Nonlocality as a benchmark for universal quantum computation in Ising anyon topological quantum computers

Mark Howard[*]

Department of Mathematical Physics, National University of Ireland, Maynooth, Ireland

Jiri Vala[†]

Department of Mathematical Physics, National University of Ireland, Maynooth, Ireland and
Dublin Institute for Advanced Studies, School of Theoretical Physics, 10 Burlington Road, Dublin, Ireland

An obstacle affecting any proposal for a topological quantum computer based on Ising anyons is that quasiparticle braiding can only implement a finite (nonuniversal) set of quantum operations. The computational power of this restricted set of operations (often called stabilizer operations) has been studied in quantum information theory, and it is known that no quantum-computational advantage can be obtained without the help of an additional nonstabilizer operation. Similarly, a bipartite two-qubit system based on Ising anyons can not exhibit nonlocality (in the sense of violating a Bell inequality) when only topologically protected stabilizer operations are performed. To produce correlations that can not be described by a local hidden variable model again requires the use of a nonstabilizer operation. Using geometric techniques, we relate the sets of operations that enable universal quantum computing (UQC) with those that enable violation of a Bell inequality. Motivated by the fact that nonstabilizer operations are expected to be highly imperfect, our aim is to provide a benchmark for identifying UQC-enabling operations that is both experimentally practical and conceptually simple. We show that any (noisy) single-qubit nonstabilizer operation that, together with perfect stabilizer operations, enables violation of the simplest two-qubit Bell inequality, can also be used to enable UQC. This benchmarking requires finding the expectation values of two distinct Pauli measurements on each qubit of a bipartite system.

PACS number(s): 03.67.Lx, 03.67.Mn, 03.65.Yz

I. INTRODUCTION AND DEFINITIONS

A topological quantum computer (TQC) would allow quantum information to be stored in quasiparticles and manipulated by quasiparticle braiding, in a way that is inherently robust against local perturbations (noise). Many of the most promising candidates for experimental implementation of a useful TQC are those that support Ising anyons, e.g., fractional quantum Hall effect [1,2], Majorana wires [3], $p + ip$ superconductors [4], and the Kitaev honeycomb model (in the presence of a magnetic field) [5]. While there are large differences in the underlying physics of all these systems, the non-Abelian braiding statistics are such that they all possess the same computational power (i.e., they enable the same types of quantum gates and measurements when viewed as a quantum information processing device).

The model of computation (sometimes known as the Clifford computer model [6]) that we assume in this work, and the one that is relevant to Ising anyon TQCs, is one in which (i) all single-qubit Clifford gates (discussed in Sec. I A), (ii) measurements in the computational basis, and (iii) the controlled-NOT gate, are all implemented in a topologically protected (i.e., effectively perfect) manner. Collectively, we refer to this set of gates and measurements as stabilizer operations. Some proposals for TQCs based on Ising anyons use "$|a_8\rangle$ distillation" to achieve the controlled-NOT [7], while others achieve it using quasiparticle braiding [8] or nondemolitional measurements of the collective charge of

four anyons [9]. It is the latter, topologically protected, implementations of the controlled-NOT that we have in mind here.

A result in quantum information theory, i.e., the Gottesman-Knill theorem (see, e.g., [10] for a discussion), says that the set of stabilizer operations is insufficient for achieving universal quantum computation. If a (noiseless) single-qubit unitary gate from outside the Clifford group was implementable, then we would immediately have full UQC (this is the case for TQC proposals based on Fibonacci anyons). In fact, for an Ising anyon TQC, attempting to implement any nonstabilizer operation necessitates using nontopological operations, which are expected to be highly noisy. For specific target gates and noise models, one can calculate the threshold noise rate [11–13] before the power to provide UQC is lost (see, e.g., Sec. IV A for a relevant example). Rather than consider a number of different target gates (unitaries) and noise models, we examine an overall quantum operation $\mathcal{E}$, which we intend to be nonstabilizer, but which may be subject to some unknown evolution, and give an operational benchmark on the utility of $\mathcal{E}$ for UQC. Proving that an operation $\mathcal{E}$ enables UQC, when used alongside perfect stabilizer operations, reduces to the question of whether $\mathcal{E}$ can be used to produce ancillas that are suitable for magic state distillation [14,15] (MSD). MSD is a subroutine used in many proposals for fault-tolerant UQC, whereby the requisite non-Clifford gate U is implemented with the assistance of a nonstabilizer pure qubit state $|\psi_U\rangle$. Crucially, many impure (noisy) copies of $|\psi_U\rangle$ can be used to create a smaller number of purer $|\psi_U\rangle$ *using only stabilizer operations*, and this process can be iterated a number of times if necessary. However, if the ancillas $|\psi_U\rangle$ are too impure, then it becomes futile to attempt an approximation of U, regardless of

_____

[*]mark.howard@nuim.ie
[†]Jiri.Vala@nuim.ie

the number of ancillas and techniques used. A summary of the current methods and limits of MSD can be found in [16–18].

The benchmark (for identifying $\mathcal{E}$ that enable UQC) that we have just described will involve the violation of a Bell inequality by a two-qubit quantum state, wherein the underlying qubit encoding is provided for by non-Abelian Ising anyons. The topological phase supporting Ising anyons is inherently nonlocal, but for the purposes of quantum information processing, we are more interested in displaying the nonlocality of encoded qubits. Investigations of nonlocality in non-Abelian anyons have been carried out in [19,20] from a different perspective. The results here are motivated mainly by Ising anyon proposals for a TQC, but we note that any system in which magic state distillation is used to achieve fault-tolerant quantum computation assumes that stabilizer operations are effectively perfect (albeit at an encoded level within a Calderbank-Shor-Steane error-correcting code, for example). Consequently, the results presented here can be applied to any such system.

Given the restriction on topologically implementable gates, and hence on measurement directions, one quickly sees (see, e.g., Sec. I C later) that this setup precludes the possibility of performing measurements that lead to violation of a Bell inequality. In other words, all measurement outcomes can be described by a local hidden variable (LHV) theory, when we restrict the experimenter to using only topologically protected gates. Previously, we have motivated examining the set of operations $\mathcal{E}$ that can provide UQC. Similarly, we can examine the set of operations $\mathcal{E}$ that can enable violation of a Bell inequality when used in conjunction with topological operations. Here, we relate the two sets of operations with each other and show that, under certain restrictions, one is strictly included within the other. More generally, we show that any operation that can be used to violate a so-called Clauser-Horne-Shimony-Holt (CHSH) Bell inequality can also be used to provide UQC.

Note that violation of Bell inequalities has been related to quantum-computational power in other contexts [21,22], although these works were more focused on generalized entanglement and measurement-based quantum computation, respectively. While an interesting topic in its own right, our results do not claim to relate nonlocality and UQC in a broad sense. Indeed, in the Ising anyon context, if we are allowed to construct tripartite states [e.g., the Greenberger-Horne-Zeilinger (GHZ) state $(|000\rangle + |111\rangle)/\sqrt{2}$], then topologically protected operations *are* sufficient to exhibit tripartite nonlocality (e.g., Mermin's experiment [23]), while still not enabling UQC.

A. Clifford operations

The single-qubit Clifford group has 24 distinct elements and contains the Pauli group (of order 4) as a subgroup (strictly speaking, we are discussing the Clifford and Pauli groups modulo their center, which amounts to ignoring a global phase in the matrix representation of these elements). Canonical, non-Pauli elements of the Clifford group are

$$H = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad S = \begin{pmatrix} 1 & 0 \\ 0 & i \end{pmatrix} \quad (1)$$

and these two operations are actually sufficient to generate the whole group. The characteristic property of Clifford gates C is that they map Pauli operators to Pauli operators under conjugation, i.e.,

$$C\sigma_j C^\dagger \mapsto \pm\sigma_k, \quad j,k \in \{x,y,z\}.$$

When visualized as operations on the Bloch sphere, the Clifford group can be identified with the 24 possible transformations that can be constructed by combining consecutive $90°$ rotations about the x, y, and z axes.

B. Quantum operations

To describe the unknown evolution of an Ising anyon qubit, while we attempt to implement the requisite nonstabilizer operation, we use the quantum operations formalism (as described in, e.g., [10]), where $\mathcal{E}$ is a superoperator that maps input density matrices to output density matrices:

$$\mathcal{E}(\rho_{\text{in}}) = \rho_{\text{out}}.$$

A well-known tool in quantum information, the Jamiołkowski isomorphism, tells us that $\mathcal{E}$ is completely characterized by the output state $\varrho_\mathcal{E}$ of a process whereby $\mathcal{E}$ is applied to one half of a maximally entangled pair:

$$\varrho_\mathcal{E} = (\mathcal{I} \otimes \mathcal{E}) [|\Phi\rangle\langle\Phi|] \quad \left(\text{where } \Phi = \frac{|00\rangle + |11\rangle}{\sqrt{2}} \right). \quad (2)$$

Because $\varrho_\mathcal{E}$ is our preferred method for representing general quantum operations, we will sometimes use $\mathcal{E}$ and $\varrho_\mathcal{E}$ interchangeably in latter sections of the paper.

Another convenient way of expressing a quantum operation is in terms of its so-called Kraus operators $\{E_i\}$ via

$$\mathcal{E}(\rho_{\text{in}}) = \rho_{\text{out}} = \sum_i E_i \rho_{\text{in}} E_i^\dagger \quad \left(\text{where } \sum_i E_i^\dagger E_i = \mathbb{I} \right).$$

It takes 12 real parameters to completely characterize an arbitrary completely positive trace-preserving operation. A nine-parameter subset of the set of all possible operations is given by those that preserve the identity (these are often called unital channels), i.e., those for which $\mathcal{E}(\mathbb{I}) = \mathbb{I}$. For unital operations, a Kraus-type description is possible, except the operation is now a probabilistic mixture of unitaries (where each unitary U_k is applied with probability p_k)

$$\mathcal{E}(\rho_{\text{in}}) = \rho_{\text{out}} = \sum_i p_i U_i \rho_{\text{in}} U_i^\dagger \quad \left(\text{where } \sum_i p_i = 1 \right).$$

C. Nonlocal correlations with restricted operations

Here, we briefly motivate why a nonstabilizer operation is necessary to exhibit nonlocality. Consider the maximally entangled state $|\Phi\rangle = \frac{|00\rangle + |11\rangle}{\sqrt{2}}$, then its Pauli expectation values, defined as

$$\text{JK} = \text{Tr} \left(|\Phi\rangle\langle\Phi| \sigma_j \otimes \sigma_k \right), \quad j,k \in \{x,y,z\}$$

are all zero except for

$$\text{XX} = -\text{YY} = \text{ZZ} = 1.$$

The Pauli expectation values for $|\Phi\rangle$ can be recreated exactly by two spatially separated parties **A** and **B**, given that they share three random (unbiased) bits $\{r_1, r_2, r_3\} \in \{0, 1\}$, and they both obey the following set of rules [24]:

(i) Measurement in the X direction $\leftrightarrow$ **A** and **B** both output -1^{r_1}.

(ii) Measurement in the Y direction $\leftrightarrow$ **A** outputs -1^{r_2}, **B** outputs -1^{r_2+1}.

(iii) Measurement in the Z direction $\leftrightarrow$ **A** and **B** both output -1^{r_3}.

A little thought shows that a similar scheme would work for any bipartite entangled state created by stabilizer operations. That such a scheme suffices, using only shared randomness, indicates that no purely quantum mechanical effects are needed to describe such an experiment.

II. BELL INEQUALITIES

We will attempt to perform an experiment that, unlike the example of Sec. I C, is not describable by local hidden variables. The basic setup is depicted in Fig. 1, where $\mathcal{E}$ is applied to one half of a maximally entangled state $(|00\rangle + |11\rangle)/\sqrt{2}$ and then Pauli measurements are performed. We can allow for either two or three distinct Pauli measurements to be performed on each qubit (Fig. 1 depicts the scenario in which only two measurements are performed on each side), and these lead to different types of Bell inequality.

For a given bipartite state ϱ, and using a notation similar to that of Collins and Gisin [26], we can arrange the expectation values for a given nonlocality experiment with Pauli measurements in a table, i.e.,

$$\begin{pmatrix} \text{II} & \text{XI} & \text{YI} & \text{ZI} \\ \text{IX} & \text{XX} & \text{YX} & \text{ZX} \\ \text{IY} & \text{XY} & \text{YY} & \text{ZY} \\ \text{IZ} & \text{XZ} & \text{YZ} & \text{ZZ} \end{pmatrix}, \tag{3}$$

where XY is the expectation value $\text{Tr}(\varrho \sigma_x \otimes \sigma_y)$ and so on. Since each Pauli operator has eigenvalues ± 1, we can list all possible matrices of the form (3) that correspond to deterministic local configurations. A deterministic local configuration is one in which the local Pauli expectation values are extremal and, when multiplied, provide the nonlocal Pauli expectation value [e.g., where identities such as $(\text{XI})(\text{IY}) = \text{XY}$ hold, with XI, IY, XY all either ± 1]. By letting $a, b, \ldots, f \in$

$\{0, 1\}$ take on all 2^6 assignments, then the following matrix fully describes all possible local configurations:

$$\begin{pmatrix} 1 & (-1)^a & (-1)^b & (-1)^c \\ (-1)^d & (-1)^{a+d} & (-1)^{b+d} & (-1)^{c+d} \\ (-1)^e & (-1)^{a+e} & (-1)^{b+e} & (-1)^{c+e} \\ (-1)^f & (-1)^{a+f} & (-1)^{b+f} & (-1)^{c+f} \end{pmatrix}. \tag{4}$$

If an experimentalist's measured expectation values correspond *exactly* to any of the above local configurations, then these expectation values do not exhibit any nonlocality. Clearly, any realistic experiment will not obey such a strict condition, and the question of exhibiting nonlocality becomes richer. The experimentalist must now check that the measured expectation values can not be expressed as a probabilistic combination of local configurations. This naturally leads to the notion of convex geometry and bounded polyhedra in higher dimensions (these are usually called polytopes).

Ignoring the constant term in the top left, the 2^6 matrices in (4) can each be identified with a vector in $\mathbb{R}^{15}$. These local configuration vectors are called vertices and the convex hull of these vertices (i.e., the set of vectors that is expressible as a probabilistic combination of the vertices) describes a polytope in $\mathbb{R}^{15}$ [this is analogous to how the eight vertices $(\pm 1, \pm 1, \pm 1)$ describe a solid cube in $\mathbb{R}^3$]. The interior of this polytope describes all possible LHV models. Conversely, any table of expectation values (3) that does not lie inside this polytope exhibits genuine quantum nonlocality. The usual, textbook, way of identifying nonlocality is via Bell inequalities such as

$$\text{XX} + \text{XY} + \text{YX} - \text{YY} \leqslant 2. \tag{5}$$

In fact, (tight) Bell inequalities such as these are actually the defining inequalities for the bounding faces (facets) of the LHV polytope that we have just described. Software such as Avis' LRS [27] can be used to derive all the facets of LHV polytope, given the 2^6 local configuration vectors in (4) as input (and vice versa). In order to unify our notation, we rearrange the inequality (5) and then rewrite in table form as

$$2 - \text{XX} - \text{XY} - \text{YX} + \text{YY} \geqslant 0 \tag{6}$$

$$\Rightarrow \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \text{II} & \text{XI} & \text{YI} & \text{ZI} \\ \text{IX} & \text{XX} & \text{YX} & \text{ZX} \\ \text{IY} & \text{XY} & \text{YY} & \text{ZY} \\ \text{IZ} & \text{XZ} & \text{YZ} & \text{ZZ} \end{pmatrix} \geqslant 0, \tag{7}$$

where the dot product between two matrices M, N behaves like the familiar vector dot product, i.e., $M \cdot N = \text{Tr}(M^T N) = \sum_{i,j} M_{i,j} N_{i,j}$. The full list of bounding inequalities for the LHV polytope is described in a concise form in Sec. IV.

III. NONSTABILIZER OPERATIONS

In the previous section, ascertaining whether measurement results had an LHV model amounted to the question of whether the results could be expressed as a probabilistic combination of local configurations. In the context of UQC, the relevant question is now whether $\mathcal{E}$ is expressible as a probabilistic
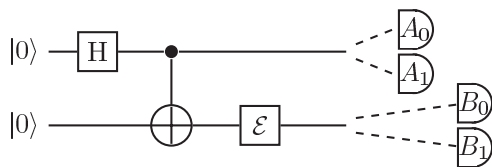

FIG. 1. A simple setup to detect nonlocality: Here, two possible measurement settings for the first (second) qubit are denoted A_i (B_j). When A_i and B_j are constrained to be Pauli operators, then $\mathcal{E}$ must be a nonstabilizer operation if any Bell inequality [e.g., that given in (5)] is to be violated.

combination of Clifford gates; for example, an operation $\mathcal{E}$ given by

$$\mathcal{E}(\rho) = \tfrac{1}{3} H \rho H^\dagger + \tfrac{1}{4} S \rho S^\dagger + \tfrac{5}{12} \rho$$

is clearly of no use in the quest to achieve UQC. The general prerequisite for $\mathcal{E}$ to be useful is

$$\mathcal{E}(\rho) \neq \sum_{i=1}^{24} p_i C_i \rho C_i^\dagger \quad \left(\text{where } \sum_{i=1}^{24} p_i = 1 \right). \quad (8)$$

If we use the Jamiołkowski isomorphism [as defined in Eq. (2)] to represent operations, then a completely equivalent condition to (8) is given by

$$\varrho_{\mathcal{E}} \neq \sum_{i=1}^{24} p_i |J_{C_i}\rangle\langle J_{C_i}| \quad \left(\text{where } \sum_{i=1}^{24} p_i = 1 \right) \quad (9)$$

with $|J_{C_i}\rangle = (\mathbb{I} \otimes C_i)|\Phi\rangle$

To test whether an unknown operation satisfies the above requirement, one is naturally led to the concept of convex polytopes once more. The 24 Clifford gates, when represented as $|J_{C_i}\rangle\langle J_{C_i}|$ and decomposed in the Pauli basis, form the 24 vertices of what is called the Clifford polytope. In fact, the Clifford operations only span a nine-dimensional subspace of $\mathbb{R}^{15}$ because the six coefficients IX, IY, IZ, XI, YI, and ZI are all identically zero when $\mathcal{E}(\mathbb{I}) = \mathbb{I}$. The Clifford gates given in (1), for example, correspond to the following two vertices:

$$v_H = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad v_S = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

The facets of the Clifford polytope [11,13,28] are bounding inequalities that partition the set of all valid $\varrho_{\mathcal{E}}$ into those that represent $\mathcal{E}$ satisfying (8) and those that do not. In other words, any operation $\mathcal{E}$ that satisfies (8) violates (at least) one of these facet inequalities. These inequalities, which we label I, can be cast in matrix form in a completely equivalent way to what was done in (7), and it turns out there are two distinct classes of facet [11], a representative example of each of which we provide in the following:

$$\mathcal{I}_\alpha = \left\{ \text{all facets of type } I_\alpha \text{ or } I_\alpha^T \right\} \quad (10)$$

$$\text{where } I_\alpha = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}, \quad (11)$$

$$\mathcal{I}_\beta = \{ \text{all facets of type } I_\beta \} \quad (12)$$

$$\text{where } I_\beta = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}. \quad (13)$$

The total number of facets required to completely characterize the boundary of the Clifford polytope is 120, which can be broken down into $|\mathcal{I}_\alpha| + |\mathcal{I}_\beta| = (24 + 24) + 72$. We are
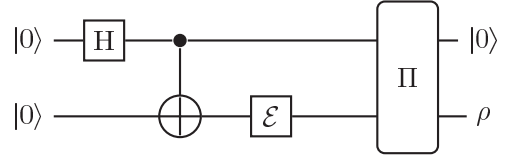

FIG. 2. A circuit to help achieve universal quantum computation via magic state distillation: Every element of this circuit except $\mathcal{E}$ is implementable using stabilizer operations. When $\mathcal{E}$ exhibits nonlocality in the setup of Fig. 1, then $\mathcal{E}$ used in the above circuit produces ancillas ρ that are useful for magic state distillation subroutine (MSD circuit not depicted). The block containing Π stands for a two-qubit Pauli measurement (e.g., parity measurement), wherein we postselect on the desired outcome.

primarily interested in facets from $\mathcal{I}_\beta$ because it was shown in [13] that an operation violating such a facet can always be used to create an ancilla ρ that is useful for magic state distillation. The circuit to create ρ is given in Fig. 2, where the choice of which two-qubit Pauli measurement Π to perform is determined by the particular facet that is violated by $\mathcal{E}$.

IV. RESULTS

A full facet description of the LHV polytope as described in Sec. II comprises 684 distinct facets, and this facet description can partitioned into three different classes, i.e,

$$\mathcal{I}_{\text{triv}} = \{ \text{all facets of type } I_{\text{triv}} \} \quad (14)$$

$$\text{where } I_{\text{triv}} = \begin{pmatrix} 1 & -1 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (15)$$

$$\mathcal{I}_{2222} = \{ \text{all facets of type } I_{2222} \} \quad (16)$$

$$\text{where } I_{2222} = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad (17)$$

$$\mathcal{I}_{3322} = \{ \text{all facets of type } I_{3322} \} \quad (18)$$

$$\text{where } I_{3322} = \begin{pmatrix} 4 & -1 & -1 & 0 \\ -1 & 1 & 1 & -1 \\ -1 & 1 & 1 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix}. \quad (19)$$

The number of distinct facets in each class is $|\mathcal{I}_{\text{triv}}| = 36$, $|\mathcal{I}_{2222}| = 72$, and $|\mathcal{I}_{3322}| = 576$. Inequalities of the type I_{triv} are considered trivial since the restrictions they impose are satisfied by any bipartite quantum state ϱ. The representative inequality I_{triv} given in (15) amounts to $\text{Tr}[(\mathbb{I} - \sigma_x) \otimes (\mathbb{I} - \sigma_x)\varrho] \geqslant 0$, for example. Inequalities of the type I_{2222} are usually known as CHSH inequalities [29] and an example of one such inequality was already discussed in the context of Eq. (7). These I_{2222} inequalities appear, in a different form closer to that of (5), in most textbook accounts of nonlocality (see, e.g., [10]). Bell inequalities of the type I_{3322} are less well known, but have a long history, appearing as early as 1981 [30] and rediscovered more recently in [26,31,32]. Finally, we note the reasoning behind the notation I_{ccdd}; the numbers $ccdd$ in the subscript refer to the number of measurement settings c available to each party, wherein each

measurement has a number d of possible outcomes (e.g., inequalities of the type I_{3322} can only arise when there are 3 possible measurement settings on each side of the bipartite state).

The I_{2222} inequalities require only two Pauli measurements on each side, whereas I_{3322} inequalities require three on each side. It turns out that the I_{2222} inequalities are sufficient for our purposes. In fact, I_{3322} inequalities are irrelevant (redundant) for a very large subset of all possible quantum operations $\mathcal{E}$. If one makes the assumption that $\mathcal{E}(\mathbb{I}) = \mathbb{I}$, which is the case for many of the most important noise models, then local expectation values for ϱ are identically zero, and coefficients of IX, IY, IZ, XI, YI, and ZI in I_{3322} facets can be ignored. With this modification, the I_{3322} inequalities are expressible as linear combinations of I_{2222} (CHSH) inequalities, and hence the I_{3322} inequalities are irrelevant. One can check that all (purely nonlocal) I_{3322} inequalities decompose into four I_{2222} inequalities, as in the following example:

$$
\begin{pmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & 1 & 0 \end{pmatrix}
$$
$$
= \frac{1}{2}\left[\begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \right.
$$
$$
\left. + \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} + \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 \end{pmatrix} \right]. \quad (20)
$$

We are now ready to state the main result:

Theorem 1. If there is an operation $\mathcal{E}$ violating a facet from $\mathcal{I}_{2222}$ [Eq. (16)] (i.e., violating a CHSH inequality), then $\mathcal{E}$ also violates a facet (of the Clifford polytope) from $\mathcal{I}_\beta$ [Eq. (12)]. Using results from [13], such $\mathcal{E}$ enable UQC (via magic state distillation) when supplemented with topologically protected operations.

Mathematically, we have the following statement:

$$
\forall \text{ valid } \mathcal{E}, \quad \forall\, I_{2222} \in \mathcal{I}_{2222}, \quad \exists\, I_\beta \in \mathcal{I}_\beta \quad (21)
$$

such that $\varrho_\mathcal{E}$ violates I_β at least as much as $\varrho_\mathcal{E}$ violates I_{2222}. Without loss of generality, we assume that the canonical representative I_{2222} [Eq. (17)] of $\mathcal{I}_{2222}$ is violated, and show that I_β [Eq. (13)] from $\mathcal{I}_\beta$ is necessarily also violated. This is a consequence of the following proposition, which is easily proved:

$$
\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \text{II} & \text{XI} & \text{YI} & \text{ZI} \\ \text{IX} & \text{XX} & \text{YX} & \text{ZX} \\ \text{IY} & \text{XY} & \text{YY} & \text{ZY} \\ \text{IZ} & \text{XZ} & \text{YZ} & \text{ZZ} \end{pmatrix}
$$
$$
\leqslant \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & -1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} \text{II} & \text{XI} & \text{YI} & \text{ZI} \\ \text{IX} & \text{XX} & \text{YX} & \text{ZX} \\ \text{IY} & \text{XY} & \text{YY} & \text{ZY} \\ \text{IZ} & \text{XZ} & \text{YZ} & \text{ZZ} \end{pmatrix}.
$$
$$
(22)
$$

Rearranging the above inequality gives

$$
\begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \text{II} & \text{XI} & \text{YI} & \text{ZI} \\ \text{IX} & \text{XX} & \text{YX} & \text{ZX} \\ \text{IY} & \text{XY} & \text{YY} & \text{ZY} \\ \text{IZ} & \text{XZ} & \text{YZ} & \text{ZZ} \end{pmatrix} \leqslant 0, \quad (23)
$$

which simply stands for

$$
-1 + \text{ZZ} \leqslant 0. \quad (24)
$$

This is always satisfied since Pauli expectation values are in the range $[-1, 1]$. To create an ancilla for magic state distillation (and consequently enable UQC), the best measurement Π to use in the circuit of Fig. 2 is the parity measurement postselected on even parity, i.e., $\Pi = \frac{1}{2}(\mathbb{I} + \sigma_z \sigma_z)$.

Note that $|\mathcal{I}_\beta| = |\mathcal{I}_{2222}| = 72$. By examining the matrix representation of these sets of facets, one sees that each $I_\beta \in \mathcal{I}_\beta$ can be paired one-to-one with each $I_{2222} \in \mathcal{I}_{2222}$ by matching facets, the nonidentity coefficients of which differ in only one position {e.g., the coefficient of ZZ in our canonical I_{2222} [Eq. (17)] is 0, whereas for our canonical I_β [Eq. (13)] the coefficient is $+1$}. This provides a recipe for identifying the relevant I_β such that Eq. (21) holds for a given I_{2222} (the CHSH inequality that we presume has been violated). Any such pair $\{I_{2222}, I_\beta\}$ will obey an inequality of the form (22). The general expression, analogous to Eq. (24), for different facets from $\mathcal{I}_{2222}$ and suitably chosen facets from $\mathcal{I}_\beta$ is

$$
-1 \pm \text{Tr}(\sigma_j \otimes \sigma_k \varrho_\mathcal{E}) \leqslant 0, \quad j, k \in \{x, y, z\} \quad (25)
$$

and this clearly always holds. In order to use $\mathcal{E}$ to obtain ancillas useful for magic state distillation, a suitable measurement Π in Fig. 2 includes one of the form $\Pi = \frac{1}{2}(\mathbb{I} \pm \sigma_j \sigma_k)$ [13].

A. Example: Phase gate subject to dephasing

The phase gate $U_z(\theta)$, defined as

$$
U_z(\theta) = \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix} = \begin{pmatrix} e^{-i\frac{\theta}{2}} & 0 \\ 0 & e^{i\frac{\theta}{2}} \end{pmatrix},
$$

is an element of the Clifford group when θ is a multiple of $\pi/2$. While technically any non-Clifford angle θ would suffice to enable UQC, one typically seeks to implement the so-called "pi-over-eight" gate $U_z(\frac{\pi}{4})$ because it possesses additional desirable properties [33]. For the case of Ising anyons, it has been proposed [34,35] to perform a phase gate using a sacklike geometry, wherein the sack contains a single qubit, and an anyonic edge current can either follow the exterior boundary of the sack or tunnel across the constriction of the sack. In [34], it was shown that, subject to certain assumptions, the overall evolution of the qubit state can be described by a superoperator $\mathcal{E}(\rho(t_0)) = \rho(t)$, where the matrix elements of the density operator change like

$$
\rho(t) = \begin{pmatrix} \rho_{00}(t_0) & e^{-\frac{s^2}{2}} e^{-i\theta} \rho_{01}(t_0) \\ e^{-\frac{s^2}{2}} e^{i\theta} \rho_{10}(t_0) & \rho_{11}(t_0) \end{pmatrix}
$$

or, equivalently, in the Kraus operator form

$$\rho(t) = E_0 \rho(t_0) E_0^\dagger + E_1 \rho(t_0) E_1^\dagger,$$

$$\text{where } E_0 = \sqrt{\frac{1 + e^{-\frac{s^2}{2}}}{2}} \begin{pmatrix} 1 & 0 \\ 0 & e^{i\theta} \end{pmatrix}, \quad (26)$$

$$E_1 = \sqrt{\frac{1 - e^{-\frac{s^2}{2}}}{2}} \begin{pmatrix} 1 & 0 \\ 0 & -e^{i\theta} \end{pmatrix}.$$

As a side note, in quantum information theory, the description of this process in terms of Kraus operators would probably be rewritten as

$$\mathcal{E}(\rho) = E_0 \rho E_0^\dagger + E_1 \rho E_1^\dagger,$$

$$\text{where } E_0 = \sqrt{1 - p}\, U_z(\theta), \quad E_1 = \sqrt{p}\, \sigma_z U_z(\theta),$$

$$\text{with } p = \frac{1 - e^{-\frac{s^2}{2}}}{2}, \quad (27)$$

which is completely equivalent. The natural interpretation is that with probability $1 - p$, the desired $U_z(\theta)$ gate is performed, and with probability p an undesired rotation $\sigma_z U_z(\theta)$ is performed. This undesired rotation is actually the worst possible noise that could be inflicted [12] (i.e., it requires the smallest noise rate p to make the overall $\mathcal{E}$ expressible as a probabilistic combination of Clifford gates).

The operation $\mathcal{E}$ of Eq. (26), when expressed as a Jamiołkowski state $\varrho_{\mathcal{E}}$ and decomposed in the Pauli basis, takes the form

$$\begin{pmatrix} \text{II} & \text{XI} & \text{YI} & \text{ZI} \\ \text{IX} & \text{XX} & \text{YX} & \text{ZX} \\ \text{IY} & \text{XY} & \text{YY} & \text{ZY} \\ \text{IZ} & \text{XZ} & \text{YZ} & \text{ZZ} \end{pmatrix}$$

$$= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & e^{-\frac{s^2}{2}}\cos[\theta] & e^{-\frac{s^2}{2}}\sin[\theta] & 0 \\ 0 & e^{-\frac{s^2}{2}}\sin[\theta] & -e^{-\frac{s^2}{2}}\cos[\theta] & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (28)$$

and so the Bell inequality given in Eq. (7) reads as

$$\text{XX} + \text{XY} + \text{YX} - \text{YY} \leqslant 2$$

$$\Rightarrow e^{-\frac{s^2}{2}} ([\cos\theta] + [\sin\theta] + [\sin\theta] - [-\cos\theta]) \leqslant 2.$$

For the optimal choice of angle $\theta = \frac{\pi}{4}$, this inequality is violated for the parameter range $0 \leqslant s < \sqrt{\ln 2}$. In terms of the simplified Kraus operators given in (27), the range of allowable noise rates p, while still giving CHSH violation, is given by $0 \leqslant p \lesssim 14\%$.

Another simple calculation shows that the facet I_β is also violated (hence, UQC is possible) for the same noise rates $0 \leqslant p \lesssim 14\%$. One can actually see this straight away by noting that Eq. (28) implies that $-1 + \text{ZZ} = 0$, which means

[with reference to Eq. (24)] that the threshold noise rate is the same for violation of I_β and I_{2222}. This will also be true for any noise model that is expressible as a probabilistic combination of rotations about the z axis (and similarly for operations comprised of rotations about x and y axes too).

B. Operations enabling UQC versus operations enabling Bell inequality violation

The noise rates for which Bell inequality violation is possible, and the noise rates for which UQC is possible, generally do not coincide. For example, a $U_z(\frac{\pi}{4})$ gate under depolarizing noise, modeled as

$$\mathcal{E}(\rho) = (1 - p)U_z\left(\frac{\pi}{4}\right)\rho U_z\left(\frac{\pi}{4}\right)^\dagger + p\frac{\mathbb{I}}{2},$$

enables UQC, but does not violate any Bell inequality for the parameter range $0.29 \lesssim p \lesssim 0.45$. In the current context, a non-Clifford operation $\mathcal{E}$ is necessary if one hopes to observe nonlocality (as shown in Sec. I C), but not always sufficient. However, if we expand the problem to allow multiple (possibly simultaneous) uses of $\mathcal{E}$ in a more complex circuit than that of Fig. 1, then it is an interesting question as to whether $\mathcal{E}$ being non-Clifford could be both a necessary and sufficient condition for the violation of a bipartite Bell inequality.

V. CONCLUSION

By using techniques from convex geometry, we can show that detection of nonlocality in the encoded qubits of an Ising anyon TQC serves as a benchmark for universal quantum computation. The strength of this approach is that it does not consider specific noise models, but rather considers general quantum operations. Given an unknown operation that has been observed to enable violation of a CHSH inequality (using two distinct Pauli measurements on each half of an entangled state), we can prescribe a way of using this operation to manufacture ancillas that are suitable for magic state distillation. When used in conjunction with topologically protected stabilizer operations, ancillas of this kind allow for universal quantum computation.

[1] G. Moore and N. Read, Nucl. Phys. B **360**, 362 (1991).

[2] P. Bonderson and A. Kitaev, Phys. Rev. Lett. **96**, 016803 (2006).

[3] J. Alicea, Y. Oreg, G. Refael, and F. von Oppen, Nat. Phys. **7**, 412 (2011).

[4] D. A. Ivanov, Phys. Rev. Lett. **86**, 268 (2001).

[5] A. Kitaev, Ann. Phys. (NY) **321**, 2 (2006).

[6] M. B. Plenio and S. Virmani, New J. Phys. **12**, 033012 (2010).

[7] Sergey Bravyi, Phys. Rev. A **73**, 042313 (2006).

[8] L. S. Georgiev, Phys. Rev. B **74**, 235112 (2006).

[9] P. Bonderson *et al.* (unpublished).

[10] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).

[11] H. Buhrman, R. Cleve, M. Laurent, N. Linden, A. Schrijver, and F. Unger, *Annual IEEE Symposium on Foundations of Computer Science* (IEEE Computer Society, Berkeley, CA, 2006), pp. 411–419.

[12] S. Virmani, S. F. Huelga, and M. B. Plenio, Phys. Rev. A **71**, 042328 (2005).

[13] W. van Dam and M. Howard, Phys. Rev. Lett. **103**, 170504 (2009).

[14] E. Knill, Nature (London) **434**, 39 (2005).

[15] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).

[16] B. Reichardt, Quantum Inf. Comput. **9**, 1030 (2009).

[17] E. T. Campbell and D. E. Browne, *Theory of Quantum Computation, Communication, and Cryptography*, edited by Andrew Childs and Michele Mosca, Lecture Notes in Computer Science, Volume 5906 (Springer, Berlin, 2009), pp. 20–32.

[18] E. T. Campbell and D. E. Browne, Phys. Rev. Lett. **104**, 030503 (2010).

[19] G. K. Brennen, S. Iblisdir, J. K. Pachos, and J. K. Slingerland, New J. Phys. **10**, 103023 (2009).

[20] Dong-Ling Deng, Chunfeng Wu, Jing-Ling Chen, and C. Oh, Phys. Rev. Lett. **105**, 060402 (2010).

[21] N. Ratanje and S.l. Virmani, Phys. Rev. A **83**, 032309 (2011).

[22] Matty J. Hoban, Earl T. Campbell, Klearchos Loukopoulos, and Dan E Browne, New J. Phys. **13**, 023014 (2011); R. Raussendorf, e-print arXiv:0907.5449.

[23] N. D. Mermin, Am. J. Phys. **58**, 731 (1990).

[24] This was first shown to us by D. Bacon. See, e.g., [25] for a similar construction.

[25] Tracey Tessier, Carlton Caves, Ivan Deutsch, Bryan Eastin, and Dave Bacon, Phys. Rev. A **72**, 032305 (2005).

[26] Daniel Collins and Nicolas Gisin, J. Phys. A: Math. Gen. **37**, 1775 (2004).

[27] D. Avis, in *Polytopes–Combinatorics and Computation (Oberwolfach Seminars)*, edited by G. Kalai and G. Ziegler (Birkhäuser, Basel, 2000), pp. 177–198.

[28] W. van Dam and M. Howard, Phys. Rev. A **83**, 032310 (2011).

[29] J. F. Clauser, M. A. Horne, A. Shimony, and R. A. Holt, Phys. Rev. Lett. **23**, 880 (1969).

[30] M. Froissart, Il Nuovo Cimento B **64**, 241 (1981).

[31] Itamar Pitowsky and Karl Svozil, Phys. Rev. A **64**, 014102 (2001).

[32] C. Sliwa, Phys. Lett. A **317**, 165 (2003).

[33] P. Boykin, Inf. Process. Lett. **75**, 101 (2000).

[34] Parsa Bonderson, David J. Clarke, Chetan Nayak, Kirill Shtengel, Phys. Rev. Lett. **104**, 180505 (2010).

[35] D. J. Clarke and K. Shtengel, Phys. Rev. B **82**, 180519 (2010).