

Non-negative subtheories and quasiprobability representations of qubits

Joel J. Wallman and Stephen D. Bartlett

Centre for Engineered Quantum Systems, School of Physics, The University of Sydney, Sydney, NSW 2006, Australia

(Received 22 March 2012; published 26 June 2012)

Negativity in a quasiprobability representation is typically interpreted as an indication of nonclassical behavior. However, this does not preclude states that are non-negative from exhibiting phenomena typically associated with quantum mechanics—the single qubit stabilizer states have non-negative Wigner functions and yet play a fundamental role in many quantum information tasks. We seek to determine what other sets of quantum states and measurements of a qubit can be non-negative in a quasiprobability distribution, and to identify nontrivial groups of unitary transformations that permute the states in such a set. These sets of states and measurements are analogous to the single qubit stabilizer states. We show that no quasiprobability representation of a qubit can be non-negative for more than two bases in any plane of the Bloch sphere. Furthermore, there is a unique set of four bases that can be non-negative in an arbitrary quasiprobability representation of a qubit. We provide an exhaustive list of the sets of single qubit bases that are non-negative in some quasiprobability distribution and are also closed under a group of unitary transformations. This list includes two nontrivial families of three bases that both include the single qubit stabilizer states as a special case. For qudits, we prove that there can be no more than 2^{d^2} states in non-negative bases of a d -dimensional Hilbert space in any quasiprobability representation. Furthermore, these bases must satisfy certain symmetry constraints, corresponding to requiring the bases to be sufficiently different from each other.

DOI: [10.1103/PhysRevA.85.062121](https://doi.org/10.1103/PhysRevA.85.062121)

PACS number(s): 03.65.Ta, 03.67.—a, 03.65.Sq

I. INTRODUCTION

As an alternative to the standard formulation of quantum theory in terms of vectors in a Hilbert space, it is possible to express quantum states, transformations, and measurements as functions on some classical state space (such as a phase space). The most common of such representations is the Wigner function [1], which represents the quantum state of a particle as a distribution on the classical phase space of the particle. To describe all states in quantum theory, this function cannot be interpreted as a probability distribution; it must take negative values. Such a description is referred to as a *quasiprobability representation* [2].

The occurrence of negative probabilities in the description of a quantum state or measurement is often thought of as a fundamental indication of “quantum-ness” [3]. Conversely, any state or measurement that can be described by non-negative (true) probabilities is sometimes said to be “classical.” For the Wigner function of particle mechanics, Hudson’s theorem shows that a pure quantum state has a non-negative Wigner function if and only if it is a Gaussian state [4,5]. The discrete Wigner function extends these results to finite-dimensional quantum systems. For finite odd dimensions, the only pure states with non-negative Wigner functions are stabilizer states [6,7]. Negative distributions, on the other hand, are nonclassical in the sense that they are contextual [8] and can serve as a resource for quantum computation [9].

However, preparing a system in a quantum state described by a non-negative Wigner function is neither a necessary or sufficient condition to say that this system is classical. It is not necessary, because the Wigner function is only one possible quasiprobability representation of quantum theory. One can construct a quasiprobability representation in which any individual state or measurement has a non-negative distribution. It is also not sufficient, as it is possible to make all quantum states have non-negative distributions, but this

forces the conditional probabilities of some measurements to take on negative values [8,10]. To say that a system has a classical description, we require the set of preparations, transformations, and measurements we are considering to all have non-negative distributions in a quasiprobability representation. Consequently, we aim to determine what *subtheories* of quantum mechanics (i.e., theories constructed from closed subsets of quantum states, transformations, and measurements) are non-negative in some quasiprobability representation, or, alternatively, can be described “classically.”

In this paper, we examine sets of *non-negative bases* of $\mathcal{H}_d$ [i.e., orthonormal bases such that both the preparations (the states in the bases) and the projective measurements (measurements in these bases) are all non-negative]. Considering bases rather than individual states is a novel approach with powerful applications, as it enables us to establish strict upper bounds on the number of such bases and constrain the relation between any set of non-negative bases. We present a complete classification of “classical” subtheories for a single qubit, where by “classical” we mean that they can be described using non-negative probabilities in some quasiprobability distribution. We also obtain an upper bound for how many states can be elements of non-negative bases in higher dimensions. We note that several interesting questions remain on the “quantum” nature of even a single qubit or qudit, as demonstrated by a range of phenomena such as the proof of contextuality for a single qubit [11], models of quantum computation that use only one “clean” qubit [12] and recent results on the distillability of magic states [9].

While our results are primarily applicable to single systems, we believe that these results may be important in the pursuit of additional classically simulatable subtheories of multiple qubits. At present, we know of very few nontrivial sets of quantum gates that can be efficiently simulated on a classical computer, such the Clifford group and matchgate

circuits. Finding other subsets of gates that are also efficiently simulatable may provide key insights into the nature of quantum computation and the origin of any advantage of quantum computation over classical computation.

The paper is structured as follows. We begin by introducing quasiprobability representations of quantum mechanics in Sec. II and prove some elementary properties that any quasiprobability representation of quantum mechanics must satisfy. We obtain an upper bound on the number of non-negative bases of $\mathcal{H}_2$ for any quasiprobability representation in Sec. III A. We consider examples of quasiprobability representations that are non-negative for subtheories of a qubit with nontrivial transformations in Sec. III B. In Sec. IV, we generalize the theorems in Sec. III A to bases of $\mathcal{H}_d$ and obtain an upper bound on the number of states that are elements of non-negative bases for any finite d . We conclude with a discussion on the implications for quantum computation in Sec. V.

II. ONTOLOGICAL MODELS AND QUASIPROBABILITY REPRESENTATIONS

In this section, we motivate the study of quasiprobability representations, and non-negative subtheories within them, from a foundational perspective based on ontological models. We present our formalism for quasiprobability representations, and prove some elementary properties that any such representation of quantum mechanics must satisfy.

One approach to explaining the predictions of quantum mechanics is to formulate an ontological model (i.e., a hidden variable model) that reproduces some or all of the measurement statistics of quantum theory. Such a model is defined over some space Λ of ontic states. Preparations of quantum states correspond to probability distributions μ over Λ . Measurements correspond to sets of conditional measurement probabilities $\xi(k|\lambda)$ of observing an outcome k given that a system is in the ontic state $\lambda \in \Lambda$, respectively. A desirable feature of such a model is noncontextuality,¹ so that each preparation as a density operator corresponds to a single probability distribution and each measurement effect corresponds to a unique conditional probability [11].

While no noncontextual ontological model can reproduce all of quantum mechanics, it may be possible to define a noncontextual ontological model for a *subtheory* (i.e., a restriction to a closed set of preparations, transformations, and measurements). Such a subtheory may still capture some of the essential phenomena of quantum mechanics. For example, the set of states, transformations, and measurements with Gaussian Wigner distributions in particle mechanics can be described by a noncontextual ontological model [15]. For finite odd dimensions, the single qudit subtheory consisting of stabilizer states and measurements has an ontological model, namely the discrete Wigner function [16] restricted to this set [6,7]. Spekkens' toy theory [17] also provides an ontological model that is in many ways analogous to the stabilizer subtheory of a single qubit. Despite being “classical,”

these models allow a variety of information processing tasks typically associated with quantum mechanics, such as quantum teleportation and dense coding; see [15,17–20].

An alternate but related approach to explaining quantum mechanical predictions is to use a quasiprobability representation. As with an ontological model, a quasiprobability representation is defined over a space Λ (often, but not necessarily, a classical phase space) which can be interpreted as the space of ontic states, and it associates preparations and measurements with distributions and conditional distributions, respectively, over Λ . A quasiprobability representation is a faithful representation of the density operators and measurement effects of quantum theory, that is, the map from operators on Hilbert space to distributions on Λ is linear and invertible. Such representations can maintain noncontextuality but nevertheless reproduce the quantum predictions because, unlike ontological models, the distributions corresponding to such preparations and measurements are allowed to be negative (thus the term “quasiprobability”). That is, nonnegativity of distributions on ontic states is the classical assumption dropped by quasiprobability representations in order to reproduce the predictions of quantum theory.

A quasiprobability representation of quantum mechanics cannot be non-negative for all preparations and measurements; this fact is equivalent to the fact that an ontological model of quantum mechanics cannot be noncontextual [8]. However, a quasiprobability representation can be non-negative for a subtheory of quantum mechanics; specifically, the preparations, transformations and measurements within the subtheory can all possess non-negative probability distributions. In such cases, the quasiprobability representation provides a noncontextual ontological model for this subtheory. That is, the existence of a noncontextual ontological model for a subtheory and a non-negative quasiprobability representation for a subtheory are equivalent notions [8]. Throughout this paper, we will restrict our language (for the most part) to that of quasiprobability representations and the possible existence of non-negative subtheories within them, although the reader should keep in mind that results for non-negative subtheories in quasiprobability representations apply identically to a perspective of noncontextual ontological models for such subtheories.

An example of a quasiprobability representation frequently used in quantum optics is the Wigner function. The subtheory of quantum mechanics consisting of Gaussian states, transformations, and measurements, is completely described by a noncontextual ontological model for which the probability distributions and conditional measurement probabilities are all non-negative [15]. This subtheory is embedded in a quasiprobability distribution, describing all possible states, transformations, and measurements, but using negative probabilities for non-Gaussian ones.

A quasiprobability representation of quantum mechanics over a space Λ is defined by two sets of Hermitian operators, $\{F(\lambda)\}$ and $\{G(\lambda)\}$, acting on $\mathcal{H}_d$ [8,21]. The sets $\{F(\lambda)\}$ and $\{G(\lambda)\}$ are dual frames for the space of operators acting on $\mathcal{H}_d$ [21]. The quasiprobability associated with a quantum state ρ is

$$\mu_\rho(\lambda) = \text{Tr}[\rho F(\lambda)] \in \mathbb{R}. \quad (1)$$

¹While ontological models exist that do not satisfy these assumptions (e.g., the models in [13,14]), we do not consider them here.

The *support* of a state ρ is the set,

$$\mathcal{S}(\rho) = \{\lambda \in \Lambda \mid \mu_\rho(\lambda) \neq 0\}. \quad (2)$$

For measurements, the quasiprobability representation of an effect E (i.e., an element of a POVM) occurring if the system is in the state λ is given by an indicator function,

$$\xi_E(\lambda) = \text{Tr}[EG(\lambda)] \in \mathbb{R}. \quad (3)$$

As a system is always in some state, we require μ to be normalized,

$$\int_{\Lambda} d\lambda \mu_\rho(\lambda) = 1. \quad (4)$$

Similarly, as some outcome of a measurement always occurs, we require

$$\sum_j \xi_{E_j}(\lambda) = 1 \quad \forall \lambda \in \Lambda, \quad (5)$$

for all POVMs $\{E_j\}$. In order to reproduce the Born rule, μ and ξ must satisfy

$$\text{Tr}(\rho E) = \int_{\Lambda} d\lambda \mu_\rho(\lambda) \xi_E(\lambda). \quad (6)$$

While negative values of μ and ξ are allowed, there may be states ρ or effects E such that $\mu_\rho(\lambda) \geq 0$ or $\xi_E(\lambda) \in [0, 1]$ for all $\lambda \in \Lambda$, respectively. Any such state or effect is referred to as *non-negative*.

In quantum mechanics, there is a one-to-one correspondence between orthonormal bases of $\mathcal{H}_d$ and projective measurements, as for any orthonormal basis $\{|\Psi_j\rangle \mid j \in \mathbb{Z}_d\}$, the set of projectors $\{\mathbf{P}(\Psi_j) \mid j \in \mathbb{Z}_d\}$ is a projective measurement, where $\mathbf{P}(\Psi) = |\Psi\rangle\langle\Psi|$. Motivated by this correspondence, we define a *non-negative basis* as an orthonormal basis $\{|\Psi_j\rangle \mid j \in \mathbb{Z}_d\}$ such that for all $k \in \mathbb{Z}_d$ and $\lambda \in \Lambda$,

$$\mu_{\Psi_k}(\lambda) \geq 0, \quad (7a)$$

$$\xi_{\mathbf{P}(\Psi_k)}(\lambda) \in [0, 1]. \quad (7b)$$

That is, each state in a non-negative basis is a non-negative state and the projective measurement corresponding to the non-negative basis is a non-negative measurement. Considering non-negative bases as corresponding to both a basis of non-negative states and a non-negative projective measurement is the fundamental tool that we will exploit to obtain the results of this paper.

We are particularly interested in full subtheories of quantum mechanics that contain preparations and measurements in a set of non-negative bases as well as a group of transformations that permute the non-negative states. For qubits, we will establish upper bounds on the number of non-negative bases and some relations any non-negative bases must satisfy. We will also consider the possible sets of non-negative bases that are closed under a group of unitary transformations that permute at least some of the bases. For qudits, we will only undertake the first task (i.e., establish upper bounds on the number of non-negative bases and relations between any non-negative bases).

We begin by establishing some properties that any quasiprobability representation must satisfy. Recall from Eq. (2) that $\mathcal{S}(\Psi)$ is the ontic support of a state $|\Psi\rangle$. The supports $\mathcal{S}(\Psi_j)$ of the states in a non-negative basis $\{|\Psi_j\rangle \mid j \in \mathbb{Z}_d\}$ must be disjoint. Furthermore, for any λ in the support

of any of the elements of a non-negative basis, the indicator functions $\xi_{\mathbf{P}(\Psi_k)}(\lambda)$ are outcome deterministic and correspond to answering the question “is λ in the support of Ψ_k ?”²

Lemma II.1. Let $\{|\Psi_j\rangle \mid j \in \mathbb{Z}_d\}$ be a non-negative basis of $\mathcal{H}_d$ in a given quasiprobability representation. Then the supports $\mathcal{S}(\Psi_j)$ for $j \in \mathbb{Z}_d$ are disjoint, and for all $k, l \in \mathbb{Z}_d$ we have

$$\xi_{\mathbf{P}(\Psi_l)}(\lambda) = \delta_{l,k} \quad \forall \lambda \in \mathcal{S}(\Psi_k). \quad (8)$$

Proof. Let $\{|\Psi_j\rangle \mid j \in \mathbb{Z}_d\}$ be a non-negative basis, and therefore $|\langle\Psi_k|\Psi_l\rangle|^2 = \delta_{l,k}$. For the preparation $|\Psi_k\rangle$ followed by a measurement in this basis, the Born rule gives for all $k, l \in \mathbb{Z}_d$,

$$\int_{\mathcal{S}(\Psi_k)} d\lambda \mu_{\Psi_k}(\lambda) \xi_{\mathbf{P}(\Psi_l)}(\lambda) = \delta_{l,k}. \quad (9)$$

As $\mu_{\Psi_k}(\lambda) \geq 0$ and is normalized, and $\xi_{\mathbf{P}(\Psi_l)}(\lambda) \in [0, 1]$, the only solution³ is as in Eq. (8). ■

Quasiprobability representations are convex linear (i.e., they are affine maps, preserving convex combinations of states and measurements) [21]. As we now show, this is a strong restriction on the possible values $\mu_{\Psi_j}(\lambda)$ that a quasiprobability representation can assign to any non-negative basis $\{|\Psi_j\rangle\}$ at a point λ . In particular, in any quasiprobability representation, for all $\lambda \in \Lambda$ there exists a unique value $q(\lambda) \geq 0$ that must be assigned to exactly one element of each non-negative basis, with all other elements of the bases being zero. This unique value is given by $q(\lambda) = d \mu_{\frac{1}{d}\mathbf{1}}(\lambda)$, that is, d times the probability that the maximally mixed state assigns to the ontic state λ .

Lemma II.2. For any quasiprobability representation, there exists a unique function $q(\lambda) \geq 0$ for all $\lambda \in \Lambda$ such that for every non-negative basis $\{|\Psi_j\rangle\}$ of $\mathcal{H}_d$, Λ can be partitioned into $d + 1$ disjoint regions $\{\mathcal{S}(\Psi_j), \Lambda_0 \mid j \in \mathbb{Z}_d\}$ such that for all $k, l \in \mathbb{Z}_d$,

$$\mu_{\Psi_k}(\lambda) = \begin{cases} \delta_{k,l} q(\lambda) & \forall \lambda \in \mathcal{S}(\Psi_l), \\ 0 & \forall \lambda \in \Lambda_0, \end{cases} \quad (10)$$

where $\Lambda_0 = \{\lambda \mid q(\lambda) = 0\}$.

Proof. Let $\{|\Psi_j\rangle\}$ be an orthonormal basis of $\mathcal{H}_d$. Then the maximally mixed state can be written as

$$\frac{1}{d}\mathbf{1} = \frac{1}{d} \sum_j \mathbf{P}(\Psi_j). \quad (11)$$

As the quasiprobability representation is convex linear,

$$\mu_{\frac{1}{d}\mathbf{1}}(\lambda) = \frac{1}{d} \sum_j \mu_{\Psi_j}(\lambda), \quad (12)$$

for all $\lambda \in \Lambda$. Set $q(\lambda) = d \cdot \mu_{\frac{1}{d}\mathbf{1}}(\lambda)$.

For any non-negative basis $\{|\Psi_j\rangle\}$, each term in the right-hand side of Eq. (12) is non-negative by definition and so $q(\lambda)$ must be non-negative. For any non-negative basis $\{|\Psi_j\rangle\}$ and any point $\lambda \in \Lambda$, at most one $\mu_{\Psi_j}(\lambda)$ can be nonzero by Lemma II.1. ■

²This result was proved in [11].

³Here and elsewhere, except possibly on a set of measure 0.

As we now show, two quantum states that are elements of non-negative bases have disjoint supports if and only if they are orthogonal quantum states.

Lemma II.3. Let $\{|\Psi_j\rangle\}$ and $\{|\Phi_j\rangle\}$ be non-negative bases in a quasiprobability distribution. Then $\mathcal{S}(\Psi_j) \cap \mathcal{S}(\Phi_k) = \emptyset$ if and only if $|\Psi_j\rangle$ and $|\Phi_k\rangle$ are orthogonal.

Proof. Let $|\Psi_k\rangle$ and $|\Phi_l\rangle$ be elements of (possibly identical) non-negative bases. By Lemma II.1 and II.2, $\xi_{\mathbf{P}(\Psi_k)}(\lambda)$ is nonzero only for λ such that $\mu_{\Psi_k}(\lambda) > 0$ or $q(\lambda) = 0$. By Lemma II.2, $\mu_{\Phi_l}(\lambda)$ can only be nonzero for λ such that $q(\lambda) > 0$. Therefore, for

$$\int_{\Lambda} d\lambda \xi_{\mathbf{P}(\Psi_k)}(\lambda) \mu_{\Phi_l}(\lambda) = |\langle \Psi_k | \Phi_l \rangle|^2 \quad (13)$$

to hold, there must exist some λ such that $\mu_{\Psi_k}(\lambda) \neq 0$ and $\mu_{\Phi_l}(\lambda) \neq 0$ if and only if $|\Psi_j\rangle$ and $|\Phi_k\rangle$ are not orthogonal. ■

Lemmas II.1–II.3 provide the basic mathematical tools that we use to provide an upper bound on the number of non-negative bases of $\mathcal{H}_d$ in any quasiprobability representation of quantum mechanics. Any quasiprobability representation over a space Λ which is non-negative when restricted to a subset of bases of $\mathcal{H}_d$ must satisfy the following: The representation of a non-negative basis $\{|\Psi_j\rangle\}$ of $\mathcal{H}_d$ must correspond to a partitioning of Λ into d disjoint regions $\{\Lambda_j\}$ such that $|\Psi_j\rangle$ assigns non-zero probability for all $\lambda \in \Lambda_j$ and 0 probability elsewhere. The measurement $\{\mathbf{P}(\Psi_j)\}$ then corresponds to determining which region the ontic state is in.

III. QUASIPROBABILITY REPRESENTATIONS OF QUBITS

We now consider the simplest case, namely, quasiprobability representations of qubits (i.e., two-dimensional quantum systems). We will first establish upper bounds on the number of non-negative bases for a single qubit. We will then construct a complete characterization of all the sets of bases that are closed under a nontrivial group of unitary transformations and can be non-negative in some quasiprobability representation.

The qubit case allows a geometrical approach, as any qubit state $\rho \in \mathcal{B}(\mathcal{H}_2)$ can be written as

$$\rho = \frac{1}{2}[\mathbf{1} + \vec{r}(\rho) \cdot \vec{\sigma}], \quad (14)$$

where $\vec{r}(\rho) \in \mathbb{R}^3$ is the Bloch vector corresponding to ρ and $\vec{\sigma}$ is the vector (X, Y, Z) of Pauli matrices. Orthogonal quantum states then correspond to antipodal Bloch vectors, that is,

$$|\langle \Psi | \Phi \rangle|^2 = 0 \Leftrightarrow \vec{r}(\Psi) = -\vec{r}(\Phi). \quad (15)$$

Another useful feature of the qubit case is that any pure single qubit state can be uniquely extended to an orthonormal basis of $\mathcal{H}_2$.

A. Maximum number of non-negative qubit bases

We begin by showing that a quasiprobability representation cannot be non-negative for more than two bases in any plane of the Bloch sphere. That is, for any quasiprobability distribution, the Bloch vectors corresponding to any three distinct non-negative bases (if three such bases exist) must be linearly independent. We will then show that any four non-negative

bases (if four such bases exist) must correspond to the vertices of a cube.

Theorem III.1. In any quasiprobability representation of $\mathcal{H}_2$, there are at most two non-negative bases in any plane of the Bloch sphere.

Proof. The proof is by contradiction. Assume μ is a quasiprobability distribution that is non-negative for three bases of $\mathcal{H}_2$ that are coplanar in the Bloch sphere. Without loss of generality, assume μ is non-negative for $\{|0\rangle, |1\rangle\}$ and denote two other non-negative bases by $\rho_1^\pm$ and $\rho_2^\pm$ and the corresponding Bloch vectors by $\pm \vec{r}_i = \pm(x_i, 0, z_i)$ for $i = 1, 2$, where $x_1, x_2 > 0$ (relabeling basis states if necessary). Then we can define a Bloch vector,

$$\vec{s} := a\vec{r}_1 + (1-a)(-\vec{r}_2) = (0, 0, 2b-1), \quad (16)$$

where

$$a = \frac{x_2}{x_1 + x_2}, \quad (17)$$

$$2b-1 = \frac{x_2 z_1 - x_1 z_2}{x_1 + x_2}.$$

Therefore the mixed state,

$$\rho = \frac{1}{2}(\mathbf{1} + \vec{s} \cdot \vec{\sigma}), \quad (18)$$

can be decomposed as a convex combination of $\{\mathbf{P}(0), \mathbf{P}(1)\}$ or $\{\rho_1^+, \rho_2^-\}$. A graphical representation of $\pm \vec{r}_i$ and $\vec{s}$ is presented in Fig. 1.

As μ is convex linear, we have

$$\begin{aligned} \mu_\rho(\lambda) &= a\mu_{\rho_1^+}(\lambda) + (1-a)\mu_{\rho_2^-}(\lambda) \\ &= b\mu_{\mathbf{P}(0)}(\lambda) + (1-b)\mu_{\mathbf{P}(1)}(\lambda), \end{aligned} \quad (19)$$

for all $\lambda \in \Lambda$.

As the three bases are distinct, neither a or b can be 0 or 1. By Lemma II.2, at any point $\lambda \in \Lambda$ such that $q(\lambda) > 0$, one state from each non-negative basis assigns probability $q(\lambda)$ to λ and the other state from each non-negative basis assigns zero probability to λ . As $\{|0\rangle, |1\rangle\}$ is a non-negative basis, the second line of Eq. (19) implies

$$\mu_\rho(\lambda) < q(\lambda) \quad \forall \lambda \in \Lambda \text{ s.t. } q(\lambda) > 0. \quad (20)$$

As ρ_1^+ and ρ_2^- are not orthogonal, there exists $\lambda' \in \mathcal{S}(\rho_1^+) \cap \mathcal{S}(\rho_2^-)$ by Lemma II.3. For this point λ' , we have $q(\lambda') > 0$. The first line of Eq. (19) gives $\mu_\rho(\lambda') = q(\lambda')$, which contradicts Eq. (20). ■

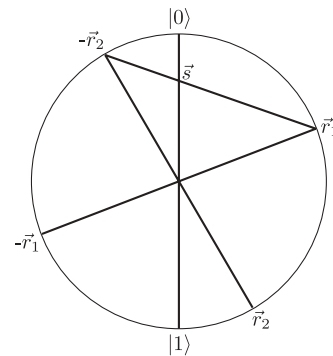


FIG. 1. Diagrammatic representation of the decomposition of ρ in terms of $\{|0\rangle, |1\rangle\}$ and $\{\rho_1^+, \rho_2^-\}$.

The proof that no three nonnegative bases can be coplanar in the Bloch sphere rests on the existence of the vector in Eq. (16). This vector must always exist as otherwise there would be three linearly independent vectors in a two-dimensional plane, which would be a contradiction.

Initially, one might expect that this proof could be generalized to rule out four non-negative bases in the full Bloch sphere (i.e., the existence of four non-negative bases would correspond to four linearly independent Bloch vectors, contradicting the dimensionality of a sphere). This intuition almost always holds. However, as we now prove, there is a single exception with a high degree of symmetry, namely, if the four bases correspond to the vertices of a cube. Therefore any set of four non-negative bases of $\mathcal{H}_2$ must correspond to the vertices of a cube, which also implies that there exists no quasiprobability representation of $\mathcal{H}_2$ with five or more non-negative bases.

Theorem III.2. If there are four non-negative bases in a quasiprobability representation of a qubit, then the Bloch vectors corresponding to these four bases must correspond to the vertices of a cube.

Proof. Assume μ is a quasiprobability distribution that is non-negative for four bases of $\mathcal{H}_2$. For convenience, we set one of the non-negative bases to $\{|0\rangle, |1\rangle\}$ and denote the Bloch vectors for three other non-negative bases by $\pm\vec{r}_j$ for $j = 1, 2, 3$. Negating the vectors $\vec{r}_j$ as necessary (which corresponds to relabeling the basis states), there exist $a, s_j \in [0, 1]$ such that

$$\sum_{j=1}^3 s_j \vec{r}_j = (0, 0, 2a - 1), \quad (21)$$

and $\sum_j s_j = 1$. This is always possible as otherwise the three vectors $\vec{r}_j \in \mathbb{R}^2$, formed by taking the x and y components of $\vec{r}_j$, would be linearly independent. Therefore,

$$\rho = s_1 \rho_1^+ + s_2 \rho_2^+ + s_3 \rho_3^+ = a \mathbf{P}(0) + (1 - a) \mathbf{P}(1), \quad (22)$$

where $\rho_{\vec{r}_j}^{\epsilon_j}$ is the density matrix corresponding to the Bloch vector $\epsilon_j \vec{r}_j$. As the bases are all distinct, none of the coefficients can be 1. Furthermore, none of the coefficients can be 0 as otherwise there would be three non-negative bases in a single plane of the Bloch sphere, contradicting Theorem III.1.

As μ is convex linear, we have

$$\begin{aligned} \mu_\rho(\lambda) &= s_1 \mu_{\rho_1^+}(\lambda) + s_2 \mu_{\rho_2^+}(\lambda) + s_3 \mu_{\rho_3^+}(\lambda) \\ &= a \mu_{\mathbf{P}(1)}(\lambda) + (1 - a) \mu_{\mathbf{P}(0)}(\lambda), \end{aligned} \quad (23)$$

for all $\lambda \in \Lambda$. As $\{|0\rangle, |1\rangle\}$ is a non-negative basis, the second line of Eq. (23) implies

$$\mu_\rho(\lambda) = \begin{cases} a q(\lambda) & \text{if } \lambda \in \mathcal{S}(|0\rangle), \\ (1 - a) q(\lambda) & \text{if } \lambda \in \mathcal{S}(|1\rangle). \end{cases} \quad (24)$$

We now show that the first line is only consistent with Eq. (24) if the four bases correspond to the vertices of a cube. To do this, we consider the intersection of the supports of ρ_k^- and ρ_l^- for different values of $k \neq l$ by considering permutations $\{j, k, l\}$ of $\{1, 2, 3\}$. As ρ_k^- and ρ_l^- are not orthogonal, there exist $\lambda_j \in \mathcal{S}(\rho_k^-) \cap \mathcal{S}(\rho_l^-)$ by Lemma II.3. As $\mu_{\rho_k^+}(\lambda_j) = 0 = \mu_{\rho_l^+}(\lambda_j)$ and the second line of Eq. (23) is strictly greater than zero for λ_j , we require $\mu_{\rho_j^+}(\lambda_j) = q(\lambda_j) >$

0 by Lemma II.2. Therefore we have $s_j = a$ or $s_j = 1 - a$. By considering all permutations $\{j, k, l\}$, we see that this must hold for all $j = 1, 2, 3$. If one of the s_j differs from the other two, then

$$1 = s_1 + s_2 + s_3 = \begin{cases} 1 + a, \\ 2 - a, \end{cases} \quad (25)$$

which only holds if $a \in \{0, 1\}$, which is a contradiction. Therefore $s_j = \frac{1}{3}$ for $j = 1, 2, 3$ and $a \in \{\frac{1}{3}, \frac{2}{3}\}$.

To show that the only Bloch vectors that satisfy this relation correspond to the vertices of a cube, consider the vectors $\vec{r}_j \in \mathbb{R}^2$ whose x and y components are the x and y components of $\vec{r}_j$. From Eq. (21) and the values of s_j , we have

$$\vec{r}_1 + \vec{r}_2 + \vec{r}_3 = 0, \quad (26)$$

which holds if and only if the $\vec{r}_j$ are, up to an arbitrary rotation,

$$\vec{r}_j = \sin \theta \left(\cos \frac{2\pi j}{3}, \sin \frac{2\pi j}{3} \right), \quad (27)$$

for some $\theta \in [0, \pi]$. Therefore the Bloch vectors are

$$\vec{r}_j = \left(\sin \theta \cos \frac{2\pi j}{3}, \sin \theta \sin \frac{2\pi j}{3}, \mu_j \cos \theta \right), \quad (28)$$

for some $\mu_j = \pm 1$. If one of the μ_j is different from the other two, then substituting Eq. (28) into Eq. (21) with $s_j = \frac{1}{3}$ and $2a - 1 = \pm \frac{1}{3}$ gives $\cos \theta = \pm 1$, which implies that the Bloch vectors are all degenerate, contradicting the assumption that they correspond to distinct bases. If the μ_j are identical, then substituting Eq. (28) into Eq. (21) gives $\cos \theta = \pm \frac{1}{3}$. These four Bloch vectors correspond to the vertices of a cube. ■

Theorems III.1 and III.2 show that in any quasiprobability representation of a qubit, there can be no more than four non-negative bases and any three non-negative bases must have linearly independent Bloch vectors. We will find that these theorems alone do not completely characterize the sets of bases that can simultaneously be non-negative in a quasiprobability representation. In particular, there are further restrictions on the sets of three bases that can simultaneously be non-negative.

B. Sets of non-negative bases that allow nontrivial transformations

We now further restrict our consideration to non-negative subtheories of a qubit that contain at least one nontrivial unitary transformation permuting the elements of the non-negative bases. Such non-negative subtheories are analogous to the subtheory of single qubit stabilizer states and measurements in the corresponding bases and single qubit Clifford transformations.

The set of Bloch vectors corresponding to a set $\mathcal{A}$ of non-negative bases can be regarded as a set of pairs of antipodal points on the surface of a (Bloch) sphere. If there are transformations that permute the elements of these bases, then they correspond to elements of the *point group* of the set of points on the surface of the sphere, which will be a discrete subgroup of $O(3)$ (except in the special case wherein there is only one non-negative basis). For the transformations to have a unitary representation, they must have determinant $+1$

[i.e., be elements of $\text{SO}(3)$]. As we are interested in subtheories of quantum mechanics, we restrict to elements of $\text{SO}(3)$.⁴

As we have proved that there can be no more than four non-negative bases in any quasiprobability distribution and that no three non-negative bases can be coplanar in the Bloch sphere, the only relevant groups are point groups that permute a set of four or fewer pairs of antipodal points. A complete list of such groups is as follows [22]:

- (1) D_∞ , the group of rotations about an axis and π flips around some orthogonal axis;
- (2) $C_2 \simeq \mathbb{Z}_2$, the cyclic group of two elements;
- (3) D_2 , the symmetry group of a rectangle (equivalent to the group of Pauli matrices);
- (4) D_3 , the symmetry group of an equilateral triangle;
- (5) D_4 , the symmetry group of a square; and
- (6) O_h , the octahedral group (equivalent to the single qubit Clifford group).

We now provide an exhaustive characterization of the sets of four or fewer bases that are non-negative in a quasiprobability representation and in addition are closed under a nontrivial group of unitary transformations. We are explicit with our construction for the cases of three and four non-negative bases; quasiprobability representations that are non-negative for one or two bases appear as subsets of those for three non-negative bases, and therefore we do not consider them separately.

1. One non-negative basis

If there is a single non-negative basis (i.e., $\{|0\rangle, |1\rangle\}$ up to unitary equivalence), then the group of transformations that permute the elements of the non-negative basis is the continuous (Lie) group D_∞ . We can therefore find a quasiprobability representation over a space consisting of two points—a classical bit—for which this basis is represented by non-negative probability distributions.

The transformations of the non-negative basis is generated by the group of $U(1)$ rotations about the z axis (which leaves the basis states invariant) and X , which corresponds to a bit flip.

2. Two non-negative bases

If there are two non-negative bases, then the point group of the four vertices $\{\pm\vec{r}_1, \pm\vec{r}_2\}$ is generically D_2 . Expressing the

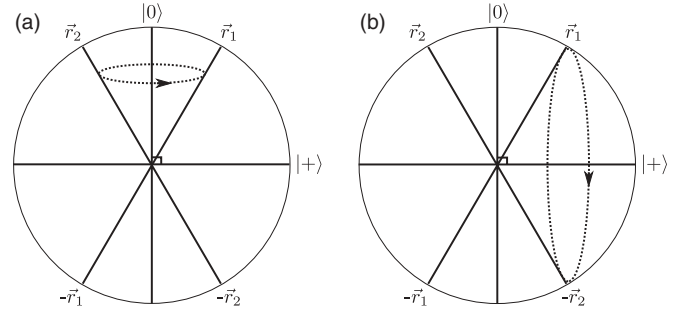


FIG. 2. Bloch vector representation of the transformations Z and X that map (a) $\vec{r}_1 \leftrightarrow \vec{r}_2$ and $-\vec{r}_1 \leftrightarrow -\vec{r}_2$ and (b) $\vec{r}_1 \leftrightarrow -\vec{r}_2$ and $-\vec{r}_1 \leftrightarrow \vec{r}_2$, respectively.

Bloch vectors as

$$\begin{aligned}\vec{r}_1 &= (\sin \theta, 0, \cos \theta), \\ \vec{r}_2 &= (-\sin \theta, 0, \cos \theta),\end{aligned}\tag{29}$$

the symmetry group is the group of Pauli matrices. The two non-negative bases and the generators of the symmetry group are illustrated in Fig. 2. For the special case defined by $\vec{r}_1 \cdot \vec{r}_2 = 0$, the four vertices correspond to the corners of a square and so the point group is D_4 , which for the above vectors is generated by X and $Y^{\frac{1}{2}}$.

3. Three non-negative bases

There are two inequivalent families of sets of three bases with nontrivial point groups, D_3 and $\mathbb{Z}_2$, respectively. Both families have special cases for which the three bases have a higher degree of symmetry. In particular, both families include a unique special case in common for which the point group is the octahedral group. For this special case, we recover the standard single qubit stabilizer states, whose point group is the single qubit Clifford group.

We wish to construct quasiprobability representations over some ontic state space Λ , in which a state ρ_j^ϵ with Bloch vector $\epsilon\vec{r}_j$ (for $\epsilon = \pm$) corresponds to a non-negative probability distribution $\mu_{\rho_j^\epsilon}(\lambda)$ over Λ . By Lemma II.2, any nonnegative basis corresponds to a bipartition of the ontic space Λ such that each element of the basis only has support on one of the partitions. Furthermore, the probability that any non-negative state Ψ assigns to an ontic state $\lambda \in \mathcal{S}(\Psi)$ must be $q(\lambda)$, which is double the probability the maximally mixed state assigns to λ . Therefore, with three non-negative bases, Λ can be partitioned into eight regions corresponding to the eight combinations of basis states that are non-negative in that region. Without loss of generality, we can integrate over these regions and, for convenience, we parametrize the ontic state space as the eight points,

$$\Lambda = \{(\gamma, k) | \gamma = \pm, k \in \mathbb{Z}_4\},\tag{30}$$

and henceforth we write $q(\lambda) = q(\gamma, k)$. Our (arbitrary) choices of six distributions over these eight ontic states are listed in Table I.

The probability distribution $q(\gamma, k)$ will have to satisfy certain constraints in order to reproduce the correct quantum mechanical predictions. The distribution must be normalized

⁴We only consider elements of $\text{SO}(3)$ rather than $\text{O}(3)$ because we aim to describe quantum mechanics, or a subtheory thereof, for which only unitary and not antiunitary transformations are allowed. However, it is possible to extend quantum mechanics to include antiunitary transformations, and if one wished to consider ontological models describing such an extended theory, then all transformations in $\text{O}(3)$ (including reflections) can be included. Because we are considering subtheories with only discrete evolution, the restriction to unitary dynamics is not necessarily well motivated. Specifically, because the evolution in a discrete theory is not required to be continuously deformable to the identity, such an extended model may be reasonable.

TABLE I. List of the six possible supports over the set of ontic states $\{(\gamma, k) | \gamma = \pm, k \in \mathbb{Z}_4\}$ for the elements of three non-negative bases with Bloch vectors $\pm \vec{r}_1, \pm \vec{r}_2, \pm \vec{r}_3$. Note that two states have disjoint supports if and only if they are orthogonal.

Bloch vector	Support of corresponding quantum state
$\vec{r}_1$	$(+, 0), (+, 1), (-, 2), (-, 3)$
$-\vec{r}_1$	$(-, 0), (-, 1), (+, 2), (+, 3)$
$\vec{r}_2$	$(+, 0), (-, 1), (+, 2), (-, 3)$
$-\vec{r}_2$	$(-, 0), (+, 1), (-, 2), (+, 3)$
$\vec{r}_3$	$(+, 0), (-, 1), (-, 2), (+, 3)$
$-\vec{r}_3$	$(-, 0), (+, 1), (+, 2), (-, 3)$

over the support of each non-negative state, that is,

$$\sum_{(\gamma, k) \in \mathcal{S}(\rho_j^+)} q(\gamma, k) = 1, \quad j = 1, 2, 3, \quad (31a)$$

$$\sum_{(\gamma, k) \in \mathcal{S}(\rho_j^-)} q(\gamma, k) = 1, \quad j = 1, 2, 3. \quad (31b)$$

These six equations give only four independent constraints.

In addition, in order for the quasiprobability representation to reproduce the quantum probabilities of preparing a system in one non-negative basis and then measuring in another non-negative basis, we also require

$$\sum_{(\gamma, k) \in \mathcal{S}(\rho_{j_1}^+) \cap \mathcal{S}(\rho_{j_2}^+)} q(\gamma, k) = \frac{1}{2}(1 + \vec{r}_{j_1} \cdot \vec{r}_{j_2}), \quad (32)$$

for all $j_1 \neq j_2$, which are obtained by substituting Eq. (8) into Eq. (6). Only three of these (say, for $j_1 > j_2$) give independent constraints. Therefore, for the cases we consider, we obtain four independent constraints from Eq. (31) and three further independent constraints from Eq. (32). With $q(\gamma, k)$ defined on eight points, the seven independent constraints ensure there will be at most a one-parameter family of quasiprobability representations for which the desired sets of states are all non-negative.

To find a quasiprobability representation of all states and measurements of a qubit that is non-negative for the above bases, we need to find operators $F(\gamma, k)$ and $G(\gamma, k)$ satisfying Eqs. (1), (3), and (6). As these operators should give a non-negative distribution for our chosen bases, we require

$$\begin{aligned} \mu_{\rho_j^\epsilon}(\gamma, k) &= \text{Tr}[F(\gamma, k)\rho_j^\epsilon] \\ &= \begin{cases} q(\gamma, k) & \text{if } (\epsilon, j) \in \mathcal{S}(\rho_j^\epsilon), \\ 0 & \text{otherwise,} \end{cases} \end{aligned} \quad (33)$$

for all $\epsilon = \pm$ and $j = 1, 2, 3$. To find such $F(\gamma, k)$ and $G(\gamma, k)$, we find vectors $\vec{d}_k$ such that

$$\vec{d}_k \cdot \vec{r}_j = \begin{cases} 1 & \text{if } (+, k) \in \mathcal{S}(\rho_j^+), \\ -1 & \text{otherwise.} \end{cases} \quad (34)$$

The operators,

$$\begin{aligned} F(\gamma, k) &= \frac{q(\gamma, k)}{2}(\mathbf{1} + \gamma \vec{d}_k \cdot \vec{\sigma}), \\ G(\gamma, k) &= \frac{1}{2}(\mathbf{1} + \gamma \vec{d}_k \cdot \vec{\sigma}), \end{aligned} \quad (35)$$

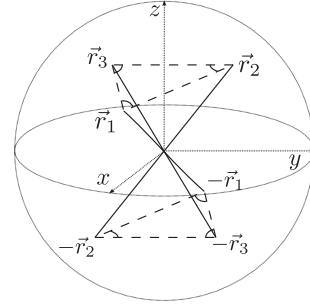


FIG. 3. Illustration of the Bloch vectors in Eq. (36), whose point group is D_3 .

then define a quasiprobability representation that is non-negative for the three bases. Note that the operators $F(\gamma, k)$ and $G(\gamma, k)$ are identical up to normalization. In what follows, we will only explicitly present the operators $F(\gamma, k)$.

We now turn separately to each of the inequivalent families and construct quasiprobability representations that are non-negative for each of the families. As both families include the single qubit stabilizer states as a special case, we will discuss this case after covering the general case for each family.

Case 1. Up to an overall unitary, the first family is the one-parameter family of three bases illustrated in Fig. 3 with

$$\begin{aligned} \vec{r}_1 &= (\sin \theta, 0, \cos \theta), \\ \vec{r}_2 &= \left(-\frac{1}{2} \sin \theta, \frac{\sqrt{3}}{2} \sin \theta, \cos \theta\right), \\ \vec{r}_3 &= \left(-\frac{1}{2} \sin \theta, -\frac{\sqrt{3}}{2} \sin \theta, \cos \theta\right), \end{aligned} \quad (36)$$

for $\theta \in (0, \pi)$. The corresponding point group is D_3 , which is generated by a $\frac{2\pi}{3}$ rotation about the z axis (which we denote by Γ) and a π rotation about the y axis (which we denote by Π). These transformations do not commute as D_3 is a non-Abelian group.

To define a quasiprobability representation for which these three bases are non-negative, we need to find a distribution $q(\gamma, k)$ satisfying Eqs. (31) and (32). The only such distributions are

$$\begin{aligned} q(+, 0) &= q_0, \\ q(+, k) &= \frac{3}{2} \sin^2 \theta - 1 + q_0, \quad k = 1, 2, 3, \\ q(-, 0) &= 2 - q_0 - \frac{9}{4} \sin^2 \theta, \\ q(-, k) &= 1 - q_0 - \frac{3}{4} \sin^2 \theta, \quad k = 1, 2, 3, \end{aligned} \quad (37)$$

where $q_0 \in [0, 1]$ is a free parameter. The requirement that all probabilities should be in the interval $[0, 1]$ implies

$$0 \leq q_0 \leq 2 - \frac{9}{4} \sin^2 \theta, \quad (38)$$

which can only be satisfied when $\sin^2 \theta \leq \frac{8}{9}$. Therefore a quasiprobability representation for which these bases are non-negative can only be defined if $\sin^2 \theta \leq \frac{8}{9}$. This constraint is not a consequence of either Theorem III.1 or III.2 and so provides an additional constraint on the set of non-negative bases.

In an ontological model, transformations are fundamentally transformations of ontic states, not epistemic states (i.e., states of knowledge). For an ontological model of a subtheory of quantum mechanics in which (pure) quantum states are

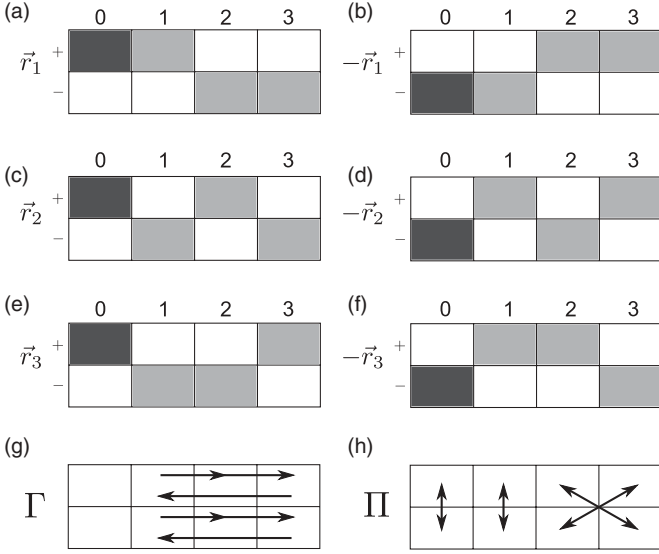


FIG. 4. (a)–(f) Probability distributions $\mu_{\rho_j^{\epsilon}}(\gamma, k)$ over eight ontic states $\{(\gamma, k) | \gamma = \pm, k \in \mathbb{Z}_4\}$, for the three bases with Bloch vectors as in Eq. (36), where the quantum states assign nonzero probability to the shaded ontic states. The shadings indicate the two nonzero probabilities q_0 (dark gray) and q_1 (light gray) that are assigned to the different ontic states. (g) and (h) Permutations of the ontic states that effect the rotations Γ and Π , respectively.

epistemic states, this means that unitary transformations must *supervene on* transformations of ontic states (i.e., must be a consequence of some underlying transformation of the ontic states). In the models we consider, we will always be able to assume that transformations of ontic states are deterministic (i.e., correspond to a permutation of the ontic state space). Conversely, some (but not all) permutations of ontic states in these models *effect* a unitary transformation. The permutations of ontic states that can effect the rotations Γ and Π are illustrated in Figs. 4(g) and 4(h), respectively. These permutations of ontic states only permute the probability distributions for non-negative states if $q(+, 0) = q(-, 0)$, which fixes

$$\begin{aligned} q(\gamma, k) &= q_0 = 1 - \frac{9}{8} \sin^2 \theta, \\ q(\gamma, k) &= \frac{3}{8} \sin^2 \theta := q_1, \quad k = 1, 2, 3, \end{aligned} \quad (39)$$

for $\gamma = \pm$, as illustrated in Figs. 4(a)–4(f). Note that the transformation Γ always supervenes on the permutation in Fig. 4(g), even if $q(+, 0) \neq q(-, 0)$.

To construct a quasiprobability distribution for which the three bases are non-negative and described by the above probability distributions, we need to find vectors $\vec{d}_k$ satisfying Eq. (34). By examining the permutation of ontic states that effects the transformation Γ , we note that $\vec{d}_k = \Gamma_R^{k-1} \vec{d}_1$ for $k = 1, 2, 3$ [where Γ_R is the fundamental (spin-1) representation of Γ], so we need only find $\vec{d}_0$ and $\vec{d}_1$, which can easily be determined to be

$$\begin{aligned} \vec{d}_0 &= (0, 0, \sec \theta), \\ \vec{d}_1 &= \left(\frac{4}{3} \csc \theta, 0, -\frac{1}{3} \sec \theta \right). \end{aligned} \quad (40)$$

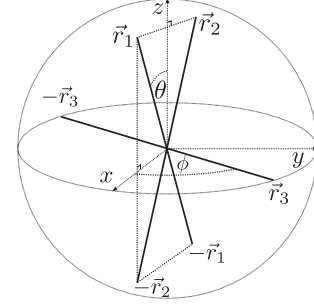


FIG. 5. Illustration of the Bloch vectors in Eq. (43), whose point group is $\mathbb{Z}_2$.

The operators $F(\gamma, k)$ are then

$$\begin{aligned} F(\gamma, 0) &= \frac{q_0}{2} \begin{pmatrix} 1 + \gamma \sec \theta & 0 \\ 0 & 1 - \gamma \sec \theta \end{pmatrix}, \\ F(\gamma, k) &= \frac{q_1}{2} \Gamma_U^{k-1} \begin{pmatrix} 1 - \frac{\gamma}{3} \sec \theta & \frac{4\gamma}{3} \csc \theta \\ \frac{4\gamma}{3} \csc \theta & 1 + \frac{\gamma}{3} \sec \theta \end{pmatrix} \Gamma_U^{1-k}, \end{aligned} \quad (41)$$

for $\gamma = \pm$ and $k = 1, 2, 3$, where

$$\Gamma_U = \begin{pmatrix} e^{\frac{2\pi i}{3}} & 0 \\ 0 & e^{-\frac{2\pi i}{3}} \end{pmatrix} \quad (42)$$

is the spin- $\frac{1}{2}$ representation of Γ . The operators in Eq. (41) define a quasiprobability representation that reproduces all of quantum mechanics for the qubit and is non-negative for the states and measurements corresponding to the three bases with Bloch vectors in Eq. (36).

Case 2. Up to an overall unitary, the second family of three bases have Bloch vectors,

$$\begin{aligned} \vec{r}_1 &= (\sin \theta, 0, \cos \theta), \\ \vec{r}_2 &= (-\sin \theta, 0, \cos \theta), \\ \vec{r}_3 &= (\cos \phi, \sin \phi, 0), \end{aligned} \quad (43)$$

for $\theta \in (0, \frac{\pi}{2})$ and $\phi \in (0, \pi)$, as illustrated in Fig. 5. For $\phi \neq \frac{\pi}{2}$, the point group of these bases is $\mathbb{Z}_2$ (i.e., the only nontrivial transformation is a π rotation about the z axis). If $\phi = \frac{\pi}{2}$, then the point group is the set of Pauli matrices, which form a representation of D_2 (or the octahedral group if $\theta = \frac{\pi}{4}$).

The only distributions $q(\gamma, k)$ satisfying Eqs. (31) and (32) are

$$\begin{aligned} q(+, 0) &= q(+, 3) := q_0, \\ q(+, 1) &= q_0 - \frac{1}{2} \cos 2\theta - \frac{1}{2} \cos \phi \sin \theta, \\ q(+, 2) &= q_0 - \frac{1}{2} \cos 2\theta + \frac{1}{2} \cos \phi \sin \theta, \\ q(-, 0) &= q(-, 3) = \frac{1}{2} - q_0 + \frac{1}{2} \cos 2\theta, \\ q(-, 1) &= \frac{1}{2} - q_0 - \frac{1}{2} \cos \phi \sin \theta, \\ q(-, 2) &= \frac{1}{2} - q_0 + \frac{1}{2} \cos \phi \sin \theta, \end{aligned} \quad (44)$$

where $q_0 \in [0, 1]$ is a free parameter. Requiring all probabilities to be in the interval $[0, 1]$, we obtain

$$\begin{aligned} 0 &\leq q(+, 1) + q(-, 1) = \sin \theta (\sin \theta - \cos \phi) \leq 2, \\ 0 &\leq q(+, 2) + q(-, 2) = \sin \theta (\sin \theta + \cos \phi) \leq 2. \end{aligned} \quad (45)$$

Therefore we require $|\cos \phi| \leq \sin \theta$ in order for the bases in Eq. (43) to be non-negative in a quasiprobability representation. This is another example of a more restrictive condition on non-negative bases than either Theorem III.1 or III.2. This condition is also sufficient, as for such θ, ϕ , we can set the distribution to

$$\begin{aligned} q_0 &:= q(\gamma, 0) = \frac{\cos^2 \theta}{2}, \\ q_1 &:= q(\gamma, 1) = \frac{\sin \theta}{2}(\sin \theta + \cos \phi), \\ q_2 &:= q(\gamma, 2) = \frac{\sin \theta}{2}(\sin \theta - \cos \phi), \end{aligned} \quad (46)$$

for $\gamma = \pm$.

As with the models for the bases in Eq. (36), we can view the unitary transformations of quantum states as supervening on permutations of ontic states. The distribution in Eq. (46) captures the symmetry of the bases and the possible transformations, as unitary transformations can only supervene on permutations of the ontic states for values of θ, ϕ for which the bases have the appropriate symmetry.

For $\phi \neq \frac{\pi}{2}$, the only nontrivial unitary transformation that permutes non-negative states is Z (i.e., a π flip about the z axis). The permutation of ontic states that effects Z is depicted in Fig. 6(g).

For $\phi = \frac{\pi}{2}$, the symmetry group is the group of Pauli matrices, generated by X and Z , which is a spin- $\frac{1}{2}$ representation of D_2 . The permutation of ontic states that effects X is depicted in Fig. 6(h). This permutation maps $(+, 1)$ to $(-, 2)$, and so we require $q(+, 1) = q(-, 2)$, which is satisfied if and only if $\phi = \frac{\pi}{2}$ in the distribution in Eq. (46).

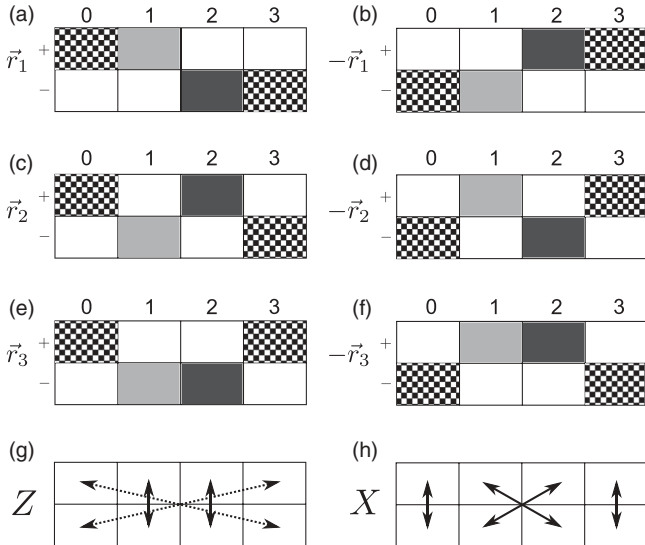


FIG. 6. (a)–(f) Probability distributions $\mu_{\rho_j^{\epsilon}}(\gamma, k)$ over eight ontic states $\{(\gamma, k) | \gamma = \pm, k \in \mathbb{Z}_4\}$, for the three bases with Bloch vectors as in Eq. (43). The shadings indicate the three nonzero probabilities, q_0 (checkered), q_1 (light gray), and q_2 (dark gray), that are assigned to the different ontic states. (g) and (h) Permutations of the ontic states that effects the unitary transformations Z and X , respectively. Note that the permutation in (h) is only valid for the special case where the bases have D_2 symmetry (i.e., when $q_1 = q_2$).

To construct a quasiprobability distribution for which these bases are non-negative, we need to find vectors $\vec{d}_k$ satisfying Eq. (34). One such set of vectors is

$$\begin{aligned} \vec{d}_0 &= (0, \csc \phi, \sec \theta), \\ \vec{d}_1 &= (\csc \theta, -\csc \phi - \cos \phi \csc \theta, 0), \\ \vec{d}_2 &= (-\csc \theta, -\csc \phi + \cos \phi \csc \theta, 0), \\ \vec{d}_3 &= (0, \csc \phi, -\sec \theta). \end{aligned} \quad (47)$$

The corresponding operators $F(\gamma, k)$ are then

$$\begin{aligned} F(\gamma, 0) &= \frac{q_0}{2}(\mathbf{1} + \gamma \csc \phi Y + \gamma \sec \theta Z), \\ F(\gamma, 1) &= \frac{q_1}{2}[\mathbf{1} + \gamma \csc \theta X - \gamma(\csc \phi + \cos \phi \csc \theta)Y], \\ F(\gamma, 2) &= \frac{q_2}{2}[\mathbf{1} - \gamma \csc \theta X - \gamma(\csc \phi - \cos \phi \csc \theta)Y], \\ F(\gamma, 3) &= \frac{q_0}{2}(\mathbf{1} + \gamma \csc \phi Y - \gamma \sec \theta Z), \end{aligned} \quad (48)$$

for $\gamma = \pm$. These operators define a quasiprobability representation that reproduces all of quantum mechanics for the qubit and is non-negative for the states and measurements corresponding to the three bases with Bloch vectors in Eq. (43).

Stabilizer states. We now consider the common special case of both families of three bases. For this special case, the three bases are equivalent to the vertices of a regular octahedron (i.e., single qubit stabilizer states) and so have a larger symmetry group, namely, the octahedral group (equivalent to the single qubit Clifford group). In the appropriate special cases, the distributions $q(\gamma, k)$ in Eq. (39) ($\sin^2 \theta = \frac{2}{3}$) and Eq. (46) ($\phi = \frac{\pi}{2}$ and $\theta = \frac{\pi}{4}$) become uniform and $q(\gamma, k) = \frac{1}{4}$ for all γ, k .

The single qubit Clifford group is generated by the Hadamard gate H and the phase gate P . These transformations supervene on the permutations of ontic states shown in Fig. 7 in the basis $\vec{r}_1 = \vec{x}$, $\vec{r}_2 = \vec{y}$, and $\vec{r}_3 = \vec{z}$. Thus, all Clifford

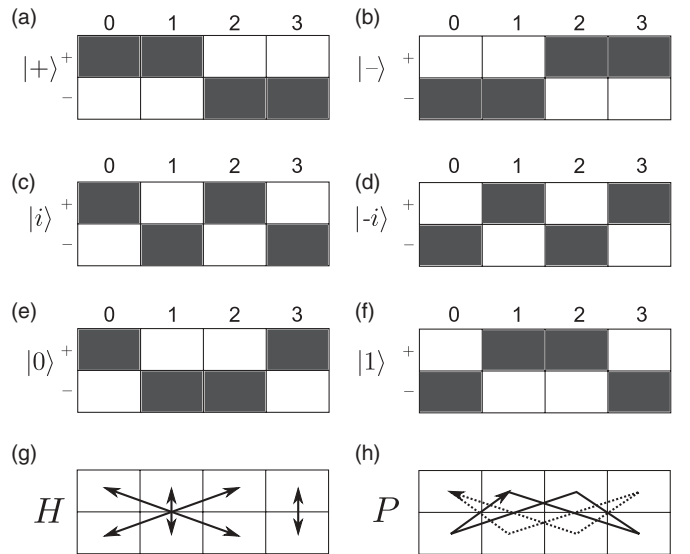


FIG. 7. (a)–(f) Probability distributions $\mu_{\rho}(\gamma, k)$ over eight ontic states $\{(\gamma, k) | \gamma = \pm, k \in \mathbb{Z}_4\}$ for the single qubit stabilizer states. The shading indicates the ontic states that are assigned probability $\frac{1}{4}$ for a given stabilizer state. (g) and (h) Permutations of the ontic states that effect the unitary transformations H and P , respectively.

transformations can be viewed as supervening on permutations of ontic states in these models.

4. General features of quasiprobability representations with three non-negative bases

We now briefly discuss some of the features of these quasiprobability representations, and how they relate to Spekkens' toy theory [17] and the discrete Wigner function [16].

Studying non-negative bases that are closed under a group of nontrivial unitary transformations has shown that Theorems III.1 and III.2 do not completely characterize the possible sets of single qubit bases that can be non-negative in some quasiprobability distribution, as neither theorem excludes the bases in Eq. (36) for $\sin^2 \theta > \frac{8}{9}$ or in Eq. (43) for $1 > |\cos \phi| > \sin \theta$, for which the bases are "close" to coplanar. This suggests that models for sets of three bases only exist when they are sufficiently far from being coplanar. However, as an exception to this, the bases in Eq. (36) still admit a noncontextual model when $\sin \theta \rightarrow 0$ (i.e., when all three bases are close to being degenerate).

Note also that whenever a quasiprobability representation can be defined such that the three bases are non-negative, then any unitary transformations that permute non-negative states can be interpreted as supervening on permutations of the ontic states.

For the remainder of this discussion, we focus on our quasiprobability representation of the stabilizer states. Our quasiprobability representation of stabilizer states is related to the standard definition of the discrete Wigner function, though our representation is defined over eight points rather than four, as follows. Note that in our quasiprobability representation, each non-negative state is uniquely defined by its distribution over the reduced phase space $\{(+, k) | k \in \mathbb{Z}_4\}$. Over this reduced phase space, the operators in either Eq. (41) or Eq. (48) can be written as

$$F(+, k) = c \sum_j \rho_j^{\epsilon_j^k} - d \mathbf{1}, \quad (49)$$

for $k \in \mathbb{Z}_4$, where ϵ_j^k is chosen such that $\rho_j^{\epsilon_j^k}$ assigns nonzero probability to the point $(+, k)$. By ignoring the points outside the reduced phase space and doing suitable renormalizations (i.e., setting $c = d = \frac{1}{2}$), we recover a typical definition of a discrete Wigner function [16]. Note that in this setting, the Clifford transformations permute the non-negative bases, but do not supervene on permutations of ontic states.

Our non-negative quasiprobability representation of preparations and measurements in the stabilizer bases, together with Clifford transformations amongst these states, also recovers those of Spekkens' toy theory [17] when restricted to this reduced phase space. In Spekkens' theory, one cannot define suitable ontic transformations on which all single qubit Clifford transformations supervene. For our theory restricted to the reduced phase space, this is also the case. For example, the permutation of ontic states that effects Π [depicted in Fig. 4(h)] mixes the values of γ and so cannot be defined on this reduced ontological space. Furthermore, antiunitary transformations can supervene on permutations of ontic states, as is the case in the toy theory.

Unlike the toy theory, our model of the stabilizer states over the full phase space of eight points allows all Clifford transformations to supervene on permutations of ontic states. One can ask the additional question of whether all permutations of ontic states that permute the non-negative states are allowed in the theory. If one requires that only unitary transformations supervene on ontic permutations, then the answer is no. However, if one allows antiunitary transformations, provided they permute the non-negative states, then all unitary and antiunitary transformations that permute the stabilizer states supervene on permutations of the ontic states. In particular, the permutation that maps $\gamma \rightarrow -\gamma$ and leaves k unchanged effects a universal NOT gate.

5. Four non-negative bases

The Bloch vectors $\frac{1}{\sqrt{3}}(\pm 1, \pm 1, \pm 1)$ corresponding to the vertices of a cube give a set of four bases, which, as we proved in Theorem III.2, is the only set of four bases that can be non-negative in any quasiprobability representation (up to unitary equivalence).

One of the quasiprobability representations that is non-negative for these states is particularly elegant due to the duality between the octahedron and the cube. The vectors $\pm \vec{d}_k$ in Eq. (40) that define a quasiprobability distribution for the six states corresponding to the vertices of a regular octahedron are, for the special case $\sin^2 \theta = \frac{2}{3}$ and up to a normalization factor, the Bloch vectors in Eq. (28) and their negatives (i.e., the Bloch vectors of the vertices of the cube that is dual to the octahedron).

The reverse also holds. Specifically, we can define a quasiprobability representation that is non-negative for the eight states corresponding to the cube by defining six operators:

$$F(\gamma, k) = \frac{1}{6}(\mathbf{1} + \sqrt{3}\gamma\sigma_k), \quad (50)$$

corresponding to the vertices of the dual octahedron, for $\gamma = \pm 1$ and $k = 1, 2, 3$, where $(\sigma_1, \sigma_2, \sigma_3) = (X, Y, Z)$. Note that all Clifford transformations (i.e., elements of the octahedral group) can be defined as permutations of the points (γ, k) because, by definition, Clifford transformations permute the matrices $\{\gamma\sigma_j | j = 1, 2, 3, \gamma = \pm 1\}$.

This quasiprobability representation is defined on an ontic state space consisting of six ontic states. The supports of the non-negative states and the permutations of ontic states corresponding to two generators of the single qubit Clifford group (the Hadamard H and phase gate P) are illustrated in Fig. 8.

Note that the bases in Eq. (36) with $\sin^2 \theta = \frac{8}{9}$ (i.e., $\cos \theta = \pm \frac{1}{3}$) are three of the four bases corresponding to the vertices of the cube, so in this limiting case, a fourth non-negative basis (with Bloch vectors parallel to the z axis) can be added. For the model in Eq. (39) in the limiting case $\sin^2 \theta = \frac{8}{9}$, $q_0 = 0$ and so only six points are assigned nonzero probability. If the points $(\pm, 0)$, to which any non-negative state assigns zero probability, are ignored, then the distributions in Figs. 4(a)–4(f) are identical (up to a relabeling of the states) to those in Figs. 8(c)–8(h).

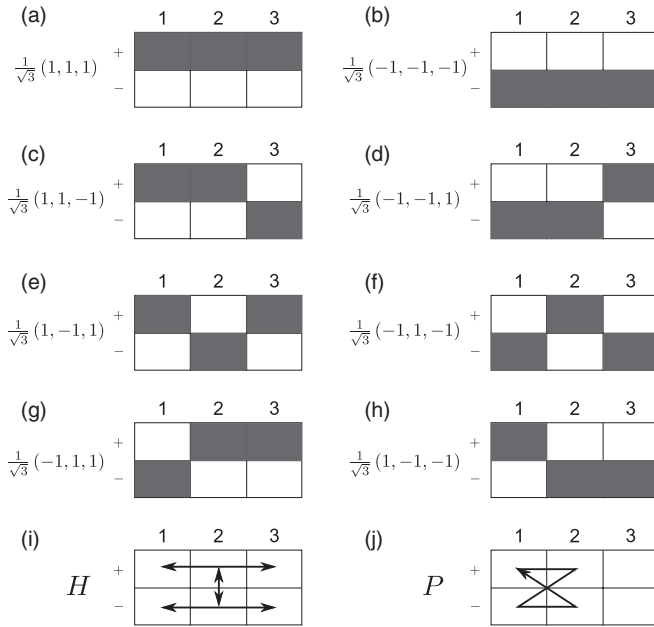


FIG. 8. (a)–(h) Probability distributions over six points $\{(\gamma, k) | \gamma = \pm, k = 1, \dots, 3\}$, for the eight quantum states with Bloch vectors $\frac{1}{\sqrt{3}}(\pm 1, \pm 1, \pm 1)$ corresponding to the vertices of the cube, where a state assigns probability $\frac{1}{3}$ to a point if it is shaded and 0 otherwise. (i) and (j) Permutation of ontic states that effect the transformations (i) H and (j) P .

6. Example of bases that do not admit a preparation noncontextual model: Icosahedron

We have provided an exhaustive list of sets of bases with nontrivial point groups that are non-negative in some quasiprobability distribution and noticed that the single qubit Clifford group (i.e., the symmetry group of a regular octahedron) plays a special role.

The octahedral group is one of two groups that permute pairs of antipodal points that are the vertices of a Platonic solid, the other being the icosahedral group. While Theorems III.1 and III.2 show that the pairs of antipodal vertices corresponding to either the icosahedron (or its dual, the dodecahedron) cannot all correspond to non-negative bases, we provide an explicit proof of this fact, which is illustrative of the general case.

The idea behind the proof is to use the Bloch vectors corresponding to the vertices of an icosahedron to obtain two decompositions of a mixed state ρ in terms of two different sets of states and then use Lemmas II.2 and II.3 to show that in any quasiprobability distribution, some of the quasiprobabilities must be negative.

The 12 vertices of an icosahedron are given by the six vectors:

$$\begin{aligned}\vec{r}_{0,\alpha} &= \frac{1}{\sqrt{1+\varphi^2}}(1, \alpha\varphi, 0), \\ \vec{r}_{1,\alpha} &= \frac{1}{\sqrt{1+\varphi^2}}(0, 1, \alpha\varphi), \\ \vec{r}_{2,\alpha} &= \frac{1}{\sqrt{1+\varphi^2}}(\alpha\varphi, 0, 1),\end{aligned}\tag{51}$$

for $\alpha = \pm$ and their negatives, where $\varphi = (1 + \sqrt{5})/2$ is the golden ratio. These vertices satisfy the relation,

$$(1 - 2a)\vec{r}_{1,+} = a\vec{r}_{0,+} - a\vec{r}_{0,-} - (1 - 2a)\vec{r}_{1,-}, \quad (52)$$

where $a = (2 + \varphi)^{-1}$. Therefore we have

$$\rho = a\rho_{0,+}^+ + a\rho_{0,-}^- + (1-2a)\rho_{1,-}^- = (1-a)\rho_{1,+}^+ + a\rho_{1,+}^-, \quad (53)$$

where $\rho_{j,\alpha}^\mu = \frac{1}{2}(\mathbf{1} + \mu \vec{r}_{j,\alpha})$. For a quasiprobability distribution, this implies

$$\begin{aligned}\mu_\rho(\lambda) &= a\mu_{\rho_{0,+}^+}(\lambda) + a\mu_{\rho_{0,-}^-}(\lambda) + (1-2a)\mu_{\rho_{1,-}^-}(\lambda) \\ &= (1-a)\mu_{\rho_{1,+}^+}(\lambda) + a\mu_{\rho_{1,+}^-}(\lambda),\end{aligned}\quad (54)$$

for all $\lambda \in \Lambda$. Assume that all the bases corresponding to the vertices of an icosahedron are non-negative. By Lemma II.3, there exists a $\lambda' \in \mathcal{S}(\rho_{1,+}) \cup \mathcal{S}(\rho_{1,-})$. By Lemma II.2, for this value of λ' , the first line is at least $(1 - 2a)q(\lambda') > 0$, while the second is exactly $aq(\lambda') > 0$. As $(1 - 2a) > a$, this is a contradiction. Therefore not all the bases can be non-negative.

IV. QUASIPROBABILITY REPRESENTATIONS OF QUDITS

For qubits, we have shown that no quasiprobability distribution can be non-negative for more than four orthonormal bases. Moreover, up to unitary equivalence, the only set of four bases that can be non-negative corresponds to the vertices of a cube. We now turn our attention to the general case and establish restrictions on the relation between non-negative bases in an arbitrary quasiprobability representation of $\mathcal{H}_d$. We will begin by constructing a mixed state ρ that plays an analogous role to the states in Eqs. (18) and (22) for qubits. We then generalize Theorems III.1 and III.2 to higher dimensions, which establishes restrictions on the relation between any set of non-negative bases. We conclude this section by obtaining an upper bound of 2^{d^2} on the number of states that are elements of non-negative bases in an arbitrary quasiprobability representation.

For higher dimensions, the relationship between states and orthonormal bases is not as straightforward as for qubits, where a state can be uniquely extended to an orthonormal basis. Moreover, for arbitrary d , there is no simple map from $\mathcal{H}_d$ to an intuitive geometric space like the Bloch sphere. Therefore we rely on abstract algebraic tools rather than geometric ones. We note that this section is quite technical relative to previous sections of this paper, and can be skipped upon first reading.

A. Linear algebra and convex geometry

We now introduce some tools from linear algebra. These tools enable us to deduce the existence of mixed states with multiple decompositions in terms of a set of non-negative bases if the bases are related in a particular way. Such states play an analogous role to the states in Eqs. (18) and (22) for qubits.

Two density matrices, $\rho, \sigma \in \mathcal{B}(\mathcal{H}_d)$, are *linearly independent* (over $\mathbb{R}$) if

$$a\rho + b\sigma = 0 \Rightarrow a = b = 0, \quad (55)$$

and are *linearly dependent* otherwise. Note that we are considering linear independence over the operator space, rather

than the space of pure states. We introduce the term *disparate* to denote a set $\mathcal{A} := \{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_N\}$ of N bases of $\mathcal{H}_d$ such that any set $\mathcal{A}'$ obtained from $\mathcal{A}$ by removing one element from each basis and including the identity is a set of linearly independent operators. That is, for all $\vec{f} \in \mathbb{Z}_d^N$ (where f_j labels the element of the j th basis that is not in $\mathcal{A}'$),

$$\sum_{\alpha, j | j \neq f_\alpha} p_{\alpha, j} \mathbf{P}(\Psi_j^\alpha) = \frac{b}{d} \mathbf{1} \Rightarrow p_{\alpha, j} = b = 0 \forall (\alpha, j). \quad (56)$$

The maximum number of disparate bases of $\mathcal{H}_d$ is $d + 1$, as each basis contributes $d - 1$ linearly independent density operators and including the identity (or the remaining element of any of the bases) gives d^2 linearly independent density operators.

We denote the *span* of a set of density matrices $\mathcal{A} = \{\rho_j\}$ over $\mathbb{R}$ by

$$\mathcal{T}(\mathcal{A}) = \left\{ \sum_j c_j \rho_j | c_j \in \mathbb{R} \right\}. \quad (57)$$

The convex hull of a set of density matrices $\mathcal{A}$ is the set,

$$\mathcal{C}(\mathcal{A}) = \left\{ \sum_j p_j \rho_j | p_j \geq 0, \sum_j p_j = 1 \right\}. \quad (58)$$

Denote by $\partial\mathcal{C}(\mathcal{A})$ the surface of $\mathcal{C}(\mathcal{A})$ with respect to $\mathcal{T}(\mathcal{A})$. Then, as we now show, if $\mathcal{A}$ is a set of disparate bases, any expansion of any $\rho \in \partial\mathcal{C}(\mathcal{A})$ as a convex combination of the elements of $\mathcal{A}$ must assign a zero coefficient to at least one element of each of the bases in $\mathcal{A}$.

Lemma IV.1. Let $\mathcal{A} = \{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_N\}$ be a set of N disparate bases. Then for all $\rho \in \partial\mathcal{C}(\mathcal{A})$, there exist $\vec{f} \in \mathbb{Z}_d^N$ and $p_{\alpha, j} \geq 0$ such that

$$\rho = \sum_{\alpha, j | j \neq f_\alpha} p_{\alpha, j} \mathbf{P}(\Psi_j^\alpha). \quad (59)$$

Proof. To prove this lemma, we prove the contrapositive. Let $\rho \in \mathcal{C}(\mathcal{A})$ be such that all the coefficients $p_{\alpha, j}$ in some decomposition,

$$\rho = \sum_{\alpha, j} p_{\alpha, j} \mathbf{P}(\Psi_j^\alpha), \quad (60)$$

are nonzero. We can define f_α such that for all $\alpha \in \mathbb{Z}_N$,

$$p_{\alpha, f_\alpha} = \min_{j \in \mathbb{Z}_d} p_{\alpha, j}. \quad (61)$$

Then we can rewrite ρ as

$$\rho = \sum_{\alpha, j | j \neq f_\alpha} (p_{\alpha, j} - p_{\alpha, f_\alpha}) \mathbf{P}(\Psi_j^\alpha) + \mathbf{1} \sum_{\alpha} p_{\alpha, f_\alpha}. \quad (62)$$

Therefore we can set all the p_{α, f_α} to be equal and strictly positive without loss of generality.

We now prove that ρ is in the interior of $\mathcal{C}(\mathcal{A})$ with respect to $\mathcal{T}(\mathcal{A})$. Let $\rho' \in \mathcal{T}(\mathcal{A})$. Then, by definition, there exists $\beta_{\alpha, j} \in \mathbb{R}$ with $\sum_{\alpha, j} \beta_{\alpha, j} = 1$ such that

$$\rho'' = \sum_{\alpha, j} \beta_{\alpha, j} \mathbf{P}(\Psi_j^\alpha). \quad (63)$$

If we set

$$\delta = \min_{\rho' \in \mathcal{T}(\mathcal{A})} \min_{\alpha, j} \beta_{\alpha, j}, \quad (64)$$

then for all $\rho' \in \mathcal{T}(\mathcal{A})$, the coefficients $\epsilon \beta_{\alpha, j} + (1 - \epsilon) p_{\alpha, j}$ in the decomposition,

$$\epsilon \rho'' + (1 - \epsilon) \rho = \sum_{\alpha, j} [\epsilon \beta_{\alpha, j} + (1 - \epsilon) p_{\alpha, j}] \mathbf{P}(\Psi_j^\alpha), \quad (65)$$

are non-negative whenever

$$\beta_{\alpha, j} \geq 0 \quad \text{or} \quad \epsilon \leq \frac{p_{\alpha, j}}{p_{\alpha, j} - \beta_{\alpha, j}}. \quad (66)$$

Therefore if we set

$$\epsilon \leq \min_{i, j} \frac{p_{i, j}}{p_{i, j} - \delta}, \quad (67)$$

then $\epsilon \rho + (1 - \epsilon) \rho' \in \mathcal{C}(\mathcal{A})$ for all $\rho' \in \mathcal{T}(\mathcal{A})$. ■

Lemma IV.1 shows that any point ρ on the surface of the convex hull of a set of disparate bases can be written as

$$\rho = \sum_{\alpha, j | j \neq f_\alpha} p_{\alpha, j} \mathbf{P}(\Psi_j^\alpha) \quad (68)$$

for some $\vec{f} \in \mathbb{Z}_d^N$ and $p_{\alpha, j} \geq 0$. Therefore for all $\Phi \in \mathcal{T}(\mathcal{A})$, the lines defined by

$$\epsilon \mathbf{P}(\Phi) + \frac{1 - \epsilon}{d} \mathbf{1} \quad (69)$$

for $\epsilon \in [0, 1]$ must intersect $\partial\mathcal{C}(\mathcal{A})$ as $\frac{1}{d} \mathbf{1}$ is in the interior of $\mathcal{C}(\mathcal{A})$ and Φ is either in $\partial\mathcal{C}(\mathcal{A})$ or not in $\mathcal{C}(\mathcal{A})$. Therefore for all $\Phi \in \mathcal{T}(\mathcal{A})$, there exists $\epsilon > 0$, $\vec{f} \in \mathbb{Z}_d^N$ and $p_{\alpha, j} \geq 0$ such that

$$\rho := \epsilon \mathbf{P}(\Phi) + \frac{1 - \epsilon}{d} \mathbf{1} = \sum_{\alpha, j | j \neq f_\alpha} p_{\alpha, j} \mathbf{P}(\Psi_j^\alpha). \quad (70)$$

The state defined in Eq. (70) can be decomposed as a convex combination of different bases in multiple ways, which allows us to use it to restrict the relation between non-negative bases in an analogous way to the states Eqs. (18) and (22) for qubits.

B. Relation between non-negative bases for qudits

We now use the state ρ defined in Eq. (70) to generalize Theorems III.1 and III.2 to qudits.

Theorem IV.2. In an arbitrary quasiprobability representation, any set of three mutually nonorthogonal non-negative bases must be disparate.

Proof. Let μ be a quasiprobability distribution that is non-negative for at least three bases $\{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_3\}$ that are not disparate. As the three bases are not disparate, then one of the states $\{\mathbf{P}(\Psi_j^2) | j \in \mathbb{Z}_d\}$ must be in the span of the other two bases. Without loss of generality, we can relabel the elements of $\{\mathbf{P}(\Psi_j^2) | j \in \mathbb{Z}_d\}$ such that $\mathbf{P}(\Psi_0^2) \in \mathcal{T}(\mathcal{A})$, where $\mathcal{A} = \{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_2\}$ is the set of the elements of the other two bases.

Therefore we can use the decomposition in Eq. (70) and relabel the bases such that $\vec{f} = (0, 0)$ to obtain

$$\epsilon \mathbf{P}(\Psi_0^2) + \frac{1 - \epsilon}{d} \mathbf{1} = \sum_{\alpha=0}^1 \sum_{j=1}^d p_{0, j} \mathbf{P}(\Psi_j^\alpha) \quad (71)$$

for some $\epsilon \in (0, 1)$. As μ is convex linear, we have

$$\epsilon \mu_{\Psi_0^2}(\lambda) + (1 - \epsilon) \mu_{\frac{1}{d}}(\lambda) = \sum_{\alpha=0}^1 \sum_{j=1}^d p_{\alpha,j} \mu_{\Psi_j^\alpha}(\lambda) \quad (72)$$

for all $\lambda \in \Lambda$. As Ψ_0^0 and Ψ_0^1 are not orthogonal, there exists $\lambda' \in \mathcal{S}(\Psi_0^0) \cap \mathcal{S}(\Psi_0^1)$ by Lemma II.3. At this value of λ' , the right-hand side of Eq. (72) is 0 while the left-hand side is at least $(1 - \epsilon)q(\lambda') > 0$, yielding a contradiction. ■

Theorem IV.3. In an arbitrary quasiprobability representation, any set of four mutually nonorthogonal non-negative bases $\{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_4\}$ must either

- (1) be disparate; or
- (2) satisfy the relationship (up to a relabeling of basis states),

$$\epsilon \mathbf{P}(\Psi_0^3) + \frac{1 - \epsilon}{d} \mathbf{1} = \sum_{\alpha=0}^2 \sum_{j=1}^d p_{\alpha,j} \mathbf{P}(\Psi_j^\alpha) \quad (73)$$

for $\epsilon = \frac{1}{d+1}$.

Proof. Let μ be a quasiprobability distribution that is non-negative for at least four bases $\{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_4\}$. If the four bases are not disparate, then, relabeling the bases as necessary, we have $\mathbf{P}(\Psi_0^3) \in \mathcal{T}(\mathcal{A})$, where $\mathcal{A} = \{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_3\}$, that is, the set of the elements of the first three bases, which must be disparate by Theorem IV.2.

Therefore we can use the decomposition in Eq. (70) and relabel the bases such that $\tilde{f} = (0, 0, 0)$ to obtain

$$\epsilon \mathbf{P}(\Psi_0^3) + \frac{1 - \epsilon}{d} \mathbf{1} = \sum_{\alpha=0}^2 \sum_{j=1}^d p_{0,j} \mathbf{P}(\Psi_j^\alpha). \quad (74)$$

As μ is convex linear, we have

$$\epsilon \mu_{\Psi_0^3}(\lambda) + (1 - \epsilon) \mu_{\frac{1}{d}}(\lambda) = \sum_{\alpha=0}^2 \sum_{j=1}^d p_{\alpha,j} \mu_{\Psi_j^\alpha}(\lambda) \quad (75)$$

for all $\lambda \in \Lambda$.

As Ψ_1^3 and Ψ_j^α are not orthogonal for any $\alpha \neq 3$, there exists $\lambda' \in \mathcal{S}(\Psi_1^3) \cap \mathcal{S}(\Psi_j^\alpha)$ by Lemma II.3. At this value of λ' , the left-hand side of Eq. (75) is $\frac{1-\epsilon}{d}q(\lambda') > 0$ while the right-hand side is at least $p_{\alpha,j}q(\lambda') > 0$, so $p_{\alpha,j} \leq \frac{1-\epsilon}{d}$ for all α, j .

As Ψ_0^3 and Ψ_0^α are not orthogonal for any $\alpha \neq 3$, there exists $\lambda' \in \mathcal{S}(\Psi_0^3) \cap \mathcal{S}(\Psi_0^\alpha)$ by Lemma II.3. At this value of λ' , the left-hand side of Eq. (75) is $(\frac{1-\epsilon}{d} + \epsilon)q(\lambda') > 0$ while the right-hand side is at most $\frac{1-\epsilon}{d}2q(\lambda') > 0$. Therefore $\epsilon \leq \frac{1-\epsilon}{d}$.

We now want to show that for each α there must exist a k_α such that $p_{\alpha,k_\alpha} = \frac{1-\epsilon}{d}$. To do this, let $\{\alpha, \beta, \gamma\}$ be a permutation of $\{0, 1, 2\}$. Then as Ψ_0^β and Ψ_0^γ are not orthogonal, there exists $\lambda_\alpha \in \mathcal{S}(\Psi_0^\beta) \cap \mathcal{S}(\Psi_0^\gamma)$. At λ_α , the right-hand side of Eq. (75) is $p_{\alpha,k_\alpha}q(\lambda_\alpha) \leq \frac{1-\epsilon}{d}q(\lambda_\alpha)$ for some value of k_α . The left-hand side is at least $\frac{1-\epsilon}{d}q(\lambda_\alpha)$, so we have that there exists k_α such that $p_{\alpha,k_\alpha} = \frac{1-\epsilon}{d}$. By considering all permutations of $\{0, 1, 2\}$, we see that this holds for all α .

We can now use this same approach to show that $\epsilon = \frac{1}{d+1}$. Let $\{\alpha, \beta, \gamma\}$ be a permutation of $\{0, 1, 2\}$. Then as $\Psi_{k_\beta}^\beta$ and $\Psi_{k_\gamma}^\gamma$ are not orthogonal, there exists $\lambda'_\alpha \in \mathcal{S}(\Psi_{k_\beta}^\beta) \cap \mathcal{S}(\Psi_{k_\gamma}^\gamma)$. At λ'_α ,

the right-hand side of Eq. (75) is at least $\frac{1-\epsilon}{d}2q(\lambda'_\alpha)$. The only way the right-hand side can be equal to the left-hand side is if $\lambda'_\alpha \in \mathcal{S}(\Psi_0^3)$ and $\epsilon = \frac{1-\epsilon}{d}$, that is, $\epsilon = \frac{1}{d+1}$. ■

We now show that in any quasiprobability distribution, any number N of non-negative bases must still satisfy a symmetry constraint. However, as N increases, this constraint becomes less restrictive.

Theorem IV.4. In an arbitrary quasiprobability representation, any set of $N > 4$ mutually nonorthogonal non-negative bases $\{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_N\}$ must either

- (1) be disparate; or
- (2) satisfy the relationship (up to a relabeling of basis states),

$$\epsilon \mathbf{P}(\Psi_0^3) + \frac{1 - \epsilon}{d} \mathbf{1} = \sum_{\alpha=0}^{N-2} \sum_{j=1}^d p_{\alpha,j} \mathbf{P}(\Psi_j^\alpha), \quad (76)$$

for some $\epsilon \leq \frac{N-3}{N-3+d}$.

Proof. Let μ be a quasiprobability distribution that is non-negative for at least N bases $\{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_N\}$ that are not disparate. Then, relabeling the bases as necessary, we have $\mathbf{P}(\Psi_0^M) \in \mathcal{T}(\mathcal{A})$, where $M = N - 1$ and $\mathcal{A} = \{\mathbf{P}(\Psi_j^\alpha) | j \in \mathbb{Z}_d, \alpha \in \mathbb{Z}_M\}$, that is, the set of the elements of the first M bases.

Therefore we can use the decomposition in Eq. (70) and relabel the bases such that $\tilde{f}$ is the zero vector to obtain

$$\epsilon \mathbf{P}(\Psi_0^M) + \frac{1 - \epsilon}{d} \mathbf{1} = \sum_{\alpha \in \mathbb{Z}_M} \sum_{j=1}^d p_{0,j} \mathbf{P}(\Psi_j^\alpha). \quad (77)$$

As μ is convex linear, we have

$$\epsilon \mu_{\Psi_0^M}(\lambda) + (1 - \epsilon) \mu_{\frac{1}{d}}(\lambda) = \sum_{\alpha \in \mathbb{Z}_M} \sum_{j=1}^d p_{\alpha,j} \mu_{\Psi_j^\alpha}(\lambda) \quad (78)$$

for all $\lambda \in \Lambda$. As Ψ_1^M and Ψ_k^α are not orthogonal for all $\alpha \in \mathbb{Z}_M$ and $k \in \mathbb{Z}_d$, there exists $\lambda' \in \mathcal{S}(\Psi_1^M) \cap \mathcal{S}(\Psi_k^\alpha)$ by Lemma II.3. At this value of λ' , the left-hand side of Eq. (78) is $\frac{1-\epsilon}{d}q(\lambda') > 0$ while the right-hand side is at least $p_{\alpha,k}q(\lambda')$. Therefore $p_{\alpha,k} \leq \frac{1-\epsilon}{d}$ for all $\alpha \in \mathbb{Z}_M$ and $k \in \mathbb{Z}_d$.

As Ψ_0^M and Ψ_0^α are not orthogonal for all $\alpha \in \mathbb{Z}_M$, there exists $\lambda'' \in \mathcal{S}(\Psi_0^M) \cap \mathcal{S}(\Psi_0^\alpha)$ by Lemma II.3. At this value of λ'' , the left-hand side of Eq. (78) is $(\frac{1-\epsilon}{d} + \epsilon)q(\lambda'') > 0$ while the right-hand side is at most $(M - 1)\frac{1-\epsilon}{d}q(\lambda'')$. ■

C. Upper bound on the number of non-negative bases for qudits

Theorems IV.2–IV.4 provide strong constraints on the relation between any set of non-negative bases in a quasiprobability representation. However, it is unclear how to use these theorems to obtain an upper bound on the number of non-negative bases in a quasiprobability representation. In order to obtain an upper bound (which will not be tight), we change tack and exploit the fact that for all $\lambda \in \Lambda$ there exists an operator $F(\lambda)$ acting on $\mathcal{H}_d$ such that

$$\mu_\rho(\lambda) = \text{Tr}[\rho F(\lambda)] \quad (79)$$

for all $\rho \in \mathcal{B}(\mathcal{H}_d)$. This will enable us to show that there are no more than 2^{d^2} states that are elements of a non-negative basis in any quasiprobability distribution, without requiring

that the bases are mutually nonorthogonal. To obtain this bound, we note that any density matrix can be written as a linear combination of the $F(\lambda)$ [14], so the $F(\lambda)$ must be a basis for the space of operators acting on $\mathcal{H}_d$.

Theorem IV.5. For any quasiprobability representation of $\mathcal{H}_d$, there are no more than 2^{d^2} states that are elements of non-negative bases.

Proof. Let $\{\Pi_\alpha | \alpha \in \mathbb{Z}_{d^2}\}$ be a trace-orthonormal basis of $\mathcal{B}(\mathcal{H}_d)$ and let $\{F(\lambda_\beta) | \beta \in \mathbb{Z}_{d^2}\}$ be a set of d^2 linearly independent operators, which must exist as the set $\{F(\lambda)\}$ is a basis for the space of operators acting on $\mathcal{H}_d$. Then, for all $\rho \in \mathcal{B}(\mathcal{H}_d)$ and $\beta \in \mathbb{Z}_{d^2}$, we can write

$$\begin{aligned}\rho &= \sum_{\alpha \in \mathbb{Z}_{d^2}} g_\alpha(\rho) \Pi_\alpha, \\ F(\lambda_\beta) &= \sum_{\alpha \in \mathbb{Z}_{d^2}} f_{\alpha,\beta} \Pi_\alpha,\end{aligned}\tag{80}$$

where two states $\rho, \rho' \in \mathcal{B}(\mathcal{H}_d)$ have the same coefficients g_α for all α if and only if $\rho = \rho'$. Therefore we can rewrite Eq. (79) as

$$\mu_\rho(\lambda_\beta) = \sum_{\alpha} f_{\alpha,\beta} g_\alpha(\rho).\tag{81}$$

As the $\{F(\lambda_\beta) | \beta \in \mathbb{Z}_{d^2}\}$ are linearly independent, $f_{\alpha,\beta}$ must be invertible. Therefore for any set of values $\{\mu(\lambda_\beta) | \beta \in \mathbb{Z}_{d^2}\}$, there can be at most one state ρ such that $\mu_\rho(\lambda_\beta) = \mu(\lambda_\beta)$ for all $\beta \in \mathbb{Z}_{d^2}$.

From Lemma II.2, any state that is an element of a non-negative basis can only assign one of two values to any point $\lambda \in \Lambda$, namely, 0 or $q(\lambda)$. Therefore there are only 2^{d^2} possible sets of values of μ over $\{\lambda_\beta | \beta \in \mathbb{Z}_{d^2}\}$ that correspond to elements of non-negative bases. ■

The bound on the number of states that are elements of a non-negative basis in Theorem IV.5 is not tight. For example, not all vectors g correspond to a valid density operator. In particular, no quantum state can have $g_\alpha = 0$ for all α . Furthermore, if an element of a non-negative basis assigned nonzero probability to more than $d^2 - d + 1$ of the points $\{\lambda_\beta\}$, then as the elements of a non-negative basis have disjoint support, at least one of the other elements of a non-negative basis would have to assign zero probability to all of the points $\{\lambda_\beta\}$ and so would have $g_\alpha = 0$ for all α . Therefore all non-negative states assign nonzero probability to between 1 and $d^2 - d + 1$ of the points $\{\lambda_\beta\}$.

However, even accounting for this does not substantially decrease the upper bound. Furthermore, for qubits, we proved in Theorem III.2 that no more than eight states can be elements of a non-negative basis. The upper bound from Theorem IV.5 is 16 states, and even excluding the combinations of μ discussed above only reduces the upper bound to 14 states.

V. DISCUSSION AND CONCLUSION

We have shown that for any quasiprobability representation of a qubit, any three non-negative bases cannot be coplanar in the Bloch sphere (i.e., they must be disparate). Moreover, if there are four non-negative bases, then they must correspond to the vertices of a cube circumscribed by the Bloch sphere.

We provided an exhaustive list of all “classical” subtheories of a qubit that include states, measurements, and nontrivial transformations. These cases revealed several interesting features. Both families of three bases that are permuted by a nontrivial group of unitary transformations can only be defined in a noncontextual model when they are sufficiently “far” from being coplanar. However, there is an exception to this behavior, as the bases in Eq. (36) are non-negative in some quasiprobability representation even when $\theta \rightarrow 0$ (i.e., the three bases are almost degenerate). We have also found that whenever a subtheory of qubit states and measurements are non-negative in some quasiprobability representation, there exists a quasiprobability representation in which all unitary transformations that permute non-negative states correspond to a permutation of the ontic states.

While we have primarily focused on the qubit case, we have also shown that the results for qubits directly generalize in that any three mutually nonorthogonal nonnegative bases in a quasiprobability distribution must be disparate and any four or more mutually nonorthogonal non-negative bases must either be disparate or satisfy a symmetry constraint. In this sense, quantum states and measurements with a small amount of complementarity can be quite difficult to model in a classical theory. In addition, we have obtained an upper bound of 2^{d^2} on the number of states that are elements of a non-negative basis.

We conclude with some discussion of the implications of our results for quantum computation, and some future research directions. While our results have been presented in the context of single qudits, they are equally applicable to multiple qudit systems. Our upper bound on the number of non-negative basis states of a qudit, although quite loose, suggests that universal quantum computation leads to negativity in *any* quasiprobability distribution. This matches the intuition obtained from the specific case of the single qudit discrete Wigner function [23].

While our higher-dimensional results can be applied to quasiprobability representations of multiple qubit systems, it is not clear how this approach accords with classical simulations of quantum systems. In particular, multiqubit stabilizers can be efficiently simulated classically [24] and yet do not correspond to a set of non-negative bases in any quasiprobability representation. To see this, note that stabilizer states and X and Y measurements (i.e., in bases corresponding to stabilizer states) can lead to violations of a Bell inequality [25] and so cannot admit a locally causal model.

A natural way of generalizing a quasiprobability representation for a single qubit to one for multiple qubits is to take tensor products of the operators $\{F(\lambda)\}$ and $\{G(\lambda)\}$ that define the single qubit quasiprobability representation via Eqs. (1) and (3). By construction, such a quasiprobability representation will be non-negative for all tensor products of the single qubit states with non-negative distributions, but may also be non-negative for other bases that include entangled states. For the entangled states to be accessible in a classical subtheory, there must be some unitary that permutes non-negative bases and maps a non-negative product basis to a non-negative basis that contains an entangled state. Such a unitary can only be viewed as supervening on a permutation of ontic

states (which could always be done for a single qubit) if it leaves the set of tensor products of the $\{F(\lambda)\}$ invariant under conjugation. Unfortunately, it is unclear whether such unitaries exist for any set of operators that define our single qubit quasiprobability representations, although based on the results of Ref. [26] we have some evidence to suggest that they do not.

ACKNOWLEDGMENTS

We acknowledge helpful discussions with Chris Ferrie, Chris Fuchs, Markus Mueller, and Rob Spekkens, and financial support from the Australian Research Council via the Centre of Excellence in Engineered Quantum Systems (EQuS), Project No. CE11001013.

-
- [1] E. Wigner, in *Perspectives in Quantum Theory*, edited by W. Yourgrau and A. Van der Merwe (MIT Press, Cambridge, 1971), p. 2536.
 - [2] C. Ferrie, *Rep. Prog. Phys.* **74**, 116001 (2011).
 - [3] A. Kenfack and K. Życzkowski, *J. Opt. B* **6**, 396 (2004).
 - [4] R. L. Hudson, *Rep. Math. Phys.* **6**, 249 (1974).
 - [5] F. Soto and P. Claverie, *J. Math. Phys.* **24**, 97 (1983).
 - [6] D. Gross, *J. Math. Phys.* **47**, 122107 (2006).
 - [7] D. Gross, *Appl. Phys. B* **86**, 367 (2007).
 - [8] R. W. Spekkens, *Phys. Rev. Lett.* **101**, 020401 (2008).
 - [9] V. Veitch, C. Ferrie, and J. Emerson, [arXiv:1201.1256](#).
 - [10] C. Ferrie and J. Emerson, *J. Phys. A: Math. Gen.* **41**, 352001 (2008).
 - [11] R. W. Spekkens, *Phys. Rev. A* **71**, 052108 (2005).
 - [12] P. W. Shor and S. P. Jordan, *Quantum Inf. Comp.* **8**, 681 (2008).
 - [13] E. G. Beltrametti and S. Bugajski, *J. Phys. A: Math. Gen.* **28**, 3329 (1995).
 - [14] A. Montina, *Phys. Rev. Lett.* **97**, 180401 (2006).
 - [15] S. D. Bartlett, T. Rudolph, and R. W. Spekkens, [arXiv:1111.5057](#).
 - [16] K. S. Gibbons, M. J. Hoffman, and W. K. Wootters, *Phys. Rev. A* **70**, 062101 (2004).
 - [17] R. W. Spekkens, *Phys. Rev. A* **75**, 032110 (2007).
 - [18] C. Weedbrook, S. Pirandola, R. Garcia-Patron, N. J. Cerf, T. C. Ralph, J. H. Shapiro, and S. Lloyd, *Rev. Mod. Phys.* **84**, 621 (2012).
 - [19] C. H. Bennett, G. Brassard, C. Crépeau, R. Jozsa, A. Peres, and W. K. Wootters, *Phys. Rev. Lett.* **70**, 1895 (1993).
 - [20] C. H. Bennett and S. J. Wiesner, *Phys. Rev. Lett.* **69**, 2881 (1992).
 - [21] C. Ferrie and J. Emerson, *New J. Phys.* **11**, 063040 (2009).
 - [22] M. Tinkham, *Group Theory and Quantum Mechanics* (Dover Publications, New York, 2003).
 - [23] C. Cormick, E. F. Galvão, D. Gottesman, J. P. Paz, and A. O. Pittenger, *Phys. Rev. A* **73**, 012301 (2006).
 - [24] S. Aaronson and D. Gottesman, *Phys. Rev. A* **70**, 052328 (2004).
 - [25] N. D. Mermin, *Phys. Rev. Lett.* **65**, 1838 (1990).
 - [26] S. Clark, R. Jozsa, and N. Linden, [arXiv:quant-ph/0701103](#).