Hidden variable model for quantum computation with magic states on qudits of any dimension

Michael Zurel[1,2], Cihan Okay[3], Robert Raussendorf[2,4], and Arne Heimendahl[5]

[1]Department of Physics and Astronomy, University of British Columbia, Vancouver, Canada
[2]Stewart Blusson Quantum Matter Institute, University of British Columbia, Vancouver, Canada
[3]Department of Mathematics, Bilkent University, Ankara, Turkey
[4]Institute of Theoretical Physics, Leibniz University Hannover, Hannover, Germany
[5]Department of Mathematics and Computer Science, University of Cologne, Cologne, Germany

It was recently shown that a hidden variable model can be constructed for universal quantum computation with magic states on qubits. Here we show that this result can be extended, and a hidden variable model can be defined for quantum computation with magic states on qudits with any Hilbert space dimension. This model leads to a classical simulation algorithm for universal quantum computation.

1 Introduction

The field of quantum computation has seen an explosion of interest in recent years. It is widely believed that the era of quantum advantage is upon us and that we are entering the realm of so-called Noisy Intermediate-Scale Quantum (NISQ) computation. This view is evidenced by the impressive performance of quantum devices in recent hardware demonstrations [1–10].

But in spite of the age of the field and the recent surge in interest, a key question at the heart of quantum computation remains without an entirely satisfying answer: what is the essential quantum resource that provides the speedup of quantum computation over classical computation? This is clearly an important question as its resolution could inform the development of quantum hardware and the design of quantum computer architectures.

One inroad to approaching this question comes from quantum computation with magic states (QCM) [11, 12]. QCM—a universal model of quantum computation closely related to the circuit model—is one of the leading candidates for scalable fault-tolerant quantum computation [13]. In QCM, the allowed operations are restricted to a subset of unitary gates forming the Clifford group, as well as arbitrary Pauli measurements. These operations by themselves are not universal for quantum computation. In fact, any quantum circuit consisting of only these operations can be simulated efficiently on a classical computer [14, 15], and so with these operations alone no quantum computational speedup is possible. Universality is restored in QCM through the inclusion of additional nonstabilizer quantum states to the input of the circuit. Therefore, this model allows us to refine the question posed above. Instead of asking broadly "which nonclassical resources are required for a quantum computational speedup?", we can focus on the quantum states and ask "which states could provide a speedup in QCM?"

A partial answer to this question is provided by the study of quasiprobability representations like the Wigner function [16]. The Wigner function is the closest quantum mechanical counterpart to the classical notion of a probability distribution over a phase space, but unlike a probability distribution it can take negative values making it a *quasi*probability function. Accordingly, negativity in the Wigner function has traditionally been considered an indicator distinguishing classically behaving quantum states from those that exhibit genuinely quantum features [17, 18]. When adapted to finite-dimensional quantum mechanics, the setting relevant for quantum computation, quantum states are represented by a discrete Wigner function [19–25]—a quasiprobability function over a finite set (a generalized phase space) usually satisfying certain constraints [26–28].

Veitch et al. [29] showed that a necessary condition for a quantum computational speedup in QCM on odd-prime-dimensional qudits—quantum systems with odd prime Hilbert space dimension—is that the discrete Wigner function of the input state of the quantum circuit must take negative values (this result is easily extended to QCM on qudits with any odd dimension [30]). In particular, the amount of negativity in the discrete Wigner function quantifies the cost of classical simulation of a quantum computation [31] with simulation being efficient if the Wigner function is nonnegative everywhere. Since nonnegativity of the discrete Wigner function also implies the existence of a classical (noncontextual) hidden variable model (HVM) describing the computation [32, 33], this proves that two traditional notions of nonclassicality for quantum systems—Wigner negativity and failure of a classical HVM description—herald a quantum computational advantage over classical computation. This aligns with work that shows contextuality is required for quantum advantage in other settings [34–36].

The usual discrete Wigner function cannot be used to extend this result to QCM on even-dimensional qudits, including arguably the most important case—QCM on qubits [28, 37–39]. That said, similar necessary conditions for quantum advantage in QCM have been proven based on other quasiprobability representations [40–45]. In all cases, negativity is required in the representation of states or measurements in order to describe universal quantum computation.

Recently, a hidden variable model was defined which bucks this trend by representing all quantum states, operations, and measurements relevant for QCM on qubits using only classical (nonnegative) probabilities [46]. This model is structurally similar to previous quasiprobability representations (modulo absence of negativity) and leads to a classical simulation method for universal quantum computation based on sampling from the defining probability distributions. In this paper we show that this result can be significantly extended in that a nonnegative hidden variable model can be constructed for quantum computation with magic states on qudits of any dimension. We also show that many of the properties of the qubit HVM also apply in the qudit case, for example, this model leads to a classical simulation algorithm for quantum computation, and it subsumes previously defined quasiprobability representations. Note that, although this model can simulate any quantum computation, the simulation is presumably inefficient in general.

The remainder of this paper is structured as follows. In Section 2 we review some background material on quantum computation with magic states and the stabilizer formalism for qudits. In Section 3 we define the hidden variable model alluded to above. In Section 4 we present a classical simulation algorithm for quantum computation with magic states based on sampling from the probability distributions that define the model. In Section 5 we characterize a subset of the hidden variables of the model. Finally, in Section 6 we consider some properties of the qubit HVM [47] which extend to the qudit case. We conclude with a discussion of the significance of these results in Section 7.

2 Preliminaries

The setting of this paper is quantum computation with magic states (QCM) [11, 12] on multiple qudits, i.e., on multiple d-level quantum systems. This is a universal model of quantum computation in which computation proceeds through the application a sequence of Clifford gates and Pauli measurements on an initially prepared "magic state". Formally, for a system of n d-dimensional qudits, the allowed measurements are associated with elements of the generalized Pauli group, $P = \langle \omega, X_k, Z_k \mid 1 \leq k \leq n \rangle$ where $\omega = \exp(2\pi i/d)$ is a primitive d^{th} root of unity and the local Pauli operators are the d-dimensional generalization of the standard qubit Pauli operators [48], given by

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

Here $\mathbb{Z}_d$ is the ring of integers modulo d. With overall phases modded out we have $\mathcal{P} := P/\mathcal{Z}(P) \cong \mathbb{Z}_d^{2n}$ and we can parametrize the Pauli operators by $a = (a_Z, a_X) \in \mathbb{Z}_d^n \times \mathbb{Z}_d^n =: E$ as

$$T_a = \mu^{\phi(a)} Z(a_Z) X(a_X) \tag{2}$$

with $Z(a_Z) = \bigotimes_{k=1}^n Z^{a_Z[k]}$, $X(a_X) = \bigotimes_{k=1}^n X^{a_X[k]}$. Here $\mu = \omega$ when d is odd, and $\mu = \sqrt{\omega}$ when d is even. The phase function ϕ can be chosen freely subject to the constraint $(T_a)^d = \mathbb{1}$ for all $a \in E$. This constraint forces the eigenvalues of the operators to be in $\{\omega^j \mid j \in \mathbb{Z}_d\}$. For concreteness we can choose the phase function to be

$$\phi(a) = \begin{cases} -\langle a_Z | a_X \rangle \cdot 2^{-1} & \text{if } d \text{ is odd} \\ -\langle a_Z | a_X \rangle & \text{if } d \text{ is even} \end{cases} \tag{3}$$

where the inner product $\langle a_Z | a_X \rangle := \sum_{k=1}^n a_Z[k] a_X[k]$ is computed mod d if d is odd, and mod $2d$ if d is even.

The gates of the model are the Clifford gates, which are drawn from the normalizer of the Pauli group in the unitary group (up to overall phases):

$$\mathcal{C}\ell = \mathcal{N}(\mathcal{P})/U(1). \tag{4}$$

The Clifford gates are not required for quantum computational universality of this model since they can always be propagated past the Pauli measurements, conjugating them into other Pauli measurements [40, 49]. After they are propagated past the final Pauli measurements they can be dropped since they no longer affect the statistics of the measurements. In the following we include the Clifford gates anyway for completeness.

The last primitive required for QCM is the preparation of so-called "magic" input states. It is these states which allow for the universality of QCM. If the set of input states in QCM were restricted to only include n-qudit stabilizer states then the model would not be universal, and in fact any circuit of this type could be efficiently simulated on a classical computer. This is the result of the Gottesman-Knill theorem [14, 48]. Magic states are any nonstabilizer states which allow for universality in the QCM model.

See Appendix A for more background on Pauli measurements, Clifford gates, and the stabilizer formalism for qudits.

Before proceeding we need to introduce some additional notation. The symplectic inner product $[\cdot,\cdot] : E \times E \longrightarrow \mathbb{Z}_d$ defined by

$$[a,b] := \langle a_Z | b_X \rangle - \langle a_X | b_Z \rangle \tag{5}$$

tracks the commutator of the generalized Pauli operators in the sense

$$[T_a, T_b] := T_a T_b T_a^{-1} T_b^{-1} = \omega^{[a,b]} \mathbb{1}. \tag{6}$$

Because of this correspondence we will say that elements $a, b \in E$ commute when $[a,b] = 0$. The Pauli group with phases modded out forms a normal subgroup of the Clifford group such that $\mathcal{C}\ell/\mathcal{P} \cong \mathrm{Sp}(E)$ is the group of symplectic transformations on E. The Clifford group acts on the Pauli group by conjugation as

$$g T_a g^\dagger = \omega^{\tilde{\Phi}_g(a)} T_{S_g(a)} \quad \forall g \in \mathcal{C}\ell \; \forall a \in E \tag{7}$$

where $S_g \in \mathrm{Sp}(E)$ is a symplectic map, and the function $\tilde{\Phi}_g : E \to \mathbb{Z}_d$ tracks the extra phases that get picked up.

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

$$T_a T_b = \omega^{-\beta(a,b)} T_{a+b}. \tag{8}$$

The functions $\tilde{\Phi}$ and β have a cohomological interpretation elucidated in Ref. [50] (also see Refs. [28, 45]).

Recall from Ref. [45] a few definitions.

Definition 1 *A set $\Omega \subset E$ is* closed under inference *if for every pair of elements $a, b \in \Omega$ satisfying $[a,b] = 0$, it holds that $a + b \in \Omega$. The* closure under inference *of a set $\Omega \subset E$, denoted $\bar{\Omega}$, is the smallest subset of E which contains Ω and is closed under inference.*

Definition 2 *A set $\Omega \subset E$ is called* noncontextual *if there exists a noncontextual value assignment for the closure of that set, i.e., there exists a function $\gamma : \bar{\Omega} \to \mathbb{Z}_d$ such that for all $a, b \in \Omega$ with $[a,b] = 0$, γ satisfies*

$$\gamma(a) + \gamma(b) - \gamma(a+b) = -\beta(a,b). \tag{9}$$

A set which is both closed under inference and noncontextual we call cnc for short.

For any isotropic subgroup $I \subset E$, i.e. any subgroup I of E on which the symplectic product is identically zero, and any noncontextual value assignment $r : I \to \mathbb{Z}_d$, the operator

$$\Pi_I^r = \frac{1}{|I|} \sum_{b \in I} \omega^{-r(b)} T_b \tag{10}$$

is the projector onto the simultaneous $+1$-eigenspace of the operators $\{\omega^{-r(b)} T_b \mid b \in I\}$. This represents a measurement of the Pauli observables labeled by I yielding measurement outcomes $\omega^{r(b)}$. In particular, for a single Pauli measurement $I = \langle a \rangle$ is generated by a single element $a \in E$, and when $|I| = d^n$, Π_I^r is a projector onto a stabilizer state.

3 Hidden variable model

In this section we define a hidden variable model that represents all components of quantum computation with magic states by a family of probability distributions. This is in contrast to previous quasiprobability representations which required negativity in the representation of either the states or the operations of QCM in order to represent universal quantum computation. The main result of this section is Theorem 1.

Let $\mathrm{Herm}(\mathcal{H})$ be the space of Hermitian operators on n-qudit Hilbert space $\mathcal{H} \cong \mathbb{C}^{d^n}$, $\mathrm{Herm}_1(\mathcal{H})$ be the affine subspace of this space obtained by fixing the trace of the operators to be 1, and let $\mathrm{Herm}_1^{\succeq 0}(\mathcal{H})$ be the subset of $\mathrm{Herm}_1(\mathcal{H})$ consisting of positive semidefinite operators. Let $\mathcal{S}$ denote the set of pure n-qudit stabilizer states. The state space of the hidden variable model is based on the set

$$\Lambda = \{X \in \mathrm{Herm}_1(\mathcal{H}) | \, \mathrm{Tr}(|\sigma\rangle\langle\sigma| \, X) \geq 0 \; \forall \, |\sigma\rangle \in \mathcal{S}\}. \tag{11}$$

The elements of Λ are much like density operators in that they are Hermitian operators with unit trace, but unlike density operators they are not necessarily positive semidefinite. In order to define the hidden variable model we first need to establish some basic properties of Λ.

Lemma 1 *For any number of qudits $n \in \mathbb{N}$ of any dimension $d \in \mathbb{N}$, (i) Λ is convex, and (ii) Λ is compact.*

The proof of Lemma 1 is in Appendix B.

Λ can be interpreted as a subset of a real affine space defined by the intersection of a finite number of linear inequalities, i.e., it is a polyhedral set. Since Λ is compact, it is a polytope, and so by the Minkowski-Weyl theorem [51] it can equivalently be described as the convex hull of finitely many vertices. Let $\{A_\alpha \mid \alpha \in \mathcal{V}\}$ denote the (finite) set of vertices of Λ. Then we have the following result, which is a generalization of [30, Theorem 1] to qudits of arbitrary Hilbert space dimension d.

Theorem 1 *For any number $n \in \mathbb{N}$ of qudits with any Hilbert space dimension $d \in \mathbb{N}$,*

1. *For any quantum state $\rho \in Herm_1^{\succeq 0}(\mathcal{H})$, there is a probability function $p_\rho : \mathcal{V} \to \mathbb{R}_{\geq 0}$ such that*
$$\rho = \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) A_\alpha. \tag{12}$$

2. *For any vertex A_α of Λ and any Clifford operation $g \in \mathcal{C}\ell$, $gA_\alpha g^\dagger =: A_{g\cdot\alpha}$ is a vertex of Λ.*

3. *For update under Pauli measurements it holds that for any isotropic subgroup $I \subset E$, any noncontextual value assignment $r : I \to \mathbb{Z}_d$, and any vertex A_α,*
$$\Pi_I^r A_\alpha \Pi_I^r = \sum_{\alpha' \in \mathcal{V}} q_{\alpha,I}(\alpha', r) A_{\alpha'}, \tag{13}$$
where $q_{\alpha,I}(\alpha', r) \geq 0$ for all $\alpha' \in \mathcal{V}$ and $\sum_{\alpha',r} q_{\alpha,I}(\alpha', r) = 1$.

4. *The Born rule takes the form*
$$\mathrm{Tr}(\Pi_I^r \rho) = \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) Q_I(r \mid \alpha) \tag{14}$$
where Q is given by
$$Q_I(r \mid \alpha) = \sum_{\alpha' \in \mathcal{V}} q_{\alpha,I}(\alpha', r). \tag{15}$$

This theorem defines a hidden variable model which represents all of the primitives of quantum computation with magic states using only probability distributions for states and measurements, and probabilistic update rules for dynamics. It has a similar structure to previous hidden variable models based on quasiprobability representations [23–25, 30, 31, 33, 40, 43, 45, 46, 52], with a key difference being that in this model every state can be represented by a probability distribution, no negativity is required. Unlike the model of Beltrametti and Bugajski [53] which requires a hidden variable for each pure quantum state, this model has a finite number of hidden variables for any number of qudits.

The proof of Theorem 1 requires the following lemma.

Lemma 2 *The polytope Λ has the following properties:*

1. *For any density operator $\rho \in Herm_1^{\succeq 0}(\mathcal{H})$ representing a physical n-qudit quantum state, $\rho \in \Lambda$.*
2. *For any $X \in \Lambda$ and any Clifford operation $g \in \mathcal{C}\ell$, $gXg^\dagger \in \Lambda$,*
3. *For any $X \in \Lambda$, any isotropic subgroup $I \subset E$, and any noncontextual value assignment $r : I \to \mathbb{Z}_d$, if $\mathrm{Tr}(\Pi_I^r X) > 0$ then*

$$\frac{\Pi_I^r X \Pi_I^r}{\mathrm{Tr}(\Pi_I^r X)} \in \Lambda. \tag{16}$$

Proof of Lemma 2. We will prove the three properties of the lemma in order. For the first property note that for any density operator $\rho \in \mathrm{Herm}_1^{\succeq 0}(\mathcal{H})$, ρ is positive semidefinite. Therefore, for any pure quantum state $|\psi\rangle$, $\mathrm{Tr}(|\psi\rangle\langle\psi|\,\rho) \geq 0$. This holds in particular for any pure stabilizer state. Therefore, ρ satisfies all of the defining inequalities of the polytope in eq. (11) and so $\rho \in \Lambda$.

For the second property, let $X \in \Lambda$ and $g \in \mathcal{C}\ell$. Then for any stabilizer state $|\sigma\rangle \in S$,

$$\mathrm{Tr}\Big(|\sigma\rangle\langle\sigma|\,(gXg^\dagger)\Big) = \mathrm{Tr}\Big((g^\dagger|\sigma\rangle\langle\sigma|\,g)X\Big) \tag{17}$$

$$= \mathrm{Tr}(|\sigma'\rangle\langle\sigma'|\,X) \geq 0. \tag{18}$$

Here the first equality follows from the cyclic property of the trace, the second equality from the fact that Clifford operations map stabilizer states to stabilizer states (see Lemma 6 in Appendix A), and the last inequality from the assumption $X \in \Lambda$.

Now we can prove the third property of Lemma 2. Let $I, J \subset E$ be isotropic subgroups with noncontextual value assignments $r : I \to \mathbb{Z}_d$ and $s : J \to \mathbb{Z}_d$. By Lemma 7 in Appendix A,

$$\Pi_I^r \Pi_J^s \Pi_I^r = \delta_{r|_{I\cap J}, s|_{I\cap J}} \frac{|J \cap I^\perp|}{|J|} \Pi^{r\star s}_{I+(J\cap I^\perp)} \tag{19}$$

where $\delta_{r|_{I\cap J}, s|_{I\cap J}}$ is equal to one if r and s agree on the intersection $I \cap J$, and it is zero otherwise. Here $r \star s$ is the unique noncontextual value assignment on the set $I + (J \cap I^\perp)$ such that $r \star s|_I = r$ and $r \star s|_{J\cap I^\perp} = s|_{J\cap I^\perp}$. For any $X \in \Lambda$ and any Pauli projector Π_I^r, if $\mathrm{Tr}(\Pi_I^r X) > 0$, then for any projector onto a stabilizer state Π_J^s,

$$\mathrm{Tr}\left(\Pi_J^s \frac{\Pi_I^r X \Pi_I^r}{\mathrm{Tr}(\Pi_I^r X)}\right) = \frac{\mathrm{Tr}((\Pi_I^r \Pi_J^s \Pi_I^r)X)}{\mathrm{Tr}(\Pi_I^r X)} \tag{20}$$

$$= \delta_{r|_{I\cap J}, s|_{I\cap J}} \frac{|J \cap I^\perp|}{|J|} \frac{\mathrm{Tr}\Big(\Pi^{r\star s}_{I+(J\cap I^\perp)} X\Big)}{\mathrm{Tr}(\Pi_I^r X)} \geq 0. \tag{21}$$

Here the first line follows from linearity and the cyclic property of the trace. The last inequality follows from the fact that by Lemma 5 in Appendix A, $\Pi^{r\star s}_{I+(J\cap I^\perp)}$ can be written as a conic combination of projectors onto stabilizer states, and from the assumption $X \in \Lambda$. Therefore, for any $X \in \Lambda$, any Pauli projector Π^r_I, and any stabilizer state $|\sigma\rangle \in \mathcal{S}$, if $\mathrm{Tr}(\Pi^r_I X) > 0$, then

$$\mathrm{Tr}\left(|\sigma\rangle\langle\sigma| \frac{\Pi^r_I X \Pi^r_I}{\mathrm{Tr}(\Pi^r_I X)}\right) \geq 0 \tag{22}$$

and so $\Pi^r_I X \Pi^r_I / \mathrm{Tr}(\Pi^r_I X) \in \Lambda$. This proves the third statement of the lemma. □

We can now prove the main result of this section.

Proof of Theorem 1. We will prove the four statements of the theorem in order. First, as shown in Lemma 2, Λ contains all density matrices corresponding to physical n-qudit quantum states. Therefore, by the Krein-Milman theorem [51] any state can be written as a convex combination of the vertices of Λ. This is the first statement of the theorem.

The second property from Lemma 2 shows that for any Clifford operation $g \in \mathcal{C}\ell$ and any vertex A_α of Λ, we have $A_{g\cdot\alpha} := gA_\alpha g^\dagger \in \Lambda$. It remains to show that $A_{g\cdot\alpha}$ is a vertex of Λ.

Let $\mathcal{S}_\alpha = \{|\sigma\rangle \in \mathcal{S} \mid \mathrm{Tr}(|\sigma\rangle\langle\sigma| A_\alpha) = 0\}$ be the set of stabilizer states with projectors orthogonal to vertex A_α with respect to the Hilbert-Schmidt inner product. By Theorem 18.1 of Ref. [54], since A_α is a vertex of Λ, A_α is the unique solution in $\mathrm{Herm}_1(\mathcal{H})$ of the system

$$\begin{cases} \mathrm{Tr}(|\sigma\rangle\langle\sigma| X) = 0 & \forall\, |\sigma\rangle \in \mathcal{S}_\alpha \\ \mathrm{Tr}(|\sigma\rangle\langle\sigma| X) \geq 0 & \forall\, |\sigma\rangle \in \mathcal{S} \setminus \mathcal{S}_\alpha. \end{cases} \tag{23}$$

For any stabilizer state $|\sigma\rangle \in \mathcal{S}_\alpha$,

$$\mathrm{Tr}(|\sigma\rangle\langle\sigma| X) = \mathrm{Tr}\Big(g|\sigma\rangle\langle\sigma| g^\dagger g X g^\dagger\Big). \tag{24}$$

Therefore, under conjugation by $g \in \mathcal{C}\ell$, solutions to the system eq. (23) are mapped bijectively to solutions of the system

$$\begin{cases} \mathrm{Tr}(|\sigma\rangle\langle\sigma| X) = 0 & \forall\, |\sigma\rangle \in \mathcal{S}_{g\cdot\alpha} \\ \mathrm{Tr}(|\sigma\rangle\langle\sigma| X) \geq 0 & \forall\, |\sigma\rangle \in \mathcal{S} \setminus \mathcal{S}_{g\cdot\alpha} \end{cases} \tag{25}$$

where $S_{g\cdot\alpha} := \{g|\sigma\rangle \mid |\sigma\rangle \in S_\alpha\}$. In particular, $A_{g\cdot\alpha} := gA_\alpha g^\dagger$ is the unique solution to this system, so by Theorem 18.1 of Ref. [54], it is a vertex of Λ.

For the third statement of Theorem 1, let A_α be a vertex of Λ and Π^r_I be any Pauli projector. We have two cases. (I) First, if $\mathrm{Tr}(\Pi^r_I A_\alpha) = 0$, then $\Pi^r_I A_\alpha \Pi^r_I$ is zero as an operator. To see this, consider the inner product $\mathrm{Tr}(T_a \Pi^r_I A_\alpha \Pi^r_I)$ for any Pauli operator T_a, $a \in E$. Here we have three subcases: (i) if $a \in I$, then

$$T_a \Pi^r_I = \frac{1}{|I|}\sum_{b\in I} \omega^{-r(b)} T_a T_b \tag{26}$$

$$= \frac{1}{|I|}\sum_{b\in I} \omega^{-r(b)-\beta(a,b)} T_{a+b} \tag{27}$$

$$= \frac{1}{|I|}\sum_{b\in I} \omega^{r(a)-r(a+b)} T_{a+b} \tag{28}$$

$$= \omega^{r(a)} \Pi^r_I. \tag{29}$$

Therefore,

$$\mathrm{Tr}(T_a \Pi_I^r A_\alpha \Pi_I^r) = \omega^{r(a)} \mathrm{Tr}(\Pi_I^r A_\alpha) = 0. \tag{30}$$

(ii) If $a \in I^\perp \setminus I$, then by Lemma 5 in Appendix A we have

$$\Pi_I^r = \sum_{\gamma \in \Gamma_{I,r}^{\langle a,I\rangle}} \Pi_{\langle a,I\rangle}^\gamma \tag{31}$$

where $\Gamma_{I,r}^{\langle a,I\rangle}$ is the set of noncontextual value assignments on $\langle a, I\rangle$ satisfying $\gamma|_I = r$. Multiplying this equation on the right by A_α and taking a trace we get

$$\mathrm{Tr}(\Pi_I^r A_\alpha) = \sum_{\gamma \in \Gamma_{I,r}^{\langle a,I\rangle}} \mathrm{Tr}\Big(\Pi_{\langle a,I\rangle}^\gamma A_\alpha\Big). \tag{32}$$

With Lemma 5, each projector on the right hand side can be written as a sum of projectors onto stabilizer states. Therefore, since $A_\alpha \in \Lambda$ each term on the right hand side is nonnegative. But the left hand side is zero by assumption. Therefore, each term on the right hand side is zero. I.e. $\mathrm{Tr}\Big(\Pi_{\langle a,I\rangle}^\gamma A_\alpha\Big) = 0$ for every $\gamma \in \Gamma_{I,r}^{\langle a,I\rangle}$.

We can write the spectral decomposition of the operator T_a as

$$T_a = \sum_{\gamma \in \Gamma_{I,r}^{\langle a,I\rangle}} \omega^{\gamma(a)} \Pi_{\langle a\rangle}^{\gamma|_{\langle a\rangle}}. \tag{33}$$

Then

$$\mathrm{Tr}(T_a \Pi_I^r A_\alpha \Pi_I^r) = \sum_{\gamma \in \Gamma_{I,r}^{\langle a,I\rangle}} \omega^{\gamma(a)} \mathrm{Tr}\Big(\Pi_{\langle a\rangle}^{\gamma|_{\langle a\rangle}} \Pi_I^r A_\alpha \Pi_I^r\Big) \tag{34}$$

$$= \sum_{\gamma \in \Gamma_{I,r}^{\langle a,I\rangle}} \omega^{\gamma(a)} \mathrm{Tr}\Big(\Pi_{\langle a,I\rangle}^\gamma A_\alpha\Big) = 0. \tag{35}$$

(iii) If $a \notin I^\perp$, then

$$\Pi_I^r T_a \Pi_I^r = \frac{1}{|I|^2} \sum_{b,c \in I} \omega^{-r(b)-r(c)} T_b T_a T_c \tag{36}$$

$$= \frac{1}{|I|^2} \sum_{b,c \in I} \omega^{-r(b)-r(c)+[b,a]-\beta(b,c)} T_a T_{b+c} \tag{37}$$

$$= \frac{1}{|I|^2} T_a \sum_{b,c \in I} \omega^{-r(b+c)+[b,a]} T_{b+c} \tag{38}$$

$$= \frac{1}{|I|} T_a \Pi_I^r \sum_{b \in I} \omega^{[b,a]}. \tag{39}$$

By character orthogonality, the sum in the final expression vanishes. Therefore,

$$\mathrm{Tr}(T_a \Pi_I^r A_\alpha \Pi_I^r) = \mathrm{Tr}(\Pi_I^r T_a \Pi_I^r A_\alpha) = 0. \tag{40}$$

We have $\mathrm{Tr}(T_a \Pi_I^r A_\alpha \Pi_I^r) = 0$ for every Pauli operator T_a, $a \in E$. Therefore, $\Pi_I^r A_\alpha \Pi_I^r$ is zero as an operator.

(II) Second, if $\mathrm{Tr}(\Pi_I^r A_\alpha) > 0$, then by the third statement of Lemma 2 we have $\Pi_I^r A_\alpha \Pi_I^r / \mathrm{Tr}(\Pi_I^r A_\alpha) \in \Lambda$, and so there exists a decomposition of $\Pi_I^r A_\alpha \Pi_I^r / \mathrm{Tr}(\Pi_I^r A_\alpha)$ as

a convex combination of the vertices of Λ. Therefore, there exist nonnegative coefficients $q_{\alpha,I}(\alpha', r)$ such that

$$\Pi_I^r A_\alpha \Pi_I^r = \sum_{\alpha' \in \mathcal{V}} q_{\alpha,I}(\alpha', r) A_{\alpha'}. \tag{41}$$

Taking a trace of this equation and adding the corresponding equations for all noncontextual value assignments r of I we have on the left hand side

$$\sum_r \mathrm{Tr}(\Pi_I^r A_\alpha \Pi_I^r) = \mathrm{Tr}\left[\left(\sum_r \Pi_I^r\right) A_\alpha\right] = \mathrm{Tr}(A_\alpha) = 1 \tag{42}$$

and on the right hand side

$$\sum_{\alpha', r} q_{\alpha,I}(\alpha', r) \,\mathrm{Tr}(A_{\alpha'}) = \sum_{\alpha', r} q_{\alpha,I}(\alpha', r). \tag{43}$$

Therefore, $\sum_{\alpha', r} q_{\alpha,I}(\alpha', r) = 1$. This proves the third statement of the Theorem.

Finally, we calculate

$$\begin{aligned} \mathrm{Tr}(\Pi_I^r \rho) &= \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) \,\mathrm{Tr}(\Pi_I^r A_\alpha) && (44)\\ &= \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) \sum_{\alpha' \in \mathcal{V}} q_{\alpha,I}(\alpha', r) && (45)\\ &= \sum_{\alpha \in \mathcal{V}} p_\rho(\alpha) Q_I(r \mid \alpha) && (46) \end{aligned}$$

and we obtain the fourth statement of the theorem. □

4 Classical simulation algorithm

Theorem 1 shows that all of the components of quantum computation with magic states can be described by a hidden variable model which represents all relevant states and dynamical operations by probabilities. This leads to a classical simulation algorithm for quantum computation with magic states, Algorithm 1, based on sampling from these probability distributions.

In short, a vertex label $\alpha \in \mathcal{V}$ is sampled according to the probability distribution of eq. (12) representing the input state of the quantum circuit. This vertex is then propagated through the circuit. When a Clifford gate $g \in \mathcal{C}\ell$ is encountered, we have a deterministic update rule: $\alpha \to g \cdot \alpha$, according to the second statement of Theorem 1. When a Pauli measurement $a \in E$ is encountered, the third and fourth statements of Theorem 1 give a way of determining probabilities for measurement outcomes, as well as a probabilistic update rule. That is, we sample a pair (α', r) according to the probability distribution $q_{\alpha,\langle a\rangle}$ where $\alpha' \in \mathcal{V}$ and $r : \langle a\rangle \to \mathbb{Z}_d$ is a noncontextual value assignment. Then $r(a)$ is returned as the measurement outcome and the vertex is updated as $\alpha \to \alpha'$. This process continues until the end of the circuit is reached.

Input: p_{ρ_0}

sample a point $\alpha \in \mathcal{V}$ according to the probability distribution p_{ρ_0}
while end of circuit has not been reached **do**
  if a Clifford gate $g \in \mathcal{C}\ell$ is encountered **then**
    update $\alpha \leftarrow g \cdot \alpha$
  if a Pauli measurement T_a, $a \in E$ is encountered **then**
    sample (α', r) according to the probability distribution $q_{\alpha,\langle a\rangle}$
    Output: $r(a)$ as the outcome of the measurement
    update $\alpha \leftarrow \alpha'$

Algorithm 1: One run of the classical simulation algorithm for quantum computation with magic states based on the hidden variable model of Theorem 1. The algorithm provides samples from the joint probability distribution of the Pauli measurements in a quantum circuit consisting of Clifford unitaries and Pauli measurements applied to an input state ρ_0.

A proof of the correctness of this simulation algorithm is given below.

Theorem 2 *The classical simulation algorithm, Algorithm 1, correctly reproduces the predictions of quantum theory.*

Proof of Theorem 2. Without loss of generality, a QCM circuit can be represented as a sequence $g_1, I_1, g_2, I_2, \ldots$ with $g_1, g_2, \cdots \in \mathcal{C}\ell$ specifying the Clifford unitaries to be applied, and $I_1, I_2, \cdots \subset E$ isotropic subgroups specifying the Pauli measurements to be performed. First, consider a single layer of this circuit consisting of a Clifford gate $g \in \mathcal{C}\ell$ followed by Pauli measurements corresponding to an isotropic subgroup $I \subset E$.

Using the classical simulation algorithm, Algorithm 1, the conditional probability of obtaining measurement outcomes specified by the noncontextual value assignment $r : I \to \mathbb{Z}_d$ for the measurements given the state $\alpha \in \mathcal{V}$ is $Q_I(r \mid g\cdot\alpha)$. Therefore, the probability of obtaining outcomes r given the gate g is applied to the state ρ followed by the measurements of the Pauli observables in I is given by

$$P^{(Sim)}_{\rho,g,I}(r) = \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha) Q_I(r \mid g\cdot\alpha). \tag{47}$$

The corresponding outcome probability predicted by the Born rule, $P^{(QM)}_{\rho,g,I}(r)$, is

$$\begin{aligned}
\mathrm{Tr}\left(\Pi^r_I g\rho g^\dagger\right) &= \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha)\,\mathrm{Tr}\left(\Pi^r_I g A_\alpha g^\dagger\right) &(48)\\
&= \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha)\,\mathrm{Tr}(\Pi^r_I A_{g\cdot\alpha}) &(49)\\
&= \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha) Q_I(r \mid g\cdot\alpha). &(50)
\end{aligned}$$

Here in the first line we use the expansion of ρ in the vertices of Λ, eq. (12), in the second line we use the second statement of Theorem 1, and in the last line we use the fourth statement of Theorem 1. This agrees with the outcome probability predicted by the classical simulation algorithm.

Now consider the postmeasurement state ρ'. According to quantum mechanics, the postmeasurement state is

$$\rho'^{(QM)} = \frac{\Pi^r_I g\rho g^\dagger \Pi^r_I}{\mathrm{Tr}(\Pi^r_I g\rho g^\dagger)}. \tag{51}$$

Here the numerator is

$$\Pi_I^r g\rho g^\dagger \Pi_I^r = \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha)\Pi_I^r g A_\alpha g^\dagger \Pi_I^r \tag{52}$$

$$= \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha)\Pi_I^r A_{g\cdot\alpha}\Pi_I^r \tag{53}$$

$$= \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha) \sum_{\alpha'\in\mathcal{V}} q_{g\cdot\alpha,I}(\alpha', r)A_{\alpha'} \tag{54}$$

and the denominator is

$$\mathrm{Tr}\left(\Pi_I^r g\rho g^\dagger \Pi_I^r\right) = \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha) \sum_{\alpha'\in\mathcal{V}} q_{g\cdot\alpha,I}(\alpha', r) \tag{55}$$

$$= \sum_{\alpha\in\mathcal{V}} p_\rho(\alpha) Q_I(r \mid g\cdot\alpha), \tag{56}$$

so the postmeasurement state predicted by quantum theory is

$$\rho'^{(QM)} = \frac{\sum_\alpha p_\rho(\alpha) \sum_{\alpha'} q_{g\cdot\alpha,I}(\alpha', r)A_{\alpha'}}{\sum_\alpha p_\rho(\alpha) Q_I(r \mid g\cdot\alpha)}. \tag{57}$$

Using the classical simulation algorithm, the probability of obtaining outcomes r and state $A_{\alpha'}$ given a Clifford gate g followed by measurements of the Pauli observables I on state ρ is $P_{\rho,g,I}(\alpha', r) = P_{\rho,g,I}(\alpha'|r)P_{\rho,g,I}(r)$. But $P_{\rho,g,I}(\alpha', r) = \sum_\alpha p_\rho(\alpha)P_{g,I}(\alpha', r|\alpha) = \sum_\alpha p_\rho(\alpha) q_{g\cdot\alpha,I}(\alpha', r)$, and $P_{\rho,g,I}(\alpha'|r) = p_{\rho'}(\alpha')$. Therefore, the postmeasurement state predicted by the classical simulation algorithm is

$$\rho'^{(Sim)} = \sum_{\alpha'\in\mathcal{V}} p_{\rho'}(\alpha')A_{\alpha'} \tag{58}$$

$$= \sum_{\alpha'\in\mathcal{V}} \frac{P_{\rho,g,I}(\alpha', r)}{P_{\rho,g,I}(r)} A_{\alpha'} \tag{59}$$

$$= \sum_{\alpha'\in\mathcal{V}} \frac{\sum_\alpha p_\rho(\alpha) q_{g\cdot\alpha,I}(\alpha', r)}{\sum_\alpha p_\rho(\alpha) Q_I(r \mid g\cdot\alpha)} A_{\alpha'}. \tag{60}$$

This agrees with the postmeasurement state predicted by quantum mechanics. Therefore, the classical simulation algorithm correctly reproduces the outcome probabilities and the postmeasurement state predicted by quantum mechanics for a single layer of a QCM circuit.

Now let $\rho(t)$ denote the state after t layers of the circuit. Then the argument above shows that the classical simulation algorithm correctly reproduces the Born rule probabilities $P_{\rho_0,g_{t+1},I_{t+1}}(r_{t+1} \mid r_1, r_2, \ldots, r_t)$ as well as the postmeasurement state $\rho(t+1)$. Therefore, by induction, the simulation algorithm correctly reproduces the outcome probabilities predicted by the Born rule for any QCM circuit. □

5 Partial characterization of vertices of Λ

In Ref. [45] a classical simulation algorithm for quantum computation with magic states is introduced based on a quasiprobability representation. Points in the generalized phase space over which the quasiprobability function is defined are associated with pairs (Ω, γ) where $\Omega \subset E$ is a cnc set and $\gamma : \Omega \to \mathbb{Z}_d$ is a noncontextual value assignment for Ω. For each point in phase space there is a corresponding phase space point operator defined as

$$A_\Omega^\gamma = \frac{1}{d^n} \sum_{b\in\Omega} \omega^{-\gamma(b)} T_b. \tag{61}$$

For qubits, if Ω is a maximal cnc set then phase point operators A^γ_Ω of the form eq. (61) are vertices of Λ [46, 55]. Vertices of the type eq. (61) we call cnc vertices.

A similar result holds for odd-prime-dimensional qudits. Namely, phase point operators of the form A^γ_E define facets of the stabilizer polytope [22, 29], or equivalently, by polar duality, A^γ_E are vertices of Λ. Here we show that this holds for qudits with any odd Hilbert space dimension. This is the result of the following theorem.

Theorem 3 *For any number n of qudits with any odd Hilbert space dimension d, phase space point operators of the form*

$$A^\gamma_E = \frac{1}{d^n} \sum_{b \in E} \omega^{-\gamma(b)} T_b \tag{62}$$

where $\gamma : E \longrightarrow \mathbb{Z}_d$ is a noncontextual value assignment for E are vertices of Λ.

This theorem provides a partial characterization of the vertices of Λ which define the hidden variable model of Theorem 1. When $n \geq 2$, the operators A^γ_E are exactly the phase space point operators of Gross' Wigner function for odd-dimensional qudits [23, 25, 33], therefore, the standard multi-qudit phase point operators are a subset of the hidden variables of our model (when $n = 1$, Gross' phase space point operators still have this form, but there are also other operators with this form [33]). Note that the phase point operators of the form eq. (62) only exist when the qudit Hilbert space dimension d is odd since noncontextual value assignments on E exist only when d is odd [50]. In the case of qubits, some additional classes of vertices have been characterized [50], but the vertices with the simplest description are the cnc type vertices of eq. (61).

The proof of Theorem 3 requires the following lemma.

Lemma 3 *For any vertex A_α of Λ and any Pauli operator T_a, $|\operatorname{Tr}(T_a A_\alpha)| \leq 1$.*

Proof of Lemma 3. For any $\alpha \in \mathcal{V}$ and any $a \in E$,

$$\operatorname{Tr}(T_a A_\alpha) = \sum_{s \in \Gamma^{\langle a \rangle}} \operatorname{Tr}\left(\Pi^s_{\langle a \rangle} A_\alpha\right) \omega^{s(a)} \tag{63}$$

where $\Gamma^{\langle a \rangle}$ is the set of noncontextual value assignments on $\langle a \rangle$ and $\Pi^s_{\langle a \rangle}$ is the projector onto the eigenspace of the Pauli observable T_a with eigenvalue $\omega^{s(a)}$. $\{\Pi^s_{\langle a \rangle} | s \in \Gamma^{\langle a \rangle}\}$ is a projection-valued measure, i.e. $\sum_s \Pi^s_{\langle a \rangle} = \mathbb{1}$. Therefore,

$$\sum_{s \in \Gamma^{\langle a \rangle}} \operatorname{Tr}\left(\Pi^s_{\langle a \rangle} A_\alpha\right) = \operatorname{Tr}\left[\left(\sum_{s \in \Gamma^{\langle a \rangle}} \Pi^s_{\langle a \rangle}\right) A_\alpha\right] = 1 \tag{64}$$

and so

$$|\operatorname{Tr}(T_a A_\alpha)| \leq \sum_{s \in \Gamma^{\langle a \rangle}} \left|\operatorname{Tr}\left(\Pi^s_{\langle a \rangle} A_\alpha\right) \omega^{s(a)}\right| \tag{65}$$

$$\leq \sum_{s \in \Gamma^{\langle a \rangle}} \left|\operatorname{Tr}\left(\Pi^s_{\langle a \rangle} A_\alpha\right)\right| \tag{66}$$

$$= \sum_{s \in \Gamma^{\langle a \rangle}} \operatorname{Tr}\left(\Pi^s_{\langle a \rangle} A_\alpha\right) = 1 \tag{67}$$

which proves the claimed bound. □

We can now prove Theorem 3.

Proof of Theorem 3. First we need to show that the phase point operators of eq. (62) are in Λ. This requires checking that the Hilbert-Schmidt inner product of A_E^γ with the projector onto any stabilizer state is nonnegative. For any maximal isotropic subgroup I with any noncontextual value assignment $r : I \to \mathbb{Z}_d$,

$$\operatorname{Tr}(\Pi_I^r A_E^\gamma) = \frac{1}{d^{2n}} \sum_{a\in I} \sum_{b\in E} \omega^{-r(a)-\gamma(b)} \operatorname{Tr}(T_a T_b) \tag{68}$$

$$= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)-\gamma(-a)} \tag{69}$$

$$= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)+\gamma(a)}. \tag{70}$$

To obtain the last equality, note that with the phase convention chosen in eq. (3) we have $T_{-a} = T_a^\dagger = T_a^{-1}$ and $T_0 = \mathbb{1}$. Therefore, $T_a T_{-a} = T_a T_a^{-1} = \mathbb{1} = T_{a+(-a)}$ and so by definition $\beta(a,-a) = 0$. Further, eq. (9) applied to the case $a = b = 0$ implies $\gamma(0) = -\beta(0,0) = 0$. Thus, with eq. (9), we have $\gamma(a) + \gamma(-a) - \gamma(0) = -\beta(a,-a)$, and so $\gamma(-a) = -\gamma(a)$.

Here we have two cases.

(I) If $r|_I = \gamma|_I$ then we have

$$\operatorname{Tr}(\Pi_I^r A_\Omega^\gamma) = \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)+\gamma(a)} = \frac{|I|}{d^n} = 1. \tag{71}$$

since $|I| = d^n$ by [56, Theorem 1].

(II) If $r|_I \neq \gamma|_I$ then by orthogonality of twisted characters [57], we have

$$\operatorname{Tr}(\Pi_I^r A_E^\gamma) = \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)+\gamma(a)} = 0. \tag{72}$$

Therefore, in both cases $\operatorname{Tr}(\Pi_I^r A_E^\gamma) \geq 0$. This proves that the phase point operators of eq. (62) are in Λ.

Since $A_E^\gamma \in \Lambda$, there exists an expansion of A_E^γ as a convex combination of the vertices of Λ:

$$A_E^\gamma = \sum_{\alpha\in\mathcal{V}} p(\alpha) A_\alpha. \tag{73}$$

Consider a Pauli operator T_a, $a \in E$. We have $\operatorname{Tr}(T_a A_E^\gamma) = \omega^{\gamma(a)}$ so multiplying eq. (73) by T_a and taking a trace we get

$$\omega^{\gamma(a)} = \sum_{\alpha\in\mathcal{V}} p(\alpha) \operatorname{Tr}(T_a A_\alpha) \tag{74}$$

Taking the absolute value of this equation we get on the left hand side $\left|\omega^{\gamma(a)}\right| = 1$ and on the right hand side

$$\left|\sum_{\alpha\in\mathcal{V}} p(\alpha) \operatorname{Tr}(T_a A_\alpha)\right| \leq \sum_{\alpha\in\mathcal{V}} p(\alpha) |\operatorname{Tr}(T_a A_\alpha)| \tag{75}$$

$$\leq \sum_{\alpha\in\mathcal{V}} p(\alpha) = 1. \tag{76}$$

Here the first inequality is the triangle inequality and the second inequality follows from Lemma 3. The second inequality is strict if $|\operatorname{Tr}(T_a A_\alpha)| < 1$ for any A_α with $p(\alpha) > 0$. If this were the case then this would lead to a contradiction: $1 < 1$. Therefore, $|\operatorname{Tr}(T_a A_\alpha)| = 1$ for all A_α with $p(\alpha) > 0$.

Now consider the real part of the equation above:

$$\operatorname{Re}\left[\omega^{\gamma(a)}\right] = \sum_{\alpha \in \mathcal{V}} p(\alpha) \operatorname{Re}\left[\operatorname{Tr}(T_a A_\alpha)\right]. \tag{77}$$

Since the coefficients are nonnegative and sum to one, this implies that $\operatorname{Re}\left[\operatorname{Tr}(T_a A_\alpha)\right] = \operatorname{Re}\left[\omega^{\gamma(a)}\right]$ for every α with $p(\alpha) > 0$. The same argument holds for the imaginary part. Thus, $\operatorname{Tr}(T_a A_\alpha) = \omega^{\gamma(a)}$ for all α with $p(\alpha) > 0$. I.e. each A_α that appears with nonzero weight in the expansion of A_E^γ in eq. (73) must agree with A_E^γ on the expectation of T_a for all $a \in E$. There is exactly one such operator in $\operatorname{Herm}_1(\mathcal{H})$, namely A_E^γ. Therefore, A_E^γ is a vertex of Λ. □

6 Mapping vertices of Λ_m to Λ_n

In this section we introduce a version of the Φ-map [47], that embeds the m-qudit polytope Λ_m as a subpolytope of the n-qudit polytope Λ_n where $n \geq m$, that works over qudits for an arbitrary $d \geq 2$. However, for $d \neq 2$ this map fails to map vertices of Λ_m to vertices of Λ_n.

We will regard E_m as a subgroup of E_n by identifying it with $\langle x_1, \cdots, x_m, z_1, \cdots, z_m \rangle$. We will write $E_{n-m}^{(n)}$ for the subgroup $\langle x_{m+1}, \cdots, x_n, z_{m+1}, \cdots, z_n \rangle$. These two subgroups intersect at the zero element and they generate the whole group. In this decomposition E_{n-m} is identified with $E_{n-m}^{(n)}$ via the map $x_i \mapsto x_{m+i}$ and $z_i \mapsto z_{m+i}$. Any isotropic subgroup $J \subset E_{n-m}$ will be identified with its image in $E_{n-m}^{(n)}$.

Theorem 4 *Let Π_J^r denote an $(n-m)$-qudit stabilizer projector and g denote an n-qudit Clifford unitary. The linear map*

$$\Phi_{g,J}^r : Herm\left(\mathbb{C}^{d^m}\right) \to Herm\left(\mathbb{C}^{d^n}\right), \quad X \mapsto g(X \otimes \Pi_J^r)g^\dagger \tag{78}$$

is injective and sends Λ_m to the subpolytope given by

$$\{(\Pi_{J'}^{r'} Y \Pi_{J'}^{r'})/Tr(Y \Pi_{J'}^{r'}) |\ Y \in \Lambda_n \text{ and } \operatorname{Tr}\left(Y \Pi_{J'}^{r'}\right) \neq 0\} \tag{79}$$

where $\Pi_{J'}^{r'} = g(\mathbb{1} \otimes \Pi_J^r)g^\dagger$.

This partially generalizes to qudits of arbitrary dimension a stronger result which applies only to the case $d = 2$ [50].

Proof of Theorem 4. To be able to distinguish Pauli operators we will write $T_a^{(n)}$ to indicate an n-qudit Pauli operator. With the choice of phase function ϕ in eq. (2) we have (1) $T_0 = \mathbb{1}$, (2) $\beta(a, ka) = 0$ for all $a \in E_n$ and $k \in \mathbb{Z}_d$ and (3) $\beta(a, b) = 0$ for all $a \in E_m$ and $b \in E_{n-m}^{(n)}$. It suffices to prove the theorem for the case where (g, J, r) is given by $(\mathbb{1}, J_0, r_0)$. Here $\mathbb{1}$ is the identity operator, $J_0 = \langle x_{m+1}, x_{m+2}, \cdots, x_n \rangle$, and r_0 is the value assignment defined by $r_0(x_i) = 0$ for all $1 \leq i \leq n-m$. This is because we can write

$$\Phi_{g,J}^r(X) = g\Phi_{\mathbb{1},J}^r(X)g^\dagger \tag{80}$$

$$= g(\mathbb{1} \otimes W)\Phi_{\mathbb{1},J_0}^{r_0}(X)(\mathbb{1} \otimes W^\dagger)g^\dagger \tag{81}$$

$$= V\Phi_{\mathbb{1},J_0}^{r_0}(X)V^\dagger \tag{82}$$

where W is the Clifford unitary such that $W\Pi^{r_0}_{J_0}W^\dagger = \Pi^r_J$ and $V = g(\mathbb{1}\otimes W)$. For the rest we will take $(g, J, r) = (\mathbb{1}, J_0, r_0)$. Let us write $\Phi = \Phi^{r_0}_{\mathbb{1},J_0}$ for simplicity of notation.

The map Φ sends X to the tensor product $X \otimes \Pi^{r_0}_{J_0}$, and therefore it is linear because of the properties of the tensor product operation. For $X \in \mathrm{Herm}(\mathbb{C}^{d^m})$ we can write

$$X = \frac{1}{d^m}\sum_{a\in E_m} \alpha_a T^{(m)}_a \tag{83}$$

which gives

$$\Phi(X) = X\otimes\Pi^r_J = \frac{1}{d^n}\sum_{a+b\in E_m+J} \alpha_a T^{(n)}_{a+b}. \tag{84}$$

In particular, $\Phi(X) = 0$ implies that $X = 0$, hence Φ is injective. Next we show that Φ maps Λ_m into Λ_n. First note that $\mathrm{Tr}(\Phi(X)) = \alpha_0 = 1$ since $\mathrm{Tr}(X) = 1$. For a maximal isotropic subgroup $I \subset E_n$ we compute

$$\mathrm{Tr}(\Pi^s_I\Phi(X)) = \frac{1}{d^{2n}}\sum_{c\in I}\sum_{a+b\in E_m+J} \alpha_a\omega^{s(c)}\underbrace{\mathrm{Tr}\Big(T^\dagger_c T_{a+b}\Big)}_{d^n\delta_{c,a+b}} \tag{85}$$

$$= \frac{1}{d^n}\sum_{a+b\in(E_m+J)\cap I} \alpha_a\omega^{s(a+b)} \tag{86}$$

$$= \frac{|K\cap J|}{d^n}\sum_{a\in K\cap E_m} \alpha_a\omega^{s(a)}\left(\frac{1}{|K\cap J|}\sum_{b\in K\cap J}\omega^{s(b)}\right) \tag{87}$$

$$= \delta_{r|_{K\cap J},s|_{K\cap J}}\frac{|K\cap J|}{d^n}\sum_{a\in K\cap E_m}\alpha_a\omega^{s(a)} \tag{88}$$

$$= \delta_{r|_{K\cap J},s|_{K\cap J}}\frac{|K|}{d^n}\mathrm{Tr}\Big(\Pi_{K\cap E_m,s|_{K\cap E_m}}X\Big) \geq 0 \tag{89}$$

where $K = (E_m + J)\cap I$. Therefore $\Phi(X) \in \Lambda_n$. This image is a convex polytope since the image of a convex polytope under a linear map is also a convex polytope. We want to show that the image of Φ is given by

$$\{\,\Pi^r_J Y\Pi^r_J/\,\mathrm{Tr}(Y\Pi^r_J)\,|\,Y\in\Lambda_n \;\text{ and }\; \mathrm{Tr}(Y\Pi^r_J)\neq 0\}. \tag{90}$$

First observe that this set is contained in Λ_n by Theorem 1. Writing Y in the Pauli basis $Y = (\sum_{c\in E_n}\beta_c T^{(n)}_c)/d^n$ we obtain

$$\Pi^r_J Y\Pi^r_J = \frac{1}{d^n}\sum_{c\in E_n}\beta_c\Pi^r_J T^{(n)}_c\Pi^r_J \tag{91}$$

$$= \frac{1}{d^n}\sum_{a+b\in E_m+J}\beta_{a+b}\Pi^r_J T^{(n)}_{a+b} \tag{92}$$

$$= \frac{1}{d^n}\sum_{a+b\in E_m+J}\beta_{a+b}\Pi^r_J(T^{(m)}_a\otimes T^{(n-m)}_b) \tag{93}$$

$$= \frac{1}{d^n}\sum_{a+b\in E_m+J}\beta_{a+b}T^{(m)}_a\otimes\Pi^r_J \tag{94}$$

$$= \frac{1}{d^m}\sum_{a\in E_m}\left(\frac{1}{d^{n-m}}\sum_{b\in J}\beta_{a+b}\right)T^{(m)}_a\otimes\Pi^r_J. \tag{95}$$

Defining X by setting $\alpha_a = \mathrm{Tr}(Y\Pi^r_J)(\sum_{b\in J}\beta_{a+b})/d^{n-m}$ in eq. (83) this computation shows that $\Phi(X) = (\Pi^r_J Y \Pi^r_J)/\mathrm{Tr}(Y\Pi^r_J)$. Moreover, X belongs to Λ_m since for any isotropic subgroup $I' \subset E_m$ and a value assignment s' defined on it we have that $\mathrm{Tr}\Big(X\Pi^{s'}_{I'}\Big) = \mathrm{Tr}\Big((X\Pi^{s'}_{I'})\otimes\Pi^r_J\Big) = \mathrm{Tr}\Big(\Phi(X)(\Pi^{s'}_{I'}\otimes\Pi^r_J)\Big) \geq 0$ since $\Pi^{s'}_{I'}\otimes\Pi^r_J$ is a stabilizer projector and $\Phi(X) \in \Lambda_n$. This shows that the set in eq. (90) is contained in the image of Φ. For the converse, we define Y by setting $\beta_{a+b} = \alpha_a$ for all $b \in J$. Then by the computation above we find that $\Pi^r_J Y \Pi^r_J = \Phi(X)$ with $\mathrm{Tr}(Y\Pi^r_J) = \mathrm{Tr}(\Phi(X)) = 1$. □

In contrast to the qubit case [47], the map Φ does not necessarily map a vertex of Λ_m to a vertex of Λ_n. For example, this is the case for odd qudit dimension d vertices of the form $A^\gamma_{E_m}$, as in (62).

To see this, consider $\Phi^r_{g,J}$ with $g = \mathbb{1}$. Analogously to Lemma 5 in Appendix A, the operator

$$\Phi^r_{\mathbb{1},J}(A^\gamma_{E_m}) = A^\gamma_{E_m}\otimes\Pi^r_J \tag{96}$$

can be written as a proper convex combination of vertices of Λ_n:

$$A^\gamma_{E_m}\otimes\Pi^r_J = A^\gamma_{E_m}\otimes\left(\frac{1}{\left|\Gamma^{E_{n-m}}_{J,r}\right|}\sum_{\gamma'\in\Gamma^{E_{n-m}}_{J,r}} A^{\gamma'}_{E_{n-m}}\right) \tag{97}$$

$$= \frac{1}{\left|\Gamma^{E_{n-m}}_{J,r}\right|}\sum_{\gamma'\in\Gamma^{E_{n-m}}_{J,r}} A^\gamma_{E_m}\otimes A^{\gamma'}_{E_{n-m}} \tag{98}$$

$$= \frac{1}{\left|\Gamma^{E_{n-m}}_{J,r}\right|}\sum_{\gamma'\in\Gamma^{E_{n-m}}_{J,r}} A^{\gamma\star\gamma'}_{E_n}, \tag{99}$$

where $\Gamma^{E_{n-m}}_{J,r}$ is the set of noncontextual value assignments r on E_{n-m} such that $r_{|J} = r$ and $\gamma\star\gamma_0$ is the unique value assignment on E_n satisfying $(\gamma\star\gamma_0)_{|E_m} = \gamma$ and $(\gamma\star\gamma_0)_{|E_{n-m}} = \gamma_0$.

7 Discussion

In this paper, we have presented a hidden variable model which represents all components of quantum computation with magic states using only classical (nonnegative) probabilities. In this model, magic states are represented by a probability distribution over a finite set according to eq. (12). Clifford gates are represented by a deterministic update rule—a map from the set of hidden variables to itself. Pauli measurements are represented by a probabilistic update rule—a map from hidden variables to probability distributions over the set of hidden variables according to eq. (13). This model is similar in form to many previously defined quasiprobability representations of quantum computation with magic states [23–25, 29, 40, 43, 45], but with the distinguishing feature that in our model all states can be represented by a probability distribution. No "negative probabilities" are required. This is the generalization of the hidden variable model of Ref. [46] to qudits of arbitrary local Hilbert space dimension.

Since everything is represented probabilistically, the model leads to a classical simulation algorithm for quantum computation with magic states based on sampling from the defining probability distributions. This is Algorithm 1 presented in Section 4. This algorithm is similar in structure to simulation algorithms based on sampling from quasiprobability distributions like the discrete Wigner function [29, 30, 40, 43, 45], except that those algorithms are limited in their scope. They can only simulate quantum circuits for which

the input state of the circuit is represented by a probability distribution. Since in our model any state can be represented by a probability distribution, we have no such limitation. Note that although Algorithm 1 can simulate any quantum computation, we make no claims that the simulation is efficient in general. In fact, if a quantum computational speedup over classical computation is possible at all, as many believe, then efficient classical simulation must fail.

There are, however, some important cases where the simulation is efficient. For example, in the qubit case, it is known that the (efficient) classical simulation algorithm of Ref. [45] is a special case of this more general model/simulation algorithm. This is a result of the fact that the phase point operators of eq. (61) are vertices of Λ [46, 55], these are the cnc-type vertices. In Section 5 we show that for higher-dimensional qudits as well there are vertices of Λ of cnc-type. We conjecture that the update of these vertices under Clifford gates and Pauli measurements will be efficiently computable classically. This would result in the simulation algorithm being efficient whenever the following two conditions are met: (i) the support of the probability distribution representing the input state of the circuit is restricted to vertices of this type and (ii) samples from the distribution representing the input state can be obtained efficiently.

Of course not all vertices of Λ are cnc-type. For example, Ref. [58] characterizes all vertices of the two-qubit Λ polytope. Under the action of the Clifford group there are eight orbits of vertices, only two of which are cnc-type vertices. Characterizing the remaining non-cnc-type vertices of Λ for arbitrary n and d is an open problem which could expand the scope of efficiency of Algorithm 1. This has already been partially achieved in the qubits case: Ref. [47] provides an efficient description, along with update rules under Pauli measurements, for one of the non-cnc-type orbits of the two qubit Λ polytope. As a result of the Φ-map, which maps vertices of the m-qubit polytope to vertices of the n-qubit polytope with $m < n$, this also provides a description of a non-cnc-type orbit of vertices for the Λ polytope on any number of qubits. Ref. [47] also provides a reduction of the simulation algorithm that shows that the Φ-map does not significantly increase the cost of classical simulation. Therefore, the scope of efficiency of Algorithm 1 in the qubit case goes beyond states supported on cnc-type vertices. In section 6 we show that a version of the Φ-map also holds for qudits of arbitrary dimension. We conjecture that a version of the corresponding reduction of the simulation algorithm also holds in this more general setting. This would align the model and simulation algorithm with the resource theory perspective in which preparation of stabilizer states is considered a "free" operation, along with Clifford gates and Pauli measurements [59].

We have seen evidence that the Λ polytopes could prove to be of independent interest. A subset of vertices, namely the cnc-type vertices, have already found a handful of applications [45, 55, 60]. The remaining vertices are less well studied but have proven useful in certain areas as well [47, 58]. Therefore, we conclude by proposing the family of Λ polytopes for arbitrary n and d as a subject of further study.

The main open questions regarding the Λ polytopes are "where does the efficiency of classically simulating quantum computation end?" (assuming it does), and now that we have found it's not always negativity in a quasiprobability representation, which physical property is responsible for the breakdown of efficient classical simulation?". In this paper we have shown that those questions do not only arise in the multi-qubit case, but rather in quantum computation on qudits of any dimension.

Acknowledgements

This research was undertaken thanks, in part, to funding from the Max Planck-UBC-UTokyo Center for Quantum Materials and the Canada First Research Excellence Fund, Quantum Materials and Future Technologies Program. M.Z. and R.R. are funded by the National Science and Engineering Research Council of Canada. C.O. is supported by the Air Force Office of Scientific Research under award number FA9550-21-1-0002. A.H. is partially supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under the Priority Program Compressed Sensing in Information Processing (CoSIP, project number SPP1798) and under Germany's Excellence Strategy – Cluster of Excellence Matter and Light for Quantum Computing (ML4Q) EXC 2004/1 – 390534769. This work was also supported by the Horizon Europe project FoQaCiA, GA no. 101070558.

References

[1] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, et al. "Quantum supremacy using a programmable superconducting processor". In: *Nature* 574.7779 (2019), pp. 505–510. DOI: 10.1038/s41586-019-1666-5. arXiv: 1910.11333.

[2] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, et al. "Hartree-Fock on a superconducting qubit quantum computer". In: *Science* 369.6507 (2020), pp. 1084–1089. DOI: 10.1126/science.abb9811. arXiv: 2004.04174.

[3] Frank Arute, Kunal Arya, Ryan Babbush, Dave Bacon, et al. "Observation of separated dynamics of charge and spin in the Fermi-Hubbard model". 2020. arXiv: 2010.07965.

[4] Matthew P. Harrigan, Kevin J. Sung, Matthew Neeley, Kevin J. Satzinger, et al. "Quantum approximate optimization of non-planar graph problems on a planar superconducting processor". In: *Nature Physics* 17.3 (2021), pp. 332–336. DOI: 10.1038/s41567-020-01105-y. arXiv: 2004.04197.

[5] Petar Jurcevic, Ali Javadi-Abhari, Lev S Bishop, Isaac Lauer, et al. "Demonstration of quantum volume 64 on a superconducting quantum computing system". In: *Quantum Science and Technology* 6.2 (2021), p. 025020. DOI: 10.1088/2058-9565/abe519. arXiv: 2008.08571.

[6] Andrew D. King, Juan Carrasquilla, Jack Raymond, Isil Ozfidan, et al. "Observation of topological phenomena in a programmable lattice of 1,800 qubits". In: *Nature* 560.7719 (2018), pp. 456–460. DOI: 10.1038/s41586-018-0410-x. arXiv: 1803.02047.

[7] K. Wright, K. M. Beck, S. Debnath, J. M. Amini, et al. "Benchmarking an 11-qubit quantum computer". In: *Nature Communications* 10.1 (2019), p. 5464. DOI: 10.1038/s41467-019-13534-2. arXiv: 1903.08181.

[8] Yunseong Nam, Jwo-Sy Chen, Neal C. Pisenti, Kenneth Wright, et al. "Ground-state energy estimation of the water molecule on a trapped-ion quantum computer". In: *npj Quantum Information* 6.1 (2020), p. 33. DOI: 10.1038/s41534-020-0259-3. arXiv: 1902.10171.

[9] Sonika Johri, Shantanu Debnath, Avinash Mocherla, Alexandros Singh, et al. "Nearest centroid classification on a trapped ion quantum computer". In: *npj Quantum Information* 7.1 (2021), p. 122. DOI: 10.1038/s41534-021-00456-5. arXiv: 2012.04145.

[10] Manuel S. Rudolph, Ntwali Bashige Toussaint, Amara Katabarwa, Sonika Johri, et al. "Generation of High-Resolution Handwritten Digits with an Ion-Trap Quantum Computer". In: *Physical Review X* 12.3 (2022), p. 031010. DOI: 10.1103/PhysRevX.12.031010. arXiv: 2012.03924.

[11] Daniel Gottesman, Alexei Kitaev, and John Preskill. "Encoding a qubit in an oscillator". In: *Phys. Rev. A* 64.1 (2001), p. 012310. DOI: 10.1103/PhysRevA.64.012310. arXiv: quant-ph/0008040.

[12] Sergey Bravyi and Alexei Kitaev. "Universal quantum computation with ideal Clifford gates and noisy ancillas". In: *Physical Review A* 71.2 (2005), p. 022316. DOI: 10.1103/PhysRevA.71.022316. arXiv: quant-ph/0403025.

[13] Earl T. Campbell, Barbara M. Terhal, and Christophe Vuillot. "Roads towards fault-tolerant universal quantum computation". In: *Nature* 549.7671 (2017), pp. 172–179. DOI: 10.1038/nature23460. arXiv: 1612.07330.

[14] Daniel Gottesman. "The Heisenberg representation of quantum computers". In: *Proceedings of the XXII International Colloquium on Group Theoretical Methods in Physics*. Ed. by S. P. Corney, R. Delbourgo, and P. D. Jarvis. 1999, pp. 32–43. arXiv: quant-ph/9807006.

[15] Scott Aaronson and Daniel Gottesman. "Improved simulation of stabilizer circuits". In: *Physical Review A* 70.5 (2004), p. 052328. DOI: 10.1103/PhysRevA.70.052328. arXiv: quant-ph/0406196.

[16] E. Wigner. "On the Quantum Correction For Thermodynamic Equilibrium". In: *Physical Review* 40.5 (1932), pp. 749–759. DOI: 10.1103/PhysRev.40.749.

[17] R.L. Hudson. "When is the Wigner quasi-probability density non-negative?" In: *Reports on Mathematical Physics* 6.2 (1974), pp. 249–252. DOI: 10.1016/0034-4877(74)90007-X.

[18] Anatole Kenfack and Karol yczkowski. "Negativity of the Wigner function as an indicator of non-classicality". In: *Journal of Optics B: Quantum and Semiclassical Optics* 6.10 (2004), pp. 396–404. DOI: 10.1088/1464-4266/6/10/003. arXiv: quant-ph/0406015.

[19] William K Wootters. "A Wigner-function formulation of finite-state quantum mechanics". In: *Annals of Physics* 176.1 (1987), pp. 1–21. DOI: 10.1016/0003-4916(87)90176-X.

[20] Kathleen S. Gibbons, Matthew J. Hoffman, and William K. Wootters. "Discrete phase space based on finite fields". In: *Physical Review A* 70.6 (2004), p. 062101. DOI: 10.1103/PhysRevA.70.062101. arXiv: quant-ph/0401155.

[21] Ernesto F. Galvão. "Discrete Wigner functions and quantum computational speedup". In: *Physical Review A* 71.4 (2005), p. 042302. DOI: 10.1103/PhysRevA.71.042302. arXiv: quant-ph/0405070.

[22] Cecilia Cormick, Ernesto F. Galvão, Daniel Gottesman, Juan Pablo Paz, and Arthur O. Pittenger. "Classicality in discrete Wigner functions". In: *Physical Review A* 73.1 (2006), p. 012301. DOI: 10.1103/PhysRevA.73.012301. arXiv: quant-ph/0506222.

[23] David Gross. "Hudson's theorem for finite-dimensional quantum systems". In: *Journal of Mathematical Physics* 47.12 (2006), p. 122107. DOI: 10.1063/1.2393152. arXiv: quant-ph/0602001.

[24] David Gross. "Non-negative Wigner functions in prime dimensions". In: *Applied Physics B* 86.3 (2006), pp. 367–370. DOI: 10.1007/s00340-006-2510-9. arXiv: quant-ph/0702004.

[25] David Gross. "Computational power of quantum many-body states and some results on discrete phase spaces". https://www.thp.uni-koeln.de/gross/files/diss.pdf. PhD Thesis. Institute for Mathematical Sciences, Imperial College London, 2008.

[26] R. L. Stratonovich. "On Distributions in Representation Space". In: *Zhurnal Éksperimental'noĭ i Teoreticheskoĭ Fiziki* 31.6 (1956). [English translation: *Soviet Physics JETP*, 4.6:891–898 (1957)], pp. 1012–1020. URL: http://jetp.ras.ru/cgi-bin/dn/e_004_06_0891.pdf.

[27] C. Brif and A. Mann. "A general theory of phase-space quasiprobability distributions". In: *Journal of Physics A: Mathematical and General* 31.1 (1998), pp. L9–L17. DOI: 10.1088/0305-4470/31/1/002. arXiv: quant-ph/9707010.

[28] Robert Raussendorf, Cihan Okay, Michael Zurel, and Polina Feldmann. "The role of cohomology in quantum computation with magic states". In: *Quantum* 7 (2023), p. 979. DOI: 10.22331/q-2023-04-13-979. arXiv: 2110.11631.

[29] Victor Veitch, Christopher Ferrie, David Gross, and Joseph Emerson. "Negative quasi-probability as a resource for quantum computation". In: *New Journal of Physics* 14.11 (2012), p. 113011. DOI: 10.1088/1367-2630/14/11/113011. arXiv: 1201.1256.

[30] Michael Zurel. "Hidden variable models and classical simulation algorithms for quantum computation with magic states on qubits". Master's Thesis. University of British Columbia, 2020. DOI: 10.14288/1.0394790.

[31] Hakop Pashayan, Joel J. Wallman, and Stephen D. Bartlett. "Estimating Outcome Probabilities of Quantum Circuits Using Quasiprobabilities". In: *Physical Review Letters* 115.7 (2015). DOI: 10.1103/physrevlett.115.070501. arXiv: 1503.07525.

[32] Robert W. Spekkens. "Negativity and Contextuality are Equivalent Notions of Nonclassicality". In: *Physical Review Letters* 101.2 (2008), p. 020401. DOI: 10.1103/PhysRevLett.101.020401. arXiv: 0710.5549.

[33] Nicolas Delfosse, Cihan Okay, Juan Bermejo-Vega, Dan E Browne, and Robert Raussendorf. "Equivalence between contextuality and negativity of the Wigner function for qudits". In: *New Journal of Physics* 19.12 (2017), p. 123024. DOI: 10.1088/1367-2630/aa8fe3. arXiv: 1610.07093.

[34] Robert Raussendorf. "Contextuality in measurement-based quantum computation". In: *Physical Review A* 88 (2013). DOI: 10.1103/physreva.88.022322. arXiv: 0907.5449.

[35] Mark Howard, Joel Wallman, Victor Veitch, and Joseph Emerson. "Contextuality supplies the "magic" for quantum computation". In: *Nature* 510.7505 (2014), pp. 351–355. DOI: 10.1038/nature13460. arXiv: 1401.4174.

[36] Farid Shahandeh. "Quantum computational advantage implies contextuality". 2021. arXiv: 2112.00024.

[37] Huangjun Zhu. "Permutation Symmetry Determines the Discrete Wigner Function". In: *Physical Review Letters* 116.4 (2016), p. 040501. DOI: 10.1103/PhysRevLett.116.040501. arXiv: 1504.03773.

[38] Angela Karanjai, Joel J. Wallman, and Stephen D. Bartlett. "Contextuality bounds the efficiency of classical simulation of quantum processes". 2018. arXiv: 1802.07744.

[39] David Schmid, Haoxing Du, John H. Selby, and Matthew F. Pusey. "Uniqueness of Noncontextual Models for Stabilizer Subtheories". In: *Physical Review Letters* 129.12 (2022), p. 120403. DOI: 10.1103/PhysRevLett.129.120403. arXiv: 2101.06263.

[40] Nicolas Delfosse, Philippe Allard Guerin, Jacob Bian, and Robert Raussendorf. "Wigner Function Negativity and Contextuality in Quantum Computation on Rebits". In: *Physical Review X* 5.2 (2015), p. 021003. DOI: 10.1103/PhysRevX.5.021003. arXiv: 1409.5170.

[41] Juan Bermejo-Vega, Nicolas Delfosse, Dan E. Browne, Cihan Okay, and Robert Raussendorf. "Contextuality as a Resource for Models of Quantum Computation with Qubits". In: *Physical Review Letters* 119.12 (2017). DOI: 10.1103/physrevlett.119.120505. arXiv: 1610.08529.

[42] Robert Raussendorf, Dan E. Browne, Nicolas Delfosse, Cihan Okay, and Juan Bermejo-Vega. "Contextuality and Wigner-function negativity in qubit quantum computation". In: *Physical Review A* 95.5 (2017). DOI: 10.1103/physreva.95.052334. arXiv: 1511.08506.

[43] Mark Howard and Earl Campbell. "Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing". In: *Physical Review Letters* 118.9 (2017). DOI: 10.1103/physrevlett.118.090501. arXiv: 1609.07488.

[44] Lucas Kocia and Peter Love. "Discrete Wigner formalism for qubits and noncontextuality of Clifford gates on qubit stabilizer states". In: *Physical Review A* 96.6 (2017), p. 062134. DOI: 10.1103/PhysRevA.96.062134. arXiv: 1705.08869.

[45] Robert Raussendorf, Juani Bermejo-Vega, Emily Tyhurst, Cihan Okay, and Michael Zurel. "Phase-space-simulation method for quantum computation with magic states on qubits". In: *Physical Review A* 101.1 (2020), p. 012350. DOI: 10.1103/PhysRevA.101.012350. arXiv: 1905.05374.

[46] Michael Zurel, Cihan Okay, and Robert Raussendorf. "Hidden Variable Model for Universal Quantum Computation with Magic States on Qubits". In: *Physical Review Letters* 125.26 (2020), p. 260404. DOI: 10.1103/PhysRevLett.125.260404. arXiv: 2004.01992.

[47] Cihan Okay, Michael Zurel, and Robert Raussendorf. "On the extremal points of the Λ-polytopes and classical simulation of quantum computation with magic states". In: *Quantum Information & Computation* 21.13&14 (2021). DOI: 10.26421/QIC21.13-14-2. arXiv: 2104.05822.

[48] Daniel Gottesman. "Fault-Tolerant Quantum Computation with Higher-Dimensional Systems". In: *Chaos, Solitons & Fractals* 10.10 (1999), pp. 1749–1758. DOI: 10.1016/s0960-0779(98)00218-5. arXiv: quant-ph/9802007.

[49] Sergey Bravyi, Graeme Smith, and John A Smolin. "Trading classical and quantum computational resources". In: *Physical Review Letters* 6.2 (2016), p. 021043. DOI: 10.1103/PhysRevX.6.021043. arXiv: 1506.01396.

[50] Cihan Okay, Sam Roberts, Stephen D. Bartlett, and Robert Raussendorf. "Topological proofs of contextuality in quantum mechanics". In: *Quantum Information and Computation* 17 (2017), pp. 1135–1166. DOI: 10.26421/QIC17.13-14-5. arXiv: 1701.01888.

[51] Gunter M. Ziegler. *Lectures on polytopes*. Springer-Verlag, 1995.

[52] John B DeBrota and Blake C Stacey. "Discrete Wigner functions from informationally complete quantum measurements". In: *Phys. Rev. A* 102.3 (2020), p. 032221. DOI: 10.1103/PhysRevA.102.032221. arXiv: 1912.07554.

[53] E.G. Beltrametti and S. Bugajski. "The Bell phenomenon in classical frameworks". In: *Journal of Physics A: Mathematical and General* 29.2 (1996), p. 247. DOI: 10.1088/0305-4470/29/2/005.

[54] Vašek Chvátal. *Linear Programming*. W.H. Freeman and Company, 1983.

[55] Arne Heimendahl. "The Stabilizer Polytope and Contextuality for Qubit Systems". http://www.mi.uni-koeln.de/opt/wp-content/uploads/2020/07/MT_Arne_Heimendahl.pdf. Master's Thesis. University of Cologne, 2019.

[56] Vlad Gheorghiu. "Standard form of qudit stabilizer groups". In: *Physics Letters A* 378.5-6 (2014), pp. 505–509. DOI: 10.1016/j.physleta.2013.12.009. arXiv: 1101.1519.

[57] Chuangxun Cheng. "A character theory for projective representations of finite groups". In: *Linear Algebra and its Applications* 469 (2015), pp. 230–242. DOI: 10.1016/j.laa.2014.11.027.

[58] Ben W. Reichardt. "Quantum universality by state distillation". In: *Quantum Information and Computation* 9.11&12 (2009), pp. 1030–1052. DOI: 10.26421/QIC9.11-12-7. arXiv: quant-ph/0608085.

[59] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. "The resource theory of stabilizer quantum computation". In: *New Journal of Physics* 16.1 (2014), p. 013009. DOI: 10.1088/1367-2630/16/1/013009. arXiv: 1307.7171.

[60] William M. Kirby and Peter J. Love. "Contextuality Test of the Nonclassicality of Variational Quantum Eigensolvers". In: *Physical Review Letters* 123.20 (2019). DOI: 10.1103/physrevlett.123.200501. arXiv: 1904.02260.

[61] Daniel Gottesman. "Stabilizer Codes and Quantum Error Correction". arXiv:quant-ph/9705052. PhD Thesis. California Institute of Technology, 1997. arXiv: quant-ph/9705052.

[62] Stanislaw Szarek and Guillaume Aubrun. *Alice and Bob Meet Banach: The Interface of Asymptotic Geometric Analysis and Quantum Information Theory*. American Mathematical Society, 2017.

[63] Richard A. Brualdi. *Introductory Combinatorics*. Pearson Education International, 2012.

A The qudit stabilizer formalism

The stabilizer formalism [61] describes a large family of quantum error correcting codes, as well as a broader framework for describing quantum error correction and fault-tolerant quantum computation. In this appendix we review some features of the stabilizer formalism for systems of qudits [48] which are useful for some of the proofs in the main text.

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

eigenvalues $\{\omega^{r(a)} \mid a \in I\}$ is

$$\Pi_I^r = \frac{1}{|I|} \sum_{b \in I} \omega^{-r(b)} T_b. \tag{100}$$

The dimension of this eigenspace is $d^n/|I|$ [56]. In particular, we have the following lemma.

Lemma 4 *The order of a maximal isotropic subgroup of E is d^n.*

For the proof of Lemma 4 see Theorem 1 in Ref. [56]. If I is a maximal isotropic subgroup of E, $|I| = d^n$, and there is a unique quantum state fixed by a stabilizer group S_I^r. Such states are called stabilizer states.

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

$$\Pi_I^r = \prod_k \Pi_{I_k}^{r_k}. \tag{101}$$

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

Lemma 5 *If I is a non-maximal isotropic subgroup of E and Π_I^r is a stabilizer code projector for some noncontextual value assignment $r : I \to \mathbb{Z}_d$, then for any isotropic subgroup I' containing I*

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \Pi_{I'}^{r'} = \Pi_I^r, \tag{102}$$

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

Proof of Lemma 5. The proof is obtained by adapting the proof of Ref. [45, Lemma 1]. Let I, I' be isotropic subgroups of E such that $I \subsetneq I'$. Let $r : I \to \mathbb{Z}_d$ be a noncontextual value assignment for I and $\Gamma_{I,r}^{I'}$ be the set of all noncontextual value assignments on I' satisfying

$$r'|_I = r \quad \forall r' \in \Gamma_{I,r}^{I'}. \tag{103}$$

Then

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \Pi_{I'}^{r'} = \sum_{r' \in \Gamma_{I,r}^{I'}} \frac{1}{|I'|} \sum_{a \in I'} \omega^{-r'(a)} T_a \tag{104}$$

$$= \frac{1}{|I'|} \sum_{a \in I'} \left[\sum_{r' \in \Gamma_{I,r}^{I'}} \omega^{-r'(a)} \right] T_a. \tag{105}$$

We have two cases for the inner sum in the last expression. If $a \in I$ then $r'(a) = r(a)$ for all $r' \in \Gamma_{I,r}^{I'}$. Therefore,

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \omega^{-r'(a)} = \left| \Gamma_{I,r}^{I'} \right| \omega^{-r(a)}. \tag{106}$$

In the second case, $a \notin I$. $\Gamma_{I,r}^{I'}$ is the coset of a vector space, the proof of this is analogous to the proof of Ref. [45, Lemma 2] which applies only to qubits. Therefore, by character orthogonality,

$$\sum_{r' \in \Gamma_{I,r}^{I'}} \omega^{-r'(a)} = 0. \tag{107}$$

Thus,

$$\sum_{r' \in \Gamma^{I'}_{I,r}} \Pi^{r'}_{I'} = \frac{1}{|I'|} \sum_{a \in I'} \left|\Gamma^{I'}_{I,r}\right| \delta_{a \in I} \omega^{-r(a)} T_a \tag{108}$$

$$= \frac{\left|\Gamma^{I'}_{I,r}\right|}{|I'|} \sum_{a \in I} \omega^{-r(a)} T_a. \tag{109}$$

We have $|\Gamma^{I'}_{I,r}| = |I'|/|I|$. Therefore,

$$\sum_{r' \in \Gamma^{I'}_{I,r}} \Pi^{r'}_{I'} = \frac{1}{|I|} \sum_{a \in I} \omega^{-r(a)} T_a = \Pi^r_I \tag{110}$$

which proves the lemma. □

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

$$\mathrm{SP} = \mathrm{conv}\left\{ \frac{|I|}{d^n} \Pi^r_I \mid \Pi^r_I \text{ stabilizer code projector} \right\}. \tag{111}$$

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

Lemma 6 *For any Clifford group element $g \in \mathcal{C}\ell$, and any stabilizer state $|\sigma\rangle \in \mathcal{S}$,*

$$g |\sigma\rangle \langle\sigma| g^\dagger = |\sigma'\rangle \langle\sigma'| \tag{112}$$

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

Proof of Lemma 6. The action of the Clifford group on the Pauli operators is defined in eq. (7). With this equation, for a projector onto a stabilizer state $|\sigma\rangle$ corresponding to maximal isotropic subgroup $I \subset E$ and noncontextual value assignment $r : I \to \mathbb{Z}_d$ we have

$$g |\sigma\rangle \langle\sigma| g^\dagger = g \Pi^r_I g^\dagger = \frac{1}{|I|} \sum_{a \in I} \omega^{-r(a)} g T_a g^\dagger \tag{113}$$

$$= \frac{1}{|I|} \sum_{a \in I} \omega^{-r(a) + \Phi_g(a)} T_{S_g(a)} \tag{114}$$

$$= \frac{1}{|I|} \sum_{a \in g \cdot I} \omega^{-g \cdot r(a)} T_a = \Pi^{g \cdot r}_{g \cdot I} \tag{115}$$

where $g \cdot I = \{S_g(a) \mid a \in I\}$ and $g \cdot r$ is defined by the relation

$$g \cdot r(S_g(a)) = r(a) - \Phi_g(a) \quad \forall a \in I. \tag{116}$$

In order to show that $g |\sigma\rangle \langle\sigma| g^\dagger$ is a projector onto a stabilizer state it suffices to show that $g \cdot I$ is a maximal isotropic subgroup of E and that $g \cdot r : g \cdot I \to \mathbb{Z}_d$ is a noncontextual value assignment for $g \cdot I$.

$g \cdot I$ is isotropic since I is isotropic and S_g is a symplectic operation. Also, $|g \cdot I| = |I| = d^n$ so by Lemma 4, $g \cdot I$ is a maximal isotropic subgroup.

Since $T_0 \propto \mathbb{1}$

$$gT_0g^\dagger = T_0 \tag{117}$$

so $\Phi_g(0) = 0$ for any $g \in \mathcal{C}\ell$. Therefore,

$$\omega^{-g\cdot r(0)}T_0 = \omega^{-r(0)+\Phi_g(0)}T_0 = \omega^{-r(0)}T_0 = \mathbb{1}. \tag{118}$$

For any $a, b \in g \cdot I$, there exist $c, d \in I$ such that $S_g(c) = a$ and $S_g(d) = b$. Computing the product $gT_cT_dg^\dagger$ in two different ways we have

$$\begin{aligned} gT_cT_dg^\dagger =& gT_cg^\dagger gT_dg^\dagger && (119)\\ =& \omega^{\Phi_g(c)+\Phi_g(d)}T_{S_g(c)}T_{S_g(d)} && (120)\\ =& \omega^{\Phi_g(c)+\Phi_g(d)-\beta(a,b)}T_{a+b} && (121)\end{aligned}$$

and

$$\begin{aligned} gT_cT_dg^\dagger =& \omega^{-\beta(c,d)}gT_{c+d}g^\dagger && (122)\\ =& \omega^{-\beta(c,d)+\Phi_g(c+d)}T_{a+b}. && (123)\end{aligned}$$

Therefore, $-\Phi_g(c) - \Phi_g(d) + \Phi_g(c+d) = \beta(c,d) - \beta(a,b)$ and so

$$\begin{aligned} & g\cdot r(a) + g\cdot r(b) - g\cdot r(a+b) && (124)\\ &= r(c) + r(d) - r(c+d) - \Phi_g(c) - \Phi_g(d) + \Phi_g(c+d) && (125)\\ &= -\beta(a,b). && (126)\end{aligned}$$

Thus, $g \cdot r$ satisfies eq. (9) and $\Pi^{g\cdot r}_{g\cdot I}$ is a projector onto a stabilizer state. □

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

Lemma 7 *For any isotropic subgroups $I, J \subset E$ and any noncontextual value assignments $r : I \to \mathbb{Z}_d$ and $s : J \to \mathbb{Z}_d$,*

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

$$\mathrm{Tr}(\Pi^s_J\Pi^r_I) = \frac{|I\cap J|}{|I||J|}d^n > 0 \tag{127}$$

and

$$\frac{\Pi^r_I\Pi^s_J\Pi^r_I}{\mathrm{Tr}(\Pi^s_J\Pi^r_I)} = \frac{|J\cap I^\perp|}{|J|}\Pi^{r\star s}_{I+J\cap I^\perp} \tag{128}$$

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

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

$$\mathrm{Tr}(\Pi^s_J\Pi^r_I) = 0 \quad \textit{and} \quad \Pi^r_I\Pi^s_J\Pi^r_I = 0. \tag{129}$$

Proof of Lemma 7. Let $I, J \subset E$ be isotropic subgroups with noncontextual value assignments $r : I \to \mathbb{Z}_d$ and $s : J \to \mathbb{Z}_d$.

Case 1: $r|_{I\cap J} = s|_{I\cap J}$. Let $r\star s$ denote the unique noncontextual value assignment on the set $I + J\cap I^\perp$ such that $r\star s|_I = r$ and $r\star s|_{J\cap I^\perp} = s|_{J\cap I^\perp}$. We calculate

$$\Pi_I^r\Pi_J^s\Pi_I^r = \frac{1}{|I|^2|J|}\sum_{a,c\in I}\sum_{b\in J}\omega^{-r(a)-r(c)-s(b)}T_aT_bT_c \tag{130}$$

$$= \frac{1}{|I|^2|J|}\sum_{a,c\in I}\sum_{b\in J}\omega^{-r(a)-r(c)-s(b)-\beta(a,c)+[b,c]}T_{a+c}T_b \tag{131}$$

$$= \frac{1}{|I|^2|J|}\sum_{a,c\in I}\sum_{b\in J}\omega^{-r(a+c)-s(b)+[b,c]}T_{a+c}T_b \tag{132}$$

$$= \frac{1}{|I||J|}\Pi_I^r\sum_{b\in J}\omega^{-s(b)}\left[\sum_{c\in I}\omega^{[b,c]}\right]T_b \tag{133}$$

$$= \frac{1}{|J|}\Pi_I^r\sum_{b\in J\cap I^\perp}\omega^{-s(b)}T_b \tag{134}$$

$$= \frac{1}{|I||J|}\sum_{a\in I}\sum_{b\in J\cap I^\perp}\omega^{-r(a)-s(b)-\beta(a,b)}T_{a+b} \tag{135}$$

$$= \frac{1}{|I||J|}\sum_{a\in I}\sum_{b\in J\cap I^\perp}\omega^{-r\star s(a+b)}T_{a+b}. \tag{136}$$

Each Pauli operator in the above sum appears with the same multiplicty. For an element $a\in I+J\cap I^\perp$, let $\mu(a)$ denote the number of distinct pairs $(b,c)\in I\times J\cap I^\perp$ such that $b+c=a$. We have $\mu(a)=\mu(0)$ for any $a\in I+J\cap I^\perp$. To see this, suppose $a\in I+J\cap I^\perp$ and let $(c_1,d_1),(c_2,d_2),\ldots,(c_{\mu(a)},d_{\mu(a)})$ be the distinct pairs in $I\times J\cap I^\perp$ such that $c_j+d_j=a$. Then the pairs $(c_j-c_1,d_j-d_1)\in I\times J\cap I^\perp$ for $j=2,\ldots,\mu(a)$ together with the pair $(0,0)\in I\times J\cap I^\perp$ show that $\mu(0)\geq\mu(a)$.

Now let $(c_1,d_1),(c_2,d_2),\ldots,(c_{\mu(0)},d_{\mu(0)})$ denote the $\mu(0)$ distinct pairs in $I\times J\cap I^\perp$ such that $c_j+d_j=0$, and $(c,d)\in I\times J\cap I^\perp$ be such that $c+d=a$. Then the pairs $(c_j+c,d_j+d),\ j=1,2,\ldots,\mu(0)$ show that $\mu(a)\geq\mu(0)$. Therefore, $\mu(a)=\mu(0)$ for any $a\in I+J\cap I^\perp$. Here we can see $\mu(0)=|I\cap J|$

Then we have

$$\Pi_I^r\Pi_J^s\Pi_I^r = \frac{1}{|I||J|}\sum_{a\in I}\sum_{b\in J\cap I^\perp}\omega^{-r\star s(a+b)}T_{a+b} \tag{137}$$

$$= \frac{\mu(0)}{|I||J|}\sum_{a\in I+J\cap I^\perp}\omega^{-r\star s(a)}T_a \tag{138}$$

$$= \frac{|I\cap J||I+J\cap I^\perp|}{|I||J|}\Pi^{r\star s}_{I+J\cap I^\perp}. \tag{139}$$

Since $|I+J\cap I^\perp| = |I|\cdot|J\cap I^\perp|/|I\cap J|$,

$$\Pi_I^r\Pi_J^s\Pi_I^r = \frac{|J\cap I^\perp|}{|J|}\Pi^{r\star s}_{I+J\cap I^\perp}. \tag{140}$$

Taking a trace of this equation we get

$$\begin{aligned}\operatorname{Tr}(\Pi_I^r \Pi_J^s) &= \operatorname{Tr}(\Pi_I^r \Pi_J^s \Pi_I^r) && (141)\\ &= \frac{|J \cap I^\perp|}{|J|} \operatorname{Tr}\left(\Pi_{I+J\cap I^\perp}^{r\star s}\right) && (142)\\ &= \frac{|J \cap I^\perp|}{|J|} \cdot \frac{d^n}{|I + J \cap I^\perp|} && (143)\\ &= \frac{|I \cap J|}{|I||J|} d^n > 0. && (144)\end{aligned}$$

Case 2: $r|_{I\cap J} \neq s|_{I\cap J}$. Denote by S_J^s the stabilizer group $\{\omega^{-s(b)} T_b \mid b \in J\}$ and denote by V_J^s the subspace of the Hilbert space $\mathbb{C}^{d^n}$ stabilized by S_J^s. By assumption, there exists a $c \in I \cap J$ such that $r(c) \neq s(c)$. Therefore, for any vector $|\psi\rangle \in V_J^s$,

$$\begin{aligned}\Pi_I^r |\psi\rangle &= \Pi_I^r \omega^{-s(c)} T_c |\psi\rangle && (145)\\ &= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)-s(c)} T_a T_c |\psi\rangle && (146)\\ &= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)-s(c)-\beta(a,c)} T_{a+c} |\psi\rangle && (147)\\ &= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a+c)-r(c)-s(c)} T_{a+c} |\psi\rangle && (148)\\ &= \frac{1}{d^n} \sum_{a\in I} \omega^{-r(a)-r(c)-s(c)} T_a |\psi\rangle && (149)\\ &= \omega^{-r(c)-s(c)} \Pi_I^s |\psi\rangle . && (150)\end{aligned}$$

That is, $\Pi_I^s |\psi\rangle = \omega^{-r(c)-s(c)} \Pi_I^s |\psi\rangle$. But since $r(c) \neq s(c)$, $\omega^{-r(c)-s(c)} \neq 1$ so this relation can be true only if $\Pi_I^r |\psi\rangle = 0$.

We can write

$$\Pi_J^s = \sum_{i=1}^{\dim(V_J^s)} |\psi_i\rangle \langle\psi_i| \tag{151}$$

where $\{|\psi_i\rangle \mid 1 \leq i \leq \dim(V_J^s)\}$ is a basis for V_J^s. Therefore,

$$\Pi_I^r \Pi_J^s \Pi_I^r = \sum_{i=1}^{\dim(V_J^s)} \Pi_I^r |\psi_i\rangle \langle\psi_i| \Pi_I^r = 0. \tag{152}$$

Taking a trace of this equation we get

$$\operatorname{Tr}(\Pi_I^r \Pi_J^s \Pi_I^r) = \operatorname{Tr}(\Pi_I^r \Pi_J^s) = 0. \tag{153}$$

This proves the lemma. □

B Proof of Lemma 1

To prove the lemma, we will use the concept of polar duality for objects in the affine space $\mathrm{Herm}_1(\mathcal{H})$ (see Ref. [51] for a discussion of polar duality). For a Hilbert space $\mathcal{H}$ with inner product $\langle\cdot,\cdot\rangle$ define[1] the *polar dual* of a set $P \subset \mathcal{H}$ as

$$P^* = \left\{ x \in \mathbb{R}^N \;\middle|\; \langle x, y\rangle \geq -\frac{1}{d^n} \text{ for all } y \in P \right\}. \tag{154}$$

[1] The usual definition differs slightly by an irrelevant scaling factor.

If $P = \mathrm{conv}\{v_1, \ldots, v_k\}$ is a polytope, then this simplifies to

$$P^* = \left\{ x \in \mathbb{R}^N \;\middle|\; \langle x, v_i \rangle \geq -\frac{1}{d^n},\ i, \ldots, k \right\}. \tag{155}$$

If $P \subset Q \subset \mathcal{H}$, then obviously $Q^* \subset P^*$.

In our setting, we are interested in objects living in the affine space of matrices of trace one $\mathrm{Herm}_1(\mathcal{H})$. We can project $\mathrm{Herm}_1(\mathcal{H})$ into the linear subspace of Hermitian matrices of trace zero $\mathrm{Herm}_0(\mathcal{H})$ via the transformation

$$\pi : X \in \mathrm{Herm}_1(\mathcal{H}) \to \mathrm{Herm}_0(\mathcal{H}),\ X \mapsto X - \frac{1}{d^n}\mathbb{1}. \tag{156}$$

Now, observe that for $X, Y \in \mathrm{Herm}_1(\mathcal{H})$ we have $\mathrm{Tr}(XY) \geq 0$ if and only if

$$\mathrm{Tr}(\pi(X)\pi(Y)) = \mathrm{Tr}(XY) - \mathrm{Tr}\left(X \frac{1}{d^n}\mathbb{1}\right) \geq -\frac{1}{d^n}. \tag{157}$$

Hence, by associating $\mathrm{Herm}_0(\mathcal{H})$ and $\mathrm{Herm}_1(\mathcal{H})$, we define for a set $M \subset \mathrm{Herm}_1(\mathcal{H})$

$$M^* := \{X \in \mathrm{Herm}_1(\mathcal{H}) \mid \mathrm{Tr}(XY) \geq 0 \ \text{ for all } Y \in M\}. \tag{158}$$

If $P = \mathrm{conv}\{X_1, \ldots, X_m\} \subset \mathrm{Herm}_1(\mathcal{H})$ is a polytope, then define its *polar dual* as

$$P^* = \{Y \in \mathrm{Herm}_1(\mathcal{H}) \mid \mathrm{Tr}(X_i Y) \geq 0,\ i = 1, \ldots, m\}. \tag{159}$$

Thus, $\Lambda = \mathrm{SP}^*$ for SP being the *stabilizer polytope*: $\mathrm{SP} := \mathrm{conv}\left\{|\sigma\rangle\langle\sigma| \mid \sigma \in \mathcal{S}\right\}$.

To prove that the set Λ is bounded, it will suffice to show that SP contains a set M, whose dual M^* is bounded. Additionally, we will make us of the concept of *dilation* [62, Chap. 9]: for a set $M \subset \mathrm{Herm}_1(\mathcal{H})$ define its dilation centered at the maximally mixed state via

$$c \cdot M := \left\{ \frac{1}{d^n}\mathbb{1} + c\pi(X) \mid X \in M \right\} \tag{160}$$

where $\pi : \mathrm{Herm}_1(\mathcal{H}) \to \mathrm{Herm}_0(\mathcal{H})$ is the projection that maps $X \in \mathrm{Herm}_1(\mathcal{H})$ to $X - \frac{1}{d^n}\mathbb{1}$. The dilation has the following property:

Lemma 8 *The dilation of a set $M \subset Herm_1(\mathcal{H})$ satisfies*

$$(c \cdot M)^* = \frac{1}{c} \cdot M^*. \tag{161}$$

Proof of Lemma 8. Observe that every $A, Y \in \mathrm{Herm}_1(\mathcal{H})$ can be written as

$$Y = \frac{1}{d^n}\mathbb{1} + c\pi(X), \quad A = \frac{1}{d^n}\mathbb{1} + \frac{1}{c}\pi(B)$$

for suitable $B, X \in \mathrm{Herm}_1(\mathcal{H})$. Then

$$\mathrm{Tr}(AY) = \frac{1}{d^{2n}}\mathrm{Tr}(\mathbb{1}) + \mathrm{Tr}(\pi(B)\pi(X)) \tag{162}$$

$$= \frac{1}{d^n} + \mathrm{Tr}\left(B(X - \frac{1}{d^n}\mathbb{1})\right) + \mathrm{Tr}\left(-\frac{1}{d^n}\mathbb{1}(X - \frac{1}{d^n}\mathbb{1})\right) \tag{163}$$

$$= \frac{1}{d^n} - \frac{1}{d^n} + \mathrm{Tr}(BX) \tag{164}$$

$$= \mathrm{Tr}(BX). \tag{165}$$

Hence, if $Y \in c \cdot M$ with $X \in M$, then $A \in (c \cdot M)^*$ if and only if

$$0 \leq \mathrm{Tr}(AY) = \mathrm{Tr}(BX) \quad \Longleftrightarrow \quad B \in M^*, \tag{166}$$

which is equivalent to $A \in \frac{1}{c} \cdot M^*$ □

Having introduced all necessary concepts, we will proceed with the proof of Lemma 1.

Proof of Lemma 1. By definition, Λ is a polyhedron and therefore convex and closed. To prove that Λ is bounded, the previous discussion implies that it suffices to find a set $M \subset \mathrm{SP}$ such that M^* is bounded. The object we will choose here will be a dilation of the full-dimensional simplex

$$\triangle_{\mathrm{Herm}_1(\mathcal{H})} := \mathrm{conv}\Big\{ A_\gamma := \frac{1}{d^n} \sum_{u \in E} \omega^{\gamma(u)} T_u \;\Big|\; \tag{167}$$

$$\gamma : E \to \mathbb{Z}_d,\ \gamma(u+v) = \gamma(u) + \gamma(v) \Big\}. \tag{168}$$

The simplex $\triangle_{\mathrm{Herm}_1(\mathcal{H})}$ is the Wigner simplex for d being an odd prime [29, 59]. For every d, it is a full-dimensional polytope as the convex-hull of d^{2n} affinely independent vertices A_γ in the $(d^{2n}-1)$-dimensional affine space $\mathrm{Herm}_1(\mathcal{H})$. Due to $\mathrm{Tr}(A_\gamma A_{\gamma'}) = \delta_{\gamma=\gamma'}$ for additive functions $\gamma, \gamma' : E \to \mathbb{Z}_d$, one can easily verify that the simplex $\triangle_{\mathrm{Herm}_1(\mathcal{H})}$ has the following hyperplane description:

$$\triangle_{\mathrm{Herm}_1(\mathcal{H})} = \{X \in \mathrm{Herm}_1(\mathcal{H}) \mid \mathrm{Tr}(A_\gamma X) \geq 0\}, \tag{169}$$

which makes it a self-dual simplex, i.e. $\triangle_{\mathrm{Herm}_1(\mathcal{H})} = \triangle^*_{\mathrm{Herm}_1(\mathcal{H})}$.

Now, Lemma 8 implies that for $c > 0$ the simplex $(c \cdot \triangle_{\mathrm{Herm}_1(\mathcal{H})})^*$ is bounded, since

$$(c \cdot \triangle_{\mathrm{Herm}_1(\mathcal{H})})^* = \frac{1}{c}\, \triangle^*_{\mathrm{Herm}_1(\mathcal{H})} = \frac{1}{c}\, \triangle_{\mathrm{Herm}_1(\mathcal{H})}\ . \tag{170}$$

Hence, it suffices to show that $c \cdot \triangle_{\mathrm{Herm}_1(\mathcal{H})} \subset \mathrm{SP}$ for some $c > 0$, because then

$$\Lambda = \mathrm{SP}^* \subset (c \cdot \triangle_{\mathrm{Herm}_1(\mathcal{H})})^* = \frac{1}{c}\, \triangle_{\mathrm{Herm}_1(\mathcal{H})} \tag{171}$$

implies that Λ is bounded. Therefore, we will show that dilations of the vertices of $\triangle_{\mathrm{Herm}_1(\mathcal{H})}$ are contained in SP, i.e.there is $c > 0$ such that

$$\frac{1}{d^n}\mathbb{1} + c(A_\gamma - \frac{1}{d^n}\mathbb{1}) \in \mathrm{SP} \tag{172}$$

for all additive maps γ.

To achieve this, we will write $1/d^n\mathbb{1} + c(A_\gamma - 1/d^n\mathbb{1})$ as a convex combination of normalized projectors of the form $\frac{|\langle a \rangle|}{d^n}\Pi^r_{\langle a \rangle} \in \mathrm{SP}$ for $a \in E$ with noncontextual value assignments $r : \langle a \rangle \to \mathbb{Z}_d$. Due to Corollary 1 in Appendix A, the elements $\frac{|\langle a \rangle|}{d^n}\Pi^r_{\langle a \rangle}$ are indeed contained in SP. A noncontextual value assignment on a line $\langle a \rangle$ is always additive, that is

$$r(ka) = kr(a), \quad \forall k \in \mathbb{Z}_d \tag{173}$$

because with the phase convention chosen in eq. (3), it can be checked directly that $T_a T_{ka} = T_{(k+1)a}$, and so by definition $\beta(a, ka) = 0$ for all $a \in E$. Let $C = \{a_1, \ldots, a_N\} \subset E$ be a set that "covers" E, that is

$$E = \cup_{a \in C} \langle a \rangle. \tag{174}$$

For each subset $\mathcal{I} \subset [N] := \{1, \ldots, N\}$, there is an $a_{\mathcal{I}} \in E$ such that

$$\langle a_{\mathcal{I}} \rangle := \bigcap_{k \in \mathcal{I}} \langle a_k \rangle. \tag{175}$$

First, we will write A_γ as a linear combination of stabilizer code projectors $\Pi^r_{\langle a \rangle}$. We claim that

$$A_\gamma = \frac{1}{d^n} \sum_{\mathcal{I} \subset [N]} (-1)^{|\mathcal{I}|+1} \cdot |\langle a_{\mathcal{I}} \rangle| \cdot \Pi^{\gamma|_{\langle a_{\mathcal{I}} \rangle}}_{\langle a_{\mathcal{I}} \rangle}. \tag{176}$$

Since γ is additive, the restriction $\gamma|_{\langle a_{\mathcal{I}} \rangle} : \langle a_{\mathcal{I}} \rangle \to \mathbb{Z}_d$ satisfies (173) and defines a noncontextual value assignment on $\langle a_{\mathcal{I}} \rangle$. We rewrite the right hand side of (176) in the following way:

$$\frac{1}{d^n} \sum_{\mathcal{I} \subset [N]} (-1)^{|\mathcal{I}|+1} \cdot |\langle a_{\mathcal{I}} \rangle| \cdot \Pi^{\gamma|_{\langle a_{\mathcal{I}} \rangle}}_{\langle a_{\mathcal{I}} \rangle} \tag{177}$$

$$= \frac{1}{d^n} \sum_{\mathcal{I} \subset [N]} (-1)^{|\mathcal{I}|+1} \sum_{b \in \langle a_{\mathcal{I}} \rangle} \omega^{\gamma(b)} T_b \tag{178}$$

$$= \frac{1}{d^n} \sum_{b \in E} \Big(\sum_{\mathcal{I} \subset [N] \,:\, b \in \langle a_{\mathcal{I}} \rangle} (-1)^{|\mathcal{I}|+1} \Big) \omega^{\gamma(b)} T_b. \tag{179}$$

Thus, it suffices to show that

$$\sum_{\mathcal{I} \subset [N] \,:\, b \in \langle a_I \rangle} (-1)^{|\mathcal{I}|+1} = 1 \quad \forall b \in E. \tag{180}$$

However, this is a consequence of the *inclusion-exclusion principle* [63], that is

$$1 = \delta_{b \in E} = \delta_{b \in \cup_{a \in C} \langle a \rangle} = \sum_{\mathcal{I} \subset [N]} (-1)^{|\mathcal{I}|+1} \delta_{b \in \cap_{k \in I} \langle a_k \rangle} \tag{181}$$

$$= \sum_{\mathcal{I} \subset [N]} (-1)^{|\mathcal{I}|+1} \delta_{b \in \langle a_{\mathcal{I}} \rangle} = \sum_{\mathcal{I} \subset [N] \,:\, b \in \langle a_{\mathcal{I}} \rangle} (-1)^{|\mathcal{I}|+1}. \tag{182}$$

Finally, we will show that there is $c > 0$ such that we can write $\frac{1}{d^n}\mathbb{1} + c(A_\gamma - \frac{1}{d^n}\mathbb{1})$ for every γ as a convex combination of the operators $\frac{|\langle a_{\mathcal{I}} \rangle|}{d^n} \Pi^{\gamma|\langle a_{\mathcal{I}} \rangle}_{\langle a_{\mathcal{I}} \rangle} \in \mathrm{SP}$.

Observe that the identity (176) is equivalent to

$$A_\gamma - \frac{1}{d^n}\mathbb{1} = \sum_{\mathcal{I} \subset [N]} (-1)^{|\mathcal{I}|+1} \cdot \frac{|\langle a_{\mathcal{I}} \rangle|}{d^n} \cdot \Pi^{\gamma|\langle a_{\mathcal{I}} \rangle}_{\langle a_{\mathcal{I}} \rangle^*} \tag{183}$$

with

$$\Pi^r_{\langle a \rangle^*} := \frac{1}{|\langle a \rangle|} \sum_{b \in \langle a \rangle \setminus \{0\}} \omega^{r(b)} T_b \in \mathrm{Herm}_0(\mathcal{H}),\ r : \langle a \rangle \to \mathbb{Z}_d. \tag{184}$$

Moreover, due to Lemma 5, we have

$$\mathbb{1} = \sum_r \Pi^r_{\langle a \rangle} \quad \Longrightarrow \quad 0 = \sum_r \Pi^r_{\langle a \rangle^*} \tag{185}$$

for all $a \in E$, $\frac{|\langle a \rangle|}{d^n} \Pi^r_{\langle a \rangle} \in \mathrm{SP}$, where r ranges over all noncontextual value assignments $r : \langle a \rangle \to \mathbb{Z}_d$. This implies that

$$(-1) \cdot \Pi^{\gamma|_{\langle a_{\mathcal{I}} \rangle}}_{\langle a_{\mathcal{I}} \rangle^*} = \sum_{r \neq \gamma|_{\langle a \rangle}} \Pi^r_{\langle a_{\mathcal{I}} \rangle^*}. \tag{186}$$

As a consequence, every summand in the right hand side of eq. (183) can be written as a conic combination of elements $\Pi^r_{\langle a\rangle^*}$. By properly properly rescaling, we can find $c > 0$ such that

$$c(A_\gamma - \frac{1}{d^n}\mathbb{1}) = \sum_a \alpha_a \Pi^{r_a}_{\langle a\rangle^*} \quad \text{with} \quad \alpha_a \geq 0,\ \sum_a \alpha_a = 1, \tag{187}$$

which is equivalent to

$$\frac{1}{d^n}\mathbb{1} + c(A_\gamma - \frac{1}{d^n}\mathbb{1}) = \sum_a \alpha_a (\Pi^{r_a}_{\langle a\rangle^*} + \frac{1}{d^n}\mathbb{1}) = \sum_a \alpha_a \Pi^{r_a}_{\langle a\rangle} \in \mathrm{SP}. \tag{188}$$

This proves that $c\cdot\ \triangle_{\mathrm{Herm}_1(\mathcal{H})} \subset \mathrm{SP}$ for some $c > 0$, which remained to be shown. □