Hidden variable models and classical simulation algorithms for quantum computation with magic states on qubits

by

Michael Zurel

B.Sc., The University of British Columbia, 2019

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

Master of Science

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Physics)

The University of British Columbia
(Vancouver)

October 2020

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

Hidden variable models and classical simulation algorithms for quantum computation with magic states on qubits

submitted by **Michael Zurel** in partial fulfillment of the requirements for the degree of **Master of Science** in **Physics**.

Examining Committee:

Robert Raussendorf, Physics
Supervisor

William G. Unruh, Physics
Supervisory Committee Member

Abstract

An important problem in quantum computation is to characterize the resources required for a computational speedup over classical computation. Veitch et al. [1] showed that one necessary condition for a computational speedup in the model of quantum computation with magic states is that the discrete Wigner function representing the input state of the quantum circuit must take negative values. The amount of negativity in the discrete Wigner function quantifies the complexity of classical simulation of a quantum computation with simulation being efficient if the Wigner function is nonnegative everywhere. In this sense, negativity of the Wigner quasiprobability representation serves as an indicator of quantumness from a computational perspective. However, this result only holds for systems of qudits where the local Hilbert space dimension is odd.

The first main result discussed in this thesis relates to a discrete Wigner function suitable for describing quantum computation with magic states defined for any local Hilbert space dimension [2]. When the local Hilbert space dimension is odd it subsumes the standard discrete Wigner function. When the local Hilbert space dimension is even, as a result of state-independent proofs of contextuality among multiqubit (or multi-even-dimensional-qudit) Pauli observables, the phase space over which the Wigner function is defined becomes much larger. However, for systems of qubits, the properties required for simulation of quantum computation with magic states remain. This simulation method effectively extends the result described above for odd-dimensional qudits to qubits. Although in the even-dimensional case the phase space is much larger, points in multiqubit phase space can be characterized and classified and some structure can be imparted on the phase space.

The second main result discussed here is a hidden variable model for quantum computation with magic states on qubits [3]. This model is similar in structure to quasiprobability representations like the discrete Wigner function, but unlike those representations the model is capable of representing all elements of any quantum computation—states, operations, and measurements—using only classical probabilities. No negativity is required. This calls into question the role of negativity in quasiprobability representations as an indicator of quantumness for models of quantum computation.

Lay Summary

The unintuitive nature of quantum theory has long puzzled laypeople and scientist working in the field alike. Quantum phenomena like entanglement and contextuality seem to place quantum theory well beyond the grasp of classical intuition. A commonly asked question among quantum theorists is: to what extent can we assign a classical explanation to the predictions of quantum mechanics? While this is an interesting question in and of itself, the discovery of quantum computation provides it with a potentially useful application. Often, the failure of classical explanations heralds an advantage of quantum computing over classical computing. In this sense, characterizing the features of quantum mechanics that evade classical explanation is an equivalent program to characterizing super-classical models of quantum computation. This dual characterization problem is explored in this thesis for a particular model of quantum computation.

Preface

Much of this thesis is the result of significant collaboration. Content in this thesis appears in one publised article, Ref. [2], and in one preprint, Ref. [3]. Ref. [2], entitled "Phase-space-simulation method for quantum computation with magic states on qubits", was coauthored by Robert Raussendorf, Juani Bermejo-Vega, Emily Tyhurst, Cihan Okay, and me, and it is published in Phyical Review A. Ref. [3], entitled "A hidden variable model for quantum computation with magic states on qubits", was coathored by me, Cihan Okay, and Robert Raussendorf.

Content from Ref. [2] is included in Chapter 3 of this thesis. 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 Wigner function presented in Section 3.1 and the state update rules for Clifford gates and Pauli measurements in Section 3.4. Section 3.4 of this thesis was written by Robert Raussendorf and reproduced almost verbatim from Sections V. and VII.C. of Ref. [2]. Efficiency of the classical simulation algorithm in Section 3.5 was proven by Juani Bermejo-Vega. My main contributions to this work were in the characterization of the multiqubit phase space in Sections 3.3 and 3.7, and in the computational methods and numerical results presented in Appendices A and B. Copyrighted material from Ref. [2] is included with permission.

Content from Ref. [3] is the focus of Sections 4.1 and 4.2 of Chapter 4. The initial idea for the hidden variable model defined in Section 4.1 is due to Robert Raussendorf. All coauthors contributed significantly to this work. My contributions include the proof of Theorem 7, the results of Section 4.3, numerical methods and results in Appendices A and B, and the discussion of contextuality and the Pusey-Barrett-Rudolph theorem in Chapter 5.

Table of Contents

List of Tables

List of Figures

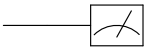
List of Algorithms

List of Abbreviations

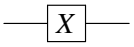
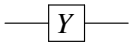
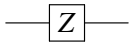
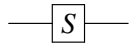
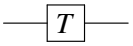
- CNC: Closed under inference and NonContextual
- CSS code: Calderbank-Shor-Steane code
- HVM: Hidden Variable Model
- PBR theorem: Pusey-Barrett-Rudolph theorem
- QCM: Quantum Computation with Magic states
- SW correspondence: Stratonovich-Weyl correspondence

List of Symbols

Quantum circuit components

- Quantum wire
- Classical wire
- Measurement of a qubit in the computational basis 

Single qubit gates

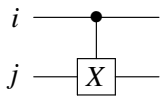
- Pauli-X gate: $X = |0\rangle\langle 1| + |1\rangle\langle 0|$ 
- Pauli-Y gate: $Y = -i|0\rangle\langle 1| + i|1\rangle\langle 0|$ 
- Pauli-Z gate: $Z = |0\rangle\langle 0| - |1\rangle\langle 1|$ 
- Hadamard gate: $H = \frac{1}{\sqrt{2}}(|0\rangle\langle 0| + |0\rangle\langle 1| + |1\rangle\langle 0| - |1\rangle\langle 1|)$ 
- 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

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

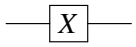
Multiqubit gates

- controlled-X gate: $CX_{i,j} = |0\rangle_i \langle 0|_i I_j + |1\rangle_i \langle 1|_i X_j$

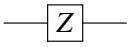


Qu*d*it gates

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

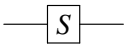


- Generalized Pauli-Z gate: $Z = \sum_{k \in \mathbb{Z}_d} \omega^k |k\rangle \langle k|$, with $\omega = \exp(2\pi i/d)$

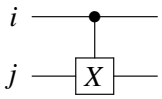


- Hadamard gate: $H = \frac{1}{\sqrt{d}} \sum_{k,\ell \in \mathbb{Z}_d} \omega^{k\ell} |k\rangle \langle \ell|$

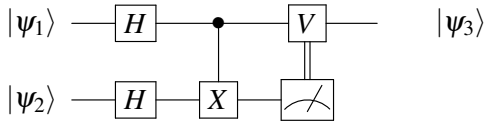
- Phase gate: $S = \sum_{k \in \mathbb{Z}_d} \omega^{k(k+1)2^{-1}} |k\rangle \langle k|$



- controlled-X gate: $CX_{i,j} = \sum_{k \in \mathbb{Z}_d} |k\rangle_i \langle k|_i X_j^k$



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.

Qubit magic states

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

Acknowledgments

I will be forever grateful to my supervisor, Dr. Robert Raussendorf, for his guidance and for his generosity with his time. No part of this project would have been possible without him. I also need to thank the other members of the UBC quantum information group including Cihan Okay, Oleg Kabernik, Paul Herringer, and Arnab Adhikary, as well as our external collaborators including Juani Bermejo-Vega and Emily Tyhurst for interesting discussions and for their contributions to the work presented here.

I would also like to thank my family and friends for their unwavering support over the years.

Chapter 1

Introduction

It is widely believed that quantum computers have the potential to solve some computational problems that are intractable for classical computers. Originally conceived in the early 1980's as a way of simulating systems which cannot be feasibly simulated classically [8, 9], quantum computation has seen an explosion of interest in recent years, in part fuelled by the discovery of quantum algorithms with broader interest beyond simulating physical systems, for example, Shor's algorithm for factoring integers [10] and Grover's algorithm for unstructured search [11]. Both of these algorithms are more efficient than any known classical algorithm for their respective tasks.

Although quantum computation as a field has been around for decades, a key question at the heart of quantum computation remains without a satisfying answer. Namely, what are the essential quantum resources that provide the speedup for quantum computation over classical computation? Characterizing these resources is an important problem with practical quantum computation likely around the corner as recent hardware demonstrations by companies like Google, IBM, D-Wave, and IonQ have shown [12–15]. Understanding the essential resources of quantum computation could drive practical research in hardware and architecture design.

Below we review some properties of quantum mechanics that have been proposed as quintessential resources for quantum computation. At the same time, we use these candidates to illustrate some aspects of the question posed above and to get a better understanding of its purpose.

An obvious candidate that comes to mind for the source of the advantage for quantum computers is the exponential blowup of the dimension of the state space with the number of qubits. Specifying an n-qubit quantum state requires specifying $O(2^n)$ complex amplitudes (Hilbert space dimension 2^n). Tracking unitary operations generally requires matrix multiplication in a 2^n-dimensional space and calculating probabilities for measurement outcomes involves computing inner products in this space. These tasks have algorithms which run in polynomial time with respect to the dimension but the dimension of the Hilbert space grows exponentially with the number n of qubits.

Let's pause here and see what we've just done. We have proposed an algorithm for the classical simulation of quantum mechanics, the most obvious one. We learn from it that the largeness of Hilbert space is a *necessary* precondition for a quantum computational speedup. That is, if the Hilbert space dimension is small, then our simulation algorithm is efficient. But is largeness of Hilbert space dimension also *sufficient* for a quantum speedup? Before answering this question, we emphasize the distinction. A necessary property is one that precludes a quantum speedup if absent, a sufficient property is one that enables speedup if present.

The ultimate goal of the "quantum resource" research programme is to identify a quantum property that is both necessary and sufficient for quantum speedup. This has to date not been achieved.

So then, how does largeness of Hilbert space dimension fail to be a sufficient condition for quantum speedup? For example, consider a quantum computer with access to n qubits but which can only prepare product states and can only implement local unitary operations and perform local measurements. In this restricted setting, specifying a state requires only $O(2n)$ amplitudes (two amplitudes for each qubit) and the dynamics (unitary operations and measurements) can be tracked for each qubit separately. Therefore, this quantum computer could be simulated efficiently on a classical computer, and thus, with local states, operations, and measurements only, no quantum computational advantage is possible.

This suggests that perhaps we should focus on entanglement as the quintessential quantum resource. Entanglement has been proposed as a quantifiable resource for quantum information tasks [16]. By the previous example, entanglement is necessary for a quantum advantage, at least for systems in pure states, and it

can be shown to be necessary and sufficient for some quantum communication tasks [17, 18] which make use of nonclassical correlations like in Bell's inequality experiments. One could ask if entanglement is a necessary and sufficient condition for a quantum computational advantage. It turns out, the answer to this question is no, entanglement is necessary but not sufficient.

There is a closed subtheory of quantum mechanics (consisting of a set of quantum states, unitary operations, and measurements) called the stabilizer subtheory. This subtheory contains plenty of entanglement, for example, Bell states are stabilizer states,[1] but by the Gottesman-Knill theorem [19], this subtheory can also be simulated efficiently on a classical computer. Therefore, entanglement is not sufficient for a quantum computational advantage. Quantum resources beyond the scope of the stabilizer subtheory are necessary for a quantum computer to exhibit a computational speedup over classical computation. By construction, nonstabilizerness is necessary, but is it also sufficient? As we'll see below, an answer to this question can be provided by another classical simulation algorithm for quantum computation based on Wigner functions [1].

But first, we take a detour and consider another candidate: contextuality. Contextuality is a feature of quantum theory which distinguishes it from classical theory in which, if you assume a measurement reveals a predetermined measurement outcome rather than producing a measurement outcome, then the outcome can depend on the context in which an observable is measured, in addition to the physical state of the system. That is, measurements can depend on which other observables are measured simultaneously.

The concept of contextuality comes from the field of quantum foundations which seeks to understand the most fundamental aspects of quantum theory and similar physical theories. E.g. to determine characteristics that distinguish quantum theory from classical theory, to characterize basic mathematical structures of quantum theory, and to consider possible extensions of quantum theory.

A common question in this field is to ask to what extent classical explanations can be applied to quantum phenomena. One form such classical explanations take is hidden variable models which explain quantum phenomena in terms of unob-

[1] The stabilizer subtheory also contains entangling unitary operations and measurements.

served (hidden) variables which align more closely with classical intuitions. This line of research has a long history going back almost to the discovery of quantum theory itself [20, 21]. It was originally motivated by the dissatisfaction of some of the founders of quantum theory with certain unintuitive aspects of the theory. The goal was to create a classical theory that could reproduce the predictions of quantum theory. This was ultimately shown to be impossible because of quantum features like nonlocality [22] and contextuality [23, 24]. As a result, these features are used as indicators of quantumness in physical systems. That is, systems exhibiting these festures could be labeled genuinely *quantum* whereas systems without these features might behave more classically. Contextuality demarks the boundary between the realms of classical and quantum physics.

Interestingly, contextuality has also been applied as an indicator of quantumness from a computational perspective: contextuality has been shown to be necessary for certain quantum computational tasks [25, 26]. But perhaps this is not so surprising as a quantum computational advantage can itself demark a classical-to-quantum transition.

Another proposed quantum computational resource is negativity in quasiprobability representations like the Wigner function [27]. The Wigner function derives from an attempt to adapt the phase-space formulation of classical statistical mechanics, which is based on probability distributions over the position-momentum phase space, to quantum mechanics. The Wigner function has many of the same properties as a probability distribution, but unlike a probability distribution it can take negative values, making it a *quasi*probability distribution.

Negativity in the Wigner function is the key feature that distinguishes it from classical statistical mechanics and that allows it to represent quantum mechanics. This motivates the use of negativity in quasiprobability distributions as quantum indicators. In fact, negativity of the Wigner quasiprobability function [27] is an indicator of quantumness that is commonly used in the fields of quantum optics and quantum statistical mechanics [28]. This framework has been adapted to the setting of finite-dimensional quantum mechanics by the introduction of so-called discrete Wigner functions for the purpose of studying quantum computation [1, 29–33].

We focus on a particular model of quantum computation called quantum com-

putation with magic states (QCM) [34]. In this model, the allowed operations are restricted to the unitary operations and the measurements of the stabilizer subtheory. As noted above, these operations are not sufficient by themselves to perform quantum computations exhibiting a computational speedup over classical computation. The power of quantum computation is restored in this model through the inclusion of extra input quantum states called *magic* states. Then the task of characterizing quantum computational resources can be reduced to characterizing the set of quantum states which are useful for quantum computations in this model.

For quantum computation with magic states on odd-dimensional qudits—d level quantum systems with d odd—any quantum circuit consisting of a sequence of stabilizer operations on an input state with nonnegative discrete Wigner function can be simulated efficiently classically [1]. Since there are states which are not stabilizer states but which have nonnegative discrete Wigner function, we can conclude from this that arbitrary nonstabilizer resources are also not sufficient for a quantum computational advantage. Wigner function negativity is also required.

It has been shown that, for the setting described above, nonnegativity of the Wigner function implies the existence of a noncontextual hidden variable model [35, 36]. Therefore, three of the quantum indicators listed so far—contextuality, Wigner function negativity, and a quantum computational speedup—agree. That is, the boundary at which the Wigner function becomes negative is the same boundary where the possibility of a noncontextual hidden variable model is eliminated, and where a quantum computational speedup becomes possible [1, 26, 35, 36]. This is a satisfying state of affairs. However, as mentioned above, it applies only to quantum computation on systems with odd local Hilbert space dimension.

With quantum computation on even-dimensional systems, including the most interesting case—quantum computation on qubits, the interaction between contextuality, negativity, and a computational speedup is more complicated. First, there is no longer a unique, natural way of defining a discrete Wigner function. No-go theorems prevent a simple extension of the odd-dimensional Wigner funtion from being applied to even dimensions, if the Wigner function is to be useful for describing quantum computation [37–39]. Furthermore, the question of contextuality is muddied by the existence of a generic (state-independent) contextuality that doesn't exist among the relevant observables on odd-dimensional qudits.

In this thesis, we discuss an extension of the odd-dimensional Wigner function that is suitable for describing quantum computation with magic states on systems with any local Hilbert space dimension [2]. This generalized Wigner function provides a classical simulation algorithm for quantum computation on qubits which shows that for QCM on qubits, as well as on odd-dimensional qudits, negativity in a quasiprobability representation is a necessary precondition for a quantum computational speedup. This realigns two of the quantum indicators above—negativity in a quasiprobability representation and a quantum computational speedup.

Contextuality plays a central role in the definition of the generalized Wigner function as points in the generalized phase space are associated with certain noncontextual sets of Pauli observables—the observables of the stabilizer subtheory. As a result of state-independent proofs of contextuality among the multiqubit Pauli observables, the phase space over which the multiqubit Wigner function is defined splinters and becomes much larger than might be expected for a generalization of the odd-dimensional phase space. This enlarging of the phase space is a necessary feature for a discrete Wigner function to be able to simulate quantum computation with magic states on even dimensional qudits [39]. Although the phase space is much larger, it maintains a certain amount of structure and points in this phase space can be characterized, classified, and counted. The characterization of these multiqubit phase space points is closely related to the characterization of noncontextual sets of multiqubit Pauli observables.

Although points in the generalized phase space are defined based on sets of Pauli observables which are explicitly noncontextual, the interpretation of the hidden variable model implied by the quasiprobability representation for qubits as contextual or noncontextual is more difficult. The necessity of some states to make use of phase points corresponding to distinct sets of noncontextual Pauli observables in their representation leaves open the possibility of contextuality of the hidden variable model. This is not suprising as the model is capable of simuating explicitly contextual systems.

We also present a hidden variable model for quantum computation with magic states on qubits [3]. This model is structurally similar to quasiprobability representations like the discrete Wigner function, but unlike those representations it is capable of representing all states, operations, and measurements relevant to quantum

computation with magic states using only classical probabilities. States are represented by probability distributions over a finite set and the dynamical ingredients of computation are represented by transition probabilities between the elements of that set.

This representation leads to a classical simulation algorithm for quantum computation with magic states that looks like a classical Markov decision process on a finite state space. Since this is a universal model of quantum computation, this means any quantum computation (and therefore all finite-dimensional quantum mechanics) can be described using only classical probabilities—no negativity is required. Note, however, that this classical simulation algorithm is not necessarily efficient for simulating arbitrary quantum computations. In fact, if a quantum computational advantage over classical computation exists then clearly this classical simulation algorithm cannot be efficient in general. This calls into question the role of negativity as an indicator of quantumness for models of quantum computation by showing that even negativity is not sufficient for a quantum computational speedup. The presence or absence of negativity can depend on the choice of quasiprobability representation.

This thesis is structured as follows: in the next chapter, Chapter 2, we review some preliminary material that will be necessary to effectively present and frame the results of the subsequent chapters. This includes quantum computation with magic states—the computational model of interest, the discrete Wigner function and associated classical simulation algorithm for quantum computation on odd-dimensional systems, and contextuality and nonclassicality of hidden variable models for finite-dimensional quantum mechanics. The generalized phase space and Wigner function and the associated simulation method for quantum computation with magic states on qubits, as well as the characterization of the multiqubit phase space are presented in Chapter 3. The hidden variable model for quantum computation with magic states on qubits is presented in Chapter 4. We close with a discussion of the implications of these results in Chapter 5. Computational methods are presented in Appendix A and some numerical results in Appendix B.

Chapter 2

Background

In this thesis, we consider methods for describing quantum computation on qudits—d level quantum systems. The states of each qudit are associated with the Hilbert space $\mathbb{C}^d$, with the computational basis an orthonormal basis $\{|s\rangle \mid s \in \mathbb{Z}_d\}$. For a system of n qudits, the full Hilbert space is $\left(\mathbb{C}^d\right)^{\otimes n}$. Later we will specialize to qubits ($d = 2$). In the quantum circuit model, computation proceeds by applying a sequence of unitary gates, elements of $U(d^n)$, to a state initialized to $|0\rangle^{\otimes n}$, followed by a measurement of each qudit in the computational basis.

We often consider models of quantum computation in which the allowed operations are restricted to a strict 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, if a qubit based quantum computer can only implement the operations in the set $\{H, T, CX\}$, it can still perform arbitrary quantum computations [4, §4.5.3].[1] That is, any unitary operation in $U(2^n)$ can be approximated to arbitrarily low error by sequential application of the gates in this set.[2] Gate sets with this property are called *universal* for quantum computation.

Another reason for restricting the allowed operations is fault-tolerance. Quantum computers are famously prone by errors caused by noise and decoherence.

[1] These gates are defined in the List of Symbols on page xvii.

[2] Formally, the topological closure of the group generated by the these operations acting on n qudits with respect to the metric induced by the operator norm is $U(2^n)$.

A foundational result of quantum computation is the quantum threshold theorem which states that if the physical error rate for each operation in a model of quantum computation can be made sufficiently small then quantum error correction techniques can be applied to make the probability of an error in computation arbitrarily small. These quantum error correcting codes require the allowed operations to have certain properties in order for the code to be able to correct errors faster than they can spread.

A large family of quantum error correcting codes similar to classical linear block codes together with a language for describing fault-tolerant quantum computation more generally is the stabilizer formalism [40]. The stabilizer formalism describes a closed subtheory of quantum mechanics consisting of a set of allowed states, unitary operations, and measurements which allow for fault-tolerant quantum computation. Although we are not concerned with the details of quantum error correction here, the stabilizer formalism provides a useful mathematical framework for characterizing superclassical models of quantum computation.

The allowed operations of the stabilizer formalism are not by themselves universal for quantum computation. In fact, the operations of the stabilizer subtheory applied to stabilizer states can be simulated efficiently on a classical computer [19, 41], therefore, there is no advantage to stabilizer quantum computation over classical computation. In order to restore universality, extra *resources* (states and/or operations) beyond the scope of the stabilizer formalism must be used. This is the basis of the resource theory of stabilizer quantum computation [42]. The characterization of the nonstabilizer resources which restore quantum computational universality to the stabilizer subtheory is a significant project, to which some results of this thesis are a contribution.

The stabilizer formalism [40] and the associated resource theory [42] will be useful in future chapters so we review it in the next section.

2.1 The stabilizer formalism

The stabilizer formalism is based on the Pauli group. First, the single-qubit Pauli group is the group generated by the Pauli matrices, $\langle X, Y, Z\rangle = \langle X, Z, iI\rangle$. The Pauli operators can be generalized to d-dimensional qudits as follows. For any

$d \in \mathbb{N}$, define the generalized Pauli operators: $X = \sum_{j=1}^{d} |j+1 \mod d\rangle\langle j|$ and $Z = \sum_{j=1}^{d} \omega^j |j\rangle\langle j|$ with $\omega = \exp(2\pi i/d)$ a primitive d^{th} root of unity.[3] Then the single-qudit generalized Pauli group can be defined as the group $\langle X, Z, \sqrt{\omega} I\rangle$. The n-qudit Pauli group, denoted $\mathscr{P}$, is the group of Pauli operators acting on n qudits. It is formed from tensor products of the elements of the one-qudit Pauli group.

The n-qudit Pauli group modulo phases, $\mathscr{P}/\langle\sqrt{\omega} I\rangle$, is isomorphic to the additive group of the free module $\mathbb{Z}_d^{2n}$. Fixing a phase convention, we can index the Pauli operators by the elements of the module:[4] for any $a = (a_Z, a_X) \in \mathbb{Z}_d^n \times \mathbb{Z}_d^n =: E$ define,

$$T_a = e^{i\phi(a)} Z(a_Z) X(a_X) \tag{2.1}$$

where $X(a_X) = \bigotimes_{k=1}^{n} X^{a_X[k]}$, and $Z(a_Z) = \bigotimes_{k=1}^{n} Z^{a_Z[k]}$. We choose the phase function $\phi : E \to \mathbb{R}$ so that the operators satisfy the constraint $(T_a)^d = I$. This constraint forces the eigenvalues of the operators to have the form ω^k with $k \in \mathbb{N}$.

The symplectic bilinear form on E, $[\cdot,\cdot] : E \times E \to \mathbb{Z}_d$, defined by

$$[a, b] := \langle a_Z | b_X \rangle - \langle a_X | b_Z \rangle \tag{2.2}$$

tracks the commutator of the Pauli operators in the sense

$$[T_a, T_b] := T_a T_b T_a^{-1} T_b^{-1} = \omega^{[a,b]} I \quad \forall a, b \in E. \tag{2.3}$$

In particular, operators T_a and T_b commute if and only if $[a,b] = 0$. Because of this correspondence between the commutator of the Pauli operators and the symplectic form on E, we will say $a, b \in E$ commute when $[a,b] = 0$.

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 $S = \langle X_1 X_2, Z_1 Z_2\rangle$ stabilizes the Bell state $|\beta_{00}\rangle = (|00\rangle +$

[3] 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. $\tilde{Z} = \sum_{j=1}^{d} j |j\rangle\langle j|$.

[4] 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 (e.g. see Ref. [32]), but for our purposes it is 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 $\mathbb{Z}_d$.

$|11\rangle)/\sqrt{2}$ since $A|\beta_{00}\rangle = |\beta_{00}\rangle$ for each $A \in S$. On the level of the index set E, a stabilizer group is specified by an isotropic subspace,[5] $\tilde{I} \subset E$, together with a function $\gamma : \tilde{I} \to \mathbb{Z}_d$ which specifies the phases of the stabilizer group elements. The stabilizer group corresponding to the pair $(\tilde{I}, \gamma)$ is $S = \left\{ \omega^{-\gamma(a)} T_a \mid a \in \tilde{I} \right\}$.

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 $\omega^{-\gamma(0)} T_{\mathbf{0}} = I$. We can define a function β that keeps track of how commuting Pauli operators compose by

$$\omega^{\beta(a,b)} T_a T_b = T_{a+b}, \quad \forall a, b \in E \text{ with } [a,b] = 0. \tag{2.4}$$

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. (2.4) gives the constraint that γ must satisfy

$$\gamma(a) + \gamma(b) - \gamma(a+b) = -\beta(a,b), \quad \forall a, b \in \tilde{I}. \tag{2.5}$$

A useful identity of the function β comes from the associativity of operator multiplication [43]. By expanding each side of the equation $T_a(T_b T_c) = (T_a T_b) T_c$ with Eq. (2.4), we get $\omega^{-\beta(a,b+c)-\beta(b,c)} T_{a+b+c} = \omega^{-\beta(a+b,c)-\beta(a,b)} T_{a+b+c}$ and so β must satisfy

$$\beta(a,b+c) + \beta(b,c) = \beta(a+b,c) + \beta(a,b) \tag{2.6}$$

for any pairwise commuting $a, b, c \in E$. The functions γ and β will be given an interpretation in section 2.7.2.

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*. 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.

The unitary operations that can be easily implemented fault-tolerantly on encoded qudits in the stabilizer formalism form the *Clifford group*. The Clifford

[5] An isotropic subspace of a symplectic space is a subspace on which the symplectic form vanishes: $[a,b] = 0$, $\forall a, b \in \tilde{I} \subset E$.

group is defined as the group of unitary operators that map Pauli operators to Pauli operators under conjugation. As global phases are inconsequential in quantum computation we can take the n-qudit Clifford group to be the quotient $\mathscr{C}\ell := \mathscr{N}(\mathscr{P})/U(1)$. That is, the normalizer of the Pauli group in the unitary group, $U(d^n)$, modulo phases.

Unfortunately, the Clifford group is not universal for quantum computation—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 [4, §10.5.2]. For qudits, the group is generated by the qudit analogues of these gates [44, 45].[6] In fact, any quantum circuit consisting of only Clifford unitaries and Pauli measurements can be efficiently simulated classically. This is the result of the so-called Gottesman-Knill theorem [19] (also see Ref. [41]). A question we can now ask is: what nonstabilizer resources can we add to the stabilizer formalism to restore universality? As alluded to above, the capability to perform T gates would suffice, but T gates cannot be directly implemented fault-tolerantly. An alternative approach to restoring universality is *quantum computation with magic states*. This approach is discussed in the next section.

2.2 Quantum computation with magic states

Quantum computation with magic states (QCM) [34] is a *universal* model of quantum computation in which the allowed operations are restricted to Clifford gates and Pauli measurements. These operations are not by themselves universal for quantum computation. Furthermore, by the Gottesman-Knill theorem, any circuit consisting of only these operations applied to a stabilizer input state can be simulated efficiently on a classical computer and so these operations alone do not allow for a quantum computational speedup over classical computation. Universality is restored in QCM through the introduction of nonstabilizer ancillary states.[7]

By allowing nonstabilizer input states as well as adaptivity—conditioning gates and measurements on the outcomes of prior measurements—we can effectively

[6] These gates are defined in the List of Symbols on page xvii.

[7] Often tensor products of one-qubit states such as $|H\rangle\langle H| := \frac{1}{2}I + \frac{1}{2\sqrt{2}}(X+Y)$ or $|T\rangle\langle T| := \frac{1}{2}I + \frac{1}{2\sqrt{3}}(X+Y+Z)$ are used. These are sufficient for restoring universality.

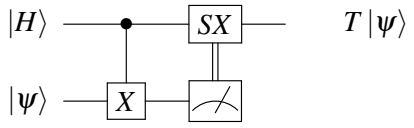


Figure 2.1: An implementation of the T gate in the framework of quantum computation with magic states. The components of this circuit are defined in the List of Symbols on page xvii. Only Clifford operations, Pauli measurements, and classical conditioning are used but with the inclusion of a nonstabilizer input state, $|H\rangle = T|+\rangle$, the effect of the circuit is a T gate applied to the second qubit. For implementations of other non-Clifford gates using different ancillas see [4, §10.6.2].

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 2.1. Since the Clifford+T gate set is universal [4, §4.5.3], this shows that the QCM model is universal. We will generally refer to circuits consisting of Clifford gates and Pauli measurements, possibly with adaptivity, as QCM circuits. States which promote the model to universality are called *magic states* or *resource states*.

In the QCM model, the main difficulty of universal fault-tolerant quantum computation is shifted from implementing gates to preparing magic input states with sufficient fidelity. Protocols for preparing certain magic states based on the stabilizer formalism exist, called magic state distillation [34] (also see Ref. [4, §10.6.3]). This is one of the leading candidates for a model of universal fault-tolerant quantum computation.

Since the operations used in the QCM model 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. This refines the question posed above: instead of broadly considering which nonstabilizer resources are required for universal quantum computation, we can focus on the states and ask: which quantum states promote QCM to universality? One approach to studying this question is through *quasiprobability representations* like the discrete Wigner function.

2.3 Wigner functions

The Wigner function is the basis of an alternative formulation of quantum mechanics discovered by Eugene Wigner in 1932 [27]. Instead of representing a state by a wave function, the state is represented by a function over the position-momentum phase space. If the position-space wavefunction of the system is $\psi(x)$, the Wigner function is defined as

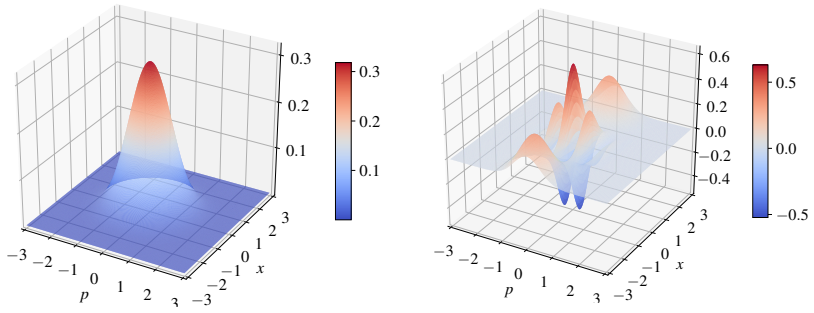
$$W(x,p) = \frac{1}{\pi\hbar}\int_{-\infty}^{\infty} dy\ \psi^*(x+y)\psi(x-y)e^{2ipy/\hbar}.$$

The Wigner function has several properties that make it structurally similar to the phase-space formulation of classical statistical mechanics which is based on probability distributions over phase space. 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:

$$\int dp\, W(x,p) = |\psi(x)|^2,$$
$$\int dx\, W(x,p) = |\varphi(p)|^2.$$

These properties make the Wigner function analogous to a joint probability distribution over the position-momentum phase space. The crucial difference that allows it to represent quantum mechanics is that, unlike a probability distribution, it can take negative values. Such functions are called *quasiprobability distributions*. Examples of the Wigner function for two particular states are shown in Fig. 2.2.

Negativity in the Wigner function is necessary for it to be able to represent quantum mechanics. It is also what distinguishes it from classical statistical mechanics. For this reason, in the field of quantum optics where the Wigner function formulation is commonly used, negativity of the Wigner function is considered an indicator of true *quantumness* [28]. States with nonnegative Wigner function are considered more classical than those for which the Wigner function takes negative values.



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

Figure 2.2: 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.

2.3.1 The Stratonovich-Weyl correspondence

In quantum optics and quantum statistical mechanics, several quasiprobability representations of quantum states have been studied, for example, the Wigner function [27], the Sudarshan-Glauber P representation [46, 47], and the Husimi Q representation [48]. They are all related by the Stratonovich-Weyl (SW) correspondence, a set of criteria that reasonable quasiprobability representations over generalized phase spaces should satisfy [49] (also see [50]).

Quasiprobabililty representations in the SW class have the form $F_A^{(s)} : X \to \mathbb{C}$ where $F_A^{(s)}$ is the quasiprobability distribution representing the operator A, X is the phase space over which the distribution is defined, and $s \in [-1, 1]$ is a real parameter. In the standard quantum optics formulation, X is the position-momentum phase space, and $s = 0$ for the Wigner function, $s = -1$ for the P function, and $s = 1$ for the Q function [50].

The SW criteria are as follows:

(SW0) *Linearity*: the map $A \rightarrowtail F_A^{(s)}$ is one-to-one and linear,

(SW1) *Reality:*[8]

$$F^{(s)}_{A^\dagger}(x) = \left(F^{(s)}_A(x)\right)^* \quad \forall x \in X,$$

(SW2) *Standardization*:

$$\int_X d\mu(x) F^{(s)}_A(x) = \mathrm{Tr}(A), \tag{2.7}$$

(SW3) *Covariance*:

$$F^{(s)}_{g\cdot A}(x) = F^{(s)}_A\left(g^{-1}\cdot x\right) \quad \forall x \in X\ \forall g \in G \tag{2.8}$$

where G is the dynamical symmetry group.

(SW4) *Traciality*:[9]

$$\int_X d\mu(x) F^{(s)}_A(x) F^{(-s)}_B(x) = \mathrm{Tr}(AB). \tag{2.9}$$

Quasiprobability representations satisfying these criteria are structurally similar and thus closely related to the original Wigner function.

2.3.2 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 [29]. There is even an essay by Richard Feynman from 1987 discussing the interpretation of negative probabilities in physics which features a discrete Wigner function for one-qubit states [51]. In 2004, a large family of discrete phase spaces and Wigner functions was defined by Gibbons, Hoffman, and Wootters based on finite fields [52]. 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.

[8] If the operator A is Hermitian, as is usually the case for our purposes, this condition ensures the quasiprobability function is real.

[9] In this sense, $F^{(-s)}$ is a dual quasiprobability representation to $F^{(s)}$.

2.4 The odd-dimensional discrete Wigner function

Gross' discrete Wigner function [32, 33, 53] is defined for systems of d-level qudits where d is odd. To start, we fix the phase convention for the generalized Pauli operators: $e^{i\phi(a)} = \omega^{-\langle a_Z|a_X\rangle 2^{-1}}$ in Eq. (2.1). Here 2^{-1} is the multiplicative inverse of 2 in $\mathbb{Z}_d$ and the inner product is the standard inner product of $\mathbb{Z}_d^n$. With this choice of phase, a direct calculation shows that 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 $a,b \in E$.

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.[10] The phase point operator corresponding to the zero vector, $0 \in V$, is

$$A_0 := \frac{1}{d^n} \sum_{v \in E} T_v. \tag{2.10}$$

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

$$A_u := T_u A_0 T_u^\dagger \quad \forall u \in V. \tag{2.11}$$

Using the commutation relation Eq. (2.3), we can rewrite the phase point operators in a more convenient form: $A_u = d^{-n} \sum_{v \in E} \omega^{[u,v]} T_v$.

These phase point operators form an orthogonal operator basis for the space of Hilbert-Schmidt operators on n-qudit Hilbert space: $\mathrm{Tr}(A_u^\dagger A_v) = d^n \delta_{u,v}$. Therefore, for operators A and B expanded in the phase 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 $\mathrm{Tr}(A^\dagger B) = d^n \sum_{u \in V} W_A(u)^* W_B(u)$. Since the phase 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

$$\rho = \sum_{u \in V} W_\rho(u) A_u. \tag{2.12}$$

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 point operators have

[10] 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.

unit trace, taking a trace of Eq. (2.12) 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 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

$$\mathrm{Tr}(E_k \rho) = \sum_{u \in V} W_\rho(u) \mathrm{Tr}(E_k A_u) = \sum_{u \in V} W_\rho(u) W_{E_k}(u). \tag{2.13}$$

The SW correspondence

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

$$B = \sum_{u \in V} W_B(u) A_u. \tag{2.14}$$

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 point operators are Hermitian, taking the Hermitian adjoint of Eq. (2.14) 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 point operators have unit trace, taking a trace of Eq. (2.14) we obtain standardization: $\mathrm{Tr}(B) = \sum_{u \in V} W_B(u)$. (SW3) The relevant dynamical symmetry group in this case is the Clifford group. Covariance of the discrete Wigner function with respect to the Clifford group is discussed in the next section. (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.

2.5 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 [1, 32, 33, 53]. 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[11] [32, 33] (also see [30]).

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 2.3 shows two examples of the discrete Wigner function for one-qutrit ($d = 3$) states.

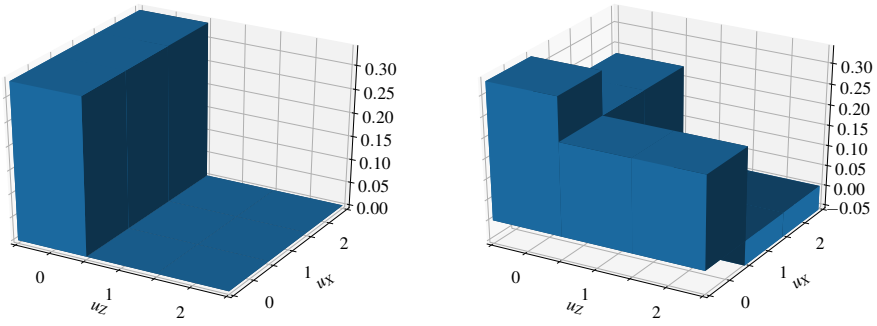
Now consider the dynamical elements of quantum computation with magic states: Clifford unitaries and Pauli measurements.

2.5.1 Clifford unitaries

Clifford unitaries map phase point operators to phase point operators under conjugation [32]. In particular, for systems of odd-dimensional qudits, the Clifford group splits as $\mathscr{C}\ell \simeq Sp(2n,\mathbb{Z}_d) \ltimes \mathbb{Z}_d^{2n}$ where $Sp(2n,\mathbb{Z}_d)$ is the symplectic group[12]

[11] This is the analogue for finite-dimensional quantum mechanics of Hudson's theorem, which applies to the original Wigner function defined in Section 2.3 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 $\psi(x) = \exp\left(-ax^2 + bx + c\right)$ [54].

[12] The symplectic group $Sp(2n,\mathbb{Z}_d)$ is the group of symplectic transformations—linear maps on the symplectic space $\mathbb{Z}_d^{2n}$ which preserve the symplectic form.



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

Figure 2.3: 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. [5]. 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.

on $\mathbb{Z}_d^{2n}$. Then Clifford unitaries can be labeled by $M \in \mathrm{Sp}(2n, \mathbb{Z}_d)$ and $b \in \mathbb{Z}_d^{2n}$ and the action of the Clifford group on phase point operators is

$$U_{M,b} A_u U_{M,b}^\dagger = A_{Mu+b}. \tag{2.15}$$

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

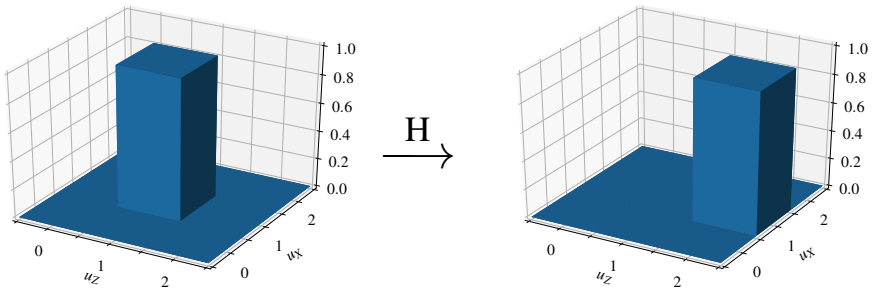
$$W_{U_{M,b}\rho U_{M,b}^\dagger}(u) = W_\rho \left(U_{M,b}^{-1} \cdot u \right) = W_\rho \left(M^{-1}(u-b) \right). \tag{2.16}$$

An example of a mapping of a qutrit phase point operator under a Clifford unitary is given in Figure 2.4.

For a more complete discussion of the interaction between the Clifford group and the discrete Wigner function see Ref. [32].

2.5.2 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, $a \in$



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

Figure 2.4: Under conjugation by the qutrit Clifford unitary H, the phase point operator $A_{(1,1)}$ maps to the phase point operator $A_{(2,1)}$. In general, phase point operators map deterministically to phase point operators under conjugation by Clifford unitaries.

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 [36]. 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 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 1 *Let A_u, $u \in V$ be a phase point operator and $\Pi_{a,s}$ be an eigenprojector corresponding to a Pauli measurement T_a, $a \in E$ giving outcome ω^s, $s \in \mathbb{Z}_d$. Then $Tr(\Pi_{a,s} A_u) = 1$ if $s = [a,u]$ and $Tr(\Pi_{a,s} A_u) = 0$ otherwise. Further, let $S_a := \{v \in E \mid [a,v] = 0\}$ and $\Gamma_{a,u} := \{w \in E \mid [w,v] = [u,v] \; \forall v \in S_a\}$. Then*

$$\Pi_{a,s} A_u \Pi_{a,s} = \frac{\delta_{s,[a,u]}}{|\Gamma_{a,u}|} \sum_{w \in \Gamma_{a,u}} A_w. \tag{2.17}$$

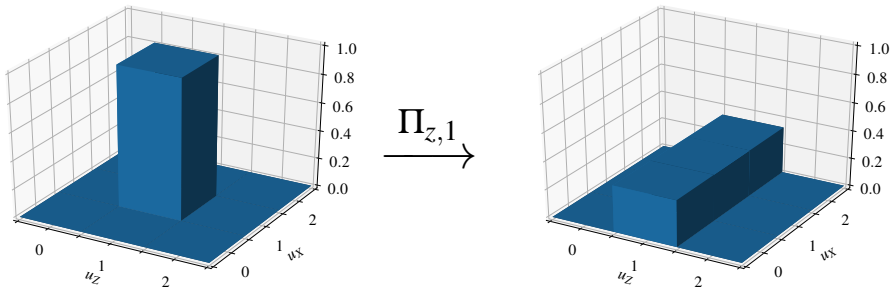
Proof of Lemma 1. By a direct calculation we obtain

$$\begin{aligned}
\Pi_{a,s}A_u\Pi_{a,s} &= \frac{1}{d^{n+2}}\sum_{v\in E}\sum_{k,\ell\in\mathbb{Z}_d}\omega^{-(k+\ell)s+[u,v]}T_a^kT_vT_a^\ell \\
&= \frac{1}{d^{n+2}}\sum_{v\in E}\sum_{k,\ell\in\mathbb{Z}_d}\omega^{-(k+\ell)s+[u,v]+(k-\ell)[a,v]2^{-1}}T_{v+(k+\ell)a} \\
&= \frac{1}{d^{n+2}}\sum_{v\in E}\left(\sum_{k,\ell\in\mathbb{Z}_d}\omega^{-(k+\ell)s+[u,v]-(k+\ell)[u,a]+(k-\ell)[a,v]2^{-1}}\right)T_v \\
&= \frac{1}{d^{n+2}}\sum_{v\in E}\delta_{[a,v],0}\left(\sum_{k,\ell\in\mathbb{Z}_d}\omega^{-(k+\ell)s+[u,v]-(k+\ell)[u,a]}\right)T_v \\
&= \frac{1}{d^{n+1}}\sum_{v\in E}\delta_{[a,v],0}\left(\sum_{k\in\mathbb{Z}_d}\omega^{[u,v]-k(s+[u,a])}\right)T_v \\
&= \frac{1}{d^{n}}\sum_{v\in E}\delta_{[a,v],0}\delta_{s,[a,u]}\omega^{[u,v]}T_v \qquad (2.18)
\end{aligned}$$

Here in the first line we use the definitions of $\Pi_{a,s}$ and A_u, in the second line we use the relation $T_aT_b = \omega^{[a,b]2^{-1}}T_{a+b}$, 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. (2.18), we have

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



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

Figure 2.5: A Pauli Z measurement on the qutrit phase 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 point operator $A_{(1,1)}$ maps to the probabilistic combination of the phase point operators $A_{(1,0)}$, $A_{(1,1)}$, and $A_{(1,2)}$ with equal weight. In general, phase point operators map to probabilistic combinations of phase point operators under Pauli measurements.

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 point operators are associated with deterministic assignments of measurement outcomes, and under Pauli measurements phase point operators map to probabilistic combinations of phase point operators. An example of a Pauli measurement on a qutrit phase point operator is given in Figure 2.5.

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

Lemma 2 *The set of positively represented states is closed under Clifford unitaries and Pauli measurements.*

Proof of Lemma 2. Suppose a state ρ is positively represented. That is, ρ can be expanded in the phase point operator basis as in Eq. (2.12) 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. (2.16)—Clifford covariance of the discrete Wigner function. Since the Wigner function $W_\rho(u)$ is nonnegative, under a Clifford unitary $U_{M,b} \in \mathscr{C}\ell$, the Wigner function $W_{U_{M,b}\rho U_{M,b}^\dagger}(u) = W_\rho(M^{-1}(u-b))$ 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

$$\begin{aligned}
W_{\frac{\Pi_{a,s}\rho\Pi_{a,s}}{\mathrm{Tr}(\Pi_{a,s}\rho)}}(u) &= d^{-n}\mathrm{Tr}\left(\frac{\Pi_{a,s}\rho\Pi_{a,s}}{\mathrm{Tr}(\Pi_{a,s}\rho)}A_u\right) \\
&= \frac{d^{-n}}{\mathrm{Tr}(\Pi_{a,s}\rho)}\mathrm{Tr}\left(\rho\Pi_{a,s}A_u\Pi_{a,s}\right) \\
&= \frac{d^{-n}}{\mathrm{Tr}(\Pi_{a,s}\rho)}\mathrm{Tr}\left(\rho\frac{\mathrm{Tr}(\Pi_{a,s}A_u)}{|\Gamma_{a,u}|}\sum_{w\in\Gamma_{a,u}}A_w\right) \\
&= \frac{d^{-n}\mathrm{Tr}(\Pi_{a,s}A_u)}{\mathrm{Tr}(\Pi_{a,s}\rho)\cdot|\Gamma_{a,u}|}\sum_{w\in\Gamma_{a,u}}\mathrm{Tr}(\rho A_w) \\
&= \frac{\delta_{s,[a,u]}}{\mathrm{Tr}(\Pi_{a,s}\rho)\cdot|\Gamma_{a,u}|}\sum_{w\in\Gamma_{a,u}}W_\rho(w). \qquad (2.19)
\end{aligned}$$

Here, in the second line we use linearity and the cyclic property of the trace, in third line Eq. (2.17), 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. □

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

Lemma 1 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. (2.16) and Eq. (2.19) 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 point is then propagated through the circuit. When a Clifford unitary $U_{M,b} \in \mathscr{C}\ell$ is encountered, the phase point is updated as $u \to Mu + b$. 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 point according to the distribution implied by Eq. (2.17). 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 2.1. A similar classical simulation algorithm that applies to a slightly different family of circuits was described first in Ref. [1].

Input: $W_{\rho_{in}} \geq 0$

sample a point $u \in V$ according to the probability distribution $W_{\rho_{in}}$
while end of circuit has not been reached **do**
  if a Clifford unitary $U_{M,b} \in \mathscr{C}\ell$ is encountered **then**
    update $u \leftarrow Mu + b$
  end if
  if a Pauli measurements T_a, $a \in E$ is encountered **then**
    Output: $\omega^{[a,u]}$
    sample w uniformly from $\Gamma_{a,u}$
    update $u \leftarrow w$
  end if
end while

Algorithm 2.1: 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 2.1 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 2.1, is similar in structure to the proof of Theorem 7 given in Chapter 4 so we leave it out.[13] To prove efficiency of the classical simulation algorithm, we must prove

[13] In fact, the proof of Theomem 7 can be applied almost without modification to prove correctness of Algorithm 2.1 with a simple reinterpretation of the symbols. In particular, $A_\alpha \to A_u$, $u \in V$ are the phase point operators, the action of the Clifford group is as $A_{U\cdot\alpha} \to A_{U_{M,b}\cdot u} = A_{Mu+b}$ with $u \in V$, $M \in Sp(2n, \mathbb{Z}_d)$, and $b \in \mathbb{Z}_d^{2n}$, and the functions Q and q have particular forms based on Lemma 1. 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 $s \in \mathbb{Z}_d$.

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 like Algorithm 2.1. 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. (2.16): 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 [32]. 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 2.1 is slightly different from the one described in Ref. [1] 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 2.1 does not apply but simulation is possible via probability estimation [55, 56] 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 $\mathfrak{N}(\rho_{in})^2/\varepsilon^2$ where

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

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*.

2.7 Hidden variable models

Attempts to formulate hidden variable models for quantum theory are almost as old as quantum theory itself [20, 21]. 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 [20, 57, 58], 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 [22]. 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.[14]

An extension of Bell's result, proven by Bell [23] and independently by Kochen and Specker [24] states that there can be no noncontextual hidden variable model for quantum theory. This result and *contextuality* are the focus of the next section.

2.7.1 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 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 [24], and independently by John Bell [23]. 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

[14]There was an earlier no-go result for hidden variable models by von Neumann in 1932 [59] but this result relied on a faulty assumption as pointed out by Hermann [60]. See Ref. [6] for a review.

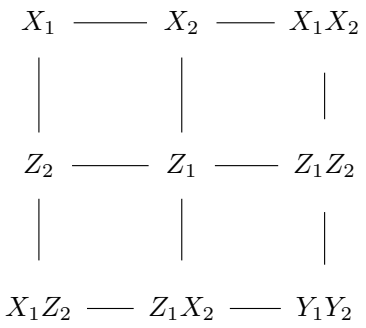


Figure 2.6: The two-qubit Pauli observables constituting the Mermin square, a set of observables which gives a simple state-independent proof of contextuality for quantum mechanics in four-dimensional Hilbert space [6]. The observables, A, B, C, in each row/column give a constraint that a deterministic 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.

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 four 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 [6]. The Mermin square consists of the nine two-qubit Pauli observables shown in Figure 2.6.

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

$$\begin{aligned} X_1 \cdot X_2 \cdot X_1X_2 &= I, \\ Z_2 \cdot Z_1 \cdot Z_1Z_2 &= I, \\ X_1Z_2 \cdot Z_1X_2 \cdot Y_1Y_2 &= I, \\ X_1 \cdot Z_2 \cdot X_1Z_2 &= I, \\ X_2 \cdot Z_1 \cdot Z_1X_2 &= I, \\ X_1X_2 \cdot Z_1Z_2 \cdot Y_1Y_2 &= -I. \end{aligned} \tag{2.20}$$

We would like to define a deterministic noncontextual value assignment for these observables. That is, a function λ 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,[15] we must have $\lambda(I) = 1$. Then, for consistency with the relations of Eq. (2.20), λ must map the Mermin square observables to $\{+1,-1\}$ subject to the constraints

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

But these six equations are inconsistent. To see this, note that if we multiply the

[15] A measurement of the identity on any state gives 1 with probability 1.

six equations together, on the left hand side there are exactly two factors of $\lambda(O)$ for each observable O so 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 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* [43].

For a more complete review of contextuality, see Ref. [6].

2.7.2 A cohomological description of contextuality

Contextuality among generalized Pauli observables can be described using group cohomology [43]. In this description, a chain complex $\mathscr{C}_* = \mathscr{C}_*(E)$ 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 $\mathscr{C}_*$ there is a cochain complex $\mathscr{C}^*$ formed from maps from edges, faces, and volumes of $\mathscr{C}_*$ to $\mathbb{Z}_d$.

Then the function β defined in Eq. (2.4) 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. (2.6) is equivalent to the statement that for any volume $(a,b,c) \in \mathscr{C}_*$, the coboundary $d\beta$ evaluated on (a,b,c) satisfies

$$d\beta(a,b,c) := \beta(a,b) + \beta(a+b,c) - \beta(b,c) - \beta(a,b+c) = 0. \tag{2.21}$$

Noncontextual value assignments can be associated with 1-cochains. The relation between a noncontextual value assignment for the Pauli observables $\lambda : \mathscr{P}/\langle\sqrt{\omega}\rangle \to \{\omega^k \mid k \in \mathbb{Z}_d\}$ and a 1-cochain $\gamma : E \to \mathbb{Z}_d$ of $\mathscr{C}^*$ 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 $(a,b) \in \mathscr{C}_*$, the coboundary $d\gamma$ evaluated on (a,b) satisfies

$$d\gamma(a,b) := \gamma(a) + \gamma(b) - \gamma(a+b) = -\beta(a,b). \tag{2.22}$$

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. [43] gives a relation between contextuality and the second cohomology group $H^2(\mathscr{C}, \mathbb{Z}_d)$. It states that if $H^2(\mathscr{C}, \mathbb{Z}_d) \ni [\beta] \neq 0$, 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), $[\beta] \neq 0$ [43].

This is a special case of a more general framework for describing contextuality using group cohomology. See Ref. [43] for details.

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

Lemma 1 shows that phase point operators are associated with deterministic assignments of measurement outcomes for the generalized Pauli observables. In particular, the phase point operator $A_u = d^{-n} \sum_{v \in V} \omega^{[u,v]} T_v$ is associated with the value assignment $T_a \rightarrowtail \omega^{-[u,a]}$. 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 [26, 35, 36].

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.[16] Then Eq. (2.13) 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 2.1, amounts to treating the hidden variable model as physical [1]. The update rules for Clifford unitaries and Pauli measurements stated in Eqs. (2.15) and (2.17) 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,[17] and thus the two traditional indicators of quantumness, negativity in the Wigner function and contextuality, agree [26, 35, 36]. With Algorithm 2.1, both are necessary conditions for a quantum computational speedup over classical computation in QCM on odd-dimensional qudits [1].

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.

[16] The phase 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.

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

2.8 Quasiprobability representations for qubits

The discrete Wigner function defined in Section 2.4 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 generalizatin of the odd-dimensional Wigner function to qubits, and we review some previously defined quasiprobability representations for quantum computation with magic states on qubits.

2.8.1 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 point operators in Section 2.4. In Eq. (2.1), 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 $T_a T_b = \omega^{[a,b]/2} T_{a+b}$. In particular, the function β is zero whenever a and b commute. Defining the phase point operators as in Eq. (2.10) and Eq. (2.11) 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 $H^2(\mathscr{C}, \mathbb{Z}_d) \ni [\beta] \neq 0$ [43].

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, I, 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 point operators defined like in Eqs. (2.10) and (2.11)

are

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

Under conjugation by the Clifford unitary H, $HA_{(0,0)}H^\dagger = \frac{1}{2}(I+X-Y+Z)$ is not a phase 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 point operators form an operator basis is Clifford covariant [37, 38]. 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 [38]. 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 [39]. 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 points 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 [61–65]. 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.

2.8.2 The eight state model

In the definition of the one-qubit phase point operators above we made a choice of phase for the Pauli operators: $e^{i\phi(a)} = i^{\langle a_Z|a_X\rangle}$ in Eq. (2.1) with the inner product taken mod 4. This choice is not unique. An equally valid choice would be $e^{i\phi(a)} = i^{-\langle a_Z|a_X\rangle}$, this choice also leads to Hermitian phase point operators of order 2. With this alternative choice, the phase 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 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 point operators, if we take all eight operators to define our phase space we get the so called eight state model [62].

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

$$\rho = \sum_{\alpha\in\{\pm\pm\pm\}} W_\rho(\alpha) A_\alpha \tag{2.23}$$

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 point operators map to phase point operators under conjugation by all one-qubit Clifford unitaries making the (nonunique) Wigner function defined by Eq. (2.23) 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.[18]

[18] We say positively representable here instead of postively represented because the decomposition Eq. (2.23) is not unique. A state being positively representable simply means there exists a

2.8.3 A discrete Wigner function for rebits

A version of the the discrete Wigner function can be defined for rebits—qubit systems in which the states are described by real density matrices [61]. Beyond restricting the admissible states to those with real density matrices, restrictions must be imposed on the allowed operations. Measurements are restricted to CSS type Pauli measurements [4, §10.4.2] and the unitary operations are restricted to CSS-ness preserving Clifford unitaries. A universal QCM model can be defined with these restrictions. The rebit Wigner function leads to an efficient classical simulation algorithm for QCM on the set of rebit states which are positively represented.

2.8.4 Stabilizer quasimixtures

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

$$\rho = \sum_{\sigma \in \mathscr{S}} s_\rho(\sigma)\sigma. \tag{2.24}$$

Since density matrices are Hermitian, the coefficients $s_\rho(\sigma)$ are real for all $\sigma \in \mathscr{S}$, and taking a trace of Eq. (2.24) we find $1 = \sum_{\sigma \in \mathscr{S}} s_\rho(\sigma)$. That is, the coefficients $s_\rho(\sigma)$ describe a quasiprobability representation of the state ρ [64].

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. (2.24) is not unique for any given state. The

decomposition of the form Eq. (2.23) in which the coefficients are nonnegative.

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

$$\mathfrak{R}_S(\rho) = \min \left\{ \sum_{\sigma \in \mathscr{S}} |s_\rho(\sigma)| \;\middle|\; \rho = \sum_{\sigma \in \mathscr{S}} s_\rho(\sigma)\sigma \right\} \tag{2.25}$$

is the parameter governing the complexity of simulation in the case where negativity in the expansion Eq. (2.24) is unavoidable [64].

Chapter 3

A generalized phase space and Wigner function for qubits

The main result of Ref. [2] is a generalized phase space and discrete Wigner function that applies to systems with any local Hilbert space dimension. This generalized Wigner function effectively unifies and extends the odd-dimensional Wigner function, the eight state model, and the quasiprobability representation based on stabilizer quasimixtures, subsuming them as special cases of a more general quasiprobability representation of quantum computation with magic states. The definition of the generalized phase space directly takes into account a prominent difference between the Pauli group in even and odd dimensions—the existence of state-independent proofs of contextuality exemplified by the Mermin square. The Wigner function provides a simulation method for quantum computation with magic states on qubits which establishes negativity in the Wigner function of the input state of a QCM circuit as a necessary condition for a quantum computational speedup in QCM on qubits. Although for qubits the phase space is enlarged, it maintains a certain amount of structure and points in the multiqubit phase space can be characterized and counted.

In this chapter, we begin by defining the generalized phase space and Wigner function. Then we consider the structure of the phase space in some particular cases, including a complete characterization of the multiqubit phase space. Next we consider how the phase space behaves under the dynamics of quantum com-

putation with magic states on qubits and describe a classical simulation algorithm for QCM on qubits based on the Wigner function. We close this chapter with a discussion of the size of the generalized phase space.

3.1 Definition of the generalized phase space

Recall, a source of difficulty in extending the definition of the odd-dimensional Wigner function to systems of multiple qubits (or systems with even local Hilbert space dimension more generally), comes from the existence of state-independent proofs of contextuality among the multiqubit Pauli observables. In the phase space defined in section 2.4, points in phase space are associated with noncontextual value assignments on the set of generalized Pauli operators (c.f. Section 2.7.3). For multiple qubits, noncontextual value assignments for the Pauli observables do not exist, as shown by the Mermin square proof of contextuality.

Even when noncontextual value assignments do not exist for the full set of Pauli observables, they can still exist for certain subsets of the Pauli observables. The definition of the generalized phase space presented here is based on subsets, $\Omega \subset E$, satisfying the following definitions.

Definition 1 *A set $\Omega \subset E$ is called closed under inference if for every $a, b \in \Omega$ such that $[a,b] = 0$, $a+b \in \Omega$.*

This is a well-defined closure property. For a set $\Omega \subset E$ which is not necessarily closed under inference, its closure under inference, denoted $\bar{\Omega}$, is the smallest set containing Ω which is closed under inference.

Definition 2 *A set $\Omega \subset E$ is called noncontextual if its closure under inference, $\bar{\Omega}$, admits a noncontextual value assignment. That is, there exists a function $\gamma : \bar{\Omega} \to \mathbb{Z}_d$ satisfying*

$$d\gamma(a,b) := \gamma(a) + \gamma(b) - \gamma(a+b) = -\beta(a,b) \tag{3.1}$$

for all $a, b \in \bar{\Omega}$ with $[a,b] = 0$.[1]

[1] For qubits, all addition in $\mathbb{Z}_2$ is mod 2 so Eq. (3.1) is equivalent to $d\gamma(a,b) := \gamma(a) + \gamma(b) + \gamma(a+b) = \beta(a,b)$. Since the primary focus of this chapter is on systems of qubits, this is the form of this equation that we use in most of this chapter.

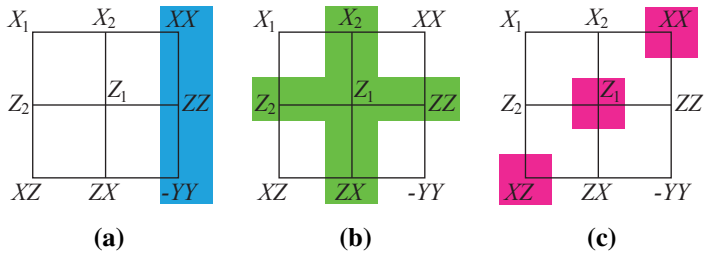


Figure 3.1: Examples of cnc sets for the Mermin square observables. The full Mermin square cannot be included in a cnc set since the Mermin square is contextual. Possible cnc sets include: (a) a single measurement context of the Mermin square, (b) two contexts of the Mermin square intersecting in one operator, and (c) three pairwise anticommuting operators. Note: for closure, each set must also contain the identity.

The name *closure under inference* comes from the fact that if two observables T_a and T_b commute then they can be measured simultaneously and the value of a measurement of their product $T_aT_b = \omega^{-\beta(a,b)}T_{a+b}$ can be inferred from the measurements of T_a and T_b individually. The definition of *noncontextuality* in Def. 2 implies there is no state-independent proof of contextuality among the Pauli observables labeled by the elements of a noncontextual set Ω.

Note that in Section 2.7.1, a noncontextual value assignment was defined as a function λ from the Pauli observables to the d^{th} roots of unity subject to certain constraints coming from the relations between the operators. Here we define a noncontextual value assignment to be a function γ from the labels of the Pauli operators to $\mathbb{Z}_d$ subject to Eq. (3.1). These are equivalent through the relation $\lambda(T_a) = \omega^{\gamma(a)}$. In the rest of this thesis we use the additive notation of Eq. (3.1). I.e., noncontextual value assignments are functions of the form described in Def. (2) and outcomes for Pauli measurements are associated with elements of $\mathbb{Z}_d$ rather than $\{\omega^k \mid k \in \mathbb{Z}_d\}$.

A set $\Omega \subset E$ which is both closed under inference and noncontextual is called *cnc*, for closed and noncontextual. Examples of cnc sets of two-qubit Pauli observables are given in Figure 3.1. With these definitions, we can define the generalized phase space.

Definition 3 *The generalized phase space $\mathscr{V}$ consists of pairs (Ω, γ) with $\Omega \subset E$ a set and $\gamma : \Omega \to \mathbb{Z}_d$ a function such that (i) Ω is closed under inference, (ii) Ω is noncontextual, (iii) $\gamma : \Omega \to \mathbb{Z}_d$ satisfies Eq. (3.1), and (iv) γ satisfies*

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

For each phase space point, $(\Omega, \gamma) \in \mathscr{V}$, there is a corresponding phase space point operator defined as

$$A^{\gamma}_{\Omega} = \frac{1}{d^n} \sum_{b \in \Omega} \omega^{-\gamma(b)} T_b. \tag{3.3}$$

These phase point operators are similar in structure to the phase point operators of the odd-dimensional discrete Wigner function defined in section 2.4 except instead of summing over all Pauli operators weighted by a noncontextual value assignment, the set $\Omega \subset E$ over which the noncontextual value assignment is defined is an additional varying parameter of the phase space and only operators in the set are included in the sum.

These phase point operators span the space of Hermitian operators. Therefore, any density matrix ρ can be expanded as

$$\rho = \sum_{(\Omega, \gamma) \in \mathscr{V}} W_{\rho}(\Omega, \gamma) A^{\gamma}_{\Omega}. \tag{3.4}$$

The coefficients in this expansion define the generalized discrete Wigner function $W_{\rho} : \mathscr{V} \to \mathbb{R}$. Since, in general, the phase point operators are not a basis, they are over-complete, the Wigner function is not unique. Any set of coefficients in an expansion of the form Eq. (3.4) could define a valid Wigner function for a state ρ but for the purpose of simulation of quantum computation discussed later in this chapter, some Wigner functions are more useful than others.

3.2 Relation to previous quasiprobability representations

3.2.1 Odd-dimensional qudits

When the local Hilbert space dimension is odd, there is no state-independent contextuality among the Pauli operators [43]. This means, the set $E = \mathbb{Z}_d^{2n}$ is cnc. The corresponding phase point operators have the form

$$A_E^\gamma = \sum_{b \in E} \omega^{-\gamma(b)} T_b$$

with $\gamma : E \to \mathbb{Z}_d$ satisfying Eqs. (3.1) and (3.2). The phase point operators defined in Section 2.4 also have this form (c.f. Section 2.7.3). Therefore, when the local Hilbert space dimension is odd, we recover the odd-dimensional discrete phase space as a subset of the generalized phase space. Note, however, that there are phase points of the generalized phase space which do not correspond to phase points in the standard odd-dimensional phase space (c.f. Ref [36]).

3.2.2 The eight state model

It is easy to check that the one-qubit set $\{0 = (0,0), x = (0,1), y = (1,1), z = (1,0)\} = \mathbb{Z}_2^2$ with $T_0 = I$, $T_x = X$, $T_y = Y$, and $T_z = Z$ the one-qubit Pauli observables is cnc. It is clearly closed under inference since all nonidentity observables pairwise anticommute so no nontrivial inference is possible. Also, since there are no nontrivial commuting pairs of observables, any function $\gamma : \mathbb{Z}_2^2 \to \mathbb{Z}_2$ satisfying the constraint Eq. (3.2) also satisfies Eq. (3.1). There are eight such functions.

The eight resulting phase point operators have the form

$$A_{\mathbb{Z}_2^2}^\gamma = \frac{1}{2} \sum_{b \in \mathbb{Z}_2^2} (-1)^{-\gamma(b)} T_b = \frac{1}{2} (I \pm X \pm Y \pm Z) .$$

These are the phase point operators of the eight state model [62] defined in section 2.8.2. Therefore, when restricted to one qubit, the generalized phase space and Wigner function is equivalent to the eight state model.

3.2.3 Stabilizer quasimixtures

As noted in section 2.1, a stabilizer group can be specified by a pair $(\tilde{I}, \gamma)$ where $\tilde{I} \subset E$ is an isotropic subspace and $\gamma : \tilde{I} \to \mathbb{Z}_d$ is a function satisfying $\omega^{-\gamma(0)} T_0 = I$ and Eq. (2.5). Isotropic subspaces are closed by definition and with the constraint on γ in Eq. (2.5), γ satisfies Eq. (3.1). Therefore, the pair $(\tilde{I}, \gamma)$ satisfies the four constraints of Def. 3 and so it specifies a point in phase space.

In the case where $\tilde{I}$ is a maximal isotropic subspace, it is easy to check that A_I^γ is a projector onto a stabilizer state. Therefore, the quasiprobability representation based on stabilizer states discussed in section 2.8.4 can be obtained by restricting the generalized phase space $\mathscr{V}$ to include only pairs $(\Omega, \gamma) \in \mathscr{V}$ where Ω is a maximal isotropic subspace of E.

Since the stabilizer quasimixture representation is obtained from a restriction of the phase space of Def. 3, it is easy to prove that all stabilizer mixtures are positively representable with respect to the generalized phase space $\mathscr{V}$, and furthermore, there are states which are not stabilizer mixtures but which are positively representable with respect to $\mathscr{V}$. See, for example, Ref. [2, §IV.D.]. This result is confirmed for up to three qubits by the numerical results presented in Appendix B.

3.2.4 Rebits

The definition of the generalized phase space can be easily modified to give a generalized phase space and Wigner function for rebits. This involves restricting to the cnc sets Ω to real Pauli matrices. Examples of real cnc sets are given in Figure 3.1. The Wigner function obtained in this way is different from the rebit Wigner function defined in Ref. [61]. In particular, the phase point operators are different, and there are fewer restrictions on the allowed operations.

For the rest of this chapter, except where otherwise stated, we specialize to systems of multiple qubits.

3.3 Characterization of the multiqubit phase space

In the odd-dimensional case and the single qubit case, the phase space has a very simple structure. For multiple qubits, the phase space is necessarily much larger

than might be expected for a generalization of the odd-dimensional phase space but it maintains some structure and the points in phase space can still be characterized. This characterization involves two parts: a characterization of multiqubit cnc sets $\Omega \subset E$, and a characterization of the noncontextual value assignments γ on those sets. We start with the sets.

Lemma 3 *For n qubits, consider an isotropic subspace $\tilde{I} \subset E$ of dimension $n-m$, with $0 \leq m \leq n$, and $\xi \leq 2m+1$ elements $a_k \in E$ that pairwise anticommute but all commute with $\tilde{I}$. Denote $I_k := \langle a_k, \tilde{I} \rangle$ for $k = 1, \ldots, \xi$. For any number n of qubits, the sets*

$$\Omega = \bigcup_{k=1}^{\xi} I_k \tag{3.5}$$

are noncontextual and closed under inference.

Proof of Lemma 3. Existence. The sets Ω of Eq. (3.5) exist for all n, m, ξ. To see this, consider the m-qubit Jordan-Wigner transforms of the Majorana Fermion operators acting on qubits 1 to m,

$$\begin{aligned} C_{2j-1} &= I_{1\ldots j-1} X_j Z_{j+1} Z_{j+2} \cdots Z_{m-1} Z_m, \\ C_{2j} &= I_{1\ldots j-1} Y_j Z_{j+1} Z_{j+2} \cdots Z_{m-1} Z_m, \end{aligned} \tag{3.6}$$

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

$$C_{2m+1} = Z_1 Z_2 \cdots Z_{m-1} Z_m. \tag{3.7}$$

Further, let $\tilde{I}$ be the isotropic subspace corresponding to a stabilizer state supported on the $n-m$ qubits numbered $m+1, \ldots, n$. Define a_k via $C_k \propto T_{a_k}$ as in Eqs. (3.6) and (3.7), for all $k = 1, \ldots, 2m+1$. 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 $\mu = 0$. W.l.o.g. 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. (3.1) on $\tilde{I}$. We now extend this function to Ω as follows. The values $\gamma(a_k)$, for $k = 1, \ldots, \xi$ 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 subspaces I_k forming Ω, and $d\gamma(a,b) = \beta(a,b)$ thus holds.

This establishes that the sets Ω of Eq. (3.5) 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 n-qubit cnc sets Ω are of the form Eq. (3.5).*

Proof of Theorem 2. Let $\Omega \subset E$ be closed under inference and noncontextual. We can partition the elements of Ω into two subsets, $\Omega = \{q_1, q_2, \ldots, q_\mu | g_1, g_2, \ldots, g_\nu\}$, where $\tilde{I} := \{g_1, g_2, \ldots, g_\nu\}$ are the elements of Ω which commute with all other elements of Ω. $\tilde{I}$ is an isotropic subspace of E since if two elements, a and b, commute with Ω then clearly their sum, $a+b$, also commutes with Ω, and $a+b \in \Omega$ since Ω is closed under inference. $\tilde{I}$ is isotropic by definition.

If all elements of Ω pairwise commute then $\Omega = \tilde{I}$ is an isotropic subspace. Isotropic subspaces are cnc. If Ω is not an isotropic subspace, then it can be written compactly as

$$\Omega = \bigcup_{k=1}^{\xi} \langle p_k, \tilde{I} \rangle \tag{3.8}$$

where $\xi \geq 2$, the cosets $p_1 + \tilde{I}, p_2 + \tilde{I}, \ldots, p_\xi + \tilde{I}$ are distinct, and $q_1, q_2, \ldots, q_\mu$ are contained in the cosets. Note that in this form, there can be no elements p_j which commutes with all of $p_1, p_2, \ldots, p_\xi$ because $\tilde{I}$ is defined to contain all such elements. Now we consider the possible commutation relations that $p_1, p_2, \ldots, p_\xi$ can have given that 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, w.l.o.g. p_1, p_2, p_3,

and p_4, with the commutation relations like those of $\{X_1, X_2, Z_1, Z_2\}$:

$$\begin{aligned} &[p_1, p_2] = [p_1, p_4] = [p_2, p_3] = [p_3, p_4] = 0, \\ &[p_1, p_3] = [p_2, p_4] = 1, \end{aligned} \tag{3.9}$$

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

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

$$\begin{aligned} &[p_1, p_2] = [p_2, p_3] = [p_3, p_4] = 0, \\ &[p_1, p_3] = [p_1, p_4] = [p_2, p_4] = 1. \end{aligned} \tag{3.10}$$

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. (3.9). Thus, it must contain a Mermin square.

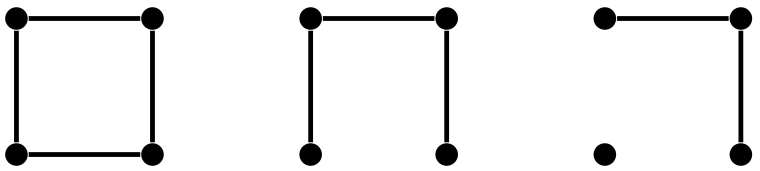
A similar argument shows that another sufficient condition for a set which is closed under inference to be contextual is that it contains four elements with the commutation relations

$$\begin{aligned} &[p_1, p_2] = [p_2, p_3] = 0, \\ &[p_1, p_3] = [p_1, p_4] = [p_2, p_4] = [p_3, p_4] = 1. \end{aligned} \tag{3.11}$$

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. (3.9).

To determine the possible commutation relations of the elements $p_1, p_2, \ldots, p_\xi$, we will look at their commutativity graph $\mathscr{G}$. That is, the undirected graph with a vertex for each of $p_1, p_2, \ldots, p_\xi$ and an edge connecting each pair of commuting vertices. Since Ω is noncontextual, the commutation relations of Eqs. (3.9), (3.10), and (3.11) provide restrictions on the possible commutation relations of the elements $p_1, p_2, \ldots, p_\xi$ of Ω. In terms of the commutativity graph $\mathscr{G}$, these correspond to forbidden induced subgraphs.[2]

[2] An induced subgraph of a graph is the graph obtained by taking a subset of the vertices of the



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

Figure 3.2: If a set Ω is cnc then no four of its elements can have the commutation relations specified by Eqs. (3.9), (3.10), or (3.11). In terms of the commutativity graph of the set these restrictions correspond to forbidden induced subgraphs. The restrictions on the commutation relations of elements of a cnc set coming from Eqs. (3.9), (3.10), and (3.11) say that the commutativity graph cannot have the graphs (a), (b), or (c) as induced subgraphs. A similar relation between these graphs and contextuality was found in Ref. [7].

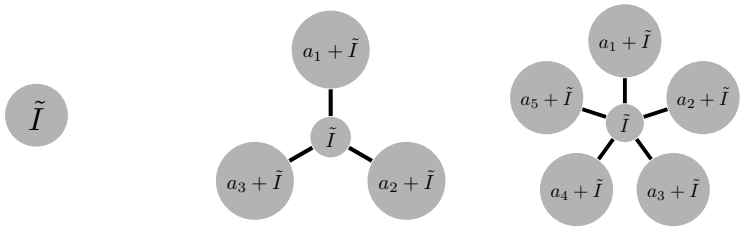
The restriction of Eq. (3.9) says that $\mathscr{G}$ cannot have a four-vertex cordless cycle (C_4) as an induced subgraph and the restriction of Eq. (3.10) says that $\mathscr{G}$ cannot have a four-vertex path (P_4) as an induced subgraph. These two forbidden induced subgraphs characterize a family of graphs known as trivially perfect graphs [66]. Therefore, since Ω is cnc, $\mathscr{G}$ must be a trivially perfect graph.

Connected trivially perfect graphs have the property that they contain a universal vertex [66].[3] If the commutativity graph $\mathscr{G}$ were connected then there would be an elements p_j which commutes with all other elements of $\{p_1, p_2, \ldots, p_\xi\}$. This is also forbidden. Therefore, $\mathscr{G}$ is disconnected.

Given that $\mathscr{G}$ is disconnected, Eq. (3.11) provides another restriction, namely, that each connected component of $\mathscr{G}$ cannot have a three-vertex path (P_3) as an induced subgraph; i.e., each connected component of $\mathscr{G}$ is a clique. The three forbidden induced subgraphs are displayed graphically in Figure (3.2).

original graph along with all of the edges connecting pairs of vertices in the subset.

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



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

Figure 3.3: The commutativity graphs of the cosets of cnc sets of the form Eq. (3.5) with parameter $\xi = 2m + 1$ and different values for m. Each node represents a coset of the isotropic subspace $\tilde{I}$. Elements pairwise commute within each node and elements in adjacent nodes pairwise commute. Elements in nonadjacent nodes pairwise anticommute.

This means we can partition the elements $\{p_1, p_2, \ldots, p_\xi\}$ into disjoint subsets

$$\begin{aligned}\{p_1, \ldots, p_\xi\} = &\{p_{1,1}, p_{1,2}, \ldots, p_{1,\xi_1}\} \\ &\cup \{p_{2,1}, p_{2,2}, \ldots, p_{2,\xi_2}\} \\ &\cup \cdots \cup \{p_{\pi,1}, p_{\pi,2}, \ldots, p_{\pi,\xi_\pi}\}\end{aligned} \tag{3.12}$$

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 under inference. Now suppose that a subset in the partition contains 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, the elements $\{p_1, p_2 \ldots, p_\xi\}$ of Eq. (3.8) pairwise anticommute. This completes the proof. □

A result equivalent to the characterization of Eq. (3.12) was discovered independently in Ref. [7] (also see [67]).

Lemma 3 and Theorem 2 effectively characterize all cnc sets of multiqubit Pauli observables that could be used in the phase space of Def. 3. A graphical representation of these sets is given in Figure 3.3.

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. When all pairs (Ω, γ) satisfying the four conditions in Def. (3) are included in the definition of the generalized phase space, the phase space contains redundancy. It will be convenient to reduce the size of the phase space by eliminating this redundancy. This is achieved by the following lemma.

Lemma 4 *A cnc set $\Omega \subset E$ is* maximal *if there is no cnc set $\tilde{\Omega} \subset E$ such that $\Omega \subsetneq \tilde{\Omega}$. Let $\mathscr{V}_M$ denote the phase space consisting of pairs (Ω, γ) where Ω is a maximal cnc set, and γ is a noncontextual value assignment on Ω, satisfying the four conditions of Def. (3). Then for any $(\tilde{\Omega}, \tilde{\gamma}) \in \mathscr{V}$ where $\tilde{\Omega}$ is not maximal, there are nonnegative coefficients $c(\Omega, \gamma) \geq 0 \; \forall (\Omega, \gamma) \in \mathscr{V}_M$ such that*

$$A^{\tilde{\gamma}}_{\tilde{\Omega}} = \sum_{(\Omega,\gamma)\in\mathscr{V}_M} c(\Omega, \gamma) A^{\gamma}_{\Omega}.$$

Further, a state ρ is positively representable with respect to the phase space $\mathscr{V}$ if and only if it is positively representable with respect to $\mathscr{V}_M$.

Proof of Lemma 4. For any cnc set $\tilde{\Omega}$ and value assignment $\tilde{\gamma}$ on $\tilde{\Omega}$, we have a phase point operator $A^{\tilde{\gamma}}_{\tilde{\Omega}}$. If $\tilde{\Omega}$ is not a maximal set, then as a result of Theorem 2, $\tilde{\Omega}$ has the form

$$\tilde{\Omega} = \bigcup_{k=1}^{\zeta} \langle a_k, \tilde{I} \rangle$$

where $\tilde{\Omega}$ is contained in a maximal cnc set

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

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 $d\gamma_0 = \beta$ and $d\gamma_1 = \beta$. Then,

$$\begin{aligned}\frac{1}{2}\left(A_\Omega^{\gamma_0} + A_\Omega^{\gamma_1}\right) &= \frac{1}{2^{n+1}} \sum_{b\in\Omega} \left((-1)^{\gamma_0(b)} + (-1)^{\gamma_1(b)}\right) T_b \\ &= \frac{1}{2^n} \sum_{b\in\tilde{\Omega}} (-1)^{\gamma(b)} T_b + \frac{1}{2^{n+1}} \sum_{b\in\Omega\setminus\tilde{\Omega}} \left((-1)^{\gamma_0(b)} + (-1)^{\gamma_1(b)}\right) T_b\end{aligned}$$

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 $g \in \tilde{I}$. Then, with all addition mod 2, we have

$$\gamma_0(b) = \beta(a_j, g) + \gamma_0(a_j) + \gamma(g) = \beta(a_j, g) + \gamma(g)$$

and

$$\gamma_1(b) = \beta(a_j, g) + \gamma_1(a_j) + \gamma(g) = \beta(a_j, g) + 1 + \gamma(g).$$

In both cases, $\gamma_1(b) = \gamma_0(b) + 1 \mod 2$. Therefore, each term in the second sum in the expression above vanishes and we have

$$\frac{1}{2}\left(A_\Omega^{\gamma_0} + A_\Omega^{\gamma_1}\right) = \frac{1}{2^n} \sum_{b\in\Omega} (-1)^{\gamma(b)} T_b = A_{\tilde{\Omega}}^{\tilde{\gamma}}.$$

If a state ρ has a nonnegative expansion with respect to $\mathscr{V}$ with positive coefficient on $A_{\tilde{\Omega}}^{\tilde{\gamma}}$, substituting the left hand side of this equation for $A_{\tilde{\Omega}}^{\tilde{\gamma}}$ does not introduce any negativity. Therefore, if a state is positively representable with respect to $\mathscr{V}$ then it is positively representable with respect to $\mathscr{V}_M$. □

As a result of this lemma, from the point of view of representing states, there is no generality lost in restricting the definition of the phase space to include only phase points (Ω, γ) where Ω is a maximal cnc set. The characterization of cnc sets satisfying this additional constraint is achieved with the following lemma.

Lemma 5 *All maximal n-qubit cnc sets Ω are of the form Eq. (3.5) with $1 \le m \le n$ and $\xi = 2m+1$.*

Proof of Lemma 5. The fact that Eq. (3.5) sets exist and are cnc is proven in Lemma 3. That there are no more cnc sets is proven in Theorem 2. Therefore,

maximal cnc sets are sets of the form Eq. (3.5) for which ξ is maximal for a given isotropic subspace $\tilde{I}$. First, if $m = 0$ then $\Omega = \tilde{I}$ is an isotropic subspace. Isotropic subspaces are not maximal cnc sets because they are always contained in Eq. (3.5) sets with parameter $m = 1$.

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 pairwise anticommuting elements a_k which complete the set are elements of $\tilde{I}^{\perp}/\tilde{I}$ where $\tilde{I}^{\perp}$ is the symplectic complement[4] of $\tilde{I}$. This is a m-dimensional symplectic space, therefore the maximal value of ξ is the largest number of pairwise anticommuting m-qubit Pauli operators. The largest sets of pairwise 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 $T_{a_k} \propto C_k$, with C_k defined in Eq. (3.6). The set $\{a_k \mid 1 \leq k \leq 2m\}$ consists of pairwise anticommuting elements. There is an element c, with $T_c \propto C_{2m+1}$ (cf. Eq. (3.7)) 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

$$[c, a_k] = 1, \quad 1 \leq k \leq 2m$$

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 constructed. Suppose $\{\tilde{a}_k \mid 1 \leq k \leq 2m\}$ is such a set. By Witt's lemma [68, §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 $v, w \in E$ (a 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 for ξ.

To complete the proof we must show that maximal sets of pairwise anticommuting elements on m qubits with cardinality 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 symplectic transformation $f : E \to E$ which maps the set to one of the form Eq. (3.6) and Eq. (3.7). Therefore, we can find $m - m'$ independent elements which commute with the set. For example, if

[4]The symplectic complement of a set $S \subset E$ is the set $S^{\perp} = \{b \in E \mid [a, b] = 0 \; \forall a \in S\}$.

$g_1, g_2, \ldots, g_{m-m'}$ are the vectors corresponding to Pauli operators $X_{m'+1}, X_{m'+2}, \ldots, X_m$, then we could take $f^{-1}(g_1), f^{-1}(g_2), \ldots, f^{-1}(g_m)$. 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. (3.5) with $1 \leq m \leq n$ and $\xi = 2m+1$. □

Lemma 3, Theorem 2, Lemma 4, and Lemma 5 fully characterize the cnc sets $\Omega \subset E$ which partially define the multiqubit phase space. To complete the characterization of the phase space, we can also characterize the noncontextual value assignments γ on these sets.

Lemma 6 *Let $\Gamma(\Omega)$ denote the set of functions $\gamma : \Omega \to \mathbb{Z}_2$ that satisfy the constraints Eqs. (3.1) and (3.2). Then for any cnc set Ω, $\Gamma(\Omega)$ is the coset of a vector space $U(\Omega)$.*

Proof of Lemma 6. 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 $c, c' \in \mathbb{Z}_2$. □

Equivalently, we can say that to specify a noncontextual value assignment γ on a cnc set Ω of the form Eq. (3.5), it suffices to specify the value of γ on the generators of Ω (i.e. on a basis for the subspace $\tilde{I} \subset E$ and on the pairwise anticommuting elements $a_1, a_2, \ldots$). Any choice of values for γ on the generators of Ω uniquely defines a noncontextual value assignment on Ω through the relation Eq. (3.1).

This completes the characterization of the multiqubit phase space.

3.4 Rules for state update under QCM dynamics

In the previous sections we have analyzed the generalized phase space $\mathscr{V}$ on which the quasiprobability function is defined. We now turn to dynamics. For our setting of QCM this concerns evolution under Clifford unitaries and Pauli measurements.

3.4.1 Clifford unitaries

Let $\mathscr{C}\ell$ be the n-qubit Clifford group. It acts on the n-qubit Pauli operators via

$$U(T_a) := UT_aU^\dagger = (-1)^{\Phi_U(a)} T_{U \cdot a}, \quad \forall U \in \mathscr{C}\ell.$$

This relation simultaneously defines the phase function Φ and the action of $\mathscr{C}\ell$ on E. It implies an action of the Clifford group on the phase point operators A_Ω^γ, which in turn induces an action on the sets Ω and the functions γ, via

$$U\left(A_\Omega^\gamma\right) = \frac{1}{2^n}\sum_{a\in\Omega}(-1)^{\gamma(a)}U(T_a) = \frac{1}{2^n}\sum_{b\in\Omega'}(-1)^{\gamma'(b)}T_b.$$

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

$$\gamma'(U\cdot a) := \gamma(a) + \Phi_U(a), \quad \forall a\in\Omega.$$

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

For use in the proof below we quote Lemma 3 from Ref. [43], which says that, for any face $(a,b)\in\Omega\times\Omega$,

$$\Phi_U\left(\partial(a,b)\right) = \beta(U\cdot a, U\cdot b) + \beta(a,b) \mod 2.$$

We have the following result.

Lemma 7 *$\mathscr{V}$ is mapped to itself undel $\mathscr{C}\ell$, and the quasiprobability function W transforms covariantly. That is, if the state ρ can be described by W_ρ through Eq. (3.4), then for any $U\in\mathscr{C}\ell$ the state $U\rho U^\dagger$ can be described by the quasiprobability function*

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

Proof of Lemma 7. First, we show that the phase space is closed under the action of $\mathscr{C}\ell$, i.e., if $(\Omega,\gamma)\in\mathscr{V}$ then $(\Omega',\gamma')\in\mathscr{V}$. The four items in Definition 3 need to be checked. (i) Ω' *is closed under inference.* Assume that $c,d\in\Omega'$, and $[c,d]=0$. Then there exist $a,b\in\Omega$ such that $c = U\cdot a$, $d = U\cdot b$ and $[a,b]=0$. Then $c+d = U\cdot a + U\cdot b = U\cdot(a+b)\in\Omega'$, since $a+b\in\Omega$ by assumption of closedness. Hence Ω' is closed under inference.

(iii) γ' *satisfies Eq. (3.1).* With the definition of γ' we have (all addition mod 2)

$$\begin{aligned} d\gamma'(U\cdot a, U\cdot b) &= d\gamma(a,b) + \phi_U(\partial(a,b)) \\ &= d\gamma(a,b) + \beta(U\cdot a, U\cdot b) + \beta(a,b) \\ &= \beta(U\cdot a, U\cdot b). \end{aligned}$$

Therein, in the third line we have used Eq. (3.1). Thus, γ' satisfies Eq. (3.1) on its domain. (ii) Ω' *is noncontextual.* With γ' we have just proved the existence of a function on Ω' that satisfies Eq. (3.1). (iv) γ' *satisfies Eq. (3.2).* Since γ satisfies Eq. (3.2), it follows $I = U(I) = U\left((-1)^{\gamma(0)}T_0\right) = (-1)^{\gamma(0)+\Phi_U(0)} = (-1)^{\gamma'(0)}T_0$. Equation (3.2) is thus satisfied for γ'. Hence, if $(\Omega,\gamma)\in\mathscr{V}$ then $(\Omega',\gamma')\in\mathscr{V}$, as claimed.

Next, we turn to the covariance of W under $\mathscr{C}\ell$. We have

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

Comparing the last expression with the expansion Eq. (3.4) for $U\rho U^\dagger$, we find that for all $U\in\mathscr{C}\ell$, the quasiprobability distribution $W_{U\rho U^\dagger}$ defined by

$$W_{U\rho U^\dagger}(\Omega,\gamma) = W_\rho(U^{-1}\cdot\Omega, U^{-1}\cdot\gamma) \tag{3.13}$$

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

3.4.2 Pauli measurements

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 $\mathrm{Comm}(a) := \{b\in E \mid [a,b]=0\}$ and $\Omega_a := \Omega\cap\mathrm{Comm}(a)$,

$$\Omega\times a := \Omega_a \cup \{a+b \mid b\in\Omega_a\}, \quad \forall a\notin\Omega. \tag{3.14}$$

Likewise, we define an update on functions γ invoking the measurement outcome s of an observable T_a, namely, $(\cdot)\times s : (\gamma : \Omega \to \mathbb{Z}_2) \rightarrowtail (\gamma \times s : \Omega \times a \to \mathbb{Z}_2)$. We define this update only for $(\Omega,\gamma) \in \mathscr{V}$, and only for $a \notin \Omega$.[5] The updated function $\gamma \times s : \Omega \times a \to \mathbb{Z}_2$ is given by

$$\gamma \times s(b) := \gamma(b), \quad \forall b \in \Omega_a, \tag{3.15a}$$

$$\gamma \times s(b) := \gamma(a+b) + s + \beta(a,b), \quad \forall a+b \in \Omega_a. \tag{3.15b}$$

The rules of Eq. (3.15) are used to formulate the update rule for phase point operators of Eq. (3.3) under Pauli measurement.

Remark. Update rules similar to Eq. (3.15) have been used previously [69] to construct a ψ-epistemic model of the multiqubit stabilizer formalism. Those rules update the value assignment in the same way but are applied under different conditions. Specifically, the update in Ref. [69] does not refer the general sets Ω satisfying the conditions of Def. 3.

Lemma 8 *Denote the projectors $\Pi_{a,s} := (I + (-1)^s T_a)/2$, and let A^γ_Ω a phase point operator defined through Eq. (3.3), with $(\Omega,\gamma) \in \mathscr{V}$ satisfying the conditions of Definition 3. Then the effect of a measurement of the Pauli observable T_a with outcome s on A^γ_Ω is*

$$\Pi_{a,s} A^\gamma_\Omega \Pi_{a,s} = \delta_{s,\gamma(a)} \frac{A^\gamma_\Omega + A^{\gamma+[a,\cdot]}_\Omega}{2}, \quad \text{if } a \in \Omega, \tag{3.16a}$$

$$\Pi_{a,s} A^\gamma_\Omega \Pi_{a,s} = \frac{1}{2} A^{\gamma\times s}_{\Omega\times a}, \quad \text{if } a \notin \Omega. \tag{3.16b}$$

Eq. (3.16) entails the update of both sets Ω and functions γ. Figure 3.4 displays the update of the set Ω shown in Figure 3.1b, under the measurement of (a) the observable X_2 with $a(X_2) \in \Omega$, and (b) the observable X_1, with $a(X_1) \notin \Omega$.

In preparation for the proof of Lemma 8 it is useful to state two relations of the function β for $d = 2$. With the definition Eq. (2.4) of β and Eq. (3.2), the operator

[5] 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 through both Eqs. (3.15a) and (3.15b), and we need to check consistency. These inferences are indeed consistent, as a consequence of Eq. (3.1). Since we do not need the case of $a \in \Omega$ subsequently, we skip the details of the argument.

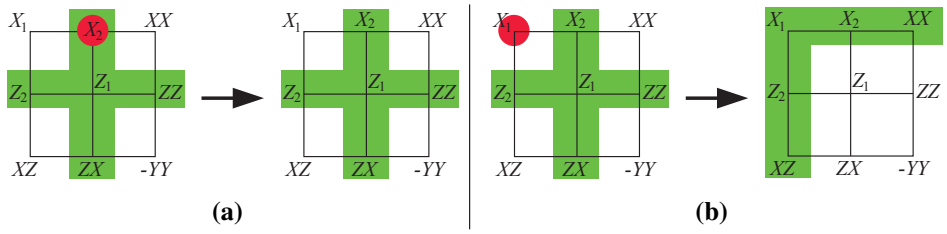


Figure 3.4: Update of a cnc set Ω in Mermin's square, under two Pauli measurements. (a) The measured observable X_2 is such that $a(X_2) \in \Omega$; hence the update proceeds by Eq. (3.16a). (b) The measured observable X_1 is such that $a(X_1) \notin \Omega$; hence the update proceeds by Eq. (3.16b).

identities $T_a T_a = I$ and $T_b I = T_b$ imply that

$$\beta(a,a) = \beta(a,0) = \gamma(0), \quad \forall a \in E. \tag{3.17}$$

Furthermore, evaluating $d\beta(a,a,0) = 0$ (see Eq. (2.21)), and using Eq. (3.17) yields

$$\beta(a,b) = \beta(a,a+b), \quad \forall a,b. \tag{3.18}$$

To prove Lemma 8 we also need the following result.

Lemma 9 *If $\Omega \subset E$ is noncontextual and closed under inference, then so is Ω_a, for all $a \in E$.*

Proof of Lemma 9. First consider closure. Assume that $c,d \in \Omega_a$ and $[c,d] = 0$. Then $c,d \in \Omega$, and $c+d \in \Omega$, 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 $\mathscr{C}(\Omega_a)$, and so can γ. Hence, $d\gamma|_{\mathscr{C}(\Omega_a)} = \beta|_{\mathscr{C}(\Omega_a)}$. Thus, Ω_a is noncontextual. □

Proof of Lemma 8. Under the measurement of T_a with outcome $s \in \mathbb{Z}_2$ we have

$$\begin{aligned} \frac{I+(-1)^s T_a}{2} A_\Omega^\gamma \frac{I+(-1)^s T_a}{2} &= \frac{I+(-1)^s T_a}{2} \frac{1}{2^n} \sum_{b \in \Omega_a} (-1)^{\gamma(b)} T_b \\ &= \frac{1}{2 \times 2^n} \sum_{b \in \Omega_a} (-1)^{\gamma(b)} T_b + \frac{(-1)^s}{2 \times 2^n} \sum_{b \in \Omega_a} (-1)^{\gamma(b)} T_a T_b. \end{aligned} \tag{3.19}$$

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. (3.19),

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

Therein, in the first line we have used Eq. (2.4), in the second line Eq. (3.1), in the third line the completeness of Ω_a under inference (Lemma 9), and the fourth line is just a relabeling of the elements of Ω_a. Inserting this result in the above expansion Eq. (3.19), we find

$$\Pi_{a,s} A^{\gamma}_{\Omega} \Pi_{a,s} = \delta_{s,\gamma(a)} \frac{1}{2^n} \sum_{b\in\Omega_a} (-1)^{\gamma(b)} T_b \tag{3.20}$$

and Eq. (3.16a) follows.

Case II: $a \notin \Omega$. Substituting $b \to a+b$ in Eq. (3.15b) gives $\gamma \times s(a+b) = \gamma(b) + s + \beta(a, a+b)$, for $b \in \Omega_a$. With Eq. (3.18) we obtain

$$\gamma \times s(a+b) = \gamma(b) + s + \beta(a,b), \quad \forall b \in \Omega_a. \tag{3.21}$$

With this, we now look at the second term in the expansion Eq. (3.19),

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

The first line above follows with Eq. (2.4), and the second with Eq. (3.21).

Considering the first term in the expansion Eq. (3.19), with Eq, (3.15a) 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. (3.19), and using the definition Eq. (3.14) of $\Omega \times a$, we obtain Eq. (3.16b). $\square$

We have so far shown how the phase point operators can be updated under measurement once. We still need to show that this update can be iterated. This requires that the phase point operators appearing on the r.h.s. of Eq. (3.16) satisfy the consistency constraints of Definition 3.

Lemma 10 *If $(\Omega, \gamma) \in \mathscr{V}$ then $(\Omega, \gamma + [a, \cdot]) \in \mathscr{V}$, for all $a \in \Omega$, and $(\Omega \times a, \gamma \times s) \in \Omega$, for all $a \notin \Omega$ and $s \in \mathbb{Z}_2$.*

Proof of Lemma 10. Statement (A): $(\Omega, \gamma + [a, \cdot]) \in \mathscr{V}$, $\forall a \in \Omega$. The set Ω does not change, and we need only to check the properties in Definition 3 that concern the function update, i.e., Eqs. (3.1) and (3.2).

Assume that $\gamma : \Omega \to \mathbb{Z}_2$ 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. (3.1).

Furthermore, assume that γ satisfies Eq. (3.2). Then $(\gamma + [a, \cdot])(0) = \gamma(0) + [a, 0] = \gamma(0)$. Hence, $\gamma + [a, \cdot]$ satisfies Eq. (3.2).

Statement (B): $(\Omega \times a, \gamma \times s) \in \mathscr{V}$, $\forall a \notin \Omega$ and $s \in \mathbb{Z}_2$. There are four items to check in Definition 3. Namely, (I) $\Omega \times a$ is closed under inference, (II) $\Omega \times a$ is noncontextual, (III) $\gamma \times s$ satisfies Eq. (3.1), and (IV) $\gamma \times s$ 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 subcases. *(i)* $c, d \in \Omega_a$. Then $c + d \in \Omega_a$, since Ω_a is closed under inference by Lemma 9. 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 9, 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 contruction 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 9, $c'+d' = c+d \in \Omega_a$. 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. (3.15a), it holds that $d(\gamma \times s)(c,d) = d\gamma(c,d) = \beta(c,d)$.

(ii) W.l.o.g. 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 $c+d$,

$$\begin{aligned} d(\gamma \times s)(f) &= \gamma \times s(c) + \gamma \times s(d) + \gamma \times s(c+d) \\ &= \gamma(c) + \gamma(d') + \gamma(c+d') + \beta(a,d) + \beta(a,c+d) \\ &= \beta(c,d') + \beta(a,d) + \beta(a,c+d) \\ &= \beta(c,d) \end{aligned}$$

Therein, in the second line we have used Eq. (3.15), in the third line Eq. (3.1), in the fourth line Eq. (3.18), and in the fourth line $d\beta(a,d,c) = 0$, cf. Eq. (2.21).

(II): Per Def. 2, $\Omega \times a$ is non-contextual if there is a function $\tau : \Omega \times a \longrightarrow \mathbb{Z}_2$ that satisfies $d\tau = \beta$. We have explicitly constructed such a function in (III) above, $\tau := \gamma \times s$.

(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. (3.15a) it follows that $\gamma \times s(0) = \gamma(0)$, and hence $\gamma \times s$ also satisfies Eq. (3.2). □

Lemma 7, covariance of the Wigner function with respect to the Clifford group, together with the update rules under Pauli measurements given in Lemma 8 with

Lemma 10 lead to the following theorem.

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

Proof of Theorem 3. Preservation positive representability under Clifford unitaries follows immediately from Lemma 7. For positivity preservation under Pauli measurements, consider a positively representable state ρ and a measurement of the Pauli observable T_a on it. Assume that the measurement outcome s can occur, $\mathrm{Tr}(\Pi_{a,s}\rho) > 0$. We have to show that under these conditions, the postmeasurement state

$$\rho' = \frac{\Pi_{a,s}\rho\Pi_{a,s}}{\mathrm{Tr}(\Pi_{a,s}\rho)}$$

is also prositively representable.

Denote $\bar{\delta}_{a\in\Omega} := 1 - \delta_{a\in\Omega}$. Then, with Lemma 8 and the state expansion Eq. (3.4) of ρ, we have

$$\rho' = \sum_{(\Omega,\gamma)\in\mathscr{V}} \frac{W_\rho(\Omega,\gamma)}{\mathrm{Tr}(\Pi_{a,s}\rho)} \left(\delta_{a\in\Omega}\delta_{s,\gamma(a)} \frac{A^\gamma_\Omega + A^{\gamma+[a,\cdot]}_\Omega}{2} + \frac{1}{2}\bar{\delta}_{a\in\Omega} A^{\gamma\times s}_{\Omega\times a} \right). \tag{3.22}$$

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

$$\begin{aligned} W_{\rho'}(\Omega',\gamma') = \sum_{(\Omega,\gamma)\in\mathscr{V}} \frac{W_\rho(\Omega,\gamma)}{2\mathrm{Tr}(\Pi_{a,s}\rho)} \big[\delta_{a\in\Omega}\delta_{s,\gamma(a)} \big(\delta_{(\Omega',\gamma'),(\Omega,\gamma)} \\ + \delta_{(\Omega',\gamma'),(\Omega,\gamma+[a,\cdot])} \big) + \bar{\delta}_{a\in\Omega}\delta_{(\Omega',\gamma'),(\Omega\times a,\gamma\times s)} \big]. \end{aligned} \tag{3.23}$$

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

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

Just like in the case of the odd-dimensional Wigner function, phase point operators map deterministically to phase point operators under conjugation by Clifford

unitaries (Lemma 7) and phase point operators map to probabilistic combinations of phase point operators under Pauli measurements (Lemma 8 with Lemma 10). These two facts lead to an efficient classical simulation algorithm for quantum computation with magic states on qubits based on the generalized Wigner function representation that applies when the input state to the quantum circuit is positively representable.

This algorithm is similar to the classical simulation method based on the odd-dimensional Wigner function, Algorithm 2.1. A point in the generalized phase space is sampled according to the Wigner function probability distribution representing the input state ρ_{in}. Then the phase point is propagated through the circuit and updated when Clifford unitaries and Pauli measurements are encountered according to the rules determined by Lemma 7 and Lemma 8. Outcomes for the Pauli measurements are returned as the phase point is propagated through. The full simulation algorithm is described in Algorithm 3.1.

Theorem 4 *For any $n \in \mathbb{N}$ and any n-qubit quantum state ρ_{in} with $W_{\rho_{in}}(\Omega,\gamma) \geq 0 \, \forall (\Omega,\gamma) \in \mathscr{V}$, the classical simulation algorithm, Algorithm 3.1, for sampling the outcomes of the Pauli measurements in a circuit consisting of Clifford unitaries and Pauli measurements applied to input state ρ_{in} agrees with the predictions of quantum mechanics. Further, when the Wigner function of the input state is nonnegative and can be efficiently sampled from, the simulation algorithm is efficient.*

The proof of correctness for a modified version of this simulation algorithm for circuits consisting of Pauli measurements only is given in Section VI.B. of Ref. [2]. The extension of the proof to include Clifford unitaries is trivial so we leave it out. The structure of the proof of correctness for Algorithm 3.1 is similar to the proof of Theorem 7 given in Chapter 4.[6] To prove efficiency, we must prove the four efficiency criteria, properties (E0)–(E3), defined in Section 2.6. First, (E0), an

[6] Like in the case of the simulation algorithm based on the odd-dimensional Wigner function (footnote on page 26), the proof of Theorem 7 can be used with only minor modification to prove correctness of Algorithm 3.1. In this case, $A_\alpha \to A^\gamma_\Omega, (\Omega,\gamma) \in \mathscr{V}$ are the multiqubit phase point operators, update under Clifford unitaries is as $A_{U\cdot\alpha} \to A^{U\cdot\gamma}_{U\cdot\Omega}$, and the functions Q and q have particular forms based on Lemma 8. Namely, $Q_a(s\,|\,\alpha) \to Q_a(s\,|\,(\Omega,\gamma)) = \delta_{a\in\Omega}\delta_{\gamma(a),s} + \frac{1}{2}\bar{\delta}_{a\in\Omega}$ and $q_{\alpha,a}(\beta,s) \to q_{(\Omega,\gamma),a}((\Omega',\gamma'),s) = \frac{1}{2}\delta_{a\in\Omega}\delta_{s,\gamma(a)}\left(\delta_{(\Omega',\gamma'),(\Omega,\gamma)} + \delta_{(\Omega',\gamma'),(\Omega,\gamma+[a,\cdot])}\right) + \frac{1}{2}\bar{\delta}_{a\in\Omega}\delta_{(\Omega',\gamma'),(\Omega\times a,\gamma\times s)}$.

Input: $W_{\rho_{in}} \geq 0$

sample a point $(\Omega, \gamma) \in \mathscr{V}$ according to the probability distribution $W_{\rho_{in}}$
while end of circuit has not been reached **do**
  if a Clifford unitary $U \in \mathscr{C}\ell$ is encountered **then**
    update $(\Omega, \gamma) \leftarrow (U \cdot \Omega, U \cdot \gamma)$
  end if
  if a Pauli measurement T_a is encountered **then**
    if $a \in \Omega$ **then**
      Output: $\gamma(a)$
      flip a coin
      if heads **then**
        update $(\Omega, \gamma) \leftarrow (\Omega, \gamma)$
      else if tails **then**
        update $(\Omega, \gamma) \leftarrow (\Omega, \gamma + [a, \cdot])$
      end if
    else if $a \notin \Omega$ **then**
      flip a coin
      if heads **then**
        $s \leftarrow 0$
      else if tails **then**
        $s \leftarrow 1$
      end if
      Output: s
      update $(\Omega, \gamma) \leftarrow (\Omega \times a, \gamma \times s)$
    end if
  end if
end while

Algorithm 3.1: One run of the classical simulation algorithm for quantum computation with magic states on qubits based on the generalized Wigner function. The algorithm provides samples from the joint probability distribution of the Pauli measurements in a quantum circuit consisting of Clifford unitraries and Pauli measurements applied to an input state ρ_{in} with $W_{\rho_{in}}(\Omega, \gamma) \geq 0$ for all $(\Omega, \gamma) \in \mathscr{V}$.

efficient description for points in phase space, follows from the characterization of the multiqubit phase space points given in Section 3.3. Specifying a phase space point with parameters n and m requires specifying $(n-m)+(2m+1) = n+m+1$ elements of $\mathbb{Z}_2^{2n}$ as generators of the cnc set Ω along with $n+m+1$ elements of $\mathbb{Z}_2$ so specify the value assignment $\gamma : \Omega \to \mathbb{Z}_2$. This requires a number of bits which is polynomial in the number n of qubits. Like in the case of Algorithm 2.1, (E1) is included as an assumption of Theorem 4. Finally, the proofs of (E2) and (E3), efficient handling of Clifford unitaries and Pauli measurements are given in Section IV.B. of Ref. [2].

When negative values in the Wigner function of the input state to a QCM circuit are unavoidable, Algorithm 3.1 fails, but classical simulation is still possible through probability estimation [55, 56]. In this case, like in the case of the odd-dimensional Wigner function and the quasiprobability representation based on stabilizer states, the complexity of classical simulation is proportional to the amount of negativity in the Wigner function. Nonuniqueness of the multiqubit Wigner function allows us to minimize the cost of classical simulation by minimizing the amount of negativity in the Wigner function. Then, similar to the robustness of magic defined in Eq. (2.25), the parameter which governs the complexity of simulation is the *phase space robustness*, $\mathfrak{R}$, defined as the solution to the optimization problem

$$\mathfrak{R}(\rho) = \min_{W} \left\{ \sum_{(\Omega,\gamma)\in\mathscr{V}} |W(\Omega,\gamma)| \;\middle|\; \rho = \sum_{(\Omega,\gamma)\in\mathscr{V}} W(\Omega,\gamma) A_\Omega^\gamma \right\}. \tag{3.24}$$

In particular, using the algorithm of Ref. [55], the number of samples required to estimate an outcome probability to a given error ε scales as $\mathfrak{R}(\rho)^2/\varepsilon^2$. For the construction of a resource theory centered around the phase space robustness, see Ref. [2, §VII.C.].

Numerical calculations of the robustness for some common magic states, along with a comparison of the phase-space-simulation method with the stabilizer quasimixture simulation method in terms of robustness and volumes of positively representable states are given in Appendix B.

To summarize, we have an efficient classical simulation algorithm for quantum

computation with magic states on qubits that applies whenever the input state to the QCM circuit is positively representable. When the input state takes negative values, classical simulation is still possible but it is generally inefficient. Therefore, we have effectively established negativity in the generalized Wigner function as a *necessary* condition for quantum computation with magic states on qubits to exhibit a quantum computational speedup over classical computation, extending the previous result that applied only to odd-dimensional qudits.

In the odd-dimensional case, contextuality serves as an indicator of a quantum advantage for QCM that agrees with wigner function negativity. In the case of qubits, the existence of state-independent contextuality makes the discussion of contextuality of the hidden variable model implied by the generalized Wigner function more complicated, as we'll see below.

3.6 The SW correspondence and contextuality

Two items regarding the generalized Wigner function have so far been omitted: (i) a justification for labeling the representation a discrete Wigner function through the Stratonovich-Weyl correspondence, and (ii) the classification of the hidden variable model implied by the generalized Wigner function as either contextual or noncontextual. It turns out, these two items are closely related. With suitable modifications, the SW criteria can be made to apply to the setting described here (see Ref. [2, §VIII.A.] for the complete discussion of how the SW criteria apply to the generalized Wigner function). One of the more significant modifications to the SW criteria that must be made is in the linearity condition.

The linearity condition ensures that the map from operators to the quasiprobability distributions representing them is one-to-one and linear. The generalized Wigner function for multiqubit states is not unique, and hence this map is not one-to-one, it is one-to-many. As a result of contextuality among multiqubit Pauli observables, this modification is required for a discrete Wigner function to be able to simulate QCM on multiple qubits.

The interpretation of the hidden variable model implied by the generalized Wigner function as contextual or noncontextual is more difficult for qubits than it is for odd-dimensional qudits. First, there is some ambiguity in how to interpret

the value assignments. For a cnc set $\Omega \subset E$ with a noncontextual value assignment $\gamma : \Omega \to \mathbb{Z}_2$, γ could be interpreted as a deterministic but partial value assignment for the multiqubit Pauli observables, or as a total but probabilistic value assignment which assigns an expectation value of 0 to observables not in Ω.

The first interpretation could define a valid noncontextual hidden variable model but only if we restrict to states which can be positively represented by phase point operators corresponding to a single cnc set Ω, and we restrict measurements to observables labeled by Ω. The assumption of the classical simulation algorithm of Section 3.5 is that states can access all phase point operators A_Ω^γ in their representation and that there is no restriction on the Pauli measurements that can be performed. This suggests the second interpretation of the value assignments as more appropriate.

With probabilistic value assignments, the original definition of contextuality cannot be easily applied. More recently, there have been generalizations of contextuality which do allow for probabilistic value assignments, for example, Spekkens notion of generalized contextuality [73]. With this generalized contextuality, it is precisely the nonuniqueness of the Wigner function (or equivalently, the modification of the linearity condition of the SW correspondence) which leaves open the door for a generalized contextuality, in particular, for Spekkens definition of preparation contextuality [73]. For a more complete discussion of the interpretation of the contextuality of this model, see Ref. [2, §VIII.B.].

3.7 Size of the generalized phase space

As shown in Section 3.2.1, when the local Hilbert space dimension is odd the generalized phase space contains the phase space of the standard odd-dimensional Wigner function defined in Section 2.4 with cardinality $|\mathscr{V}| = |\mathbb{Z}_d^{2n}| = d^{2n}$.

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 [37–39]. 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 $(\Omega, \gamma) \in \mathscr{V}$ subject to the constraints of Def. 3 with the additional constraint that Ω is maximal. Using the classification of multiqubit phase space points described in Section 3.3 we can count the number of multiqubit phase space points.

Theorem 5 *The cardinality of the n-qubit phase space is*

$$|\mathscr{V}| = \sum_{m=1}^{n} \left\{ \frac{2^{n+m^2+m+1}}{(2m+1)!} \begin{bmatrix} n \\ m \end{bmatrix}_4 \left(\prod_{k=1}^{n-m} 2^k + 1 \right) \left(\prod_{j=1}^{m} 4^j - 1 \right) \right\} \tag{3.25}$$

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 5. Per Lemma 5, 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.[7]

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 $\mathscr{A} \subset \mathscr{P}(\mathbb{Z}_2^{2m})$ be the set of all such sets, and $A = \{a_1, a_2, \ldots, a_{2m+1}\} \in \mathscr{A}$ be one particular set.

To find $|\mathscr{A}|$, consider the action of the symplectic group, $Sp(2m, \mathbb{Z}_2)$, on $\mathscr{A}$ induced by the action of $Sp(2m, \mathbb{Z}_2)$ on the elements of the sets in $\mathscr{A}$. As shown in the proof of Lemma 5, the symplectic group acts transitively on $\mathscr{A}$. Therefore, by

[7] 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 $\mathbb{Z}_2^{2n}$.

the orbit-stabilizer theorem and Lagrange's theorem [70],

$$|\mathscr{A}| = |Sp(2m,\mathbb{Z}_2)\cdot A| = [Sp(2m,\mathbb{Z}_2) \ : \ Sp(2m,\mathbb{Z})_A] = \frac{|Sp(2m,\mathbb{Z}_2)|}{|Sp(2m,\mathbb{Z}_2)_A|}$$

where $Sp(2m,\mathbb{Z}_2)_A = \{M \in Sp(2m,\mathbb{Z}_2) \,|\, M\cdot A = A\}$ is the stabilizer of A in $Sp(2m,\mathbb{Z}_2)$.

Consider a permutation σ of $\{1,2,\ldots,2m+1\}$ and a linear map that sends a_k to $a_{\sigma(k)}$ for $1 \le k \le 2m$. By Witt's lemma [68, §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 $[a_k, c] = 1,\ 1 \le k \le 2m$ and $[f(a_k), c] = 1,\ 1 \le k \le 2m$ 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. $|Sp(2m,\mathbb{Z}_2)_A| = (2m+1)!$. Also, $|Sp(2m,\mathbb{Z}_2^{2m})| = 2^{m^2}\prod_{j=1}^{m} 4^j - 1$. Thus, the number of maximal sets of pairwise anticommuting elements in $\mathbb{Z}_2^{2m}$ of size $2m+1$ is

$$N_{(ii)} = \frac{|Sp(2m,\mathbb{Z}_2)|}{|Sp(2m,\mathbb{Z}_2)_A|} = \frac{1}{(2m+1)!} 2^{m^2} \prod_{j=1}^{m} 4^j - 1.$$

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. (3.1), 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. (3.25). Then, summing m from 1 to n gives the result. □

Table 3.1 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 $W_\rho : \mathscr{V} \to \mathbb{R}_{\ge 0}$ representing ρ with $|\mathrm{Supp}(W_\rho)| \le$

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	$\lvert\mathscr{V}\rvert$ $(m=0)$	$\lvert\mathscr{V}\rvert$ $(m=1)$	$\lvert\mathscr{V}\rvert$ (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 3.1: 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 points corresponding to cnc sets of the form Eq. (3.5) 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.

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 [71, §1.6].

A similar result holds for states which are not positively representable.

Lemma 11 *For any $n \in \mathbb{N}$ and any n-qubit state ρ, there is a Wigner function $W_\rho(\Omega,\gamma)$ that minimizes the robustness Eq. (3.24) with $|Supp(W_\rho)| \leq 2^{2n}$.*

Proof of Lemma 11. The phase space robustness $\mathfrak{R}(\rho)$ of a state ρ is defined as the solution to the convex optimization problem

$$\min_{\mathbf{q}} \{||\mathbf{q}||_1 \mid M\mathbf{q} = \mathbf{b}\}$$

where $M_{ij} = \mathrm{Tr}(A_{\alpha_j} T_{a_i})$, $b_i = \mathrm{Tr}(\rho T_{a_i})$, $\{A_{\alpha_j} \mid 1 \leq j \leq |\mathscr{V}|\}$ is an enumeration of the phase point operators, and $\left\{T_{a_i} \mid 1 \leq i \leq 2^{2n}\right\}$ are the n-qubit Pauli operators. For each variable q_j in this optimization problem, we can define two new variables, $q_j^+ = \max(0, q_j)$, and $q_j^- = \max(0, -q_j)$. Then the robustness can equivalently be defined as the solution to the standard form linear program

$$\min_{\mathbf{q}} \left\{ \sum_j q_j^+ + q_j^- \;\middle|\; \tilde{M}\tilde{\mathbf{q}} = \mathbf{b},\ \tilde{\mathbf{q}} \geq 0 \right\} \tag{3.26}$$

where $\tilde{M} = \begin{bmatrix} M & -M \end{bmatrix}$ and $\tilde{\mathbf{q}} = \begin{bmatrix} \mathbf{q}^+ \\ \mathbf{q}^- \end{bmatrix}$. Since this program is feasible (any physical state can be written as an affine combination of phase point operators) and bounded (no physical state can have robustness less than 1), by the fundamental theorem of linear programming, Eq. (3.26) has a solution at a vertex of the feasible polytope. Since $\tilde{M}$ in Eq. (3.26) has dimensions $2^{2n} \times 2|\mathscr{V}|$, this means there is a solution to Eq. (3.26) where $\tilde{\mathbf{q}}$ has no more than 2^{2n} nonzero entries [72, §13.1,13.2]. □

With this lemma, we recover a familiar feature of the original discrete phase space.

Chapter 4

A hidden variable model for quantum computation on qubits

The quasiprobability representation described in Chapter 3 goes a long way toward unifying several quasiprobability representations of finite-dimensional quantum mechanics, including the odd-dimensional Wigner function [1, 32, 53], the eight state model [62], and the quasiprobability representation based on stabilizer states [64]. The conventional wisdom is that for any quasiprobability representation, negativity, either in the states or in the dynamics, is necessary to accurately represent quantum theory [35, 74–77]. It turns out that a hidden variable model can be constructed for universal quantum computation with magic states on qubits in which all elements of computation—all states, operations, and measurements—are represented using only classical probabilities—no negativity is needed [3].

The existence of this model is unintuitive. It defies precedents established by previous quasiprobability representations. Further, there are several no-go theorems suggesting that a model of this form shouldn't exists.

In this chapter, we will review the construction of this hidden variable model. We will present a classical simulation algorithm for quantum computation with magic states based on the hidden variable model. And finally, we will conclude with an explanation for why the construction works. We defer the discussion of the relation of this model to no-go theorems to Chapter 5. The main theorem of this chapter is Theorem 6 which formally defines the hidden variable model.

4.1 Definition of the hidden variable model

Let $\mathscr{O}(2^n)$ denote the trace-1 affine subspace of the vector space of 2^n-dimensional Hermitian matrices and let $\mathscr{S}$ denote the set of pure n-qubit stabilizer states represented as density matrices. The hidden variable model is based on the set

$$\Lambda := \{X \in \mathscr{O}(2^n) \mid \mathrm{Tr}(\sigma X) \geq 0,\ \forall \sigma \in \mathscr{S}\}. \tag{4.1}$$

These are the allowed "states" of the hidden variable model. The elements of Λ are Hermitian operators with unit trace much like density operators. Unlike density operators, they are not necessarily positive semidefinite. This set was studied previously in a different context in Ref. [78]. Therein, the author considers the relation of this set to the stabilizer polytope and to contextuality.

When interpreted as a subset of the real affine space $\mathscr{O}(2^n)$, the set Λ has several convenient geometric and topological properties, two of which are formalized in the following lemma:

Lemma 12 *For any number of qubits $n \in \mathbb{N}$, (i) Λ is convex, and (ii) Λ is compact.*

Proof of Lemma 12. Convexity of Λ follows immediately from the definition. For compactness, since Λ is a subset of a real vector space, by the Heine-Borel theorem [79, Theorem 2.41] it suffices to prove that Λ is closed and bounded. Closure of Λ also follows immediately from the definition. Finally, the inequalities that define Λ imply that Λ is contained in the hypercube $\{X \in \mathscr{O}(2^n) \mid \mathrm{Tr}(\Pi_{a,s}X) \geq 0\ \forall a \in E \setminus \{0\}, s \in \mathbb{Z}_2\}$. Therefore, Λ is bounded. □

As a set in a real vector space, Λ is defined as the intersection of a finite number of halfspaces, i.e. Λ is a convex polyhedral set. As shown in Lemma 12, Λ is compact, and so it is a convex polytope. Therefore, by the Minkowski-Weyl theorem [71, §1.1], an equivalent description of Λ is as the convex hull of finitely many vertices. We denote the set of vertices of the polytope by $\mathscr{A}$ with elements $A_\alpha \in \mathscr{A}$. The index set for the vertices is denoted $\mathscr{V}$ with $\alpha \in \mathscr{V} \Leftrightarrow A_\alpha \in \mathscr{A}$. Then, $\Lambda = \mathrm{ConvHull}\{A_\alpha \mid \alpha \in \mathscr{V}\}$.

The hidden variable model is similar in structure to the hidden variable models based on the odd-dimensional Wigner function described in Section 2.4 and the

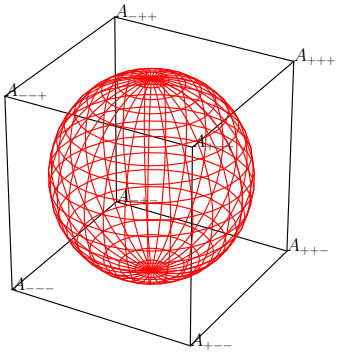


Figure 4.1: For one qubit, the polytope Λ is a cube inscribing the Bloch sphere. In the definition of Λ there is an inequality for each of the six one-qubit stabilizer states. These define the six faces of the cube. The eight vertices the cube are the eight phase point operators of the eight-state model.

generalized Wigner function of Chapter 3. It is based on a finite set of Hermitian operators (phase point operators) which we associate with points in a generalized phase space. Here, we take the index set $\mathscr{V}$ to be the generalized phase space and the vertices $A_\alpha \in \mathscr{A}$ are the generalized phase point operators.

An illustration of the polytope Λ for one qubit is shown in Figure 4.1. On one qubit, this hidden variable model is equivalent to the model defined in Chapter 3. This is not true in general. It turns out that for any number of qubits, the phase point operators A_Ω^γ defined in chapter 3 are vertices of Λ (see the Supplementary Material of Ref. [3] for the proof), but in general there are more vertices of Λ than there are generalized phase point operators A_Ω^γ.

The generalized phase space considered here has properties similar to those that made the phase spaces defined in section 2.4 and Chapter 3 useful for describing quantum computation with magic states. For example, we have the following lemma:

Lemma 13 *For any number of qubits $n \in \mathbb{N}$, Λ has the following properties:*

1. *For any $X \in \Lambda$ and any Clifford unitary $U \in \mathscr{C}\ell$, $UXU^\dagger \in \Lambda$,*

2. *For any $X \in \Lambda$, any $a \in E$, and any $s \in \mathbb{Z}_2$, if $Tr(\Pi_{a,s}X) > 0$ then*

$$X' := \frac{\Pi_{a,s}X\Pi_{a,s}}{Tr(\Pi_{a,s}X)} \in \Lambda,$$

3. *For any density matrix ρ representing a physical n-qubit quantum state, $\rho \in \Lambda$. That is, Λ contains every 2^n-dimensional positive-semidefinite Hermitian matrix with unit trace.*

Proof of Lemma 13. To prove that a 2^n-dimensional Hermitian operator is in Λ, we have to check two properties: (i) that it has unit trace (i.e. that it is in $\mathscr{O}(2^n)$), and (ii) that the Hilbert-Schmidt inner product with every stabilizer state is nonnegative. We will prove the three properties of Λ listed above in order. (1.) Let $U \in \mathscr{C}\ell$ be a Clifford unitary and $X \in \Lambda$. Clifford unitaries map stabilizer states to stabilizer states under conjugation [4, §10.5.2]. Therefore, for any $\sigma \in \mathscr{S}$, we have $\mathrm{Tr}(UXU^\dagger\sigma) = \mathrm{Tr}(XU^\dagger\sigma U) = \mathrm{Tr}(X\sigma') \geq 0$ where the last inequality follows from the fact that $\sigma' \in \mathscr{S}$ and $X \in \Lambda$. Also, $\mathrm{Tr}(UXU^\dagger) = \mathrm{Tr}(XU^\dagger U) = \mathrm{Tr}(X) = 1$. Therefore, $UXU^\dagger \in \Lambda$.

(2.) Let $a \in E$, $s \in \mathbb{Z}_2$, and $X \in \Lambda$ and suppose that $\mathrm{Tr}(X\Pi_{a,s}) > 0$. Then for any stabilizer state $\sigma \in \mathscr{S}$, we have

$$\mathrm{Tr}\left(\Pi_{a,s}X\Pi_{a,s}\sigma\right) = \mathrm{Tr}(X\Pi_{a,s}\sigma\Pi_{a,s}) = c \cdot \mathrm{Tr}(X\sigma')$$

with $c \geq 0$ and $\sigma' \in \mathscr{S}$. In particular, $c = 1$ if $(-1)^s T_a$ is in the stabilizer group of σ, $c = 0$ if $(-1)^{s+1} T_a$ is in the stabilizer group of σ, and $c = \frac{1}{2}$ otherwise [4, §10.5.3]. Therefore, since $X \in \Lambda$, $\mathrm{Tr}(X'\sigma) \geq 0$, $\forall \sigma \in \mathscr{S}$. We also have $\mathrm{Tr}(X') = 1$ by definition so $X' \in \Lambda$.

(3.) Any density matrix ρ corresponding to a physical quantum state has unit trace and has nonnegative probabilities associated with stabilizer measurements. I.e. $\mathrm{Tr}(\rho\sigma) \geq 0$ for any stabilizer state $\sigma \in \mathscr{S}$. Therefore, for any n-qubit density matrix ρ, $\rho \in \Lambda$. □

The first two properties show that the polytope Λ is closed under the dynamical operations of quantum computation with magic states—Clifford unitaries and Pauli measurements. These are the key properties, shared by the previously defined phase spaces, that allow the phase space to describe quantum computation (c.f. Lemma 2 and Theorem 3). The third property sets this phase space apart from the previous phase spaces. For vertices of Λ, property 1 of Lemma 13 can be refined:

Lemma 14 *Vertices of Λ map to vertices under conjugation by Clifford unitaries. That is, for any number of qubits $n \in \mathbb{N}$, any Clifford unitary $U \in \mathscr{C}\ell$, and any vertex $A_\alpha \in \mathscr{A}$ of Λ, $A_{U\cdot\alpha} := UA_\alpha U^\dagger$ is a vertex of Λ.*

Proof of Lemma 14. For a vertex $A_\alpha \in \mathscr{A}$ of Λ, define the set $S_\alpha = \{\sigma \in \mathscr{S} \mid \mathrm{Tr}(A_\alpha \sigma) = 0\}$. A_α is the unique solution in $\mathscr{O}(2^n)$ to the system of equations

$$\mathrm{Tr}(Y\sigma) = 0, \quad \forall \sigma \in S_\alpha.$$

In this way, the vertices A_α of Λ are uniquely specified by sets of stabilizer states S_α [80, §18].

Consider a set of stabilizer states $S \subset \mathscr{S}$ and a Clifford unitary $U \in \mathscr{C}\ell$. There is an action of U on the set S defined by $U \cdot S = \{U\sigma U^\dagger \mid \sigma \in S\}$. Under conjugation by U, solutions in $\mathscr{O}(2^n)$ of the system

$$\mathrm{Tr}(Y\sigma) = 0, \quad \forall \sigma \in S$$

are mapped bijectively to solutions of the system

$$\mathrm{Tr}(Y\sigma) = 0, \quad \forall \sigma \in U^{-1} \cdot S$$

since $\mathrm{Tr}(UYU^\dagger) = \mathrm{Tr}(YU^\dagger U) = \mathrm{Tr}(Y)$ and $\mathrm{Tr}(UYU^\dagger\sigma) = \mathrm{Tr}(Y(U^\dagger\sigma U))$. Therefore, if S_α is a set of stabilizer states specifying a vertex A_α of Λ, i.e. A_α is the unique solution to the system of equations derived from $S_\alpha \subset \mathscr{S}$, then $A_{U\cdot\alpha} := UA_\alpha U^\dagger$ is the unique solution to the system of equations derived from $U^{-1} \cdot S_\alpha \subset \mathscr{S}$. Then since $UA_\alpha U^\dagger \in \Lambda$ by property 1. of Lemma 13, $A_{U\cdot\alpha}$ is a vertex of Λ. □

The hidden variable model for quantum computation with magic states on qubits is formally defined by the following theorem.

Theorem 6 *For any number of qubits $n \in \mathbb{N}$, (i) each n-qubit quantum state ρ can be represented by a probability function $p_\rho : \mathscr{V} \to \mathbb{R}_{\geq 0}$,*

$$\rho = \sum_{\alpha \in \mathscr{V}} p_\rho(\alpha) A_\alpha. \tag{4.2}$$

(ii) Under the action of the Clifford group, $\mathscr{V}$ is mapped to itself, and the proba-

bility function p_ρ transforms covariantly. That is, for any state ρ and any Clifford unitary $U \in \mathscr{C}\ell$, the state $U\rho U^\dagger$ can be represented by the probability function defined by

$$p_{U\rho U^\dagger}(\alpha) = p_\rho(U^{-1} \cdot \alpha). \tag{4.3}$$

(iii) For state update under Pauli measurements it holds that

$$\Pi_{a,s} A_\alpha \Pi_{a,s} = \sum_{\beta \in \mathscr{V}} q_{\alpha,a}(\beta, s) A_\beta. \tag{4.4}$$

For all $a \in E$, $\alpha \in \mathscr{V}$, the $q_{\alpha,a} : \mathscr{V} \times \mathbb{Z}_2 \to \mathbb{R}_{\geq 0}$ are probability functions. (iv) 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 $P_{\rho,a}(s) = Tr(\Pi_{a,s}\rho)$ takes the form

$$Tr(\Pi_{a,s}\rho) = \sum_{\alpha \in \mathscr{V}} p_\rho(\alpha) Q_a(s \,|\, \alpha) \tag{4.5}$$

where $Q_a(s \,|\, \alpha)$ is given by

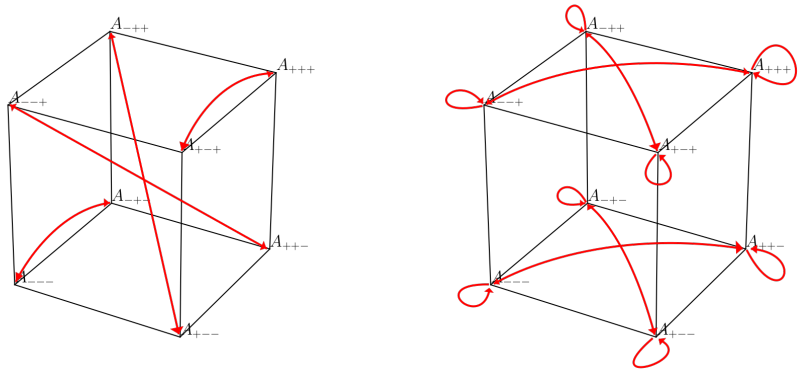
$$Q_a(s \,|\, \alpha) := \sum_{\beta \in \mathscr{V}} q_{\alpha,a}(\beta, s). \tag{4.6}$$

Hence $0 \leq Q_a(s \,|\, \alpha) \leq 1$, for all a, s, α.

Proof of Theomem 6. We will consider the four statemensts of the theorem in order. First, property (3.) of Lemma 13 shows that for any n-qubit density matrix ρ, $\rho \in \Lambda$. Therefore, by the Krein-Milman theorem, ρ can be expressed as a convex combination of the vertices of Λ as in Eq. (4.2). This proves the first statement.

Second, let $U \in \mathscr{C}\ell$ be a Clifford unitary. Lemma 14 shows that for a vertex A_α of Λ, $UA_\alpha U^\dagger =: A_{U\cdot\alpha}$ is also a vertex of Λ. Now consider a state ρ with an expansion Eq. (4.2). We have

$$\begin{aligned} U\rho U^\dagger &= \sum_{\alpha \in \mathscr{V}} p_\rho(\alpha) U A_\alpha U^\dagger \\ &= \sum_{\alpha \in \mathscr{V}} p_\rho(\alpha) A_{U\cdot\alpha} \\ &= \sum_{\alpha \in \mathscr{V}} p_\rho(U^{-1} \cdot \alpha) A_\alpha. \end{aligned}$$



(a) Update under Clifford Unitary H **(b)** Update under Pauli Z measurement

Figure 4.2: Update rules for the vertices of Λ for one qubit under (a) the Clifford unitary H and (b) a Pauli Z measurement. Under the Clifford unitary, the update is deterministic. It amounts to a permutation of the vertices of Λ. Under the Pauli measurement, the update is probabilistic. Each red arrow in the figure on the right represents a transition probability of $1/2$.

Here in the second line we use Lemma 14 and in the third line a relabeling of the terms in the sum. Comparing this expression with Eq. (4.2), since $p_\rho(\alpha)$ is a valid probability function describing ρ, this shows that $p_\rho(U^{-1}\cdot\alpha)$ is a valid probability function describing $U\rho U^\dagger$.[1]

For (iii), consider a projector $\Pi_{a,s}$ and phase point operator A_α with $\mathrm{Tr}(\Pi_{a,s}A_\alpha) > 0$. Since $X' := \Pi_{a,s}A_\alpha\Pi_{a,s}/\mathrm{Tr}(\Pi_{a,s}A_\alpha) \in \Lambda$ as shown in property (2.) of Lemma 12, X' can be expanded as a convex combination of the vertices $\mathscr{A}$. Therefore,

$$\Pi_{a,s}A_\alpha\Pi_{a,s} = \sum_{\beta\in\mathscr{V}} q_{\alpha,a}(\beta,s)A_\beta,$$

with $q_{\alpha,a}(\beta,s) \geq 0$ for all $\beta \in \mathscr{V}$ and $s \in \mathbb{Z}_2$. Fixing α and a and adding the above

[1] Note that in general the probability function p_ρ is not unique for any given state. This property does not say that $p_\rho(U^{-1}\cdot\alpha)$ is the unique probability function describing $U\rho U^\dagger$, only that it is a valid probability function describing $U\rho U^\dagger$.

equations for $s = 0$ and $s = 1$, then taking a trace we get

$$\sum_{s\in\mathbb{Z}_2}\sum_{\beta\in\mathscr{V}} q_{\alpha,a}(\beta,s) = 1.$$

Therefore, $q_{\alpha,s} : \mathscr{V}\times\mathbb{Z}_2 \to \mathbb{R}_{\geq 0}$ is a probability distribution for any $\alpha\in\mathscr{V}$ and any $a\in E$.

Finally, for statement (iv) we calculate

$$\begin{aligned}
\mathrm{Tr}(\Pi_{a,s}\rho) &= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha)\mathrm{Tr}(\Pi_{a,s}A_\alpha) \\
&= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha)\mathrm{Tr}(\Pi_{a,s}A_\alpha\Pi_{a,s}) \\
&= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha)\sum_{\beta\in\mathscr{V}} q_{\alpha,a}(\beta,s) \\
&= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha)Q_a(s\,|\,\alpha).
\end{aligned}$$

Here in the first line we use the expansion of ρ, Eq. 4.2, in the third line Eq. (4.4), and in the fourth line the definition of $Q_a(s\,|\,\alpha)$. This proves the formulation of the Born rule in the hidden variable model. □

This theorem effectively defines a hidden variable model capable of representing all elements of quantum computation with magic states using only classical (nonnegative) probabilities. States are represented by probability distributions over the vertices of Λ and the dynamics of QCM are represented by transition probabilities between the vertices. Although this representation is similar in structure to quasiprobability representations like the discrete Wigner functions in previous chapters (the discussion of contextuality and the SW correspondence mirrors that of the generalized Wigner function of Chapter 3 given in Section 3.6), it is set apart from other quasiprobability representations by the fact that no negativity is needed—it is a *probability* representation. This shows that, in general, the presence or absence of negativity depends on the particular choice of quasiprobability representation, and that to represent universal quantum computation, no negativity is required.

4.2 A classical simulation algorithm for universal quantum computation with magic states on qubits

The hidden variable model described by Theorem 6 has a probabilistic representation for the dynamical elements of QCM: Clifford unitaries and Pauli measurements. This, together with the probabilistic representation of states, is all we need to construct a classical simulation algorithm for quantum computation with magic states like those based on the discrete Wigner functions in the previous chapters. The simulation algorithm is given in Algorithm 4.1. Whereas Algorithm 2.1 and Algorithm 3.1 applied only for the strict subset of input states for which the Wigner function was nonnegative, in this case all states are positively representable. Therefore, the classical simulation algorithm, Algorithm 4.1, can be used to simulate *any* QCM circuit on *any* n-qubit input state.

Input: $p_{\rho_{in}}$
1: sample a point $\alpha \in \mathscr{V}$ according to the probability distribution $p_{\rho_{in}}$
2: **while** end of circuit has not been reached **do**
3: **if** a Clifford unitary $U \in \mathscr{C}\ell$ is encountered **then**
4: update $\alpha \leftarrow U \cdot \alpha$
5: **end if**
6: **if** a Pauli measurement T_a, $a \in E$ is encountered **then**
7: sample $(\beta, s) \in \mathscr{V} \times \mathbb{Z}_2$ according to $q_{\alpha,a}(\beta, s)$
8: **Output:** s
9: update $\alpha \leftarrow \beta$
10: **end if**
11: **end while**

Algorithm 4.1: One run of the classical simulation algorithm for quantum computation with magic states on qubits based on the hidden variable model defined by Theorem 6. The algorithm provides samples from the joint probability distribution of the Pauli measurements in a circuit consisting of Clifford unitaries and Pauli measurements applied to the input state ρ_{in}.

Theorem 7 *For any $n \in \mathbb{N}$ and any n-qubit quantum state ρ_{in}, the classical simulation algorithm of Algorithm 4.1 for sampling the outcomes of the Pauli measurements in a circuit consisting of Clifford unitaries and Pauli measurements applied*

to input state ρ_{in} agrees with the predictions of quantum mechanics.

Proof of Theorem 7. Without loss of generality, a QCM circuit can be represented as a sequence $U_1, a_1, U_2, a_2, \ldots$ with $U_1, U_2, \cdots \in \mathscr{C}\ell$ representing the Clifford unitaries to be applied and $a_1, a_2, \cdots \in E$ representing the Pauli measurements to be performed. Consider a single layer of the a circuit of this form: a Clifford unitary U followed by a Pauli measurement T_a on input state ρ. Using the classical simulation algorithm, the conditional probability of obtaining outcome s for the measurement given the state $A_\alpha \in \mathscr{A}$ is $Q_a(s|U\cdot\alpha)$. Therefore, the probability of obtaining outcome s given a measurement of T_a on state ρ as predicted by the classical simulation algorithm is

$$P^{(Sim)}_{\rho,U,a}(s) = \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha) Q_a(s|U\cdot\alpha). \tag{4.7}$$

The outcome probability predicted by the Born rule, $P^{(QM)}_{\rho,U,a}(s)$, is

$$\begin{aligned}
\mathrm{Tr}\left(\Pi_{a,s} U\rho U^\dagger\right) &= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha)\mathrm{Tr}(\Pi_{a,s} U A_\alpha U^\dagger) \\
&= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha)\mathrm{Tr}(\Pi_{a,s} A_{U\cdot\alpha}) \\
&= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha) Q_a(s|U\cdot\alpha).
\end{aligned} \tag{4.8}$$

Here in the first line we use the expansion of ρ, Eq. (4.2), in the second line we use Lemma 14, and in the third line Eqs. (4.4) and (4.6). Comparing Eq. (4.7) and Eq. (4.8), we see that the classical simulation algorithm reproduces the outcome probabilities predicted by the Born rule for a single layer of a QCM circuit.

Now we turn the the postmeasurement state ρ'. The postmeasurement state predicted by quantum mechanics is

$$\rho'^{(QM)} = \frac{\Pi_{a,s} U\rho U^\dagger \Pi_{a,s}}{\mathrm{Tr}(\Pi_{a,s} U\rho U^\dagger)}.$$

Here the numerator is

$$\begin{aligned}\Pi_{a,s}U\rho U^\dagger\Pi_{a,s} &= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha)\Pi_{a,s}UA_\alpha U^\dagger\Pi_{a,s}\\ &= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha)\Pi_{a,s}A_{U\cdot\alpha}\Pi_{a,s}\\ &= \sum_{\alpha\in\mathscr{V}} p_\rho(\alpha)\sum_{\beta\in\mathscr{V}} q_{U\cdot\alpha,a}(\beta,s)A_\beta,\end{aligned}$$

and so

$$\rho'^{(QM)} = \frac{\sum_\alpha p_\rho(\alpha)\sum_\beta q_{U\cdot\alpha,a}(\beta,s)A_\beta}{\sum_\alpha p_\rho(\alpha)Q_a(s\,|\,U\cdot\alpha)}. \tag{4.9}$$

Using the classical simulation algorithm, the probability of obtaining outcome s and state A_β given a Clifford unitary U followed by measurement of T_a on state ρ is $P_{\rho,U,a}(\beta,s) = P_{\rho,U,a}(\beta|s)P_{\rho,U,a}(s)$. But $P_{\rho,U,a}(\beta,s) = \sum_\alpha p_\rho(\alpha)P_{U,a}(\beta,s|\alpha) = \sum_\alpha p_\rho(\alpha)q_{U\cdot\alpha,a}(\beta,s)$ and $P_{\rho,U,a}(\beta|s) = p_{\rho'}(\beta)$. Therefore, the postmeasurement state according to the classical simulation algorithm is

$$\begin{aligned}\rho'^{(Sim)} &= \sum_{\beta\in\mathscr{V}} p_{\rho'}(\beta)A_\beta = \sum_{\beta\in\mathscr{V}} \frac{P_{\rho,U,a}(\beta,s)}{P_{\rho,U,a}(s)}A_\beta\\ &= \sum_{\beta\in\mathscr{V}} \frac{\sum_\alpha p_\rho(\alpha)q_{U\cdot\alpha,a}(\beta,s)}{\sum_\alpha p_\rho(\alpha)Q_a(s\,|\,U\cdot\alpha)}A_\beta.\end{aligned} \tag{4.10}$$

This agrees with Eq. (4.9) above. Therefore, the classical simulation algorithm also reproduces the postmeasurement state predicted by quantum mechanics for a single layer of a QCM circuit.

Now let $\rho(t)$ denote the state after $t-1$ layers of the circuit. Then the above shows that the classical simulation algorithm correctly reproduces the Born rule probabilities $P_{\rho,U_t,a_t}(s_t\,|\,s_1,s_2,\ldots,s_{t-1})$ 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 QCM circuit. □

Although this classical simulation algorithm can be used to simulate any quantum computation, in contrast to the earlier algorithms of Sections 2.6 and 3.5, we have not proved the computational efficiency of the Algorithm 4.1. In fact, if it were efficient then all quantum computation could be efficiently classically

simulated, and there could be no quantum computational speedup. If a quantum computational speedup exists, as is widely assumed, then this classical simulation algorithm cannot be efficient for arbitrary quantum computations.

A proof of efficiency would require we prove that the algorithm satisfies the four efficiency criteria (E0)–(E3) defined in Section 2.6. If, as we strongly suspect, Algorithm 4.1 is not efficient in general, then at least one of these criteria must fail. Inefficiency in classical simulation must come from a source other than negativity in the representation of the input state. This inefficiency could come from difficulty in obtaining samples from the probability function defining the input state, from inefficient update rules under Clifford unitaries and/or Pauli measurements, or from an inability to even specify a phase space point with a polynomial number of bits.

4.3 A ~~quasi~~-probability representation for quantum computation with magic states

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. (4.5) has the form of a law of total probability. The dynamics described by Lemma 14 and Eq. (4.4) 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.

Before continuing we need to introduce some nomenclature and notation for some classes of quantum circuits.

Definition 4 *A stabilizer circuit is a quantum circuit consisting of only Clifford unitaries and Pauli measurements. A stabilizer circuit is represented by a sequence* $C_S = (U_1, a_1, U_2, a_2, \dots)$ *with* $U_1, U_2, \cdots \in \mathscr{C}\ell$ *specifying Clifford unitaries to be applied and* $a_1, a_2, \cdots \in E$ *specifying the Pauli measurements to be performed. Denote by* $\mathfrak{C}_S$ *the set of all stabilizer circuits.*

This is the most general kind of quantum circuit allowed within the QCM

framework with all free operations in the resource theory of stabilizer quantum computation [42] accessible. Stabilizer circuits applied to magic input states allow for universal quantum computation as shown in Section 2.2. Note that since adaptivity is allowed in QCM, a measurement a_j or a unitary U_j is not necessarily fixed in advance, they could be functions of the outcomes of the prior measurements. That is, if the outcomes of the first $j-1$ measurements are $s_1, s_2, \ldots, s_{j-1} \in \mathbb{Z}_2$, then U_j is a map from $(s_1, s_2, \ldots, s_{j-1})$ to $\mathscr{C}\ell$ and a_j is a map from $(s_1, s_2, \ldots, s_{j-1})$ to E.

Definition 5 *A Pauli circuit is a circuit consisting of a sequence of (not necessarily commuting) Pauli measurements. A Pauli circuit is represented by a sequence $C_P = (a_1, a_2, \ldots)$ with $a_1, a_2, \cdots \in E$ labeling the measurements to be performed. Denote by $\mathfrak{C}_P$ the set of all Pauli circuits.*

Again, adaptivity allows the measurements chosen to be conditioned on the outcomes of prior measurements. Any Pauli circuit can be identified with a stabilizer circuit in which each Clifford unitary in the circuit is the identity. With this identification we have the inclusion $\mathfrak{C}_P \subset \mathfrak{C}_S$.

Definition 6 *A commuting circuit is a Pauli circuit in which the measurements pairwise commute. Denote by $\mathfrak{C}_I$ the set of all commuting circuits with elements $C_I \in \mathfrak{C}_I$.*

The "I" in $\mathfrak{C}_I$ is for instantaneous as this can be seen as an analogue in the magic state model of the quantum complexity class IQP (Instantaneous Quantum Polynomial time circuits) [81].

Definition 7 *For a circuit $C_S \in \mathfrak{C}_S$ (resp. $C_P \in \mathfrak{C}_P$, $C_I \in \mathfrak{C}_I$), an effect of the circuit is a pair $(C_S, \boldsymbol{s})$ (resp. $(C_P, \boldsymbol{s})$, $(C_I, \boldsymbol{s})$) where $\boldsymbol{s} \in \mathbb{Z}_2^*$ is a binary string labeling a possible sequence of outcomes of the measurements performed in the circuit. The set of all effects, denoted by $\mathfrak{E}_S$ (resp. $\mathfrak{E}_P$, $\mathfrak{E}_I$) is the subset of $\mathfrak{C}_S \times \mathbb{Z}_2^*$ where the length of the outcome sequence matches the number of measurements.*[2]

[2]Since circuits can be adaptive—gates and measurements can depend on the outcomes of prior measurements—not every effect corresponding to a given circuit need have outcome sequences $\mathbf{s} \in \mathbb{Z}_2^*$ of the same length and not every outcome sequence of any given length will necessarily correspond to a valid effect.

For an effect $(C_S, \mathbf{s})$ of a circuit $C_S \in \mathfrak{C}_S$ corresponding to measurement outcomes $\mathbf{s} = (s_1, s_2, \ldots, s_N) \in \mathbb{Z}_2^N$, we can define the operator

$$\Pi(C_S, \mathbf{s}) := \Pi_{a_N, s_N} U_N \cdots \Pi_{a_2, s_2} U_2 \Pi_{a_1, s_1} U_1. \tag{4.11}$$

The corresponding operators for circuits $C_P \in \mathfrak{C}_P$ and $C_I \in \mathfrak{C}_I$ are defined similarly except without the Clifford unitaries. It can be easily shown that, according to the Born rule, the probability of obtaining measurement outcomes $\mathbf{s}$ for the measurements of a circuit $C_S \in \mathfrak{C}_S$ applied to an input state ρ is given by

$$\mathrm{Tr}\left(\Pi(C_S, \mathbf{s}) \rho \Pi(C_S, \mathbf{s})^\dagger\right). \tag{4.12}$$

An analogous statement holds for Pauli circuit and for commuting circuits.

All three classical simulation algorithms for quantum computation with magic states discussed in this thesis so far, Algorithms 2.1, 3.1, and 4.1, apply to a set of input "states" which forms a polytope in the space $\mathscr{O}(d^n)$. In the case of Algorithm 4.1, 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 point operators, ConvHull $\{A_u \mid u \in V\}$ for Algorithm 2.1 and ConvHull $\left\{A_\Omega^\gamma \mid (\Omega, \gamma) \in \mathscr{V}\right\}$ for Algorithm 3.1.

The fundamental 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 $X \in \Lambda$, $U \in \mathscr{C}\ell$, $a \in E$, and $s \in \mathbb{Z}_d$, we have $UXU^\dagger \in \Lambda$, 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 is proven in Lemma 1, Lemma 8, and Theorem 6 part (iv) for the cases considered here. This 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 1, 7, and 8,

and in Theorem 6 parts (ii) and (iii). 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 $\mathscr{O}(2^n)$ satisfying these two properties is the set

$$\tilde{\Lambda} = \left\{ X \in \mathscr{O}(2^n) \mid \mathrm{Tr}\left(\Pi(C_S, \mathbf{s}) X \Pi(C_S, \mathbf{s})^\dagger\right) \geq 0 \; \forall (C_S, \mathbf{s}) \in \mathfrak{E}_S \right\}. \tag{4.13}$$

That is, $\tilde{\Lambda}$ satisfies these two properties, and for any set $\Omega \subset \mathscr{O}(2^n)$ satisfying these two properties, $\Omega \subset \tilde{\Lambda}$. Therefore, any polytope in $\mathscr{O}(2^n)$ 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 effects of stabilizer 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 entirely captured by the following theorem.

Theorem 8 $\tilde{\Lambda} = \Lambda$.

The rest of this section is devoted to proving this theorem. The proof of Theorem 8 requires several lemmas. First,

Lemma 15 *For any effect $(C_S, \boldsymbol{s}) \in \mathfrak{E}_S$, there is an effect $(C_P, \boldsymbol{t}) \in \mathfrak{E}_P$ such that for any $X \in \mathscr{O}(2^n)$,*

$$Tr\left(\Pi(C_S, \boldsymbol{s}) X \Pi(C_S, \boldsymbol{s})^\dagger\right) = Tr\left(\Pi(C_P, \boldsymbol{t}) X \Pi(C_P, \boldsymbol{t})^\dagger\right).$$

Lemma 15 is a result of the fact that in the model of quantum computation with magic states the Clifford unitaries are redundant. Any stabilizer circuit is equivalent to a circuit consisting of only Pauli measurements and so the Pauli measurements alone are sufficient for universal quantum computation. The proof of

this fact is given in Ref. [38, §IV.A.]. We reproduce the relevant aspect of the proof here.

Proof sketch of Lemma 15. For a stabilizer effect $(C_S, \mathbf{s}) \in \mathfrak{E}_S$, consider the operator

$$\Pi(C_S, \mathbf{s}) = \Pi_{a_N, s_N} U_N \cdots \Pi_{a_s, s_2} U_2 \Pi_{a_1, s_1} U_1.$$

with $U_1, U_2, \cdots \in \mathscr{C}\ell$, $a_1, a_2, \cdots \in E$, and $s_1, s_2, \cdots \in \mathbb{Z}_2$. Since Clifford unitaries are defined as the set of unitary operations that map Pauli operators to Pauli operators under conjugation, the Clifford unitaries in this operator can be propagated to the left giving the equivalent operator

$$\Pi(C_S, \mathbf{s}) = U \Pi_{\bar{a}_N, \bar{s}_N} \cdots \Pi_{\bar{a}_2, \bar{s}_2} \Pi_{\bar{a}_1, \bar{s}_1}$$

with $U \in \mathscr{C}\ell$, $\bar{a}_1, \bar{a}_2, \cdots \in E$, and $\bar{s}_1, \bar{s}_2, \cdots \in \mathbb{Z}_2$. Then

$$\begin{aligned} \mathrm{Tr}\left(\Pi(C_S, \mathbf{s}) X \Pi(C_S, \mathbf{s})^\dagger\right) &= \mathrm{Tr}\left(U \Pi_{\bar{a}_N, \bar{s}_N} \cdots \Pi_{\bar{a}_2, \bar{s}_2} \Pi_{\bar{a}_1, \bar{s}_1} X \Pi_{\bar{a}_1, \bar{s}_1} \Pi_{\bar{a}_2, \bar{s}_2} \cdots \Pi_{\bar{a}_N, \bar{s}_N} U^\dagger\right) \\ &= \mathrm{Tr}\left(\Pi_{\bar{a}_N, \bar{s}_N} \cdots \Pi_{\bar{a}_2, \bar{s}_2} \Pi_{\bar{a}_1, \bar{s}_1} X \Pi_{\bar{a}_1, \bar{s}_1} \Pi_{\bar{a}_2, \bar{s}_2} \cdots \Pi_{\bar{a}_N, \bar{s}_N} U^\dagger U\right) \\ &= \mathrm{Tr}\left(\Pi(C_P, \mathbf{t}) X \Pi(C_P, \mathbf{t})^\dagger\right) \end{aligned}$$

where $(C_P, \mathbf{t}) \in \mathfrak{E}_P$ is the effect corresponding to the circuit $C_P = (\bar{a}_1, \bar{a}_2, \dots)$ giving outcomes $\bar{s}_1, \bar{s}_2, \dots, \bar{s}_N$. □

Lemma 16 *For any effect $(C_P, \boldsymbol{s}) \in \mathfrak{E}_P$, there exists a set of coefficients $c(C_I, \boldsymbol{t})$ with $c(C_I, \boldsymbol{t}) \geq 0 \; \forall (C_I, \boldsymbol{t}) \in \mathfrak{E}_I$ and $\sum_{(C_I, \boldsymbol{t})} c(C_I, \boldsymbol{t}) = 1$ such that for any $X \in \mathscr{O}(2^n)$ we have*

$$Tr\left(\Pi(C_P, \boldsymbol{s}) X \Pi(C_P, \boldsymbol{s})^\dagger\right) = \sum_{(C_I, \boldsymbol{t}) \in \mathfrak{E}_I} c(C_I, \boldsymbol{t}) Tr\left(\Pi(C_I, \boldsymbol{t}) X \Pi(C_I, \boldsymbol{t})^\dagger\right).$$

The proof of Lemma 16 requires the following additional lemma:

Lemma 17 *Let T_a be a Pauli observable and*

$$S = \langle (-1)^{\gamma(a_1)} T_{a_1}, (-1)^{\gamma(a_2)} T_{a_2}, \dots, (-1)^{\gamma(a_m)} T_{a_m} \rangle$$

be a stabilizer group. Suppose that T_a does not commute with one of the generators

T_{a_k}. Then there exists a n-qubit Clifford unitary $U \in \mathscr{C}\ell$ such that $USU^\dagger$ and $UT_aU^\dagger$ act only on qubits $1,2,\ldots,m$.

Proof of lemma 17. The stabilizer group S can be identified with an isotropic subspace $\tilde{I} = \langle a_1, a_2, \ldots, a_m \rangle \subset E$ with corresponding noncontextual value assignment $\gamma : \tilde{I} \to \mathbb{Z}_2$. Without loss of generality, we can assume that $[a_1, a] \neq 0$ and $[a_k, a] = 0$ for $k \in \{2, \ldots, m\}$. By Witt's lemma [68, §20], the linear map that sends a_j to z_j for $1 \leq j \leq m$ and a to x_1 can be extended to a symplectic map $M : E \to E$. Then since $\mathrm{Sp}(2n, \mathbb{Z}_2) \simeq \mathscr{C}\ell/\mathscr{P}$, there exists a Clifford unitary $U \in \mathscr{C}\ell$ such that $UT_bU^\dagger = \pm T_{Mb}$ for all $b \in E$. In particular, $UT_aU^\dagger = \pm T_{x_1}$ and $UT_{a_j}U^\dagger = \pm T_{z_j}$. □

Proof of Lemma 16. We will prove by induction that for any effect $(C_P, \mathbf{s}) \in \mathfrak{E}_P$ and any $X \in \mathscr{O}(2^n)$,

$$\Pi(C_P, \mathbf{s}) X \Pi(C_P, \mathbf{s})^\dagger = \sum_{(C_I, \mathbf{t}) \in \mathfrak{E}_I} c(C_I, \mathbf{t}) U(C_I, \mathbf{t}) \Pi(C_I, \mathbf{t}) X \Pi(C_I, \mathbf{t})^\dagger U(C_I, \mathbf{t})^\dagger \quad (4.14)$$

with $U(C_I, \mathbf{t}) \in \mathscr{C}\ell$, $c(C_I, \mathbf{t}) \geq 0 \; \forall (C_I, \mathbf{t}) \in \mathfrak{E}_I$, and $\sum_{(C_I, \mathbf{t})} c(C_I, \mathbf{t}) = 1$. The base case is where C_P consists of a single Pauli measurements. In this case $C_P \in \mathfrak{C}_I$ so Eq. (4.14) holds trivially.

Suppose Eq. (4.14) holds for any effect $(C_P, \mathbf{s})$ up to depth D. For any effect $\left(C_P^{(D+1)}, \tilde{\mathbf{s}}\right)$ of depth $D+1$, $\Pi\left(C_P^{(D+1)}, \tilde{\mathbf{s}}\right) = \Pi_{a,s} \Pi\left(C_P^{(D)}, \mathbf{s}\right)$ where $\left(C_P^{(D)}, \mathbf{s}\right)$ has depth D. Then

$$\begin{aligned} \Pi\left(C_P^{(D+1)}, \tilde{\mathbf{s}}\right) X \Pi\left(C_P^{(D+1)}, \tilde{\mathbf{s}}\right)^\dagger &= \Pi_{a,s} \Pi\left(C_P^{(D)}, \mathbf{s}\right) X \Pi\left(C_P^{(D)}, \mathbf{s}\right)^\dagger \Pi_{a,s} \\ &= \sum_{(C_I, \mathbf{t}) \in \mathfrak{E}_I} c(C_I, \mathbf{t}) \Pi_{a,s} U(C_I, \mathbf{t}) \Pi(C_I, \mathbf{t}) X \Pi(C_I, \mathbf{t})^\dagger U(C_I, \mathbf{t})^\dagger \Pi_{a,s}. \end{aligned} \quad (4.15)$$

A generic term in this sum, up to a nonnegative constant factor, has the form

$$\Pi_{a,s} U \Pi(C_I, \mathbf{t}) X \Pi(C_I, \mathbf{t})^\dagger U^\dagger \Pi_{a,s} \quad (4.16)$$

with $(C_I, \mathbf{t}) \in \mathfrak{E}_I$, $a \in E$, $s \in \mathbb{Z}_2$, and $U \in \mathscr{C}\ell$. Consider the operator $\Pi_{a,s} U \Pi(C_I, \mathbf{t})$. Since Clifford group operations are defined by the property that they map Pauli operators to Pauli operators under conjugation, the unitary U can be propagated to

the other side of the projector $\Pi_{a,s}$ giving a different projector:

$$\Pi_{a,s} U \Pi(C_I, \mathbf{t}) = U \Pi_{a',s'} \Pi(C_I, \mathbf{t}). \tag{4.17}$$

Now we have four cases:

Case (1): If $(-1)^{s'} T_{a'}$ is an element of the stabilizer group $S = \langle (-1)^{t_1} T_{b_1}, (-1)^{t_2} T_{b_2}, \dots \rangle$ corresponding to the effect $(C_I, \mathbf{t})$ where $b_1, b_2, \dots \in E$ are the measurements in the circuit C_I and $t_1, t_2, \dots \in \mathbb{Z}_2$ are the measurement outcomes, then

$$U \Pi_{a',s'} \Pi(C_I, \mathbf{t}) = U \Pi(C_I, \mathbf{t}).$$

Case (2): If $(-1)^{s'+1} T_{a'}$ is an element of the stabilizer group S, then

$$U \Pi_{a',s'} \Pi(C_I, \mathbf{t}) = 0.$$

Case (3): If $T_{a'}$ commutes with $T_{b_1}, T_{b_2}, \dots$ but neither of the conditions required for case (1) or case (2) hold, then

$$U \Pi_{a',s'} \Pi(C_I, \mathbf{t}) = U \Pi(C_I', (\mathbf{t}, s'))$$

where $(C_I', (\mathbf{t}, s')) \in \mathfrak{E}_I$ and $C_I' \in \mathfrak{C}_I$ is the circuit obtained by appending the measurement a' to the circuit C_I.

Case (4): The last case is where $T_{a'}$ anticommutes with at least one element of S. In this case, by lemma 17, there exists a Clifford unitary $V \in \mathscr{C}\ell$ such that

$$V \Pi_{a',s'} V^\dagger = \Pi_{\tilde{a},\tilde{s}} \otimes \mathbb{I}_{n-m} \quad \text{and} \quad V \Pi(C_I, \mathbf{t}) V^\dagger = \Pi(\tilde{C}_I, \tilde{\mathbf{t}}) \otimes \mathbb{I}_{n-m}$$

where $\Pi_{\tilde{a},\tilde{s}}$ and $\Pi(\tilde{C}_I, \tilde{\mathbf{t}})$ act only on the first m qubits where m is the dimension of the isotropic subspace corresponding to the stabilizer group S. $\Pi(\tilde{C}_I, \tilde{\mathbf{t}})$ is a product of m pairwise commuting Pauli projectors acting on m qubits, i.e., it is a projection onto a stabilizer state so it can be written as $\Pi(\tilde{C}_S, \tilde{\mathbf{t}}) = |\sigma\rangle \langle\sigma|$ for some m-qubit

stabilizer state $|\sigma\rangle$. Then

$$\begin{aligned} U\Pi_{a',s'}\Pi(C_I,\mathbf{t}) &= U\Pi_{a',s'}V^\dagger V\Pi(C_I,\mathbf{t}) \\ &= UV^\dagger(\Pi_{\tilde{a},\tilde{s}}\otimes\mathbb{I}_{n-m})(|\sigma\rangle\langle\sigma|\otimes\mathbb{I}_{n-m})V \\ &= UV^\dagger(\Pi_{\tilde{a},\tilde{s}}|\sigma\rangle\langle\sigma|\otimes\mathbb{I}_{n-m})V. \end{aligned}$$

Since $T_{\tilde{a}}$ anticommutes with an element of the stabilizer of the stabilizer state $|\sigma\rangle$, a measurement of $T_{\tilde{a}}$ on the state $|\sigma\rangle$ would yield 0 or 1 with equal probability and the postmeasurement state is another stabilizer state depending on the outcome of the measurement [4, §10.5.3]. Therefore, $\Pi_{\tilde{a},\tilde{s}}|\sigma\rangle = \frac{1}{\sqrt{2}}|\sigma_{\tilde{s}}\rangle$ for some m-qubits stabilizer state $|\sigma_{\tilde{s}}\rangle$. Since the Clifford group acts transitively on the set of stabilizer states, there exists a m-qubit Clifford group operation W such that $W|\sigma\rangle = |\sigma_{\tilde{s}}\rangle$. Then we have

$$\begin{aligned} U\Pi_{a',s'}\Pi(C_I,\mathbf{t}) &= \frac{1}{\sqrt{2}}UV^\dagger(|\sigma_{\tilde{s}}\rangle\langle\sigma|\otimes\mathbb{I}_{n-m})V \\ &= \frac{1}{\sqrt{2}}UV^\dagger(|\sigma_{\tilde{s}}\rangle\langle\sigma|W^\dagger W\otimes\mathbb{I}_{n-m})V \\ &= \frac{1}{\sqrt{2}}UV^\dagger(|\sigma_{\tilde{s}}\rangle\langle\sigma_{\tilde{s}}|\otimes\mathbb{I}_{n-m})(W\otimes\mathbb{I}_{n-m})V. \end{aligned}$$

Finally, the n-qubit Clifford group operations $(W\otimes\mathbb{I}_{n-m})$ and V can be propagated to the other side of the projector giving another projector $\Pi(\bar{C}_I,\bar{\mathbf{t}})$ with $\bar{C}_I\in\mathfrak{C}_I$. I.e.

$$U\Pi_{a',s'}\Pi(C_I,\mathbf{t}) = \frac{1}{\sqrt{2}}UV^\dagger(W\otimes\mathbb{I}_{n-m})V\Pi(\bar{C}_I,\bar{\mathbf{t}}). \tag{4.18}$$

To summarize,

$$U\Pi_{a',s'}\Pi(C_I,\mathbf{t}) = \begin{cases} U\Pi(C_I,\mathbf{t}) & \text{if } (-1)^{s'}T_{a'}\in S \\ 0 & \text{if } (-1)^{s'+1}T_{a'}\in S \\ U\Pi(C_I',(\mathbf{t},s')) & \text{if } [T_{a'},S]=0 \text{ and } \pm T_{a'}\notin S \\ \frac{1}{\sqrt{2}}UV^\dagger(W\otimes\mathbb{I}_{n-m})V\Pi(\bar{C}_I,\bar{\mathbf{t}}) & \text{otherwise.} \end{cases} \tag{4.19}$$

Putting this back into Eq. (4.15), we can see that $\Pi\left(C_S^{(D+1)}\right)$ can be written in the form Eq. (4.14). Thus, by the principle of induction, Eq. (4.14) holds for any effect of any Pauli circuit $C_P \in \mathfrak{C}_P$.

Taking a trace of eq. (4.14), we get

$$\begin{aligned}\mathrm{Tr}\left(\Pi(C_P,\mathbf{s})\rho\Pi(C_P,\mathbf{s})^\dagger\right) &= \sum_{(C_I,\mathbf{t})\in\mathfrak{E}_I} c(C_I,\mathbf{t})\mathrm{Tr}\left(U(C_I,\mathbf{t})\Pi(C_I,\mathbf{t})X\Pi(C_I,\mathbf{t})^\dagger U(C_I,\mathbf{t})^\dagger\right)\\ &= \sum_{(C_I,\mathbf{t})\in\mathfrak{E}_I} c(C_I,\mathbf{t})\mathrm{Tr}\left(\Pi(C_I,\mathbf{t})X\Pi(C_I,\mathbf{t})^\dagger\right).\end{aligned}$$

This completes the proof. □

Proofs of similar propositions to Lemma 16 are given in Refs. [82] and [83].

The final lemma required is the following.

Lemma 18 *For any effect* $(C_I,\boldsymbol{s}) \in \mathfrak{E}_I$*,*

$$\Pi(C_I,\boldsymbol{s}) = \sum_{\sigma\in\mathscr{S}} c(\sigma)\sigma$$

with $0 \le c(\sigma) \le 1\ \forall\sigma \in \mathscr{S}$ *and* $\sum_{\sigma\in\mathscr{S}} c(\sigma) = 1$.

Proof sketch of Lemma 18. Let $S = \langle (-1)^{s_1}T_{b_1}, (-1)^{s_2}T_{b_2}, \dots\rangle$ be the stabilizer group corresponding to the measurements $b_1, b_2, \dots \in E$ of the circuit C_I giving outcomes $s_1, s_2, \dots \in \mathbb{Z}_2$. S can be identified with a pair $(\tilde{I},\tilde{\gamma})$ where $\tilde{I} = \langle b_1, b_2, \dots\rangle \subset E$ is an isotropic subspace and $\tilde{\gamma}: \tilde{I} \to \mathbb{Z}_2$ is a noncontextual value assignment on $\tilde{I}$. Then $\Pi(C_I,\mathbf{s}) = A_{\tilde{I}}^{\tilde{\gamma}}$.

Let $I \subset E$ be a maximal isotropic subspace containing $\tilde{I}$ and let Γ denote the set of noncontextual value assignments on I which agree with $\tilde{\gamma}$ when restricted to $\tilde{I}$. Then

$$A_{\tilde{I}}^{\tilde{\gamma}} = \frac{1}{|\Gamma|}\sum_{\gamma\in\Gamma} A_I^\gamma.$$

This follows from character orthogonality (see Appendix A of Ref. [2] for the details of the proof). □

We can now prove the main theorem of this section.

Proof of Theorem 8. The inclusion $\tilde{\Lambda} \subseteq \Lambda$ is clear from the definitions of $\tilde{\Lambda}$ and Λ. It remains only to prove $\Lambda \subseteq \tilde{\Lambda}$. Supppose $X \in \Lambda$. Then by definition $\mathrm{Tr}(X) = 1$ and $\mathrm{Tr}(\sigma X) \geq 0$ for all $\sigma \in \mathscr{S}$. Let $(C_S, \mathbf{s}) \in \mathfrak{E}_S$ be a stabilizer effect. Then

$$\begin{aligned}
\mathrm{Tr}\left(\Pi(C_S,\mathbf{s})X\Pi(C_S,\mathbf{s})^\dagger\right) &= \mathrm{Tr}\left(\Pi(C_P,\tilde{\mathbf{s}})X\Pi(C_P,\tilde{\mathbf{s}})^\dagger\right) \\
&= \sum_{(C_I,\mathbf{t})\in\mathfrak{E}_I} c(C_I,\mathbf{t})\mathrm{Tr}\left(\Pi(C_I,\mathbf{t})X\Pi(C_I,\mathbf{t})^\dagger\right) \\
&= \sum_{(C_I,\mathbf{t})\in\mathfrak{E}_I} c(C_I,\mathbf{t})\mathrm{Tr}\left(\Pi(C_I,\mathbf{t})X\right) \\
&= \sum_{\sigma\in\mathscr{S}} c(\sigma)\mathrm{Tr}(\sigma X) \geq 0
\end{aligned}$$

with $(C_P, \tilde{\mathbf{s}}) \in \mathfrak{E}_P$, $c(C_I, \mathbf{t}) \geq 0 \; \forall (C_I, \mathbf{t}) \in \mathfrak{E}_I$, and $c(\sigma) \geq 0 \; \forall \sigma \in \mathscr{S}$. Here the first line follows from Lemma 15, the second line from Lemma 16, the third line from the fact that the measurements in each $C_I \in \mathfrak{C}_I$ pairwise commute, the fourth line from Lemma 18, and the last inequality follows from the fact $X \in \Lambda$.

Therefore, $X \in \Lambda$ implies $X \in \tilde{\Lambda}$, and thus $\Lambda \subseteq \tilde{\Lambda}$. □

Theorem 8 shows that the set $\tilde{\Lambda}$, which has a natural connection to quantum computation with magic states, can be used to defined a hidden variable model for QCM, and further, that the resulting hidden variable model is equivalent to the one defined by Theorem 6.

Chapter 5

Discussion

In this chapter, we will summarize the results presented thus far, and we will discuss some open problems relating to them. Since these open problems are not the main focus of this thesis, we keep the discussion here informal.

In this thesis, we have discussed several quasiprobability representations for finite-dimensional quantum mechanics and considered their application to the study of quantum computation with magic states. The generalized Wigner function defined in Chapter 3 unifies several of these representations, including the odd dimensional discrete Wigner function [32, 53] and previously defined quasiprobability representations for qubits like the eight state model [62] and the representation based on stabilizer quasimixtures [64]. It provides a discrete Wigner function that applies to any number of qudits of any dimension, where previously defined discrete Wigner functions applied to quantum computation only in limited cases. The classical simulation algorithm for quantum computation with magic states on qubits descibed in Section 3.5 establishes negativity in the Wigner function of the input state of a QCM circuit as a necessary condition for quantum comptutation with magic states to exhibit a computational speedup over classical computation, extending the previous result of Ref. [1] which applied only to odd-dimensional qudits.

In Chapter 4, we have defined a hidden variable model for quantum computation with magic states which represents all n-qubit quantum states, as well as all QCM dynamics, in terms of (nonnegative) probabilities. This model is struc-

turally similar to a quasiprobability representations of finite-dimensional quantum mechanics, the key difference between this representation and previous quasiprobability representations being the lack of negativity—it could be called a probability representation. This hidden variable model leads to a classical simulation algorithm for quantum computation with magic states that can simulate any QCM circuit applied to any input quantum state, though we make no claims of efficiency of this classical simulation algorithm and we strongly suspect that it is not efficient in general. In fact, since this classical simulation algorithm applies to any quantum computation, if a quantum computational advantage for quantum computation over classical computation exists, as is widely believed, then the classical simulation algorithm must be inefficient in general.

The existence of a classical seeming hidden variable model for a universal model of quantum computation is unintuitive. In previous quasiprobility representations applied to quantum computation with magic states such as the discrete Wigner functions, states would be partitioned into two subsets. The positively representable states would be considered more classical as these are the states for which a hidden variable model applies and further, these are the states for which QCM can be efficiently simulated [1, 2]. The states which cannot be positively represented are simultaneously the states for which the hidden variable model fails, and those for which a quantum computational speedup in QCM is possible. With the hidden variable model of Theorem 6, we no longer have this partitioning of quantum states—all states are positively representable. Therefore, negativity in a quasiprobability representation alone is not sufficient for a quantum computational speedup as the presence or absence of negativity can depend on the choice of quasiprobability representation.

The unintuitive nature of the existence of this hidden variable model is further evidenced by its near misses with many no-go theorems. For example, the Pusey-Barrett-Rudolph (PBR) theorem [84], which states that, under certain assumptions, there can be no ψ-epistemic model for quantum theory. With Theorem 6, we have a ψ-epistemic model. This apparent contradiction is resolved by the fact that the model of Theorem 6 considers only sequences of Pauli measurements, it does not allow more general measurements. This restriction does not affect the universal-

ity of quantum computation. Also, the model does not satisfy the assumption of preparation independence required for the PBR 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 models of quantum computation where, from a resource theory perspective, the free states and operations are the stabilizer states and stabilizer operations [42], not local states and local operations. Further, the memory lower bound of Karanjai, Wallman, and Bartlett [39] shows that no hidden variable model which spawns a classical simulation algorithm for quantum computation with magic states on qubits like that of Algorithm 4.1 can satisfy the assumption of preparation independence.

There is another result which states, under certain assumptions, there can be no quasiprobability representation for quantum theory that is entirely nonnegative [74, 75, 85]. Also, under similar assumptions, a nonnegative quasiprobability representation is equivalent to a noncontextual hidden variable model [35]. Since quantum mechanics of multiple qubits is known to be contextual, there should be no nonnegative quasiprobability representation for multiqubit states. The resolution here is in the fact that these results apply to quasiprobability representations for which that map from states to quasiprobability distributions is one-to-one and linear. This map for the model defined by Theorem 6 is not one-to-one and linear since the probability function p_ρ is not unique (the discussion of the SW criteria and contextuality for the hidden variable model of Chapter 4 mirrors that of the quasiprobability representation of Chapter 3 given in Section 3.6). Therefore, it is the same modification made to the SW criteria that allowed for contextuality of the model defined in Chapter 3 which allows the model of Chapter 4 to skirt these no-go results.

To summarize, we have presented two models for describing quantum computation with magic states on qubits. Both models lead to classical simulation algorithms for quantum computation. Algorithm 3.1 is efficient, but it applies only to QCM circuits on the strict subset of input states which are positively representable. There is no such restriction on the applicability of Algorithm 4.1, it can simulate any quantum computation, but it is not necessarily efficient. Both of these models have intricate structure, as partially elucidated in Section 3.3 and Section 4.3,

which is closely related to the geometry of polytopes in spaces of operators. These models also open several questions surrounding their relation to contextuality, negativity, quantum computation, and geometry which could serve as the basis for future study.

Bibliography

[1] Victor Veitch, Christopher Ferrie, David Gross, and Joseph Emerson. Negative quasi-probability as a resource for quantum computation. *New Journal of Physics*, 14(11):113011, 2012. doi:10.1088/1367-2630/14/11/113011. → pages iii, 3, 4, 5, 19, 25, 27, 34, 73, 94, 95

[2] Robert Raussendorf, Juani Bermejo-Vega, Emily Tyhurst, Cihan Okay, and Michael Zurel. Phase-space-simulation method for quantum computation with magic states on qubits. *Physical Review A*, 101(1):012350, 2020. doi:10.1103/PhysRevA.101.012350. → pages iii, vi, 6, 40, 45, 63, 65, 66, 67, 92, 95, 118

[3] Michael Zurel, Cihan Okay, and Robert Raussendorf. A hidden variable model for universal quantum computation with magic states on qubits, 2020. URL https://arxiv.org/abs/2004.01992. arXiv:2004.01992. → pages iv, vi, 6, 73, 75

[4] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000. → pages xi, 8, 12, 13, 38, 76, 91, 110

[5] Hussain Anwar, Earl T Campbell, and Dan E Browne. Qutrit magic state distillation. *New Journal of Physics*, 14(6):063006, 2012. doi:10.1088/1367-2630/14/6/063006. → pages xi, 20

[6] N. David Mermin. Hidden variables and the two theorems of john bell. *Reviews of Modern Physics*, 65(3):803–815, 1993. doi:10.1103/RevModPhys.65.803. → pages xii, 29, 30, 31, 32

[7] William M. Kirby and Peter J. Love. Contextuality test of the nonclassicality of variational quantum eigensolvers. *Physical Review Letters*, 123(20), 2019. doi:10.1103/physrevlett.123.200501. → pages xiii, 49, 50

[8] Yuri I. Manin. *Vychislimoe i nevychislimoe*. Sov. Radio., 1980. → page 1

[9] Richard P Feynman. Simulating physics with computers. *International Journal of Theoretical Physics*, 21(6):467–488, 1982. doi:10.1007/BF02650179. → page 1

[10] P. W. Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings, 35th Annual Symposium on Foundations of Computer Science*, pages 124–134, 1994. → page 1

[11] Lov K. Grover. A fast quantum mechanical algorithm for database search. In *Proceedings, 28th Annual ACM Symposium on the Theory of Computing (STOC)*, pages 212–219, 1996. → page 1

[12] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, Joseph C. Bardin, Rami Barends, Rupak Biswas, Sergio Boixo, Fernando G. S. L. Brandao, David A. Buell, Brian Burkett, Yu Chen, Zijun Chen, Ben Chiaro, Roberto Collins, William Courtney, Andrew Dunsworth, Edward Farhi, Brooks Foxen, Austin Fowler, Craig Gidney, Marissa Giustina, Rob Graff, Keith Guerin, Steve Habegger, Matthew P. Harrigan, Michael J. Hartmann, Alan Ho, Markus Hoffmann, Trent Huang, Travis S. Humble, Sergei V. Isakov, Evan Jeffrey, Zhang Jiang, Dvir Kafri, Kostyantyn Kechedzhi, Julian Kelly, Paul V. Klimov, Sergey Knysh, Alexander Korotkov, Fedor Kostritsa, David Landhuis, Mike Lindmark, Erik Lucero, Dmitry Lyakh, Salvatore Mandrà, Jarrod R. McClean, Matthew McEwen, Anthony Megrant, Xiao Mi, Kristel Michielsen, Masoud Mohseni, Josh Mutus, Ofer Naaman, Matthew Neeley, Charles Neill, Murphy Yuezhen Niu, Eric Ostby, Andre Petukhov, John C. Platt, Chris Quintana, Eleanor G. Rieffel, Pedram Roushan, Nicholas C. Rubin, Daniel Sank, Kevin J. Satzinger, Vadim Smelyanskiy, Kevin J. Sung, Matthew D. Trevithick, Amit Vainsencher, Benjamin Villalonga, Theodore White, Z. Jamie Yao, Ping Yeh, Adam Zalcman, Hartmut Neven, and John M. Martinis. Quantum supremacy using a programmable superconducting processor. *Nature*, 574(7779):505–510, 2019. doi:10.1038/s41586-019-1666-5. → page 1

[13] Petar Jurcevic, Ali Javadi-Abhari, Lev S. Bishop, Isaac Lauer, Daniela F. Bogorin, Markus Brink, Lauren Capelluto, Oktay Günlük, Toshinaro Itoko, Naoki Kanazawa, Abhinav Kandala, George A. Keefe, Kevin Kruslich, William Landers, Eric P. Lewandowski, Douglas T. McClure, Giacomo Nannicini, Adinath Narasgond, Hasan M. Nayfeh, Emily Pritchett, Mary Beth Rothwell, Srikanth Srinivasan, Neereja Sundaresan, Cindy Wang, Ken X. Wei, Christopher J. Wood, Jeng-Bang Yau, Eric J. Zhang, Oliver E.

Dial, Jerry M. Chow, and Jay M. Gambetta. Demonstration of quantum volume 64 on a superconducting quantum computing system, 2020. arXiv:2008.08571.

[14] Andrew D. King, Juan Carrasquilla, Jack Raymond, Isil Ozfidan, Evgeny Andriyash, Andrew Berkley, Mauricio Reis, Trevor Lanting, Richard Harris, Fabio Altomare, Kelly Boothby, Paul I. Bunyk, Colin Enderud, Alexandre Fréchette, Emile Hoskinson, Nicolas Ladizinsky, Travis Oh, Gabriel Poulin-Lamarre, Christopher Rich, Yuki Sato, Anatoly Yu. Smirnov, Loren J. Swenson, Mark H. Volkmann, Jed Whittaker, Jason Yao, Eric Ladizinsky, Mark W. Johnson, Jeremy Hilton, and Mohammad H. Amin. Observation of topological phenomena in a programmable lattice of 1,800 qubits. *Nature*, 560(7719):456–460, 2018. doi:10.1038/s41586-018-0410-x.

[15] K. Wright, K. M. Beck, S. Debnath, J. M. Amini, Y. Nam, N. Grzesiak, J.-S. Chen, N. C. Pisenti, M. Chmielewski, C. Collins, K. M. Hudek, J. Mizrahi, J. D. Wong-Campos, S. Allen, J. Apisdorf, P. Solomon, M. Williams, A. M. Ducore, A. Blinov, S. M. Kreikemeier, V. Chaplin, M. Keesan, C. Monroe, and J. Kim. Benchmarking an 11-qubit quantum computer. *Nature Communications*, 10(1):5464, 2019. doi:10.1038/s41467-019-13534-2. → page 1

[16] Guifré Vidal. Efficient classical simulation of slightly entangled quantum computations. *Physical Review Letters*, 91(14):147902, 2003. doi:10.1103/PhysRevLett.91.147902. → page 2

[17] Artur K. Ekert. Quantum cryptography based on bell's theorem. *Phys. Rev. Lett.*, 67(6):661–663, 1991. doi:10.1103/PhysRevLett.67.661. → page 3

[18] Charles H. Bennett and Stephen J. Wiesner. Communication via one- and two-particle operators on einstein-podolsky-rosen states. *Phys. Rev. Lett.*, 69 (20):2881–2884, 1992. doi:10.1103/PhysRevLett.69.2881. → page 3

[19] Daniel Gottesman. The heisenberg representation of quantum computers. In S. P. Corney, R. Delbourgo, and P. D. Jarvis, editors, *Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*, pages 32–43, 1999. URL https://arxiv.org/abs/quant-ph/9807006. arXiv:quant-ph/9807006. → pages 3, 9, 12

[20] Louis de Broglie. La nouvelle dynamique des quanta. In Hendrik Antoon Lorentz, editor, *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., pages 105–132. Gauthier-Villars et cie, 1928. → pages 4, 28

[21] A. Einstein, B. Podolsky, and N. Rosen. Can quantum-mechanical description of physical reality be considered complete? *Physical Review*, 47 (10):777–780, 1935. doi:10.1103/PhysRev.47.777. → pages 4, 28

[22] J. S. Bell. On the Einstein Podolsky Rosen paradox. *Physics Physique Fizika*, 1(3):195–200, 1964. doi:10.1103/PhysicsPhysiqueFizika.1.195. → pages 4, 29

[23] John S. Bell. On the problem of hidden variables in quantum mechanics. *Reviews of Modern Physics*, 38(3):447–452, 1966. doi:10.1103/RevModPhys.38.447. → pages 4, 29

[24] S. Kochen and E. Specker. The problem of hidden variables in quantum mechanics. *Journal of Mathematics and Mechanics*, 17(1):59–87, 1968. doi:10.1512/iumj.1968.17.17004. → pages 4, 29

[25] Robert Raussendorf. Contextuality in measurement-based quantum computation. *Physical Review A*, 88, 2013. doi:10.1103/physreva.88.022322. → page 4

[26] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. Contextuality supplies the "magic" for quantum computation. *Nature*, 510 (7505):351–355, 2014. doi:10.1038/nature13460. → pages 4, 5, 33, 34

[27] Eugene Wigner. On the quantum correction for thermodynamic equilibrium. *Physical Review*, 40(5):749–759, 1932. doi:10.1103/PhysRev.40.749. → pages 4, 14, 15

[28] Anatole Kenfack and Karol Życzkowski. Negativity of the wigner function as an indicator of non-classicality. *Journal of Optics B: Quantum and Semiclassical Optics*, 6(10):396–404, 2004. doi:10.1088/1464-4266/6/10/003. → pages 4, 14

[29] William K Wootters. A wigner-function formulation of finite-state quantum mechanics. *Annals of Physics*, 176(1):1–21, 1987. doi:10.1016/0003-4916(87)90176-X. → pages 4, 16

[30] Ernesto F. Galvão. Discrete wigner functions and quantum computational speedup. *Physical Review A*, 71(4):042302, 2005. doi:10.1103/PhysRevA.71.042302. → page 19

[31] Cecilia Cormick, Ernesto F. Galvão, Daniel Gottesman, Juan Pablo Paz, and Arthur O. Pittenger. Classicality in discrete wigner functions. *Physical Review A*, 73(1):012301, 2006. doi:10.1103/PhysRevA.73.012301.

[32] D. Gross. Hudson's theorem for finite-dimensional quantum systems. *Journal of Mathematical Physics*, 47(12):122107, 2006. doi:10.1063/1.2393152. → pages 10, 17, 19, 20, 27, 73, 94

[33] D. Gross. Non-negative wigner functions in prime dimensions. *Applied Physics B*, 86(3):367–370, 2006. doi:10.1007/s00340-006-2510-9. → pages 4, 17, 19

[34] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Physical Review A*, 71(2):022316, 2005. doi:10.1103/PhysRevA.71.022316. → pages 5, 12, 13

[35] Robert W. Spekkens. Negativity and contextuality are equivalent notions of nonclassicality. *Physical Review Letters*, 101(2):020401, 2008. doi:10.1103/PhysRevLett.101.020401. → pages 5, 33, 34, 73, 96

[36] Nicolas Delfosse, Cihan Okay, Juan Bermejo-Vega, Dan E Browne, and Robert Raussendorf. Equivalence between contextuality and negativity of the wigner function for qudits. *New Journal of Physics*, 19(12):123024, 2017. doi:10.1088/1367-2630/aa8fe3. → pages 5, 21, 33, 34, 44

[37] Huangjun Zhu. Permutation symmetry determines the discrete wigner function. *Physical Review Letters*, 116(4):040501, 2016. doi:10.1103/PhysRevLett.116.040501. → pages 5, 36, 67

[38] Robert Raussendorf, Dan E. Browne, Nicolas Delfosse, Cihan Okay, and Juan Bermejo-Vega. Contextuality and wigner-function negativity in qubit quantum computation. *Physical Review A*, 95(5), 2017. doi:10.1103/physreva.95.052334. → pages 36, 88

[39] Angela Karanjai, Joel J. Wallman, and Stephen D. Bartlett. Contextuality bounds the efficiency of classical simulation of quantum processes, 2018. URL http://arxiv.org/abs/1802.07744. arXiv:1802.07744. → pages 5, 6, 36, 67, 96

[40] Daniel Gottesman. *Stabilizer Codes and Quantum Error Correction*. PhD thesis, California Institute of Technology, 1997. URL https://arxiv.org/abs/quant-ph/9705052. arXiv:quant-ph/9705052. → page 9

[41] Scott Aaronson and Daniel Gottesman. Improved simulation of stabilizer circuits. *Physical Review A*, 70(5):052328, 2004. doi:10.1103/PhysRevA.70.052328. → pages 9, 12, 68

[42] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. The resource theory of stabilizer quantum computation. *New Journal of Physics*, 16(1):013009, 2014. doi:10.1088/1367-2630/16/1/013009. → pages 9, 85, 96

[43] Cihan Okay, Sam Roberts, Stephen D. Bartlett, and Robert Raussendorf. Topological proofs of contextuality in quantum mechanics. *Quantum Information and Computation*, 17:1135–1166, 2017. → pages 11, 32, 33, 35, 44, 55

[44] Daniel Gottesman. Fault-tolerant quantum computation with higher-dimensional systems. *Chaos, Solitons & Fractals*, 10(10): 1749–1758, 1999. doi:10.1016/s0960-0779(98)00218-5. → page 12

[45] Sean Clark. Valence bond solid formalism ford-level one-way quantum computation. *Journal of Physics A: Mathematical and General*, 39(11): 2701–2721, 2006. doi:10.1088/0305-4470/39/11/010. → page 12

[46] E. C. G. Sudarshan. Equivalence of semiclassical and quantum mechanical descriptions of statistical light beams. *Physical Review Letters*, 10(7): 277–279, 1963. doi:10.1103/PhysRevLett.10.277. → page 15

[47] Roy J. Glauber. Coherent and incoherent states of the radiation field. *Physical Review*, 131(6):2766–2788, 1963. doi:10.1103/PhysRev.131.2766. → page 15

[48] Kôdi Husimi. Some formal properties of the density matrix. *Proceedings of the Physico-Mathematical Society of Japan. 3rd Series*, 22(4):264–314, 1940. doi:10.11429/ppmsj1919.22.4_264. → page 15

[49] R. L. Stratonovich. On distributions in representation space. *Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki*, 31(6):1012–1020, 1956. English translation: *Soviet Physics JETP*, 4(6):891–898, 1957. → page 15

[50] C Brif and A Mann. A general theory of phase-space quasiprobability distributions. *Journal of Physics A: Mathematical and General*, 31(1): L9–L17, 1998. doi:10.1088/0305-4470/31/1/002. → page 15

[51] Richard P. Feynman. Negative probability. In F. David Peat and Basil Hiley, editors, *Quantum Implications: Essays in Honour of David Bohm*, pages 235–248. Routledge & Keegan Paul Ltd., 1987. URL https://cds.cern.ch/record/154856/files/pre-27827.pdf. → page 16

[52] Kathleen S. Gibbons, Matthew J. Hoffman, and William K. Wootters. Discrete phase space based on finite fields. *Physical Review A*, 70(6): 062101, 2004. doi:10.1103/PhysRevA.70.062101. → page 16

[53] 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. → pages 17, 19, 73, 94

[54] R.L. Hudson. When is the wigner quasi-probability density non-negative? *Reports on Mathematical Physics*, 6(2):249 – 252, 1974. doi:https://doi.org/10.1016/0034-4877(74)90007-X. → page 19

[55] Hakop Pashayan, Joel J. Wallman, and Stephen D. Bartlett. Estimating outcome probabilities of quantum circuits using quasiprobabilities. *Physical Review Letters*, 115(7), 2015. doi:10.1103/physrevlett.115.070501. → pages 27, 65

[56] Hakop Pashayan, Stephen D. Bartlett, and David Gross. From estimation of quantum probabilities to simulation of quantum circuits. *Quantum*, 4:223, 2020. doi:10.22331/q-2020-01-13-223. → pages 27, 65

[57] David Bohm. A suggested interpretation of the quantum theory in terms of ”hidden” variables. i. *Physical Review*, 85(2):166–179, 1952. doi:10.1103/PhysRev.85.166. → page 28

[58] David Bohm. A suggested interpretation of the quantum theory in terms of ”hidden” variables. ii. *Physical Review*, 85(2):180–193, 1952. doi:10.1103/PhysRev.85.180. → page 28

[59] John von Neumann. *Mathematische Grundlagen der Quantenmechanik*. Springer-Verlag, 1932. English translation: *Mathematical Foundations of Quantum Mechanics*, Princeton University Press, 1955. → page 29

[60] Grete Hermann. Die naturphilosophischen grundlagen der quantenmechanik (anzug). *Abhandlungen der Freis'schen Schule*, 6:75–152, 1935. → page 29

[61] Nicolas Delfosse, Philippe Allard Guerin, Jacob Bian, and Robert Raussendorf. Wigner function negativity and contextuality in quantum computation on rebits. *Physical Review X*, 5(2):021003, 2015. doi:10.1103/PhysRevX.5.021003. → pages 36, 38, 45

[62] Joel J. Wallman and Stephen D. Bartlett. Non-negative subtheories and quasiprobability representations of qubits. *Physical Review A*, 85(6):062121, 2012. doi:10.1103/PhysRevA.85.062121. → pages 37, 44, 73, 94, 115

[63] Lucas Kocia and Peter Love. Discrete wigner formalism for qubits and noncontextuality of clifford gates on qubit stabilizer states. *Physical Review A*, 96(6):062134, 2017. doi:10.1103/PhysRevA.96.062134.

[64] Mark Howard and Earl Campbell. Application of a resource theory for magic states to fault-tolerant quantum computing. *Physical Review Letters*, 118(9), 2017. doi:10.1103/physrevlett.118.090501. → pages 38, 39, 73, 94, 117, 119

[65] John B. DeBrota and Blake C. Stacey. Discrete wigner functions from informationally complete quantum measurements, 2019. URL https://arxiv.org/abs/1912.07554. arXiv:1912.07554. → page 36

[66] Martin Charles Golumbic. Trivially perfect graphs. *Discrete Mathematics*, 24(1):105–107, 1978. doi:10.1016/0012-365X(78)90178-4. → page 49

[67] William M. Kirby and Peter J. Love. Classical simulation of noncontextual pauli hamiltonians, 2020. URL https://arxiv.org/abs/2002.05693. arXiv:2002.05693. → page 50

[68] Michael Aschbacher. *Finite Group Theory*. Cambridge University Press, 2000. → pages 53, 70, 89

[69] Piers Lillystone and Joseph Emerson. A contextual ψ-epistemic model of the n-qubit stabilizer formalism, 2019. URL https://arxiv.org/abs/1904.04268. arXiv:1904.04268. → page 57

[70] David S. Dummit and Richard M. Foote. *Abstract Algebra*. John Wiley and Sons, Inc., 2004. → page 70

[71] Gunter M. Ziegler. *Lectures on polytopes*. Springer-Verlag, 1995. → pages 71, 74, 114

[72] Jorge Nocedal and Stephen J. Wright. *Numerical Optimization*. Springer Science+Business Media, LLC., 2 edition, 2006. → page 72

[73] R. W. Spekkens. Contextuality for preparations, transformations, and unsharp measurements. *Physical Review A*, 71(5):052108, 2005. doi:10.1103/PhysRevA.71.052108. → page 67

[74] Christopher Ferrie and Joseph Emerson. Frame representations of quantum mechanics and the necessity of negativity in quasi-probability representations. *Journal of Physics A: Mathematical and Theoretical*, 41 (35):352001, 2008. doi:10.1088/1751-8113/41/35/352001. → pages 73, 96

[75] Christopher Ferrie and Joseph Emerson. Framed hilbert space: hanging the quasi-probability pictures of quantum theory. *New Journal of Physics*, 11(6): 063040, 2009. doi:10.1088/1367-2630/11/6/063040. → page 96

[76] Huangjun Zhu. Quasiprobability representations of quantum mechanics with minimal negativity. *Physical Review Letters*, 117(12):120404, 2016. doi:10.1103/PhysRevLett.117.120404.

[77] John B. DeBrota, Christopher A. Fuchs, and Blake C. Stacey. Symmetric informationally complete measurements identify the irreducible difference between classical and quantum systems. *Physical Review Research*, 2(1): 013074, 2020. doi:10.1103/PhysRevResearch.2.013074. → page 73

[78] Arne Heimendahl. *The Stabilizer Polytope and Contextuality for Qubit Systems*. Master's thesis, University of Cologne, 2019. → page 74

[79] Walter Rudin. *Principles of Mathematical Analysis*. McGraw-Hill, Inc., 3 edition, 1976. → page 74

[80] Vašek Chvátal. *Linear Programming*. W.H. Freeman and Company, 1983. → pages 77, 115

[81] Dan Shepherd and Michael J. Bremner. Temporally unstructured quantum computation. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 465(2105):1413–1439, 2009. doi:10.1098/rspa.2008.0443. → page 85

[82] Michael Beverland, Earl Campbell, Mark Howard, and Vadym Kliuchnikov. Lower bounds on the non-clifford resources for quantum computations. *Quantum Science and Technology*, 5(3):035009, 2020. doi:10.1088/2058-9565/ab8963. → page 92

[83] Emily Tyhurst. *Cost of contextuality quantified by a classical simulation of Mermin's square*. Bachelor's thesis, University of British Columbia, 2017. → page 92

[84] Matthew F. Pusey, Jonathan Barrett, and Terry Rudolph. On the reality of the quantum state. *Nature Physics*, 8(6):475–478, May 2012. doi:10.1038/nphys2309. → page 95

[85] Christopher Ferrie. Quasi-probability representations of quantum theory with applications to quantum information science. *Reports on Progress in Physics*, 74(11):116001, 2011. doi:10.1088/0034-4885/74/11/116001. → page 96

[86] Donald E Knuth. Subspaces, subsets, and partitions. *Journal of Combinatorial Theory, Series A*, 10(2):178 – 180, 1971. doi:10.1016/0097-3165(71)90022-7. → page 110

[87] Robert Koenig and John A. Smolin. How to efficiently select an arbitrary clifford group element. *Journal of Mathematical Physics*, 55(12):122202, 2014. doi:10.1063/1.4903507. → page 111

[88] Markus Heinrich and David Gross. Robustness of magic and symmetries of the stabiliser polytope. *Quantum*, 3:132, 2019. doi:10.22331/q-2019-04-08-132. → page 112

[89] Karol Życzkowski and Hans-Jürgen Sommers. Induced measures in the space of mixed quantum states. *Journal of Physics A: Mathematical and General*, 34(35):7111–7125, 2001. doi:10.1088/0305-4470/34/35/335. → pages 113, 114

[90] Hans-Jürgen Sommers and Karol Życzkowski. Statistical properties of random density matrices. *Journal of Physics A: Mathematical and General*, 37(35):8457–8466, 2004. doi:10.1088/0305-4470/37/35/004. → page 113

[91] V Al Osipov, H-J Sommers, and K Życzkowski. Random bures mixed states and the distribution of their purity. *Journal of Physics A: Mathematical and Theoretical*, 43(5):055302, 2010. doi:10.1088/1751-8113/43/5/055302. → page 114

[92] Karol Życzkowski, Karol A. Penson, Ion Nechita, and Benoît Collins. Generating random density matrices. *Journal of Mathematical Physics*, 52 (6):062201, 2011. doi:10.1063/1.3595693. → page 113

[93] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *Journal of Mathematical Physics*, 6(3):440–449, 1965. doi:10.1063/1.1704292. → page 114

[94] Stuart G. Hoggar. 64 lines from a quaternionic polytope. *Geometriae Dedicata*, 69(3):287–289, 1998. doi:10.1023/A:1005009727232. → page 119

[95] Patrick Rall, Daniel Liang, Jeremy Cook, and William Kretschmer. Simulation of qubit quantum circuits via pauli propagation. *Physical Review A*, 99(6):062337, 2019. doi:10.1103/PhysRevA.99.062337. → page 119

Appendix A

Computational methods

A.1 Generating multiqubit phase point operators

For the purpose of the numerical calculations in Appendix B, it is important to be able to enumerate all phase points of the multiqubit generalized phase space $\mathscr{V}$ defined in Chapter 3. Points in $\mathscr{V}$ are associated with pairs (Ω, γ) where $\Omega \subset E$ is a cnc set and $\gamma : \Omega \to \mathbb{Z}_2$ is a noncontextual value assignment on that set. The task of enumerating all such pairs can be broken down into two steps: (i) enumerating cnc sets, and (ii) enumerating their noncontextual value assignments.

First, the classification of multiqubit phase space points provided by Lemma 3, Theorem 2, and Lemma 5 gives a concise description of all maximal cnc sets. They are sets of the form

$$\Omega = \bigcup_{k=1}^{2m+1} \langle a_k, \tilde{I} \rangle$$

where $\tilde{I} = \langle g_1, g_2, \ldots, g_{n-m} \rangle$ is a $n-m$-dimensional isotropic subspace of E and $a_1, a_2, \ldots, a_{2m+1} \in \tilde{I}^{\perp}/\tilde{I}$ pairwise anticommute. Given a cnc set Ω with generators $g_1, g_2, \ldots, g_{n-m}, a_1, a_2, \ldots, a_{2m+1}$, all possible nonontextual value assignments on Ω are obtained by choosing the value of γ freely on the generators and then determining γ on the rest of Ω through the relation Eq. (3.1).

It remains to enumerate all cnc sets Ω. This task can itself be broken down into two steps. To enumerate all cnc sets with parameters n and m, it suffices to (i) enumerate all $n-m$-dimensional isotropic subspaces $\tilde{I}$ of E, and (ii) enumerate

all sets A of pairwise anticommuting elements of $\mathbb{Z}_2^{2m}$ with cardinality $2m+1$. As a result of the isomorphism $\tilde{I}^{\perp}/\tilde{I} \simeq \mathbb{Z}_2^{2m}$, for each $\tilde{I} \subset E$ we can find a symplectic map $M : \mathbb{Z}_2^{2m} \to \tilde{I}^{\perp}/\tilde{I}$. Using a fixed symplectic map $M_{\tilde{I}}$ for each isotropic subspace $\tilde{I}$, a pair $(\tilde{I}, M_{\tilde{I}} \cdot A)$ uniquely specifies a cnc set Ω. Subroutines for each of the tasks described here are given in below.

A.1.1 A canonical basis for isotropic subspaces

Enumerating phase point operators requires a subroutine which enumerates isotropic subspaces of $\mathbb{Z}_2^{2n}$. To this end, we can use a canonical basis for subspaces of vector spaces over finite fields. One such canonical basis comes from the rows of a matrix in upper triangular form [86]. For example, a four dimensional subspace of the vector space $\mathbb{F}_d^8$ is obtained from the row space of the matrix

$$\begin{bmatrix} A_{11} & 0 & A_{13} & 0 & 0 & A_{16} & 1 & 0 \\ A_{21} & 0 & A_{23} & 0 & 1 & 0 & 0 & 0 \\ A_{31} & 0 & A_{33} & 1 & 0 & 0 & 0 & 0 \\ A_{41} & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

with different values for $A_{11}, A_{13}, A_{16}, A_{21}, A_{23}, A_{31}, A_{33}, A_{41} \in \mathbb{F}_d$ giving different subspaces. To enumerate all subspaces of $\mathbb{Z}_2^{2n}$ it suffices to enumerate upper triangular matrices with entries in $\mathbb{Z}_2$. Then to obtain an enumeration of isotropic subspaces we can filter out matrices whose rows do not pairwise commute.

An alternative approach to enumerating isotropic subspaces could be based on a canonical form for stabilizer codes (e.g. [4, §10.5.7]). The advantage to this approach is that every basis described in this standard form generates an isotropic subspace. There is no need to filter out nonisotropic subspaces. On the other hand, there are multiple bases in this standard form which generate the same isotropic subspace. Since for our purposes it is important to obtain a unique set of generators for each isotropic subspace, and it is generally easier to check if a set of generators pairwise commute than to remove redundancy by checking that a set of bases generate distinct isotropic subspaces, for few qubits the method above is more practical.

A.1.2 Maximal anticommuting sets in $\mathbb{Z}_2^{2m}$

There are many equivalent ways of generating all maximal sets of pairwise anticommuting vectors of size $2m+1$ in $\mathbb{Z}_2^{2m}$. Since the symplectic group acts transitively on these sets, one method proceeds by acting with the symplectic group on the set defined by Eqs. (3.6) and (3.7). The size of the symplectic group $Sp(2m,\mathbb{Z}_2)$ grows very quickly with the number of qubits m so this method quickly becomes intractible.[1] Alternatively, since the number of these sets is known (c.f. proof of Theorem 5), a brute force search is also possible for few qubits.

A.1.3 The symplectic Gram-Schmidt procedure

A useful subroutine for generating phase point operators is the symplectic Gram-Schmidt procedure [87]. The symplectic Gram-Schmidt procedure takes as input a set of vectors S in the symplectic vector space $\mathbb{Z}_2^{2n}$ and returns a symplectic basis for the span of $S/(S\cap S^{\perp})$. The procedure is described in Algorithm A.1. At each step of the algorithm, either an element of $S\cap S^{\perp}$ is removed from S, or a symplectic pair (e,f) is picked out from S and added to the basis and S is updated to a new set S' such that $S'\cup\{e,f\}$ has the same span as the S at the previous iteration, $[e,s]=[f,s]=0\ \forall s\in S'$, and $|S'|\leq|S|-2$.

The Choose(S) function used in Algorithm A.1 picks an arbitrary element from the set S. An element can be chosen at random, or to make the algorithm deterministic, we can order vectors in $\mathbb{Z}_2^{2n}$ as binary strings and have the Choose function return the least element in the set S with respect to this order.

For a $n-m$-dimensional isotropic subspace $\tilde{I}\subset E$, using $\tilde{I}^{\perp}$ as input to the symplectic Gram-Schmidt procedure gives a basis for $\tilde{I}^{\perp}/\tilde{I}$. The map that takes the standard symplectic basis of $\mathbb{Z}_2^{2m}$ to the basis of $\tilde{I}^{\perp}/\tilde{I}$ defines a symplectic map $M:\mathbb{Z}_2^{2m}\to\tilde{I}^{\perp}/\tilde{I}$.

[1] The size of the symplectic group is

$$|Sp(2m,\mathbb{Z}_2)|=2^{m^2}\prod_{k=1}^{m}4^k-1.$$

Input: $S \subset \mathbb{Z}_2^{2n}$

basis $\leftarrow \emptyset$
repeat
  $e \leftarrow \text{Choose}(S)$
  if $[a,e] = 0 \; \forall a \in S$ **then**
    $S \leftarrow S \setminus \{e\}$
  else
    $f \leftarrow \text{Choose}(\{a \in S \mid [e,a] = 1\})$
    basis $\leftarrow$ basis $\cup \{(e,f)\}$
    $S \leftarrow \{a + [a,f]e + [a,e]f \mid a \in S\}$
  end if
until $|S| = 0$

Output: basis

Algorithm A.1: The symplectic Gram-Schmidt procedure. This algorithm takes as input a set of vectors $S \subset \mathbb{Z}_2^{2n}$ and returns a symplectic basis for the span of $S/(S \cap S^{\perp})$.

A.2 Numerical calculation of robustness

The robustness of magic, defined in Eq. (2.25), and the phase space robustness, defined in Eq. (3.24), are both defined as the solutions to convex optimization problems. Each of these problems can be transformed into a linear program (as in the proof of Lemma 11) and then solved with a linear programming solver. For the robustness calculations presented in Appendix B, we use the software package Gurobi.[2]

As demonstrated in table 3.1, the size of the phase space and the number of stabilizer states grow very quickly with the number of qubits making the optimization problems intractible even for few qubits. For certain states which share symmetries with the phase space, this shared symmetry can be exploited to reduce the number of variables needed in the robustness calculations. This method was used in Ref. [88] to compute the robustness of magic for copies of the states $|H\rangle$ and $|T\rangle$ on up to 9 and 10 qubits. Since the symmetry group used in this reduction, a subgroup of the Clifford group, is also a symmetry of the phase space of Chapter 3, this method could also be used in the calculations of the phase space robustness.

[2] gurobi.com

A.3 Generating random quantum states

Some of the numerical results presented in Appendix B require the ability to generate random quantum states. For systems of qubits and systems of rebits we need to be able to generate both pure states distributed uniformly according to the Fubini-Study measure[3] and mixed states distributed uniformly according to the Hilbert-Schmidt measure[4] for systems of qubits and rebits. Algorithms for each of these ensembles exist [89–92].

The primitive task in all four cases is to generate a random matrix X with normally distributed entries. Then the state returned is the normalized density matrix

$$\rho = \frac{XX^\dagger}{\mathrm{Tr}(XX^\dagger)}. \tag{A.1}$$

The nature of the state and the measure according to which it is distributed depend only on how we choose X.

We have four cases of interest:

1. To generate a random pure state of a D-dimensional quantum system distributed according to the Fubini-Study measure, we choose X to have dimensions $D \times 1$ with elements distributed according to a complex standard normal distribution [89].

2. To generate a random pure state of a D-dimensional quantum system with a real density matrix distributed according to the Fubini-Study measure we choose X to have dimensions $D \times 1$ with entries distributed according to a real standard normal distribution [89].

3. To generate a random mixed state of a D-dimensional quantum system distributed according to the Hilbert-Schmidt measure, we choose X to have

[3] The Fubini-Study measure is the measure induced by the Fubini-Study metric. It is the unique measure on the set of pure quantum states which is invariant with respect to the action of the unitary group. It can equivalently be defined as the measure on pure quantum states induced by the Haar measure on the unitary group. I.e. a Fubini-Study distributed random state is the result of a Haar distributed random unitary applied to a fixed reference state [89].

[4] The Hilbert-Schmidt measure is the measure induced by the Hilbert-Schmidt metric. The average purity of Hilbert-Schmidt distributed random states scales like $1/D$ where D is the dimension of the quantum system [90].

dimensions $D \times D$ with entries distributed according to a complex normal distribution, i.e. X is taken from the complex Ginibre ensemble [93]. This is equivalent to generating a pure state of a composite system in the Hilbert space $\mathbb{C}^D \otimes \mathbb{C}^D$ distributed according to the Fubini-Study measure and tracing out the first subsystem [89].

4. To generate a random mixed state of a D-dimensional quantum system with a real density matrix distributed according to the Hilbert-Schmidt measure, we choose X to have dimensions $D \times D+1$ with entries distributed according to a real standard normal distribution [91].

A.4 Enumerating vertices of Λ

The polytope Λ from Chapter 4 is defined as the intersection of a finite set of half-spaces:

$$\Lambda := \{X \in \mathscr{O}(2^n) \mid \mathrm{Tr}(\sigma X) \geq 0 \; \forall \sigma \in \mathscr{S}\}.$$

This is known as the H-representation of a polytope. Since the polytope is compact (c.f. Lemma 12), by the Minkowski-Weyl theorem, the polytope can equivalently be described as the convex hull of its finitely many vertices [71, §1.1]:

$$\Lambda = \mathrm{ConvHull}\{A_\alpha \mid \alpha \in \mathscr{V}\}.$$

This is called the V-represenation of the polytope. As the probability representation of quantum states and dynamics in the hidden variable model described by Theorem 6 is in terms of the vertices of the polytope it is useful to be able to convert between these representations. This is a standard problem in computational geometry known as the vertex enumeration problem with several dedicated software packages (e.g. lrslib[5] and cddlib[6]).

A numerical enumeration of the vertices suggests that the number of vertices of Λ grows extremely quickly with the number n of qubits. Using lrslib and cddlib, a complete enumeration of the vertices has so far only been possible for up to two

[5] cgm.cs.mcgill.ca/ avis/C/lrs.html
[6] github.com/cddlib/cddlib

qubits. For one qubit Λ has 8 vertices (the phase point operators of the eight state model [62]) and for two qubits Λ has 22320 vertices.

In addition to enumerating the vertices of Λ, it is important to be able to find a probability function $p_\rho : \mathscr{V} \to \mathbb{R}_{\geq 0}$ representing a state ρ. This can be achieved via linear programming. In particular, a valid probability function representing the state ρ can be obtained as the solution to the linear program

$$\min_{p,u,v} \left\{ \sum_{a \in E} u_a + v_a \;\middle|\; \rho = \sum_{a \in E} (u_a - v_a) T_a + \sum_{\alpha \in \mathscr{V}} p(\alpha) A_\alpha,\ u_a, v_a \geq 0 \ \forall a \in E, p(\alpha) \geq 0 \ \forall \alpha \in \mathscr{V} \right\}.$$

This is known as the auxillary problem in linear programming [80], or phase 1 of the two phase simplex method. An initial dictionary for this linear program is obtained by taking $u_a = \max(0, 2^{-n}\mathrm{Tr}(\rho T_a))$ and $v_a = \max(0, -2^{-n}\mathrm{Tr}(\rho T_a))\ \forall a \in E$, and $p(\alpha) = 0\ \forall \alpha \in \mathscr{V}$.

If among the vertices of Λ, there is a subset $\mathscr{V}^* \subset \mathscr{V}$ of preferred vertices, we can find a probability distribution p_ρ which maximizes the weight of the vertices in this set by instead solving the linear program

$$\max_{p} \left\{ \sum_{\alpha \in \mathscr{V}^*} p(\alpha) \;\middle|\; \rho = \sum_{\alpha \in \mathscr{V}} p(\alpha) A_\alpha,\ p(\alpha) \geq 0 \ \forall \alpha \in \mathscr{V} \right\}.$$

In these two linear programs, there is a variable for each vertex of Λ. The number of variables needed for the problems can be reduced (and hence the complexity of finding p_ρ can be reduced) if the state ρ lies on the boundary of the polytope Λ.

To see this, consider an expansion of a state ρ of the form Eq. (4.2). Multiplying Eq. (4.2) by a stabilizer state $\sigma \in \mathscr{S}$ and taking a trace we have

$$\mathrm{Tr}(\sigma\rho) = \sum_{\alpha \in \mathscr{V}} p_\rho(\alpha) \mathrm{Tr}(\sigma A_\alpha).$$

Since the coefficients $p_\rho(\alpha)$ in the expansion of ρ are all nonnegative, if the left hand side of this equation is zero, then for each $\alpha \in \mathscr{V}$ with $p_\rho(\alpha) > 0$ we must have $\mathrm{Tr}(\sigma A_\alpha) = 0$. That is, any vertex $A_\alpha \in \mathscr{A}$ that can appear with nonzero weight in the expansion of a state ρ must must be orthogonal to any stabilizer state that is orthogonal to ρ with respect to the Hilbert-Schmidt inner product. If ρ is

on the boundary of Λ then there is a nonempty set of stabilizer states orthogonal to ρ and so for the purposes of solving the linear programs above we can restrict the variables to vertices in the set $\{\alpha \in \mathscr{V} \mid \mathrm{Tr}(\sigma A_\alpha) = 0 \; \forall \sigma \in \mathscr{S}_\rho\}$ where $\mathscr{S}_\rho := \{\sigma \in \mathscr{S} \mid \mathrm{Tr}(\sigma\rho) = 0\}$.

Appendix B

Numerical comparison of QCM simulation methods

B.1 Robustness values for common magic states

One method of comparing the phase space simulation method of Chapter 3 with the previous simulation method based on decompositions of states into stabilizer states [64] is to compare the parameter that quantifies the cost of classical simulation for some common magic states. For the method based on stabilizer quasimixtures this is the robustness of magic, $\mathfrak{R}_S$, defined in Eq. 2.25 [64]. For the phase space method the corresponding parameter is the phase space robustness, $\mathfrak{R}$, defined in Eq. (3.24). Some common magic states used in QCM are tensor products of

n	$\mathfrak{R}_S$	$\mathfrak{R}$
1	1.4142	1.0000
2	1.7476	1.0000
3	2.2190	1.2828

(a) State: $|H\rangle\langle H|^{\otimes n}$

n	$\mathfrak{R}_S$	$\mathfrak{R}$
1	1.7321	1.0000
2	2.2321	1.0000
3	3.0981	1.3849

(b) State: $|T\rangle\langle T|^{\otimes n}$

Table B.1: Robustness of magic, $\mathfrak{R}_S$, and phase space robustness, $\mathfrak{R}$, for n copies of the single-qubit magic states $|H\rangle\langle H| = \frac{1}{2}I + \frac{1}{2\sqrt{2}}(X+Y)$ and $|T\rangle\langle T| = \frac{1}{2}I + \frac{1}{2\sqrt{3}}(X+Y+Z)$.

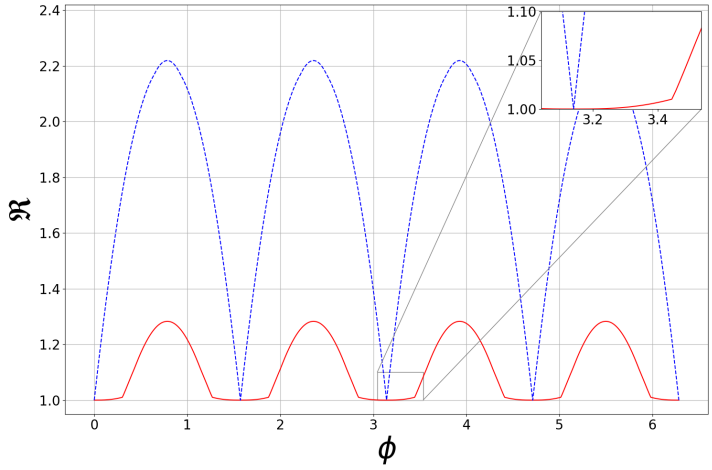


Figure B.1: Robustness of magic (blue) and phase space robustness (red) calculated numerically for the magic state $|H(\phi)\rangle^{\otimes 3}$ as a function of ϕ where $|H(\phi)\rangle = (|0\rangle + e^{-i\phi}|1\rangle)/\sqrt{2}$. At $\phi = 0, \pi/2, \pi, 3\pi/2$, and 2π, $|H(\phi)\rangle$ is a stabilizer state and so $\mathfrak{R}_S = \mathfrak{R} = 1$. For all other states plotted, $\mathfrak{R} < \mathfrak{R}_S$.

the one-qubit states $|H\rangle\langle H| = \frac{1}{2}I + \frac{1}{2\sqrt{2}}(X+Y)$ and $|T\rangle\langle T| = \frac{1}{2}I + \frac{1}{2\sqrt{3}}(X+Y+Z)$. Numerically calculated values for the robustness of magic and phase space robustness for copies of these magic states are given in Table B.1.

From Table B.1 we can see that the phase space robustness is consistently lower than the robustness of magic. In fact, it can be proven that for any state ρ, $\mathfrak{R}(\rho) \leq \mathfrak{R}_S(\rho) \leq (4n+1)\mathfrak{R}(\rho)$ (see Appendix C of Ref. [2] for the proof). To illustrate this point further, in Fig. B.1, we plot numerically calculated values of the robustness of magic and the phase space robustness for the state $|H(\phi)\rangle^{\otimes 3}$ with $|H(\phi)\rangle = (|0\rangle + e^{-i\phi}|1\rangle)/\sqrt{2}$ as a function of ϕ for $\phi \in [0, 2\pi]$. Since stabilizer states are positively representable by both methods, the robustness of magic and phase space robustness agree on the stabilizer states occuring at $\phi = 0, \pi/2, \pi, \ldots$. Away from these stabilizer states, the phase space robustness is strictly smaller than the robustness of magic.

Finally, we can consider the states with maximal robustness of magic for a given number of qubits. On one qubit, the robustness of magic is maximal for the state $|T\rangle$. Robustness of magic and phase space robustness for this state are given in

the first row of Table B.1b. For two-qubit equally-weighted states, the robustness of magic is maximal for the state $|max_2\rangle = (|00\rangle + |01\rangle + |10\rangle + i|11\rangle)/2$ with $\mathfrak{R}_S(|max_2\rangle\langle max_2|) = 2.2$ [64]. The phase space robustness of this state is 1, i.e., this state is positively representable.

For three qubits, the robustness of magic is maximized by the Hoggar state [64],

$$|\text{Hoggar}\rangle = \frac{1}{\sqrt{6}}\begin{bmatrix} 1+i & 0 & -1 & 1 & -i & 1 & 0 & 0 \end{bmatrix},$$

a fiducial state of a SIC-POVM [94]. The robustness of magic for this state is $\mathfrak{R}_S(|\text{Hoggar}\rangle\langle\text{Hoggar}|) = 3.8$, and the phase space robustness is $\mathfrak{R}(|\text{Hoggar}\rangle\langle\text{Hoggar}|) = 1.8$.

B.2 Positively representable states

Another method of comparing the phase space simulation algorithm with the simulation algorithm based on stabilizer quasimixtures is to compare the sets of states for which the simulation algorithms are efficient. That is, to compare the states which are positively representable by each method. For an additional comparison, in this section we also consider states for which another simulation method called Schrödinger propagation is efficient [95]. This method is based on decompositions of states into Pauli operators, and the set of states for which this method is efficient are called hyper-octahedral states.

To get an estimate of the volume of states which are positively representable in each method, we sample states distributed uniformly, and compute the fraction of states sampled which are positively representable. We consider four cases: pure state distributed according to the Fubini-Study measure and mixed states distributed according to the Hilbert-Schmidt measure for systems of qubits and for systems of rebits. Estimates of the volume of positively representable states for each of these cases on up to three qubits are given in Table B.2. For one-qubit and two-qubit states, the estimates are based on 100,000 samples, and for three-qubit states the estimates are based on 10,000 samples.

From Table B.2, we can see that by this metric the phase space method generally outperforms the stabilizer quasimixture method and the Schrödinger prop-

$n\backslash m$	$\mathscr{V}$, all m	$\mathscr{V}$, $m=1$	$\mathscr{V}$, $m=0$	hy.oct.
1	1.000	1.000	0.000	0.000
2	0.980	0.980	0.000	0.000
3	0.000	0.000	0.000	0.000

(a) Pure qubit states

$n\backslash m$	$\mathscr{V}$, all m	$\mathscr{V}$, $m=1$	$\mathscr{V}$, $m=0$	hy.oct.
1	1.000	1.000	0.318	0.318
2	1.000	1.000	0.009	0.568
3	1.000	1.000	0.000	0.897

(b) Mixed qubit states

$n\backslash m$	$\mathscr{V}$, all m	$\mathscr{V}$, $m=1$	$\mathscr{V}$, $m=0$	hy.oct.
1	1.000	1.000	0.000	0.000
2	1.000	1.000	0.000	0.000
3	0.001	0.001	0.000	0.000

(c) Pure rebit states

$n\backslash m$	$\mathscr{V}$, all m	$\mathscr{V}$, $m=1$	$\mathscr{V}$, $m=0$	hy.oct.
1	1.000	1.000	0.637	0.637
2	1.000	1.000	0.144	0.924
3	1.000	1.000	0.001	0.999

(d) Mixed rebit states

Table B.2: Fraction of states positively representable by various quasiprobability representations. The second column uses the representation based on the generalized phase space of Chapter 3, the third column is obtained by restricting the phase space to cnc sets with parameter $m = 1$, the fourth column is the fraction of stabilizer mixtures, and the last column is the fraction of hyper-octahedral states.

agation method. In the second last column of Tables B.2a and B.2c, we can see that 0% of pure states sampled are stabilizer mixtures. This is true for any number of qubits. The set of pure states which are positively representable as stabilizer quasimixtures are the finitely many pure stabilizer states, a set with measure zero. On the other hand, the set of pure states which are positively representable by the phase space method has nonzero measure.

It is also worth noting that for the cases of one qubit, one rebit, and two rebits, it

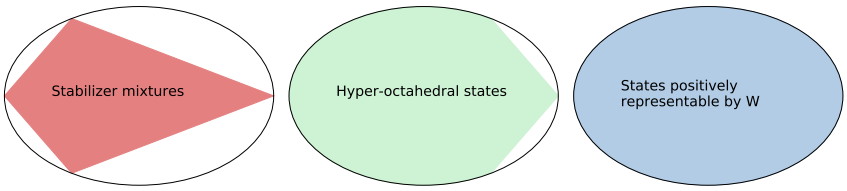


Figure B.2: States in the cross section of two-qubit states $\rho(x,y)$ defined by Eq. B.1 which are positively representable with respect to (a) decompositions in stabilizer states, (b) hyper-octahedral states, and (c) the phase space of Chapter 3.

can be proven that all states are positively representable by the phase space method. Enumerating the vertices of the polytope Λ defined in Chapter 4 shows that in these three cases the vertices of Λ coincide precisely with the phase point operators of Chapter 3. Since all states are contained in Λ (c.f. Lemma 13), this shows that all states are positively representable. This is not true in general, for two qubits for example, there are more vertices of Λ than there are two-qubit phase point operators.

Table B.2 also shows the fraction of states which are positively representable by the phase space of Chapter 3 obtained using only cnc sets with parameter $m = 1$ in Eq. 3.5. Comparing this with the fraction of positively representable states with respect to the full phase space, we see that most of the benefit of the phase space method is obtained from only these $m = 1$ sets. There are states which are positively representable with respect to the full phase space but which are not positively representable with respect to the restricted $m = 1$ phase space, but these states represent a small fraction of positively representable states.

As a final illustration of the difference in the simulation methods in terms of positive representability of states, in Fig. B.2 we plot the portion of two-qubit states which are positively representable by each of the methods for the cross section of two-qubit states of the form

$$\rho(x,y) = \frac{I}{4} I_1 I_2 + x(X_1X_2 + Z_1Z_2 - Y_1Y_2) + y(Z_1 + Z_2). \tag{B.1}$$