,c)+[b,c]} T_{a+c} T_b && \text{(B.24)} \\ =& \frac{1}{|I|^2|J|} \sum_{a,c\in I} \sum_{b\in J} \omega^{-r(a+c)-s(b)+[b,c]} T_{a+c} T_b && \text{(B.25)} \\ =& \frac{1}{|I||J|} \Pi_I^r \sum_{b\in J} \omega^{-s(b)} \left[\sum_{c\in I} \omega^{[b,c]} \right] T_b && \text{(B.26)} \\ =& \frac{1}{|J|} \Pi_I^r \sum_{b\in J\cap I^\perp} \omega^{-s(b)} T_b && \text{(B.27)} \\ =& \frac{1}{|I||J|} \sum_{a\in I} \sum_{b\in J\cap I^\perp} \omega^{-r(a)-s(b)-\beta(a,b)} T_{a+b} && \text{(B.28)} \\ =& \frac{1}{|I||J|} \sum_{a\in I} \sum_{b\in J\cap I^\perp} \omega^{-r\star s(a+b)} T_{a+b}. && \text{(B.29)} \end{aligned}$$

Each Pauli operator in the above sum appears with the same multiplicty. For an element $a \in I + J \cap I^\perp$, let $\mu(a)$ denote the number of distinct pairs $(b,c) \in I \times J \cap I^\perp$ such that $b+c = a$. We have $\mu(a) = \mu(0)$ for any $a \in I+J\cap I^\perp$. To see this, suppose $a \in I+J\cap I^\perp$ and let $(c_1,d_1),(c_2,d_2),\ldots,(c_{\mu(a)},d_{\mu(a)})$ be the distinct pairs in $I \times J \cap I^\perp$ such that $c_j + d_j = a$. Then the pairs $(c_j - c_1, d_j - d_1) \in I \times J \cap I^\perp$ for $j = 2,\ldots,\mu(a)$ together with the pair $(0,0) \in I \times J \cap I^\perp$ show that $\mu(0) \geq \mu(a)$.

Now let $(c_1,d_1),(c_2,d_2),\ldots,(c_{\mu(0)},d_{\mu(0)})$ denote the $\mu(0)$ distinct pairs in $I \times J \cap I^\perp$ such that $c_j + d_j = 0$, and $(c,d) \in I \times J \cap I^\perp$ be such that $c + d = a$. Then the pairs $(c_j + c, d_j + d)$, $j = 1,2,\ldots,\mu(0)$ show that $\mu(a) \geq \mu(0)$. Therefore, $\mu(a) = \mu(0)$ for any $a \in I + J \cap I^\perp$. Here we can see $\mu(0) = |I \cap J|$

Then we have

$$\Pi_I^r \Pi_J^s \Pi_I^r = \frac{1}{|I||J|} \sum_{a\in I} \sum_{b\in J\cap I^\perp} \omega^{-r\star s(a+b)} T_{a+b} \tag{B.30}$$

$$= \frac{\mu(0)}{|I||J|} \sum_{a\in I+J\cap I^\perp} \omega^{-r\star s(a)} T_a \tag{B.31}$$

$$= \frac{|I\cap J||I+J\cap I^\perp|}{|I||J|} \Pi^{r\star s}_{I+J\cap I^\perp}. \tag{B.32}$$

Since $|I + J \cap I^\perp| = |I| \cdot |J \cap I^\perp|/|I \cap J|$,

$$\Pi_I^r \Pi_J^s \Pi_I^r = \frac{|J\cap I^\perp|}{|J|} \Pi^{r\star s}_{I+J\cap I^\perp}. \tag{B.33}$$

Taking a trace of this equation we get

$$\operatorname{Tr}(\Pi_I^r \Pi_J^s) = \operatorname{Tr}(\Pi_I^r \Pi_J^s \Pi_I^r) \tag{B.34}$$

$$= \frac{|J\cap I^\perp|}{|J|} \operatorname{Tr}\big(\Pi^{r\star s}_{I+J\cap I^\perp}\big) \tag{B.35}$$

$$= \frac{|J\cap I^\perp|}{|J|} \cdot \frac{d^n}{|I+J\cap I^\perp|} \tag{B.36}$$

$$= \frac{|I\cap J|}{|I||J|} d^n > 0. \tag{B.37}$$

Case 2: $r|_{I\cap J} \neq s|_{I\cap J}$. Denote by S_J^s the stabilizer group $\{\omega^{-s(b)} T_b \mid b \in J\}$ and denote by V_J^s the subspace of the Hilbert space $\mathbb{C}^{d^n}$ stabilized by S_J^s. By assumption, there exists a $c \in I \cap J$ such that $r(c) \neq s(c)$. Therefore, for any vector $|\psi\rangle \in V_J^s$,

$$\Pi_I^r |\psi\rangle = \Pi_I^r \omega^{-s(c)} T_c |\psi\rangle \tag{B.38}$$

$$= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)-s(c)} T_a T_c |\psi\rangle \tag{B.39}$$

$$= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)-s(c)-\beta(a,c)} T_{a+c} |\psi\rangle \tag{B.40}$$

$$= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a+c)-r(c)-s(c)} T_{a+c} |\psi\rangle \tag{B.41}$$

$$= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)-r(c)-s(c)} T_a |\psi\rangle \tag{B.42}$$

$$= \omega^{-r(c)-s(c)} \Pi_I^s |\psi\rangle . \tag{B.43}$$

That is, $\Pi_I^s |\psi\rangle = \omega^{-r(c)-s(c)} \Pi_I^s |\psi\rangle$. But since $r(c) \neq s(c)$, $\omega^{-r(c)-s(c)} \neq 1$ so this relation can be true only if $\Pi_I^r |\psi\rangle = 0$.

We can write

$$\Pi_J^s = \sum_{i=1}^{\dim(V_J^s)} |\psi_i\rangle \langle\psi_i| \tag{B.44}$$

where $\{|\psi_i\rangle \mid 1 \leq i \leq \dim(V_J^s)\}$ is a basis for V_J^s. Therefore,

$$\Pi_I^r \Pi_J^s \Pi_I^r = \sum_{i=1}^{\dim(V_J^s)} \Pi_I^r |\psi_i\rangle \langle\psi_i| \Pi_I^r = 0. \tag{B.45}$$

Taking a trace of this equation we get

$$\mathrm{Tr}(\Pi_I^r \Pi_J^s \Pi_I^r) = \mathrm{Tr}(\Pi_I^r \Pi_J^s) = 0. \tag{B.46}$$

This proves the lemma. □

B.2 Proofs for Chapter 3

B.2.1 Proof of Lemma 8

Proof of Lemma 8. The "if" direction follows immediately from the fact $\mathcal{V}_M \subset \mathcal{V}$. We only need to prove the "only if" direction.

Let $(\Omega, \gamma) \in \mathcal{V}$ and $\tilde{\Omega}$ be a maximal cnc set such that $\Omega \subsetneq \tilde{\Omega}$. As a result of Theorem 2 of Chapter 3, Ω and $\tilde{\Omega}$ have the forms

$$\Omega = \bigcup_{k=1}^{\zeta} \langle a_k, \tilde{I}\rangle \text{ and } \tilde{\Omega} = \bigcup_{k=1}^{\xi} \langle a_k, \tilde{I}\rangle$$

where $\zeta < \xi$, $\tilde{I} \subset E$ is an isotropic subspace, and $a_1, \ldots, a_\xi$ pairwise anticommute. We define two noncontextual value assignments γ_0 and γ_1 on $\tilde{\Omega}$ as follows: $\gamma_0(b) = \gamma_1(b) = \gamma(b)$ for all $b \in \Omega$, and for each $b \in \{a_{\zeta+1}, \ldots, a_\xi\}$, $\gamma_0(b) = 0$ and $\gamma_1(b) = 1$. The values of γ_0, γ_1 on the remaining elements of $\tilde{\Omega}$ can be determined through Eq. (3.5). Then

$$\begin{aligned}\frac{1}{2}\left(A_{\tilde{\Omega}}^{\gamma_0} + A_{\tilde{\Omega}}^{\gamma_1}\right) &= \frac{1}{2^{n+1}} \sum_{b\in\tilde{\Omega}} \left((-1)^{\gamma_0(b)} + (-1)^{\gamma_1(b)}\right) T_b \\ &= A_\Omega^\gamma + \frac{1}{2^{n+1}} \sum_{b\in\tilde{\Omega}\setminus\Omega} \left((-1)^{\gamma_0(b)} + (-1)^{\gamma_1(b)}\right) T_b. \end{aligned} \tag{B.47}$$

Each $b \in \tilde{\Omega} \setminus \Omega$ can be written as $b = a_j + g$ for some $j \in \{\zeta + 1, \ldots, \xi\}$ and $g \in \tilde{I}$. Then, with all addition mod 2, we have

$$\begin{aligned}\gamma_0(b) &= \beta(a_j, g) + \gamma_0(a_j) + \gamma(g) = \beta(a_j, g) + \gamma(g), \\ \gamma_1(b) &= \beta(a_j, g) + \gamma_1(a_j) + \gamma(g) = \beta(a_j, g) + 1 + \gamma(g).\end{aligned}$$

In particular, $\gamma_1(b) = \gamma_0(b) + 1 \mod 2$ and therefore each term in the sum in Eq. (B.47)

vanishes and we are left with $\frac{1}{2}(A^{\gamma_0}_{\tilde{\Omega}} + A^{\gamma_1}_{\tilde{\Omega}}) = A^{\gamma}_{\Omega}$.

Then, if a state ρ has a nonnegative expansion in phase space point operators $A^{\gamma}_{\Omega}, (\Omega, \gamma) \in \mathcal{V}$, for each A^{γ}_{Ω} that appears with nonnegative weight in the expansion corresponding to a nonmaximal set Ω, $\frac{1}{2}(A^{\gamma_0}_{\tilde{\Omega}} + A^{\gamma_1}_{\tilde{\Omega}})$ with $(\tilde{\Omega}, \gamma_0), (\tilde{\Omega}, \gamma_1) \in \mathcal{V}_M$ can be substituted without introducing any negativity. Therefore, ρ is also positively representable with respect to $\mathcal{V}_M$. □

B.2.2 Proof of Lemma 14

Proof of Lemma 14. Statement (A): $(\Omega, \gamma + [a, \cdot]) \in \mathcal{V}$, $\forall a \in \Omega$. The set Ω does not change, and we only need to check the properties in Def. 7 that concern the function update, i.e., Eqs. (3.5), and (3.2).

Assume that $\gamma : \Omega \longrightarrow \mathbb{Z}_2$ satisfies $d\gamma = \beta$ on Ω, i.e. $d\gamma(f) = \beta(f)$ for all faces $f \in F(\Omega)$. Consider any such face, with its boundary ∂f consisting of the edges c, d and $c + d$. By definition of $F(\Omega)$ it holds that $c, d, c + d \in \Omega$. Then, with all addition mod 2,

$$\begin{aligned} d(\gamma + [a, \cdot])(f) =& d\gamma(f) + [a, \cdot](\partial f) \\ =& d\gamma(f) + [a, c] + [a, d] + [a, c + d] \\ =& d\gamma(f) \\ =& \beta(f). \end{aligned}$$

Thus, $\gamma + [a, \cdot]$ satisfies Eq. (3.5).

Furthermore, assume that γ satisfies Eq. (3.2). Then, $(\gamma + [a, \cdot])(0) = \gamma(0) + [a, 0] = \gamma(a)$. Hence, $\gamma + [a, \cdot]$ satisfies Eq. (3.2).

Statement (B): $(\Omega \times a, \gamma \times s_a) \in \mathcal{V}$, $\forall a \notin \Omega$ and $s_a \in \mathbb{Z}_2$. There are four items to check in Def. 7, namely (I) $\Omega \times a$ is closed under inference, (II) $\Omega \times a$ is noncontextual, (III) $\gamma \times s_a$ satisfies Eq. (3.5), and (IV) $\gamma \times s_a$ satisfies Eq. (3.2).

(I): Consider $c, d \in \Omega \times a$, with $[c, d] = 0$, and denote $c' = c + a$, $d' = d + a$. There are three sub-cases. *(i)* $c, d \in \Omega_a$. Then, $c + d \in \Omega_a$, since Ω_a is closed under inference by Lemma 13. Thus, $c + d \in \Omega \times a$.

(ii) $c \in \Omega_a$, $d \notin \Omega_a$. By construction of $\Omega \times a$, $d' \in \Omega_a$. Thus, $c + d = c + (d' + a) = (c + d') + a$. Now, since $[c, d] = 0$ by assumption and $[c, a] = 0$ ($c \in \Omega_a$) it follows that $[c, d'] = 0$. Since Ω_a is closed by Lemma 13, it holds that $c + d' \in \Omega_a$. By construction of $\Omega \times a$, $c + d = (c + d') + a \in \Omega \times a$.

(iii) $c, d \notin \Omega_a$. By construction of $\Omega \times a$, $c', d' \in \Omega_a$. Thus, $c + d = (c' + a) + (d' + a) = c' + d'$, and further $[c', d'] = 0$. Since Ω_a is closed under inference by Lemma 13, $c' + d' = c + d \in \Omega_a$. Thus, $c + d \in \Omega \times a$.

Thus in all three cases, $c, d \in \Omega \times a$, with $[c, d] = 0$, implies $c + d \in \Omega \times a$. Hence, $\Omega \times a$ is closed under inference.

(III): Assume that $d\gamma = \beta$ on Ω, and consider a triple of edges $c, d, c + d \in \Omega \times a$ with $[c, d] = 0$. Then, either (i) all or (ii) one of these edges are in the component Ω_a.

(i) $c, d, c + d \in \Omega_a$. Since $\Omega_a \subset \Omega$ and with Eq. (3.20a), it holds that $d(\gamma \times s_a)(c, d) = d\gamma(c, d) = \beta(c, d)$.

(ii) Without loss of generality assume that $c \in \Omega_a$ and $d, c+d \notin \Omega_a$, and denote $c' = c+a$, $d' = d+a$ as before. Then, for the face $f = (c, d)$ with boundary ∂f consisting of the edges

c, d and $c+d$,

$$\begin{aligned} d(\gamma \times s_a)(f) =& \gamma \times s_a(c) + \gamma \times s_a(d) + \gamma \times s_a(c+d) \\ =& \gamma(c) + \gamma(d') + \gamma(c+d') + \beta(a,d) + \beta(a,c+d) \\ =& \beta(c,d') + \beta(a,d) + \beta(a,c+d) \\ =& \beta(c,d). \end{aligned}$$

Therein, in the second line we have used Eq. (3.20), in the third line Eq. (3.5), in the fourth line Eq. (3.23), and in the fourth line $d\beta(a,d,c) = 0$, cf. Eqs. (2.7) and (3.6).

(II): Per Def. 6, $\Omega \times a$ is noncontextual if there is a function $\Gamma : \Omega \times a \longrightarrow \mathbb{Z}_2$ that satisfies $d\Gamma = \beta$. We have explicitly constructed such a function in (III) above, $\Gamma := \gamma \times s_a$.

(IV): Assume that $\gamma : \Omega \longrightarrow \mathbb{Z}_2$ satisfies Eq. (3.2). Since $0 \in \Omega_a$ for all cnc sets Ω, with Eq. (3.20a) it follows that $\gamma \times s_a(0) = \gamma(0)$, and hence $\gamma \times s_a$ also satisfies Eq. (3.2). □

B.2.3 Proof of Lemma 16

Proof of Lemma 16. Recall from Lemma 10 that each set Ω can be written in the form $\Omega = \bigcup_{k=1}^{\xi(\Omega)} I_k$, where each I_k is an isotropic subspace, $I_k = \langle a_k, \tilde{I} \rangle$, $a_k \in E$. Therefore, for all $(\Omega, \gamma) \in \mathcal{V}$, it holds that

$$A_\Omega^\gamma = \left(\sum_{k=1}^{\xi(\Omega)} A_{I_k}^{\gamma|_{I_k}} \right) - (\xi(\Omega) - 1) A_{\tilde{I}}^{\gamma|_{\tilde{I}}}.$$

Therein, the phase space point operators appearing on the r.h.s. are all of the type $m = 0$, i.e., they correspond to stabilizer states. The Wigner function $\delta_{(\Omega,\gamma)}$ representing the operator A_Ω^γ can thus be expanded as

$$\delta_{(\Omega,\gamma)} = \left(\sum_{k=1}^{\xi(\Omega)} \delta_{(I_k, \gamma|_{I_k})} \right) - (\xi(\Omega) - 1) \delta_{(\tilde{I}, \gamma|_{\tilde{I}})}.$$

Denote by $\| \cdot \|_1$ the 1-norm of the expansion in terms of phase space point operators A_Ω^γ, and by $\| \cdot \|_{1,S}$ the 1-norm of the expansion in terms of (density matrices of) stabilizer states. With the last equation, the triangle inequality, $\|\delta_{(I_k,\gamma|_{I_k})}\|_{1,S} = \|\delta_{(\tilde{I},\gamma|_{\tilde{I}})}\|_{1,S} = 1$, and $\xi(\Omega) \leq 2n+1$ for all cnc sets Ω (cf. Lemma 10), it follows that

$$\|\delta_{(\Omega,\gamma)}\|_{1,S} \leq 4n+1. \tag{B.48}$$

Now, for any given state ρ consider the optimal representation W_ρ, i.e., the one with minimal norm $\|W_\rho\|_1$. Then,

$$\begin{aligned}\mathfrak{R}_S(\rho) \leq & \|W_\rho\|_{1,S} \\ = & \left\| \sum_{(\Omega,\gamma)\in\mathcal{V}} W_\rho(\Omega,\gamma)\delta_{(\Omega,\gamma)} \right\|_{1,S} \\ \leq & (4n+1) \sum_{(\Omega,\gamma)\in\mathcal{V}} |W_\rho(\Omega,\gamma)| \\ = & (4n+1)\|W_\rho\|_1 \\ = & (4n+1)\mathfrak{R}(\rho).\end{aligned}$$

Therein, in the first line we have an inequality because the representation W_ρ of ρ is optimized for $\|W_\rho\|_1$, not necessarily for $\|W_\rho\|_{1,S}$. The third line follows by the triangle inequality and Eq. (B.48), and the fifth line holds as an equality because W_ρ, per assumption, was chosen to minimize $\|W_\rho\|_1$.

This proves the right half of Eq. (3.33). The left half, $\mathfrak{R}(\rho) \leq \mathfrak{R}_S(\rho)$, follows from the fact that all stabilizer states correspond to phase space point operators of type $m = 0$. Hence, an expansion in terms of stabilizer states induces an expansion in terms of phase space point operators A_Ω^γ, with the same nonzero coefficients. □

B.3 Proofs for Chapter 4

B.3.1 Proof of Lemma 19

Theorem 1 in [1] classifies the maximal cnc sets. For the present purpose it may be rephrased as

Lemma 45. *If a subset of E_n is closed under inference and does not contain a Mermin square then it is noncontextual.*

Proof sketch for Lemma 45. Theorem 1 of [1] classifies the subsets of E_n that are closed under inference and do not contain a Mermin square. They all turn out to be noncontextual. □

Proof of Lemma 19. Pick an n, any pair (Ω,γ). A_Ω^γ has unit trace, and, as shown in [1], satisfies $\mathrm{Tr}(|\sigma\rangle\langle\sigma|A_\Omega^\gamma) \geq 0$. Therefore, $A_\Omega^\gamma \in \Lambda_n$, and A_Ω^γ has an expansion

$$A_\Omega^\gamma = \sum_{\beta\in\mathcal{V}_n} p_{\Omega,\gamma}(\beta) A_\beta, \tag{B.49}$$

where $p_{\Omega,\gamma}(\beta) \geq 0$, $\forall\beta$, and $\sum_\beta p_{\Omega,\gamma}(\beta) = 1$. Thus, $p_{\Omega,\beta}$ is a probability distribution. Henceforth, we consider any A_β for which $p_{\Omega,\gamma}(\beta) > 0$.

Now pick an $a \in \Omega$ and consider $\mathrm{Tr}\left(T_a A_\Omega^\gamma\right)$. With Eq. (4.11), it holds that $(-1)^{\gamma(a)} = \sum_\beta p_{\Omega,\gamma}(\beta)\langle T_a\rangle_\beta$. Since $p_{\Omega,\beta}$ is a probability distribution and $|\langle T_a\rangle_\beta| \leq 1$ for all β, it follows

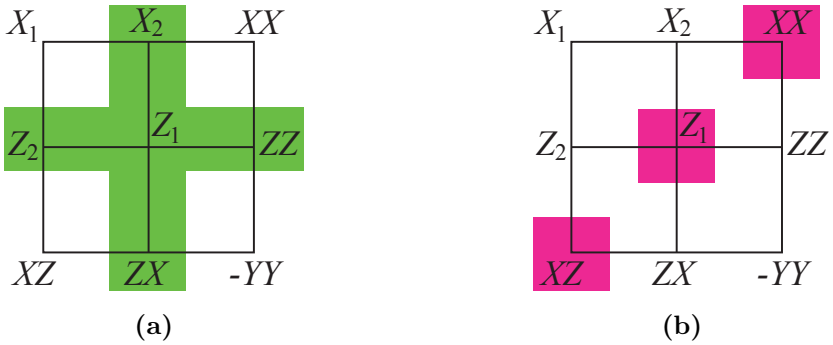


Figure B.1: Two possibilities for the set $\Omega \cap M$, shown in color.

that

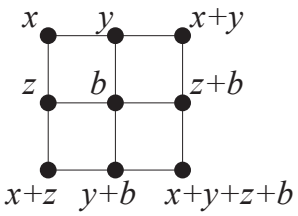
$$\langle T_a \rangle_\beta = (-1)^{\gamma(a)}, \ \forall\, \beta \text{ with } p_{\Omega,\gamma}(\beta) > 0.$$

That is, every phase space point operator that appears on the rhs. of Eq. (B.49) with nonzero coefficient agrees with A^γ_Ω on the expectation values $\langle T_a \rangle$ for all $a \in \Omega$.

Now we turn to the expectation values for $b \notin \Omega$. Any set $\tilde{\Omega} \subset E_n$ that is closed under inference and contains both Ω and b is contextual, by the maximality of Ω. By Lemma 45, any such $\tilde{\Omega}$ contains a Mermin square M, and furthermore $b \in M$.

Since M is closed under inference, so is $\Omega \cap M$. Also, since Ω is maximal, $\Omega \cap M$ is maximal in M. Up to permutations of rows and columns, there are two possibilities for $\Omega \cap M$, which are displayed in Fig. B.1.

Case (a). For any b there exists a triple $\{x, y, z\} \subset M \backslash b$ such that $[x, y] = [x, z] = [b, y] = [b, z] = 0$, $[x, b] = [y, z] \neq 0$. We have the following Mermin square:



Therein, Mermin's contradiction to the existence of a noncontextual HVM is encapsulated in the operator relation $(T_x T_y)(T_z T_b) = -(T_x T_z)(T_y T_b)$.

We chose the following phase conventions.

$$\begin{aligned} T_{x+y} &= T_x T_y, \ T_{z+b} = T_z T_b, \\ T_{x+z} &= T_x T_z, \ T_{y+b} = T_y T_b, \end{aligned} \tag{B.50}$$

and

$$\begin{aligned} T_{x+y+z+b} &= T_{x+z} T_{y+b}, \\ T_{x+y+z+b} &= -T_{x+y} T_{z+b}. \end{aligned} \tag{B.51}$$

Recall that with the first part of the proof $\langle T_j \rangle_\beta = (-1)^{\gamma(j)}$, for $j = x, y, z$. Now assume

that $\langle T_b \rangle_\beta = \nu$, with $-1 \leq \nu \leq 1$. Now, with Eq. (B.50)

$$\begin{aligned} \langle T_{x+y} \rangle_\beta &= (-1)^{\gamma(x)+\gamma(y)},\ \langle T_{x+z} \rangle_\beta = (-1)^{\gamma(x)+\gamma(z)}, \\ \langle T_{y+b} \rangle_\beta &= \nu(-1)^{\gamma(y)},\ \langle T_{z+b} \rangle_\beta = \nu(-1)^{\gamma(z)}. \end{aligned}$$

Therefore, with Eq. (B.51),

$$\begin{aligned} \langle T_{x+y+z+b} \rangle_\beta &= \nu(-1)^{\gamma(x)+\gamma(y)+\gamma(z)} \\ &= -\nu(-1)^{\gamma(x)+\gamma(y)+\gamma(z)}. \end{aligned}$$

This is satisfiable only if $\nu = 0$, and hence $\langle T_b \rangle_\beta = 0$.

Case (b). The argument is analogous to case (a), and we do not repeat it here.

By the above case distinction, for any $b \in E_n \backslash \Omega$ either case (a) or (b) applies, and each way the consequence is that $\langle T_b \rangle_\beta = 0$. Therefore, any phase space point operator A_β that appears on the rhs of Eq. (B.49) with nonzero $p_{\Omega,\gamma}(\beta)$ agrees with A^γ_Ω on *all* expectation values of Pauli observables; hence $A^\gamma_\Omega = A_\beta$ for all such β.

Now assume there exists no such A_β. Taking the trace of Eq. (B.49) yields $1 = 0$; contradiction. Hence, there must exist a β such that $A^\gamma_\Omega = A_\beta$, for all (Ω, γ). □

B.4 Proofs for Chapter 5

B.4.1 Proof of Lemma 23

Lemma 23 is a restatement of Theorem 44 in [31]. The entire chapter in [31] to which Theorem 44 belongs is written under the assumption that d is odd. We restate the proof here, to clarify that for this particular lemma the assumption of odd d is not needed.

Proof of Lemma 23. Denote by ν_k, $k \in \mathbb{Z}_d$, a character of $\mathbb{Z}_d$, $\nu_k(x) = \omega^{kx}$, for all $x \in \mathbb{Z}_d$; and by the same symbol the vector $\nu_k = (1, \omega^k, \omega^{2k}, .., \omega^{(d-1)k})^T$. For any $d \times d$ matrix M defined in Eq. (5.45) it holds that

$$\begin{aligned} [M\nu_k]_y &= \textstyle\sum_x f(x-y)\omega^{kx} \\ &= \textstyle\sum_x f(x)\omega^{kx}\omega^{ky} \\ &= d\,\hat{f}(k)\,\omega^{ky}, \end{aligned}$$

and hence ν_k is an eigenvector of M with eigenvalue $d\,\hat{f}(k)$. Since there are d characters ν_k, the matrix M is diagonal in their basis. All its eigenvalues are nonnegative if and only if $\hat{f}$ is nonnegative. □

B.5 Proofs for Chapter 6

B.5.1 Proof of Theorem 17

As a preparation for the update rules of $A^\gamma_{I,\Omega}$ we consider certain linear combinations of operators in $\mathrm{Herm}_1(\mathbb{C}^2)$ and express them as probabilistic mixtures of the vertices of Λ_1. We will denote the vertices in Λ_1 by A^α_E where α is a value assignment on $E = E_1$. Other points

in Λ_1 that we are interested in are the projectors $\Pi_v^{s_v}$ where $v = x, y, z$, where $y = x + z$, and $s_v = 0, 1$.

Lemma 46. *We have the following decompositions for the points in Λ_1. Let $v, w \in E$ be nonzero distinct elements.*

1. *Let α_0 and α_1 be value assignments on E defined by $\alpha_0(v) = \alpha_1(v) = s_v$, $\alpha_1(w) = 1 + \alpha_0(w)$ and $\alpha_1(v + w) = 1 + \alpha_0(v + w)$. Then*

$$\Pi_v^{s_v} = \frac{A_E^{\alpha_0} + A_E^{\alpha_1}}{2}.$$

2. *Let α_0 and α_1 be value assignments on E defined by $\alpha_0(v) = \alpha_1(v) = s_v$, $\alpha_0(w) = \alpha_1(w) = s_w$ and $\alpha_1(v + w) = 1 + \alpha_0(v + w)$. Then*

$$\Pi_v^{s_v} + \frac{\Pi_w^{s_w} - \Pi_w^{s_w+1}}{2} = \frac{A_E^{\alpha_0} + A_E^{\alpha_1}}{2}.$$

3. *Let α' be value assignment on E defined by $\alpha'(v) = \alpha(v)$, $\alpha'(w) = 1 + \alpha(w)$ and $\alpha'(v + w) = 1 + \alpha(v + w)$. Then*

$$\frac{2\Pi_v^{\alpha(v)} + A_E^{\alpha}}{3} = \frac{2A_E^{\alpha} + A_E^{\alpha'}}{3}.$$

4. *Let α' be value assignment defined in (3). Then*

$$2\Pi_v^{\alpha(v)} - A_E^{\alpha} = A_E^{\alpha'}.$$

5. *Let α_i, $i = 0, 1, 2$, be defined by $\alpha_i(v) = s_v$ for all i, $\alpha_0(w) = \alpha_1(w) = 1+\alpha_2(w) = s_w$, $\alpha_1(v + w) = \alpha_2(v + w) = 1 + \alpha_0(v + w)$. Then*

$$\Pi_v^{s_v} + \frac{\Pi_w^{s_w} - \Pi_w^{s_w+1}}{4} = \frac{2A_E^{\alpha_0} + A_E^{\alpha_1} + A_E^{\alpha_2}}{4}.$$

Proof of Lemma 46. Verification is straightforward: We write each side of the equations in the Pauli basis and compare. □

Let us label the maximal isotropics in $\langle a \rangle^\perp$ by $I_v = \langle \tilde{v}, a \rangle$, $I_w = \langle \tilde{w}, a \rangle$ and $I_{v+w} = \langle \tilde{v} + \tilde{w}, a \rangle$. We introduce a map

$$\Lambda_1 \to \Lambda_2 \tag{B.52}$$

determined by $A_E^{\alpha} \mapsto A_{\langle a \rangle^\perp}^{\tilde{\alpha}}$ where $\tilde{\alpha}$ is specified by $\tilde{\alpha}(a) = s_a$, $\tilde{\alpha}(b) = \alpha(b)$ for $b = \tilde{v}, \tilde{w}, \tilde{v}+\tilde{w}$. In the following proof we label the maximal isotropic subspaces of $\langle a \rangle^\perp$ by I_v, I_w, I_{v+w}.

Proof of Theorem 17. In Case I we have

$$\Pi_{a,s} A_{I,\Omega}^{\gamma} \Pi_a^s = \delta_{s_a,\gamma(a)} \left(A_I^{\gamma} + \frac{1}{4}(A_{\tilde{\Omega}}^{\tilde{\gamma}'} - A_{\tilde{\Omega}}^{\tilde{\gamma}''}) \right)$$

where we used $\tilde{\gamma}'(a) = \tilde{\gamma}''(a) = \gamma(a)$ since $a = v + w$ for some $v, w \in \Omega - \{0\}$. In this case I and $\tilde{\Omega}$ are maximal isotropics in $\langle a \rangle^\perp$ intersecting at $\langle a \rangle$. (1.a) follows immediately by calculating $\Pi_a^s A_{I,\Omega}^\gamma \Pi_a^s$. For (1.b) we observe that Lemma 46 part (5) and Eq. (B.52) imply that

$$\frac{\Pi_a^s A_{I,\Omega}^\gamma \Pi_a^s}{\mathrm{Tr}\Big(A_{I,\Omega}^\gamma \Pi_a^s\Big)} = A_I^\gamma + \frac{1}{4}(A_{\tilde{\Omega}}^{\tilde{\gamma}'} - A_{\tilde{\Omega}}^{\tilde{\gamma}''}) = \frac{2A_{\langle a \rangle^\perp}^{\tilde{\alpha}_0} + A_{\langle a \rangle^\perp}^{\tilde{\alpha}_1} + A_{\langle a \rangle^\perp}^{\tilde{\alpha}_2}}{4}$$

where $I_v = I$, $I_w = \tilde{\Omega}$, $s_v = \gamma(\tilde{v})$, $s_w = \tilde{\gamma}'(\tilde{w})$,

$$\begin{aligned} \tilde{\alpha}_0(a) = \tilde{\alpha}_1(a) = \tilde{\alpha}_2(a) &= s_a \\ \tilde{\alpha}_0(\tilde{v}) = \tilde{\alpha}_1(\tilde{v}) = \tilde{\alpha}_2(\tilde{v}) &= \gamma(\tilde{v}) \\ \tilde{\alpha}_0(\tilde{w}) = \tilde{\alpha}_1(\tilde{w}) = 1 + \tilde{\alpha}_2(\tilde{w}) &= \tilde{\gamma}'(\tilde{w}) \\ 1 + \tilde{\alpha}_0(\tilde{v} + \tilde{w}) = \tilde{\alpha}_1(\tilde{v} + \tilde{w}) &= \tilde{\alpha}_2(\tilde{v} + \tilde{w}). \end{aligned} \tag{B.53}$$

Case II follows from

$$\frac{\Pi_a^s A_{I,\Omega}^\gamma \Pi_a^s}{\mathrm{Tr}\Big(A_{I,\Omega}^\gamma \Pi_a^s\Big)} = \frac{1}{2} A_{I\times a}^{\gamma\times a} + \frac{1}{4}\left(\delta_{s_a,\tilde{\gamma}'(a)} A_{\tilde{\Omega}}^{\tilde{\gamma}'} - \delta_{s_a,1+\tilde{\gamma}'(a)} A_{\tilde{\Omega}}^{\tilde{\gamma}''}\right)$$

since $\tilde{\gamma}''(v) = \tilde{\gamma}'(v) + 1$ for $v \in \Omega - \{0\}$. Here $\tilde{\Omega} = \langle a \rangle^\perp$. For (2.a) Lemma 46 part (4) and Eq. (B.52) imply that

$$\frac{\Pi_a^s A_{I,\Omega}^\gamma \Pi_a^s}{\mathrm{Tr}\Big(A_{I,\Omega}^\gamma \Pi_a^s\Big)} = \frac{1}{2} A_{I\times a}^{\gamma\times a} - \frac{1}{4} A_{\tilde{\Omega}}^{\tilde{\gamma}''} = A_{\langle a \rangle^\perp}^{\tilde{\alpha}'}$$

where $I_v = I \times a$, I_w is one of the other two isotropics,

$$\begin{aligned} \tilde{\alpha}'(a) &= s_a \\ \tilde{\alpha}'(\tilde{v}) &= \tilde{\gamma}''(\tilde{v}) \\ \tilde{\alpha}'(\tilde{w}) &= 1 + \tilde{\gamma}''(\tilde{w}) \\ \tilde{\alpha}'(\tilde{v} + \tilde{w}) &= 1 + \tilde{\gamma}''(\tilde{v} + \tilde{w}). \end{aligned} \tag{B.54}$$

For (2.a) Lemma 46 part (4) and Eq. (B.52) imply that

$$\frac{\Pi_a^s A_{I,\Omega}^\gamma \Pi_a^s}{\mathrm{Tr}\Big(A_{I,\Omega}^\gamma \Pi_a^s\Big)} = \frac{1}{2} A_{I\times a}^{\gamma\times a} + \frac{1}{4} A_{\tilde{\Omega}}^{\tilde{\gamma}'} = \frac{2A_{\langle a \rangle^\perp}^{\tilde{\alpha}} + A_{\langle a \rangle^\perp}^{\tilde{\alpha}'}}{3}$$

where $I_v = I \times a$, I_w is one of the other two isotropics, $\tilde{\alpha} = \tilde{\gamma}'$,

$$\begin{aligned} \tilde{\alpha}'(a) &= s_a \\ \tilde{\alpha}'(\tilde{v}) &= \tilde{\gamma}'(\tilde{v}) \\ \tilde{\alpha}'(\tilde{w}) &= 1 + \tilde{\gamma}'(\tilde{w}) \\ \tilde{\alpha}'(\tilde{v} + \tilde{w}) &= 1 + \tilde{\gamma}'(\tilde{v} + \tilde{w}). \end{aligned} \tag{B.55}$$

Case III follows from the following computation

$$\Pi_a^s A_{I,\Omega}^\gamma \Pi_a^s = \frac{1}{2} A_{I\times a}^{\gamma\times a} + \frac{\delta_{s_a,\tilde{\gamma}'(a)}}{4}\left(A_{\tilde{\Omega}}^{\tilde{\gamma}'} - A_{\tilde{\Omega}}^{\tilde{\gamma}''}\right).$$

We have $\tilde{\gamma}'(a) = \tilde{\gamma}''(a)$ since a can be written as $a = v + w$ with $v, w \in \Omega - \{0\}$. In this case I and $\tilde{\Omega}$ are maximal isotropics contained in $\langle a\rangle^\perp$ intersecting at $\langle a\rangle$. For (3.a) Lemma 46 part (1) and Eq. (B.52) imply that

$$\frac{\Pi_a^s A_{I,\Omega}^\gamma \Pi_a^s}{\mathrm{Tr}\left(A_{I,\Omega}^\gamma \Pi_a^s\right)} = \frac{1}{2} A_{I\times a}^{\gamma\times a} = \frac{A_{\langle a\rangle^\perp}^{\tilde{\alpha}_0} + A_{\langle a\rangle^\perp}^{\tilde{\alpha}_1}}{2}$$

where $I_v = I \times a$, $I_w = \tilde{\Omega}$, $s_v = \gamma(\tilde{v})$,

$$\begin{aligned} \tilde{\alpha}_0(a) = \tilde{\alpha}_1(a) &= s_a \\ \tilde{\alpha}_0(\tilde{v}) = \tilde{\alpha}_1(\tilde{v}) &= \gamma(\tilde{v}) \\ \tilde{\alpha}_1(\tilde{w}) &= 1 + \tilde{\alpha}_0(\tilde{w}) \\ \alpha_1(\tilde{v} + \tilde{w}) &= 1 + \tilde{\alpha}_0(\tilde{v} + \tilde{w}) \end{aligned} \tag{B.56}$$

For (3.b) Lemma 46 part (2) and Eq. (B.52) imply that

$$\frac{\Pi_a^s A_{I,\Omega}^\gamma \Pi_a^s}{\mathrm{Tr}\left(A_{I,\Omega}^\gamma \Pi_a^s\right)} = \frac{1}{2} A_{I\times a}^{\gamma\times a} + \frac{1}{4}(A_{\tilde{\Omega}}^{\tilde{\gamma}'} - A_{\tilde{\Omega}}^{\tilde{\gamma}''}) = \frac{A_{\langle a\rangle^\perp}^{\tilde{\alpha}_0} + A_{\langle a\rangle^\perp}^{\tilde{\alpha}_1}}{2}$$

where $I_v = I \times a$, $I_w = \tilde{\Omega}$, $s_v = \gamma(\tilde{v})$, $s_w = \tilde{\gamma}'(\tilde{w})$,

$$\begin{aligned} \tilde{\alpha}_0(a) = \tilde{\alpha}_1(a) &= s_a \\ \tilde{\alpha}_0(\tilde{v}) = \tilde{\alpha}_1(\tilde{v}) &= \gamma(\tilde{v}) \\ \tilde{\alpha}_0(\tilde{w}) = \tilde{\alpha}_1(\tilde{w}) &= \tilde{\gamma}'(\tilde{w}) \\ 1 + \tilde{\alpha}_0(\tilde{v} + \tilde{w}) &= \tilde{\alpha}_1(\tilde{v} + \tilde{w}). \end{aligned} \tag{B.57}$$

Note that the case where $a \in I$ and $a \in \Omega$ does not occur since $I \cap \Omega = 0$. □

B.6 Proofs for Chapter 7

B.6.1 Proof of Lemma 27

To prove the lemma, we will use the concept of polar duality for objects in the affine space $\mathrm{Herm}_1(\mathcal{H})$ (see Ref. [161] for a discussion of polar duality). For a Hilbert space $\mathcal{H}$ with inner product $\langle\cdot,\cdot\rangle$ define[1] the *polar dual* of a set $P \subset \mathcal{H}$ as

$$P^* = \left\{ x \in \mathbb{R}^N \;\middle|\; \langle x, y\rangle \geq -\frac{1}{d^n} \text{ for all } y \in P \right\}. \tag{B.58}$$

[1]The usual definition differs slightly by an irrelevant scaling factor.

If $P = \mathrm{conv}\{v_1, \ldots, v_k\}$ is a polytope, then this simplifies to

$$P^* = \left\{ x \in \mathbb{R}^N \;\middle|\; \langle x, v_i \rangle \geq -\frac{1}{d^n},\ i, \ldots, k \right\}. \tag{B.59}$$

If $P \subset Q \subset \mathcal{H}$, then obviously $Q^* \subset P^*$.

In our setting, we are interested in objects living in the affine space of matrices of trace one $\mathrm{Herm}_1(\mathcal{H})$. We can project $\mathrm{Herm}_1(\mathcal{H})$ into the linear subspace of Hermitian matrices of trace zero $\mathrm{Herm}_0(\mathcal{H})$ via the transformation

$$\pi : X \in \mathrm{Herm}_1(\mathcal{H}) \to \mathrm{Herm}_0(\mathcal{H}),\ X \mapsto X - \frac{1}{d^n}\mathbb{1}. \tag{B.60}$$

Now, observe that for $X, Y \in \mathrm{Herm}_1(\mathcal{H})$ we have $\mathrm{Tr}(XY) \geq 0$ if and only if

$$\mathrm{Tr}(\pi(X)\pi(Y)) = \mathrm{Tr}(XY) - \mathrm{Tr}\left(X \frac{1}{d^n}\mathbb{1} \right) \geq -\frac{1}{d^n}. \tag{B.61}$$

Hence, by associating $\mathrm{Herm}_0(\mathcal{H})$ and $\mathrm{Herm}_1(\mathcal{H})$, we define for a set $M \subset \mathrm{Herm}_1(\mathcal{H})$

$$M^* := \{ X \in \mathrm{Herm}_1(\mathcal{H}) \mid \mathrm{Tr}(XY) \geq 0 \ \text{ for all } Y \in M \}. \tag{B.62}$$

If $P = \mathrm{conv}\{X_1, \ldots, X_m\} \subset \mathrm{Herm}_1(\mathcal{H})$ is a polytope, then define its *polar dual* as

$$P^* = \{ Y \in \mathrm{Herm}_1(\mathcal{H}) \mid \mathrm{Tr}(X_i Y) \geq 0,\ i = 1, \ldots, m \}. \tag{B.63}$$

Thus, $\Lambda = \mathrm{SP}^*$ for SP being the *stabilizer polytope*: $\mathrm{SP} := \mathrm{conv}\,\{ |\sigma\rangle\langle\sigma| \mid \sigma \in \mathcal{S} \}$.

To prove that the set Λ is bounded, it will suffice to show that SP contains a set M, whose dual M^* is bounded. Additionally, we will make us of the concept of *dilation* [213, Chap. 9]: for a set $M \subset \mathrm{Herm}_1(\mathcal{H})$ define its dilation centered at the maximally mixed state via

$$c \cdot M := \left\{ \frac{1}{d^n}\mathbb{1} + c\pi(X) \mid X \in M \right\} \tag{B.64}$$

where $\pi : \mathrm{Herm}_1(\mathcal{H}) \to \mathrm{Herm}_0(\mathcal{H})$ is the projection that maps $X \in \mathrm{Herm}_1(\mathcal{H})$ to $X - \frac{1}{d^n}\mathbb{1}$. The dilation has the following property:

Lemma 47. *The dilation of a set $M \subset Herm_1(\mathcal{H})$ satisfies*

$$(c \cdot M)^* = \frac{1}{c} \cdot M^*. \tag{B.65}$$

Proof of Lemma 47. Observe that every $A, Y \in \mathrm{Herm}_1(\mathcal{H})$ can be written as

$$Y = \frac{1}{d^n}\mathbb{1} + c\pi(X), \quad A = \frac{1}{d^n}\mathbb{1} + \frac{1}{c}\pi(B)$$

for suitable $B, X \in \mathrm{Herm}_1(\mathcal{H})$. Then

$$\mathrm{Tr}(AY) = \frac{1}{d^{2n}} \mathrm{Tr}(\mathbb{1}) + \mathrm{Tr}(\pi(B)\pi(X)) \tag{B.66}$$

$$= \frac{1}{d^n} + \mathrm{Tr}\Big(B(X - \frac{1}{d^n}\mathbb{1})\Big) + \mathrm{Tr}\Big(-\frac{1}{d^n}\mathbb{1}(X - \frac{1}{d^n}\mathbb{1})\Big) \tag{B.67}$$

$$= \frac{1}{d^n} - \frac{1}{d^n} + \mathrm{Tr}(BX) \tag{B.68}$$

$$= \mathrm{Tr}(BX). \tag{B.69}$$

Hence, if $Y \in c \cdot M$ with $X \in M$, then $A \in (c \cdot M)^*$ if and only if

$$0 \leq \mathrm{Tr}(AY) = \mathrm{Tr}(BX) \quad \Longleftrightarrow \quad B \in M^*, \tag{B.70}$$

which is equivalent to $A \in \frac{1}{c} \cdot M^*$ □

Having introduced all necessary concepts, we will proceed with the proof of Lemma 27.

Proof of Lemma 27. By definition, Λ is a polyhedron and therefore convex and closed. To prove that Λ is bounded, the previous discussion implies that it suffices to find a set $M \subset \mathrm{SP}$ such that M^* is bounded. The object we will choose here will be a dilation of the full-dimensional simplex

$$\triangle_{\mathrm{Herm}_1(\mathcal{H})} := \mathrm{conv}\Big\{ A_\gamma := \frac{1}{d^n} \sum_{u \in E} \omega^{\gamma(u)} T_u \;\Big|\; \tag{B.71}$$

$$\gamma : E \to \mathbb{Z}_d,\ \gamma(u+v) = \gamma(u) + \gamma(v) \Big\}. \tag{B.72}$$

The simplex $\triangle_{\mathrm{Herm}_1(\mathcal{H})}$ is the Wigner simplex for d being an odd prime [19, 42]. For every d, it is a full-dimensional polytope as the convex-hull of d^{2n} affinely independent vertices A_γ in the $(d^{2n}-1)$-dimensional affine space $\mathrm{Herm}_1(\mathcal{H})$. Due to $\mathrm{Tr}\big(A_\gamma A_{\gamma'}\big) = \delta_{\gamma=\gamma'}$ for additive functions $\gamma, \gamma' : E \to \mathbb{Z}_d$, one can easily verify that the simplex $\triangle_{\mathrm{Herm}_1(\mathcal{H})}$ has the following hyperplane description:

$$\triangle_{\mathrm{Herm}_1(\mathcal{H})} = \{X \in \mathrm{Herm}_1(\mathcal{H}) \mid \mathrm{Tr}(A_\gamma X) \geq 0\}, \tag{B.73}$$

which makes it a self-dual simplex, i.e. $\triangle_{\mathrm{Herm}_1(\mathcal{H})} = \triangle^*_{\mathrm{Herm}_1(\mathcal{H})}$.

Now, Lemma 47 implies that for $c > 0$ the simplex $(c \cdot \triangle_{\mathrm{Herm}_1(\mathcal{H})})^*$ is bounded, since

$$(c \cdot \triangle_{\mathrm{Herm}_1(\mathcal{H})})^* = \frac{1}{c}\, \triangle^*_{\mathrm{Herm}_1(\mathcal{H})} = \frac{1}{c}\, \triangle_{\mathrm{Herm}_1(\mathcal{H})}\ . \tag{B.74}$$

Hence, it suffices to show that $c \cdot \triangle_{\mathrm{Herm}_1(\mathcal{H})} \subset \mathrm{SP}$ for some $c > 0$, because then

$$\Lambda = \mathrm{SP}^* \subset (c \cdot \triangle_{\mathrm{Herm}_1(\mathcal{H})})^* = \frac{1}{c}\, \triangle_{\mathrm{Herm}_1(\mathcal{H})} \tag{B.75}$$

implies that Λ is bounded. Therefore, we will show that dilations of the vertices of $\triangle_{\mathrm{Herm}_1(\mathcal{H})}$

are contained in SP, i.e.there is $c > 0$ such that

$$\frac{1}{d^n}\mathbb{1} + c(A_\gamma - \frac{1}{d^n}\mathbb{1}) \in \mathrm{SP} \tag{B.76}$$

for all additive maps γ.

To achieve this, we will write $1/d^n\mathbb{1} + c(A_\gamma - 1/d^n\mathbb{1})$ as a convex combination of normalized projectors of the form $\frac{|\langle a\rangle|}{d^n}\Pi^r_{\langle a\rangle} \in \mathrm{SP}$ for $a \in E$ with noncontextual value assignments $r : \langle a\rangle \to \mathbb{Z}_d$. Due to Corollary 1, the elements $\frac{|\langle a\rangle|}{d^n}\Pi^r_{\langle a\rangle}$ are indeed contained in SP. A noncontextual value assignment on a line $\langle a\rangle$ is always additive, that is

$$r(ka) = kr(a), \quad \forall k \in \mathbb{Z}_d \tag{B.77}$$

because with the phase convention chosen in Eq. (2.11), it can be checked directly that $T_a T_{ka} = T_{(k+1)a}$, and so by definition $\beta(a, ka) = 0$ for all $a \in E$. Let $C = \{a_1, \ldots, a_N\} \subset E$ be a set that "covers" E, that is

$$E = \cup_{a\in C}\langle a\rangle. \tag{B.78}$$

For each subset $\mathcal{I} \subset [N] := \{1, \ldots, N\}$, there is an $a_\mathcal{I} \in E$ such that

$$\langle a_\mathcal{I}\rangle := \bigcap_{k\in\mathcal{I}}\langle a_k\rangle. \tag{B.79}$$

First, we will write A_γ as a linear combination of stabilizer code projectors $\Pi^r_{\langle a\rangle}$. We claim that

$$A_\gamma = \frac{1}{d^n}\sum_{\mathcal{I}\subset[N]} (-1)^{|\mathcal{I}|+1} \cdot |\langle a_\mathcal{I}\rangle| \cdot \Pi^{\gamma|_{\langle a_\mathcal{I}\rangle}}_{\langle a_\mathcal{I}\rangle}. \tag{B.80}$$

Since γ is additive, the restriction $\gamma|_{\langle a_\mathcal{I}\rangle} : \langle a_\mathcal{I}\rangle \to \mathbb{Z}_d$ satisfies (B.77) and defines a noncontextual value assignment on $\langle a_\mathcal{I}\rangle$. We rewrite the right hand side of (B.80) in the following way:

$$\frac{1}{d^n}\sum_{\mathcal{I}\subset[N]} (-1)^{|\mathcal{I}|+1} \cdot |\langle a_\mathcal{I}\rangle| \cdot \Pi^{\gamma|_{\langle a_\mathcal{I}\rangle}}_{\langle a_\mathcal{I}\rangle} \tag{B.81}$$

$$= \frac{1}{d^n}\sum_{\mathcal{I}\subset[N]} (-1)^{|\mathcal{I}|+1} \sum_{b\in\langle a_\mathcal{I}\rangle} \omega^{\gamma(b)} T_b \tag{B.82}$$

$$= \frac{1}{d^n}\sum_{b\in E}\Big(\sum_{\mathcal{I}\subset[N]\,:\,b\in\langle a_\mathcal{I}\rangle} (-1)^{|\mathcal{I}|+1}\Big)\omega^{\gamma(b)} T_b. \tag{B.83}$$

Thus, it suffices to show that

$$\sum_{\mathcal{I}\subset[N]\,:\,b\in\langle a_I\rangle} (-1)^{|\mathcal{I}|+1} = 1 \quad \forall b \in E. \tag{B.84}$$

However, this is a consequence of the *inclusion-exclusion principle* [214], that is

$$1 = \delta_{b\in E} = \delta_{b\in\bigcup_{a\in C}\langle a\rangle} = \sum_{\mathcal{I}\subset[N]} (-1)^{|\mathcal{I}|+1}\delta_{b\in\bigcap_{k\in I}\langle a_k\rangle} \tag{B.85}$$

$$= \sum_{\mathcal{I}\subset[N]} (-1)^{|\mathcal{I}|+1}\delta_{b\in\langle a_{\mathcal{I}}\rangle} = \sum_{\mathcal{I}\subset[N]\,:\,b\in\langle a_{\mathcal{I}}\rangle} (-1)^{|\mathcal{I}|+1}. \tag{B.86}$$

Finally, we will show that there is $c > 0$ such that we can write $\frac{1}{d^n}\mathbb{1} + c(A_\gamma - \frac{1}{d^n}\mathbb{1})$ for every γ as a convex combination of the operators $\frac{|\langle a_{\mathcal{I}}\rangle|}{d^n}\Pi_{\langle a_{\mathcal{I}}\rangle}^{\gamma|_{\langle a_{\mathcal{I}}\rangle}} \in \mathrm{SP}$.

Observe that the identity (B.80) is equivalent to

$$A_\gamma - \frac{1}{d^n}\mathbb{1} = \sum_{\mathcal{I}\subset[N]} (-1)^{|\mathcal{I}|+1} \cdot \frac{|\langle a_{\mathcal{I}}\rangle|}{d^n} \cdot \Pi_{\langle a_{\mathcal{I}}\rangle^*}^{\gamma|_{\langle a_{\mathcal{I}}\rangle}} \tag{B.87}$$

with

$$\Pi_{\langle a\rangle^*}^r := \frac{1}{|\langle a\rangle|} \sum_{b\in\langle a\rangle\setminus\{0\}} \omega^{r(b)} T_b \in \mathrm{Herm}_0(\mathcal{H}),\ r : \langle a\rangle \to \mathbb{Z}_d. \tag{B.88}$$

Moreover, due to Lemma 3, we have

$$\mathbb{1} = \sum_r \Pi_{\langle a\rangle}^r \quad \Longrightarrow \quad 0 = \sum_r \Pi_{\langle a\rangle^*}^r \tag{B.89}$$

for all $a \in E$, $\frac{|\langle a\rangle|}{d^n}\Pi_{\langle a\rangle}^r \in \mathrm{SP}$, where r ranges over all noncontextual value assignments $r : \langle a\rangle \to \mathbb{Z}_d$. This implies that

$$(-1) \cdot \Pi_{\langle a_{\mathcal{I}}\rangle^*}^{\gamma|_{\langle a_{\mathcal{I}}\rangle}} = \sum_{r\neq\gamma|_{\langle a\rangle}} \Pi_{\langle a_{\mathcal{I}}\rangle^*}^r. \tag{B.90}$$

As a consequence, every summand in the right hand side of Eq. (B.87) can be written as a conic combination of elements $\Pi_{\langle a\rangle^*}^r$. By properly properly rescaling, we can find $c > 0$ such that

$$c(A_\gamma - \frac{1}{d^n}\mathbb{1}) = \sum_a \alpha_a \Pi_{\langle a\rangle^*}^{r_a} \quad \text{with} \quad \alpha_a \geq 0,\ \sum_a \alpha_a = 1, \tag{B.91}$$

which is equivalent to

$$\frac{1}{d^n}\mathbb{1} + c(A_\gamma - \frac{1}{d^n}\mathbb{1}) = \sum_a \alpha_a(\Pi_{\langle a\rangle^*}^{r_a} + \frac{1}{d^n}\mathbb{1}) = \sum_a \alpha_a \Pi_{\langle a\rangle}^{r_a} \in \mathrm{SP}. \tag{B.92}$$

This proves that $c\cdot\triangle_{\mathrm{Herm}_1(\mathcal{H})} \subset \mathrm{SP}$ for some $c > 0$, which remained to be shown. □

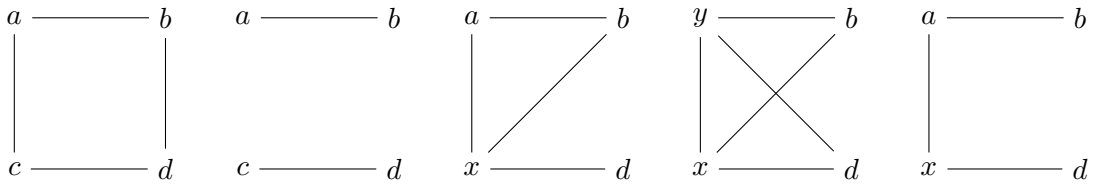


Figure B.2: Orthogonality graphs used in the proofs of Lemma 48, and Lemma 49.

B.7 Proofs for Chapter 8

B.7.1 Proofs of lemmas 31 and 32

To prove the lemmas, we will require some additional definitions and statements. For a set $\Omega \subset E$, we define the (undirected) *orthogonality graph* of Ω to be the graph $G(\Omega) = (\Omega, E)$ with vertex set Ω and edge set $E = \{\{a, b\} \mid [a, b] = 0\}$, i.e., vertices in $G(\Omega)$ are adjacent if and only if the corresponding elements of Ω are orthogonal.

One crucial observation is the following. Suppose you are given a set $\Omega = \{a, b, c, d\} \subset E$ such that the orthogonality graph $G(\Omega)$ is a 4-cycle (see Figure B.2a). If the qudit Hilbert space dimension is $d = 2$, i.e., the underlying field is $\mathbb{Z}_2$, then the orthogonal closure $\overline{\Omega}$ is the Mermin square [1, 72, 149]. In contrast, if the underlying field is $\mathbb{Z}_d$ and d is an odd prime, then $\overline{\Omega}$ is the 4-dimensional subspace spanned by $\{a, b, c, d\}$. This is the result of the following lemma.

Lemma 48. *If $\Omega = \{a, b, c, d\}$ and $G(\Omega)$ is a 4-cycle, then $\overline{\Omega} = \langle a, b, c, d\rangle$.*

The proof of this Lemma is based on the proof of Lemma 1 in Ref. [51].

Proof of Lemma 48. The orthogonality relations are depicted in Figure B.2a. Since $\langle a\rangle, \langle b\rangle, \langle c\rangle, \langle d\rangle \subset \overline{\Omega}$, without loss of generality we may assume that $[a, c] = [b, d] = 1$. Any point $v \in \langle a, b, c, d\rangle$ can be written as

$$v = (\alpha a + \delta d) + (\beta b + \gamma c) \tag{B.93}$$

with $\alpha, \beta, \gamma, \delta \in \mathbb{Z}_d$. Then since $[a, b] = [a, c] = [b, d] = [c, d] = 0$, it holds that $[\alpha a + \delta d, \beta b + \gamma c] = 0$, and so it suffices to show that the planes spanned by nonorthogonal elements are contained in $\overline{\Omega}$, i.e., $\langle a, d\rangle, \langle b, c\rangle \subset \overline{\Omega}$. But this follows immediately from the arguments in the proof of Lemma 1 in Ref. [51]. □

We also need the following additional lemma.

Lemma 49. *Assume that $\Omega = \{a, b, c, d\}$ such that $\dim(\langle a, b\rangle) = \dim(\langle c, d\rangle) = 2$ and $G(\Omega)$ is given by Figure B.2b. Then $\overline{\Omega}$ has the form Eq. (8.12), or $\overline{\Omega}$ contains a subset M such that $G(M)$ is a 4-cycle.*

Proof of Lemma 49. Since $\dim(\langle c, d\rangle) = 2$, there is a nontrivial $x \in \langle c, d\rangle \subset \overline{\Omega}$, such that $[a, x] = 0$. Now we distinguish two cases. For the first case, assume $[b, x] = 0$, and so the orthogonality graph of $\{a, b, x, d\}$ is Figure B.2c. Then there is a nontrivial $y \in \langle a, x\rangle \subset \overline{\Omega}$

Figure B.3: Orthogonality graphs used in the proof of Lemma 31.

such that $[d, y] = 0$, and defining $I := \langle x, y\rangle$ we have $b, d \in I^\perp$ (see Figure B.2d for the orthogonality relations). Then the closure of Ω is given by

$$\overline{\Omega} = \langle b, I\rangle \cup \langle d, I\rangle. \tag{B.94}$$

In the second case, assume $[b, x] \neq 0$ (see Figure B.2e for the relevant orthogonality relations). Since $\dim(\langle x, d\rangle) = 2$ there is a nontrivial $y \in \langle x, d\rangle = \langle c, d\rangle$ such that $[b, y] = 0$. Furthermore, as $a^\perp \cap \langle c, d\rangle = \langle x\rangle$, it holds that $[a, y] \neq 0$. Then for $\Omega' = \{a, b, x, y\}$ we have $\overline{\Omega} = \overline{\Omega'}$ and the graph $G(\Omega')$ is a 4-cycle. □

Proof of Lemma 31. Let $\Omega \subset E$ be a subspace. If $v \in \Omega$, there is nothing to show, so assume $v \notin \Omega$. Clearly,

$$\langle v, \Omega \cap v^\perp\rangle \subseteq \overline{\Omega \cup \{v\}} \tag{B.95}$$

and if $\Omega \subset v^\perp$ we have equality in (B.95) and $\overline{\Omega \cup \{v\}}$ is a subspace. Now assume that there is an $x \in \Omega$ such that $[v, x] \neq 0$. Since $\dim(\Omega \cap v^\perp) \geq \dim(\Omega) - 1$, it follows that $\Omega = \langle x, \Omega \cap v^\perp\rangle$. Now we consider two cases:

(I) If $\Omega \cap v^\perp$ is an isotropic subspace, then $H := \langle v, \Omega \cap v^\perp\rangle$ is isotropic as well. Then $\dim(H \cap x^\perp) = \dim(H) - 1$ and $H \cap x^\perp \subset v^\perp$ and we can write the closure of Ω as

$$\overline{\Omega} = \langle x, H \cap x^\perp\rangle \cup \langle v, H \cap x^\perp\rangle \tag{B.96}$$

so $\overline{\Omega}$ has the form of Eq. (8.12).

(II) If $\Omega \cap v^\perp$ is not an isotropic subspace, then there are $a, b \in \Omega \cap v^\perp$ such that $[a, b] \neq 0$. If $a, b \in x^\perp$, then $G(\{a, b, v, x\})$ is a 4-cycle (see Figure B.3a), and we apply Lemma 48. Otherwise, since $\dim(\langle a, b\rangle) = 2$ we may assume without loss of generality that $[a, x] = 0$ and $[b, x] \neq 0$ (Figure B.3b). Then we are precisely in case (2) of the proof of Lemma 49 implying that $\overline{\{a, b, v, x\}} = \langle a, b, v, x\rangle$. Consequently, $\langle v, x\rangle \subset \overline{\Omega \cup \{v\}}$. As x was chosen arbitrarily in $\Omega \setminus v^\perp$, it follows $\overline{\Omega \cup \{v\}} = \langle \Omega, v\rangle$. □

Finally, to prove Lemma 32, we need one more observation, which follows from Lemma 31.

Corollary 17. *If $\Omega \subset E$ is a subspace and contains a, b, c, d such that $G(\{a, b, c, d\})$ is a 4-cycle (Figure B.2a), then $\overline{\Omega \cup \{v\}} = \langle \Omega, v\rangle$ for all $v \in E$.*

Proof of Corollary 17. It suffices to prove that $\Omega \cap v^\perp$ is not isotropic because then case (II) in the proof of Lemma 31 can be applied. Therefore, let $I \subset \Omega$ be an isotropic subspace. As the largest isotropic subspace contained in $\langle a, b, c, d\rangle$ has dimension 2, it follows that

$\dim(\langle a,b,c,d\rangle \cap I) \leq 2$. Since $\langle a,b,c,d\rangle$ is a 4-dimensional subspace contained in Ω and $I \subset \Omega$, it follows $\dim(I) \leq \dim(\Omega) - 2$. However, $\Omega \cap v^\perp$ is a $(\dim(\Omega)-1)$-dimensional subspace, hence, $\Omega \cap v^\perp$ is not isotropic. □

Proof of Lemma 32. Let

$$\Omega = \bigcup_{k=1}^{\xi} \langle a_k, I\rangle \tag{B.97}$$

where I is an isotropic subspace, $a_k \in I^\perp$, and $[a_i, a_j] \neq 0$ for $i \neq j$. We assume that $\xi > 1$, otherwise we are in the situation of Lemma 31. As before, there is nothing to show if $v \in \Omega$, so suppose $v \notin \Omega$. We have two cases.

(I) First, assume $I \subset v^\perp$. Here we have two subcases.

(I.1) If $[v, a_k] \neq 0$ for all $k \in \{1,\dots,\xi\}$, then set $a_{\xi+1} = v$ and we have

$$\overline{\Omega \cup \{v\}} = \bigcup_{k=1}^{\xi+1} \langle a_k, I\rangle. \tag{B.98}$$

(I.2) Otherwise, without loss of generality we can assume that $[a_1, v] = 0$. In this case there is a nontrivial $\tilde{v} \in \langle a_1, v\rangle \subset \overline{\Omega}$ such that $[a_2, \tilde{v}] = 0$, and we can replace v by $\tilde{v}$. Here again we have two subcases.

(I.2.a) If $[a_2, \tilde{v}] = \cdots = [a_\xi, \tilde{v}] = 0$, then $\tilde{I} := \langle \tilde{v}, I\rangle$ is an isotropic subspace and

$$\overline{\Omega \cup \{v\}} = \bigcup_{k=1}^{\xi} \langle a_k, \tilde{I}\rangle. \tag{B.99}$$

(I.2.b) If $[\tilde{v}, a_k] \neq 0$ for some $k \in \{3,\dots,\xi\}$, without loss of generality $[\tilde{v}, a_3] = 0$. Then there is a nontrivial $v' \in \langle a_1, v\rangle = \langle a_1, \tilde{v}\rangle$ such that $[a_3, v'] = 0$. Since $[\tilde{v}, a_2] = 0$ and $v' = \alpha a_1 + \beta\tilde{v} \in \langle a_1, \tilde{v}\rangle$ for any $\alpha, \beta \in \mathbb{Z}_d$, it follows that

$$[v', a_2] = \beta[a_1, a_2] \neq 0. \tag{B.100}$$

As $\tilde{v}, v'$ are contained in the isotropic subspace $\langle a_1, v\rangle$ the orthogonality graph $G(\{\tilde{v}, v', a_2, a_3\})$ is given by Fig. B.4a.

Applying case (2) of Lemma 49 and Lemma 48 shows that

$$\langle \tilde{v}, v', a_2, a_3\rangle = \overline{\{\tilde{v}, v', a_2, a_3\}} \subset \overline{\Omega \cup \{v\}}, \tag{B.101}$$

and therefore, since $\tilde{v}, v', a_2, a_3 \in I^\perp$, the subspace $U' = \langle I, \tilde{v}, v', a_2, a_3\rangle$ is contained in $\overline{\Omega}$.

To conclude this subcase, observe that

$$\overline{\Omega} = \overline{U' \cup \{a_1\} \cup \{a_4\} \cup \cdots \cup \{a_\xi\}}. \tag{B.102}$$

Now consider $\overline{U' \cup \{h_1\}}$ and use the fact that we can construct a 4-cycle in $\langle \tilde{v}, v', h_2, h_3\rangle \subset U'$. Thus, we are able to apply Corollary 17 to get $\overline{U' \cup \{a_1\}} = \langle U', a_1\rangle$ and iteratively we obtain

$$\overline{\Omega} = \overline{U' \cup \{a_1\} \cup \{a_4\} \cdots \cup \{a_\xi\}} = \langle I, a_1, \dots, a_\xi, v\rangle. \tag{B.103}$$

Figure B.4: Orthogonality graphs used in the proof of Lemma 32.

(II) For the second case, suppose that $I \not\subseteq v^\perp$. Let $u \in I$ with $[u, v] \neq 0$. Since $a_1, a_2 \in I^\perp \subset u^\perp$ we can find a nontrivial $\tilde{a}_1 \in \langle u, a_1 \rangle$ and a $\tilde{a}_2 \in \langle u, a_2 \rangle$ such that $[\tilde{a}_1, v] = [\tilde{a}_2, v] = 0$. Since $[\tilde{a}_1, \tilde{a}_2] = [a_1, a_2] \neq 0$, the orthogonality graph $G(u, v, \tilde{a}_1, \tilde{a}_2)$ is given in Figure B.4b and therefore a 4-cycle. Hence, by Lemma 48, it follows that $\overline{\Omega \cup \{v\}}$ contains the subspace $\langle \tilde{a}_1, \tilde{a}_2, u, v \rangle$. Now we can iteratively add the remaining elements of I and the cosets $a_3 + I, \ldots a_\xi + I$ to $\langle \tilde{a}_1, \tilde{a}_2, u, v \rangle$ and apply Corollary 17 to obtain $\overline{\Omega} = \langle I, a_1, \ldots, a_\xi, v \rangle$. □