Bound States for Magic State Distillation in Fault-Tolerant Quantum Computation

Earl T. Campbell and Dan E. Browne

Department of Physics and Astronomy, University College London, Gower Street, London, WC1E 6BT, United Kingdom
(Received 13 August 2009; revised manuscript received 4 December 2009; published 22 January 2010)

Magic state distillation is an important primitive in fault-tolerant quantum computation. The magic states are pure nonstabilizer states which can be distilled from certain mixed nonstabilizer states via Clifford group operations alone. Because of the Gottesman-Knill theorem, mixtures of Pauli eigenstates are not expected to be magic state distillable, but it has been an open question whether all mixed states outside this set may be distilled. In this Letter we show that, when resources are finitely limited, nondistillable states exist outside the stabilizer octahedron. In analogy with the bound entangled states, which arise in entanglement theory, we call such states bound states for magic state distillation.

The significant noise and decoherence in quantum systems means that harnessing these systems for computational tasks must be performed fault tolerantly [1,2]. In a wide variety of setups only a limited set of gates, known as the Clifford group, are implemented in a manifestly fault-tolerant manner. Examples include some anyonic topological quantum computers [3–5], postselected quantum computers [6,7], and measurement based topological quantum computers [8]. This motivates the problem of when such devices, with practically error free Clifford gates, may be promoted to a full quantum computer. The celebrated Gottesman-Knill theorem shows that a Clifford circuit acting on stabilizer states—simultaneous eigenstates of several Pauli operators—can be efficiently simulated by a classical computer [9]. However, given a resource of pure nonstabilizer states, we can implement gates outside the Clifford group. For example, a qubit in an eigenstate of the Hadamard enables one to implement a $\pi/8$ phase gate that when supplementing the Clifford group gives a dense covering of all unitary operations [10], and so enables universal quantum computation.

Preparation of nonstabilizer states would usually require a non-Clifford operation, so in this context, one would require that even noisy copies of these states enable high fidelity quantum computation. Bravyi and Kitaev [10] showed that this can be achieved. Coining the term magic state distillation, they showed that most mixed nonstabilizer states can be distilled via Clifford group circuits to fewer copies of a lower entropy state, reaching in the limit of infinite iterations a pure nonstabilizer magic state. However, the protocols they presented do not succeed for all mixed nonstabilizer states. Bravyi and Kitaev were not satisfied by the ambiguous status of these states and concluded that "The most exciting open problem is to understand the computational power of the model in [this] region of parameters." Either all nonstabilizer states are efficiently distillable by an undiscovered protocol, or there exist nonstabilizer states that are impossible to distill. Such undistillable states we call bound states for magic state distillation, in analogy with bound states in entanglement

distillation [11]. Here we make progress by showing that bound states exist for a very broad class of protocols. By showing that a single round of a finite sized protocol will not improve these states, it follows that repeating such a protocol, even with an infinite number of iterations, will also have no benefit. Hence, we explain why all known protocols fail to distill some states.

The single-qubit stabilizer states, for which the Gottesman-Knill theorem applies, are the six pure stabilizer states (the eigenstates of $\pm X$, $\pm Y$, and $\pm Z$) and any incoherent mixture of these. In the Bloch sphere, this convex set with 6 vertices forms the stabilizer octahedron partially shown in Fig. 1(a). Single-qubit states have density matrices:

$$\rho(f, \mathbf{a}) = [\mathbb{1} + (2f - 1)(a_X X + a_Y Y + a_Z Z)]/2, \quad (1)$$

where $\mathbf{a} = (a_X, a_Y, a_Z)$ is a unit vector, and f is the fidelity with respect to the pure state $|\psi_{\mathbf{a}}\rangle\langle\psi_{\mathbf{a}}| = (\mathbb{1} + a_X X + a_Y Y + a_Z Z)/2$. Stabilizer states satisfy

$$|2f - 1|(|a_X| + |a_Y| + |a_Z|) \le 1, \quad (2)$$

where the equality holds for states on the surface of the octahedron, and we denote the fidelity of such surface states as $f_{\mathbf{a}}^S$, which is unique assuming $f \ge 1/2$.

Prior protocols for magic state distillation [7,10,12,13] increase fidelity towards eigenstates of Clifford gates, such as the Hadamard H and the T gate [14]. These eigenstates have $\mathbf{a}_H = (1, 0, 1)/\sqrt{2}$ and $\mathbf{a}_T = (1, 1, 1)/\sqrt{3}$, with $f = 1$ for ideal magic states. Given the ability to prepare a mixed nonstabilizer state ρ, we can perform an operation called polarization, or twirling, that brings ρ onto a symmetry axis of the octahedron. For example, by randomly applying $\mathbb{1}$, T, or $T^\dagger$, we map $\rho \to \rho(f, \mathbf{a}_T)$.

Bravyi and Kitaev proposed the following protocol [10] for $|T\rangle$ state distillation: (1) prepare five copies of $\rho(f, \mathbf{a}_T)$; (2) measure the four stabilizers of the five-qubit error correcting code; (3) if all measurements give $+1$, the protocol succeeds and the encoded state is decoded into a single-qubit state, and otherwise restart. Upon a successful implementation of this protocol the output qubit has a
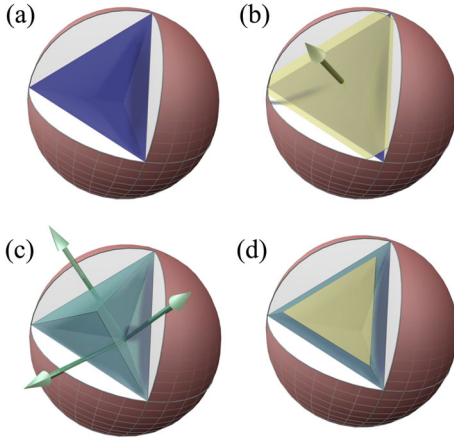
FIG. 1 (color online).   One octant of the Bloch sphere with various regions and directions shown. (a) The set of stabilizer states in one octant. Each octant is identical, with all stabilizer states forming an octahedron. (b) The plane for the distillation threshold for the 5 qubit code, with the direction of $\mathbf{a}_T$ shown. The threshold plane is parallel to the underlying face of the stabilizer octahedron, but is displaced by a small gap. (c) The threshold planes for the STEANE code, with each plane differing by local Clifford gates. These planes meet the stabilizer octahedron at its edges. The three vectors are axes of H-like gates, e.g. $\mathbf{a}_H$; (d) The combined region of states distilled by either the 5 qubit code or the STEANE code. This region only touches the stabilizer octahedron at its edges, and no other known protocol is tight in any other direction.

FIG. 2 (color online).   The performance of magic state distillation of (a) the STEANE code for distilling states in a H-like direction; (b) the five-qubit code distilling states in a T-like direction. Notice that both functions are continuous, and that in (b) an input state on the octahedron surface, $f = f^S_{\mathbf{a}_T}$, will output a state below the surface. Consequently, there is also a region above the surface where the output is a stabilizer state, and hence, not an improvement on the initial state.

fidelity $F(f)$ plotted in Fig. 2(b). Provided the initial fidelity is greater than some threshold, a successful implementation yields a higher fidelity. This protocol has a nontight threshold, and exhibits a gap between the threshold and the set of stabilizer states. Because the initial state was twirled onto the T axis, the threshold forms a plane in the Bloch sphere (see Fig. 1). In contrast, Reichardt has proposed a protocol that does have a tight threshold for distillation of $\rho(f, \mathbf{a}_H)$ states in a H-like direction [12]. His protocol is similar to above, but uses seven qubits each attempt and measures the six stabilizers of the STEANE code [1]. In Fig. 2(a) we show the performance of this protocol, where there is no threshold gap. When the initial mixture is not of the form $\rho(f, \mathbf{a}_H)$, we twirl the initial mixture onto the H axis. Hence, the threshold forms a plane for each H-like direction (see Fig. 1). Although the protocol is tight in directions crossing an octahedron edge, the protocol fails to distill some mixed states just above the octahedron faces, and so is not tight in all directions. Even the combined region of states distilled by all known protocols still leaves a set of states above the octahedron faces, whose distillability properties are unknown.

Here we show that for all size n protocols there is a region of bound states above the octahedron faces. More formally, we are considering all states $\rho(f, \mathbf{a}_P)$ where $\mathbf{a}_P$ has all positive (nonzero) components. Having all components as nonzero excludes states above octahedron edges. Considering only states in the positive octant is completely
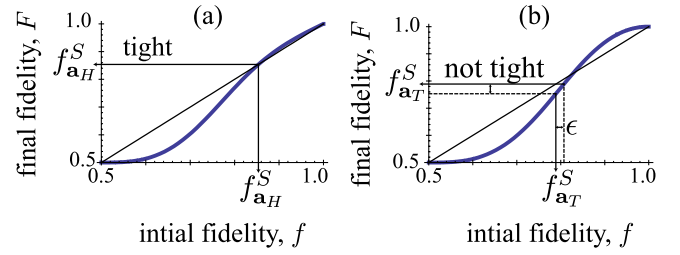
general as Clifford gates enable movement between octants. Many copies of bound states cannot be used to improve on a single copy, and below we formalize the idea of not improved and state our main result.

Definition 1.—We say ρ' is not an improvement on $\rho(f, \mathbf{a}_P)$, when ρ' is a convex mixture of $C_i \rho(f, \mathbf{a}_P) C_i^\dagger$ and stabilizer states, where C_i are Clifford group gates.

Theorem 1.—Consider a device capable of ideal Clifford gates, preparation of stabilizer states, classical feedforward and Pauli measurements. For any protocol on this device that takes $\rho(f, \mathbf{a}_P)^{\otimes n}$ and outputs a single qubit ρ', there exists an $\epsilon > 0$ such that ρ' is not an improvement on $\rho(f, \mathbf{a}_P)$ for $f \leq f^S_{\mathbf{a}_P} + \epsilon$.

Theorem 1 covers a wide class of protocols, which attain a fidelity that is upper bounded by a narrower class of protocols [15], such that Theorem 1 follows from:

Theorem 2.—Consider all protocols that follow these steps: (i) prepare $\rho(f, \mathbf{a}_P)^{\otimes n}$; (ii) measure the $n - 1$ generators of an n qubit stabilizer code $\mathcal{S}_{n-1}$ with one logical qubit; (iii) postselect on all "+1" measurement outcomes; (iv) decode the stabilizer code and output the logical qubit as the single-qubit state ρ'. For all such protocols there exists an $\epsilon > 0$ such that ρ' is not an improvement on $\rho(f, \mathbf{a}_P)$ for $f \leq f^S_{\mathbf{a}_P} + \epsilon$.

Prior protocols, such as those based on the STEANE code and 5 qubit code, are covered explicitly by Theorem 2. Here we use the structure of stabilizer codes to prove Theorem 2, with Theorem 1 following directly from the results of [15], where such distillation protocols are shown to have equal efficacy with more general Clifford protocols. It is crucial to consider the implication of these theorems when an n-qubit protocol is iterated m times. When a single round provides no improvement on the initial resource, the input into the second round will only differ by Clifford group operations, and hence our theorem applies to the second, and all subsequent, rounds. Hence, repeated iteration cannot be used to circumvent our theorem. Before proving these theorems, we derive a pair of powerful lemmas that identify bound states.

Lemma 1.—Consider n copies of an octahedron surface state $\rho(f_{\mathbf{a}_P}^S, \mathbf{a}_P)$ projected onto the code space of $\mathcal{S}_{n-1}$ and then decoded. If the output qubit is in the octahedron interior, then there exists an $\epsilon > 0$ such that for $f \leq f_{\mathbf{a}_P}^S + \epsilon$ the same projection on $\rho(f, \mathbf{a}_P)^{\otimes n}$ also projects onto a mixed stabilizer state.

This Lemma 1 follows directly from the dependence of the output on f, which for finite n is always continuous. We can observe Lemma 1 at work in Fig. 2(b). Our next lemma identifies when octahedron surface states are projected into the octahedron interior. Before stating this we must establish some notation. An initial state $\rho(f_{\mathbf{a}_P}^S, \mathbf{a}_P)^{\otimes n}$ is an ensemble of pure stabilizer states:

$$\rho(f_{\mathbf{a}_P}^S, \mathbf{a}_P)^{\otimes n} = \sum_{\mathbf{g} \in \{X,Y,Z\}^n} q_{\mathbf{g}} |\Psi_{\mathbf{g}}\rangle\langle\Psi_{\mathbf{g}}|, \qquad (3)$$

where $|\Psi_{\mathbf{g}}\rangle$ is stabilized, $g|\Psi_{\mathbf{g}}\rangle = |\Psi_{\mathbf{g}}\rangle$, by the group $\mathcal{G}_{\mathbf{g}}$ generated by $\mathbf{g} = (g_1, g_2, \ldots g_n)$. The operator g_i is X_i, Y_i, or Z_i, with i labeling the qubit on which it acts. Each contribution has a weighting $q_{\mathbf{g}} = \Pi_i[a_{g_i}/(a_X + a_Y + a_Z)]$. Measuring the generators of $\mathcal{S}_{n-1}$ and postselecting on "+1" outcomes, projects onto the code space of $\mathcal{S}_{n-1}$ with projector $P = \sum_{s \in \mathcal{S}_{n-1}} s/2^{n-1}$, producing

$$\frac{P\rho(f_{\mathbf{a}_P}^S, \mathbf{a}_P)^{\otimes n} P}{\mathrm{tr}[P\rho(f_{\mathbf{a}_P}^S, \mathbf{a}_P)^{\otimes n} P]} = \sum_{\mathbf{g} \in \{X,Y,Z\}^n} q_{\mathbf{g}}' |\Psi_{\mathbf{g}}'\rangle\langle\Psi_{\mathbf{g}}'|, \qquad (4)$$

with projected terms $|\Psi_{\mathbf{g}}'\rangle$ of new weighting $q_{\mathbf{g}}'$. Each $|\Psi_{\mathbf{g}}'\rangle$ has its stabilizer generated by $(G_{\mathbf{g}}, s_1, s_2, \ldots, s_{n-1})$, where $G_{\mathbf{g}}$ is an independent generator that (a) was present in the initial group $G_{\mathbf{g}} \in \mathcal{G}_{\mathbf{g}}$, and (b) commutes with the measurement stabilizers $G_{\mathbf{g}}\mathcal{S}_{n-1} = \mathcal{S}_{n-1}G_{\mathbf{g}}$. In other words, it must be equivalent to one of six logical Pauli operators of the code space. We denote the set of logical operators as $\mathcal{L}$, and its elements $\pm X_L$, $\pm Y_L$, and $\pm Z_L$, and so $G_{\mathbf{g}} \in \mathcal{L} \cdot \mathcal{S}_{n-1}$. This defines a decoding via the Clifford map, $X_L \to X_1$ and $Z_L \to Z_1$. Since there are only six distinct logical states, we can combine many terms in Eq. (4):

$$\frac{P_\rho(f_{\mathbf{a}_P}^S, \mathbf{a}_P)^{\otimes n} P}{\mathrm{tr}[P_\rho(f_{\mathbf{a}_P}^S, \mathbf{a}_P)^{\otimes n} P]} = \sum_{L \in \mathcal{L}} q_L |\Psi_L\rangle\langle\Psi_L|, \qquad (5)$$

where $|\Psi_L\rangle$ has stabilizer generators $(L, s_1, s_2, \ldots, s_{n-1})$. The new weighting is $q_L = \sum q_{\mathbf{g}}'$ with the sum taken over all $\mathbf{g}$ that generate $\mathcal{G}_{\mathbf{g}}$ containing an element $G_{\mathbf{g}} \in L \cdot \mathcal{S}_{n-1}$. We can now state the next lemma:

Lemma 2.—Given n copies of $\rho(f_{\mathbf{a}_P}^S, \mathbf{a}_P)$ projected into the code space of $\mathcal{S}_{n-1}$ and decoded, the output qubit is in the octahedron interior if there exist any two pure states in the initial ensemble, $|\Psi_{\mathbf{g}}\rangle$ and $|\Psi_{\mathbf{g}'}\rangle$ [defined in Eq. (3)], such that both (i) the projected pure states are orthogonal, so that $L \in \mathcal{G}_{\mathbf{g}}$ and $-sL \in \mathcal{G}_{\mathbf{g}'}$ where $L \in \mathcal{L}$ and $s \in \mathcal{S}_{n-1}$; and (ii) upon projection $|\Psi_{\mathbf{g}}\rangle$ and $|\Psi_{\mathbf{g}'}\rangle$ do not vanish, so $q_{\mathbf{g}}' \neq 0$ and $q_{\mathbf{g}'}' \neq 0$.

We prove this lemma by contradiction. From Eq. (2), and $(2f - 1)a_L = (q_L - q_{-L})$, surface states satisfy

$$|q_{X_L} - q_{-X_L}| + |q_{Y_L} - q_{-Y_L}| + |q_{Z_L} - q_{-Z_L}| = 1, \quad (6)$$

and we assume to the contrary that the projected state has this form. Since $q_{\pm L}$ are non-negative reals, we have $|q_L - q_{-L}| = q_L + q_{-L} - 2\min(q_L, q_{-L})$, where $\min(q_L, q_{-L})$ is the minimum of q_L and q_{-L}. Along with the normalization condition, $\sum_L q_L = 1$, this entails

$$\min(q_{X_L}, q_{-X_L}) + \min(q_{Y_L}, q_{-Y_L}) + \min(q_{Z_L}, q_{-Z_L}) = 0.$$

Since all terms are positive, no cancellations can occur and so every term must vanish, hence $\min(q_L, q_{-L}) = 0$, $\forall L$. However, conditions (i) and (ii) of the lemma entail that there exists a nonvanishing $\min(q_L, q_{-L})$, as $q_L \geq q_{\mathbf{g}}' \neq 0$ and $q_{-L} \geq q_{\mathbf{g}'}' \neq 0$. Having arrived at this contradiction, we conclude the falsity of the assumption that the projected state remains on the octahedron surface, and so must be in the octahedron interior. This proves Lemma 2, and we now show that Lemma 2 applies to all stabilizer reductions that do not trivially take $\rho(f, \mathbf{a}_P)^{\otimes n} \to C_i \rho(f, \mathbf{a}_P) C_i^\dagger$.

Our proof continues by finding canonical generators for the code $\mathcal{S}_{n-1}$. A related method has been used to prove that all stabilizer states are local Clifford equivalent to a graph state [16], and we review this first. All stabilizer states have a stabilizer $\mathcal{S}_n$ with n generators. Each generator is a tensor product of n single-qubit Pauli operators. This can be visualized as an n by n matrix with elements that are Pauli operators, each row a generator and each column a qubit. Different, yet equivalent, generators are produced by row multiplication, via which we can produce a canonical form. In this form column i has a nontrivial Pauli operator A_i that appears on the diagonal, and all other operators in that column are either the identity or another operator B_i. Note that A_i and B_i compose a third nontrivial Pauli $A_iB_i = i(-1)^{\gamma_i}C_i$ with $\gamma_i = 0, 1$. Hence, all stabilizer states differ from some graph state by only local Cliffords that map $(A_i, B_i) \to (X_i, Z_i)$.

A code, $\mathcal{S}_{n-1}$, has one less generator than the number of qubits, and so more columns than rows. We can apply the diagonalization procedure on an $n - 1$ by $n - 1$ submatrix, to bring this submatrix into canonical form. Hence, we can find generators of $\mathcal{S}_{n-1}$ such that

$$s_j = (-1)^{\alpha_j} A_j \left(\prod_{k \neq j, n} B_k^{\beta_{k,j}}\right) T_{j,n}, \qquad (7)$$

where the variables $\beta_{k,j} = 0, 1$ denote whether B_k or $\mathbb{1}_k$ is present, and $\alpha_j = 0, 1$ defines the phase. With the nth column out of canonical form, this leaves the nth qubit operator $T_{j,n}$ unspecified. However, if all these generators have $T_{j,n} = \mathbb{1}_n$, then the protocol is trivial and projects $n - 1$ qubits into a known stabilizer state and the last qubit untouched, and so no improvement is made for any f. Hence, herein we assume the nontrivial case; in particular, we assume stabilizer $T_{n-1,n} \neq \mathbb{1}_n$. Since we can always relabel qubits this is completely general. Furthermore, we can define $T_{n-1,n} = A_n$. Now we are ready to define a

logical operator in the code space of $\mathcal{S}_{n-1}$:

$$Z_L = \left(\prod_{1 \leq j \leq n-2} B_j^{\zeta_j} \right) B_{n-1} B_n, \qquad (8)$$

where the variables $\zeta_j = 0, 1$ are uniquely fixed by commutation relations $Z_L s_j = s_j Z_L$. Note that Z_L has some inbuilt freedom as B_n is not fixed other than that $B_n \neq A_n$, $\mathbb{1}_n$, which is equivalent to free choice of γ_n in the expression $A_n B_n = i(-1)^{\gamma_n} C_n$. Now we enquire whether the final state contains two terms stabilized by Z_L and $-sZ_L$, respectively, hence satisfying the conditions for Lemma 2. If we consider the product of Z_L and s_{n-1}, and choose $\gamma_n = a_{n-1} + \gamma_{n-1}$ mod 2, we have

$$- s_{n-1} Z_L = \left(\prod_{1 \leq k \leq n-2} B_k^{\beta_{k,n-1} + \zeta_k} \right) C_{n-1} C_n. \qquad (9)$$

Our choice of γ_n ensures a minus sign on the left hand side, which aids in finding $|\Psi_{\mathbf{g}}\rangle$ and $|\Psi_{\mathbf{g}'}\rangle$ that satisfy our lemma by being stabilized by $G_{\mathbf{g}} = Z_L$ and $G_{\mathbf{g}'} = -s_{n-1} Z_L$, respectively. This criterion is fulfilled when

$$\mathbf{g} = (B_1, B_2, \ldots, B_{n-2}, B_{n-1}, B_n),$$
$$\mathbf{g}' = (B_1, B_2, \ldots, B_{n-2}, C_{n-1}, C_n).$$

These states only vanish under projection, $q'_{\mathbf{g}}, q'_{\mathbf{g}'} = 0$, if they are stabilized by the negative of some element of the code $\mathcal{S}_{n-1}$. To prove they do not vanish, we first observe that every element of $G_{\mathbf{g}}$ and $G_{\mathbf{g}'}$ has either $\mathbb{1}_j$ or B_j acting on qubit j, for all $j = 1, 2, \ldots n - 2$. The only elements of $\mathcal{S}_{n-1}$ for which this is true are $\mathbb{1}$ and s_{n-1}, but s_{n-1} has $A_{n-1} A_n$ acting on the last two qubits and neither $G_{\mathbf{g}}$ or $G_{\mathbf{g}'}$ contain any such element.

Using a canonical form of the generators of $\mathcal{S}_{n-1}$, we have shown that nontrivial codes always satisfy the conditions of Lemma 2. That is, all nontrivial code space projections take many surface states into the octahedron interior. From the continuity expressed by Lemma 1, this entails the existence of a finite region of nonstabilizer states that are also projected into the octahedron. Hence, all n-copy protocols do no improve on a single copy for some region of bound states above the octahedron faces, completing the proof. This does not contradict known tight thresholds in edge directions, as these directions have $\mathbf{a}$ with one zero component.

Although our proof holds for protocols using fixed and finite n copies of $\rho(f, \mathbf{a}_P)$, we could conceive of a protocol that varies n. If this varying-n protocol has an n-dependent threshold, $f_{\mathbf{a}_P}^T(n)$, and $f_{\mathbf{a}_P}^T(n) \to f_{\mathbf{a}_P}^S$ as $n \to \infty$, then its threshold would be arbitrarily suppressible. Repeated iterations of a protocol, or equivalently employing concatenation of a single-qubit code, will not change the threshold. However, one could consider a broader class of protocols consisting of iterates that act on p qubits and output q qubits (for $p > q > 1$) followed by a final round outputting a single qubit. Such protocols map n qubits to 1 qubit, with

n growing each iterate, but with only p qubits involved in each iterate. This implies that multiqubit output iterates may suppress the threshold effectively, and are worth further study. Currently, no such protocol is known. As such, in the asymptotic regime, bound magic states may not exist. However, numerical evidence so far indicates that smaller codes tend to produce better thresholds than larger codes. Nevertheless, the theorem does not rule out infinite cases from attaining a tight threshold. In the regime of finite resources, bound states do exist, and it is interesting ask what computational power Clifford circuits acting on such states possess. Can we find methods of efficiently classically simulating bound states, or can bound states be exploited in algorithms that offer a speedup over classical computation?

Furthermore, our proof assumes a protocol acting on identical copies, which invites study into whether our results extend to nonidentical copies. In particular, following the analogy with entanglement distillation, we speculate that bound magic states may be distillable via "catalysis," where some nonconsumed distillable resource activates the distillation [11]. Finally, we note that noisy Clifford gates can also enable quantum computation [17,18], and we conjecture that a similar theorem will apply to a class of noisy Clifford gates analogous to states just above the octahedron faces.

[1] A. Steane, Proc. R. Soc. A **452**, 2551 (1996).
[2] D. Gottesman, Phys. Rev. A **57**, 127 (1998).
[3] G. Moore and N. Read, Nucl. Phys. **B360**, 362 (1991).
[4] S. Lloyd, Quant. Info. Proc. **1**, 13 (2002).
[5] B. Douçot and J. Vidal, Phys. Rev. Lett. **88**, 227005 (2002).
[6] E. Knill, Nature (London) **434**, 39 (2005).
[7] B. W. Reichardt, Algorithmica **55**, 517 (2009).
[8] R. Raussendorf, J. Harrington, and K. Goyal, New J. Phys. **9**, 199 (2007).
[9] D. Gottesman, Ph.D. thesis, Caltech, arXiv:quant-ph/9705052.
[10] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[11] P. Horodecki, M. Horodecki, and R. Horodecki, Phys. Rev. Lett. **82**, 1056 (1999); D. Jonathan and M. B. Plenio, Phys. Rev. Lett. **83**, 3566 (1999).
[12] B. W. Reichardt, Quant. Info. Proc. **4**, 251 (2005).
[13] B. W. Reichardt, arXiv:quant-ph/0608085.
[14] The T gate performs, $TXT^\dagger = Y$, $TYT^\dagger = Z$.
[15] E. T. Campbell and D. E. Browne, arXiv:0908.0838.
[16] M. Van den Nest, J. Dehaene, and B. D. Moor, Phys. Rev. A **69**, 022316 (2004).
[17] M. B. Plenio and S. Virmani, arXiv:0810.4340.
[18] M. Howard and W. van Dam, Phys. Rev. Lett. **103**, 170504 (2009).