Application of a Resource Theory for Magic States to Fault-Tolerant Quantum Computing

Mark Howard* and Earl Campbell†
Department of Physics and Astronomy, University of Sheffield, Sheffield S3 7RH, United Kingdom
(Received 11 October 2016; published 3 March 2017)

Motivated by their necessity for most fault-tolerant quantum computation schemes, we formulate a resource theory for magic states. First, we show that robustness of magic is a well-behaved magic monotone that operationally quantifies the classical simulation overhead for a Gottesman-Knill-type scheme using ancillary magic states. Our framework subsequently finds immediate application in the task of synthesizing non-Clifford gates using magic states. When magic states are interspersed with Clifford gates, Pauli measurements, and stabilizer ancillas—the most general synthesis scenario—then the class of synthesizable unitaries is hard to characterize. Our techniques can place nontrivial lower bounds on the number of magic states required for implementing a given target unitary. Guided by these results, we have found new and optimal examples of such synthesis.

DOI: 10.1103/PhysRevLett.118.090501

Quantum resource theories attempt to capture what is quintessentially quantum in a piece of technology. For example, entanglement is the relevant resource for quantum cryptography and communication. The resource framework for entanglement finds practical application in bounding the efficiency of entanglement distillation protocols. An abundance of other resource theories have been related to various aspects of quantum theory [1–8]. Once a quantum computer is made fault-tolerant, some computational operations become relatively easy, and some more difficult, leading to a natural resource picture called the magic state model [9,10] (although, alternative routes to fault-tolerant universality exist [11]). Preparation of stabilizer states and implementation of Clifford unitaries and Pauli measurements constitute free resources. Difficult operations include preparation of magic states, a supply of which is necessary in order to promote the easier operations to a universal gate set. With only free resources, the computation can be efficiently classically simulated, whereas with a liberal supply of pure magic states, universal quantum computation is unlocked. For qudit (d-level) quantum computers with odd d, a resource theory of magic (or equivalently contextuality with respect to stabilizer measurements [12,13]) has been developed [7,14,15]. This relies on a well-behaved discrete Wigner function [16], which, in turn, relies on quirks of odd dimensional Hilbert space. Here, we address the most practically important case by quantifying the magic for multiqubit systems, relating this resource measure to simulation complexity and applying the resource theory to the practical problem of gate synthesis.

The canonical magic state is $|H\rangle = (|0\rangle + e^{i\pi/4}|1\rangle)/\sqrt{2}$, which enables application of a single-qubit unitary $T = \mathrm{diag}(1, e^{i\pi/4})$ [9,10]. A circuit composed of elements from the Clifford $+\,T$ gate set acting on the standard computational basis input suffices for universal quantum computation. Such a circuit can be classically simulated, but in a time that scales exponentially in the number of T gates [17]. Faster simulation algorithms were recently discovered that relate the simulation complexity to the stabilizer rank [18–20], a measure of magic for pure states. Such techniques do not naturally adapt to mixed magic states, and stabilizer rank is qualitatively very different to the magic measure we establish here. For quantum computations using qudits with odd dimension, the discrete Wigner function provides a quasiprobability distribution, and Pashayan *et al.* [21] showed that the negativity quantifies the simulation complexity. Here, we provide a general simulation scheme, which can be naturally applied to mixed-state qubit quantum computations using any kind of ancillary magic state (e.g., a multiqubit magic state enabling a Toffoli gate). Furthermore, for many problems, our approach is competitive with comparable schemes based on stabilizer rank [18,19].

Because of the high price assigned to $|H\rangle$ states and, hence, T gates, it behooves us to find Clifford $+\,T$ circuit implementations of quantum algorithms that are parsimonious in their use of T gates. The topic of circuit synthesis has made tremendous progress in recent years [22–30], compared with Solovay-Kitaev type constructions that were for a long time the standard benchmark. Developments include identifying special algebraic forms for all gates that can be unitarily synthesized over the Clifford $+\,T$ gate set [24] or over the smaller CNOT $+\,T$ gate set [22,23]. However, circuit synthesis need not be a purely unitary process and, more generally, may be aided by ancillary stabilizer states, measurements, and classical feedforward. There are hints that general synthesis can be significantly more powerful [26,27,30,31], though the paradigm is not well understood. Our resource framework helps with this problem by establishing nontrivial lower bounds on the number of $|H\rangle$ states, or equivalently T gates, required for the general synthesis scenario. This allows us to identify several circuits as optimal. Such resource-theoretic tools work for any kind of state, not just $|H\rangle$, but they are particularly well motivated for magic states from the third level of the Clifford hierarchy, e.g., Toffoli resource states. We note how general synthesis has a curious relationship to Clifford equivalence of magic states. From this vantage point, we discover several new examples of general synthesis protocols with resource savings over previous unitary synthesis methods.

The Supplemental Material [32] (text, which includes Refs. [33–37], and files) contain numerous additional computations. Results include identifying the most robust states, most robust gates, and classification of all three and four qubit diagonal gates from the third level of the Clifford hierarchy. Intriguingly, one maximally robust state is the Hoggar [38] fiducial state.

Robustness of Magic.—Vidal and Tarrach [8] showed that the amount of separable noise that makes an entangled state become separable is an entanglement monotone, which they called robustness. The basic principle can be adapted for use in other resource theories with a set of free states. Denoting $\mathcal{S}_n = \{\sigma_i\}$ as the set of pure n-qubit stabilizer states, we define the robustness of magic (ROM) as

$$\text{(ROM)} \quad \mathcal{R}(\rho) = \min_x \left\{ \sum_i |x_i|; \rho = \sum_i x_i \sigma_i \right\}. \tag{1}$$

Decompositions of the form $\sum_i x_i \sigma_i$ are called stabilizer pseudomixtures. We have $\sum_i x_i = 1$, but x_i may be negative, and so they provide a quasiprobability distribution. The optimization in (1) can be rewritten in terms of a linear system as

$$\mathcal{R}(\rho) = \min \|x\|_1 \text{ subject to } Ax = b, \tag{2}$$

where $\|x\|_1 = \sum_i |x_i|$, $b_i = \mathrm{Tr}(P_i \rho)$ and $A_{j,i} = \mathrm{Tr}(P_j \sigma_i)$ where P_j is the jth Pauli operator. For example, consider the single-qubit magic state $|H\rangle = (|0\rangle + e^{i\pi/4}|1\rangle)/\sqrt{2}$, then in the Pauli operator basis

$$A = \begin{array}{c} \langle \mathbb{1} \rangle \\ \langle X \rangle \\ \langle Y \rangle \\ \langle Z \rangle \end{array} \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{pmatrix},$$

and the solution of (2) is $x = (\sqrt{2}, 0, 1, 1-\sqrt{2}, 0, 0)/2$ implying $\mathcal{R}(|H\rangle) = \sqrt{2}$. There are a number of freely available solvers for linear programs [39,40], which are efficient in the size of A. From our formulation of the problem, it is clear that $\min_{Ax=b}\|x\|_1$ is feasible and bounded. Consequently, strong duality holds, i.e.,

$$\min_{Ax=b} \|x\|_1 = \max_{\|A^T y\|_\infty \le 1} -b^T y, \tag{3}$$

and the aforementioned solvers can provide a certificate y of optimality [41]. Despite the theoretical efficiency of the linear programing problem, the number of stabilizer states in $\mathcal{P}_n$ scales superexponentially with n, so that $|\mathcal{S}_n| = 2^n \prod_{j=1}^n (2^j + 1)$ [16]. Practically, we are limited to $1 \le n \le 5$ qubits. We have made available the corresponding A matrices in the Supplemental Material [32].

Robustness of magic possesses all the desirable qualities of a resource theoretic measure (see Supplemental Material [32] for proofs of the following). For a mixed stabilizer state, we find $x_i > 0$ entailing $\sum_i |x_i| = \sum_i x_i = 1$. For a nonstabilizer state, at least one x_i is negative, and then, ROM must exceed unity. Therefore, ROM is faithful. Crucially, ROM is nonincreasing under stabilizer operations, the free set of operations in the resource theory. Finally, ROM is submultiplicative,

$$\mathcal{R}(\rho_1 \otimes \rho_2) \le \mathcal{R}(\rho_1)\mathcal{R}(\rho_2), \tag{4}$$

which follows by using the minimal stabilizer pseudomixtures for ρ_1 and ρ_2 to construct a not-necessarily-minimal stabilizer pseudomixture for $\rho_1 \otimes \rho_2$. Useful lower bounds on $\mathcal{R}(\rho_1 \otimes \rho_2)$ can also be obtained (see Supplemental Material [32]).

Resource theoretic frameworks are commonplace in quantum information theory but do not always directly lend themselves to an operational meaningful interpretation or to useful applications in solving relevant problems. In the next section, we show how ROM quantifies the exponential simulation overhead for a version of the Gottesman-Knill protocol where nonstabilizer ancillas ρ can be added to an, otherwise, stabilizer circuit. The subsequent section discusses ROM's application to the task of implementing non-Clifford operations in an economical way.

Robustness quantifies classical simulation overhead.—The Gottesman-Knill (GK) theorem shows that, for any stabilizer circuit written as a superoperator $\mathcal{E}$ and pure stabilizer state σ_i, we can efficiently sample from the outcome of a Pauli measurement on $\mathcal{E}(\sigma_i)$. By collecting many samples, we can estimate the expectation value to any desired accuracy. If the input state is a probabilistic ensemble of stabilizer states, the Gottesman-Knill theorem still holds provided we can efficiently sample from the ensemble. A Clifford + T circuit suffering very heavy noise can be simulated in this way [42–44]. Here, we provide a simulation algorithm for estimating an expectation value $P_\rho \coloneqq \mathrm{tr}[P\mathcal{E}(\rho)]$ where $\mathcal{E}$ is a stabilizer operation. The simulation time cost scales with $\sum_i |x_i| \ge \mathcal{R}(\rho)$ where x_i are the quasiprobabilities used (which may be suboptimal). First, as in [21], we use the quasiprobability distribution x_i to form the probability distribution

$p_i = |x_i|/\sum_i |x_i|$. We sample an i value from this probability distribution and use GK to obtain an eigenvalue $m = \pm 1$ for the Pauli measurement on $\mathcal{E}(\sigma_i)$. Our simulation outputs not m, but $M = \mathrm{sign}(x_i) m \sum_i |x_i|$ where $\mathrm{sign}(x_i) = 1$ for $x_i > 0$ and $\mathrm{sign}(x_i) = -1$, otherwise. Notice that each run outputs $\pm \sum_i |x_i|$ and not ± 1, which leads to a larger variance of our random variable. We repeat this sampling process many times and find the mean value of M, which gives an unbiased estimator of P_ρ. The Hoeffding inequalities show that, for random variables bounded in the interval $[-\sum_i |x_i|, +\sum_i |x_i|]$, N samples will estimate the mean to within δ of the actual mean with probability exceeding $1 - \epsilon$ where $\epsilon = 2\exp[-N\delta^2/2(\sum_i |x_i|)^2]$. In other words, the desired accuracy is guaranteed by using

$$N = \frac{2}{\delta^2}\left(\sum_i |x_i|\right)^2 \ln\left(\frac{2}{\epsilon}\right) \tag{5}$$

samples. Using an optimal stabilizer pseudomixture, the number of samples scales quadratically in the robustness, though the robustness typically scales exponentially in the number of magic states. For each of these samples, the GK scheme requires a polynomial amount of time, provided we know how to efficiently sample from the quasiprobability distribution.

Nonstabilizer circuits.—For any quantum circuit, we can find an equivalent gadgetized version [9,18,19] over the Clifford plus T gate set; all uses of T are replaced with the standard state injection circuit whereby a $|H\rangle$ state is entangled with a data qubit and subsequently measured out (see Fig. 1 for an example). The T gadget is just one example from an infinitely large family of similar gadgets. All diagonal gates from the third level of the Clifford hierarchy—the set of gates that map Pauli operators to Clifford gates under conjugation—are also suitable for gadgetization. These gates are sufficient for promoting Clifford gates to universality and have the added property that access to the state $|U\rangle = U|+\rangle^{\otimes k}$ allows for deterministic implementation of the gate U [45,46], as depicted in Fig. 1. This family includes important multiqubit gates such as the control-control-Z (CCZ), which is the Clifford equivalent to the Toffoli, and control-S (CS) where $S = T^2$. Therefore, a quantum circuit $C_1 U_1 C_2 U_2 \ldots U_N C_{N+1}$ where C_i is the Clifford equivalent to a stabilizer circuit consuming the resource $|U\rangle$ where $U \coloneqq \otimes U_j$. We remark that diagonal, third-level gates are exactly synthesizable from CNOT and T gates [22,23,47,48].

Large resource states.—For large resource states, the exact robustness may be difficult to determine. However, instead of using the optimal robustness, we use stabilizer pseudomixtures built up from constant sized blocks of qubits, here, limited to five qubits per block. For instance, given $t = bm$ copies of the $|H\rangle$ state, we break it into b blocks of m qubits $(|H\rangle^{\otimes m})^{\otimes b}$ and work with a pseudomixture whose sample complexity scales as $\sum_i |x_i| = \mathcal{R}(|H\rangle^{\otimes m})^b = [\mathcal{R}(|H\rangle^{\otimes m})^{1/m}]^t$.

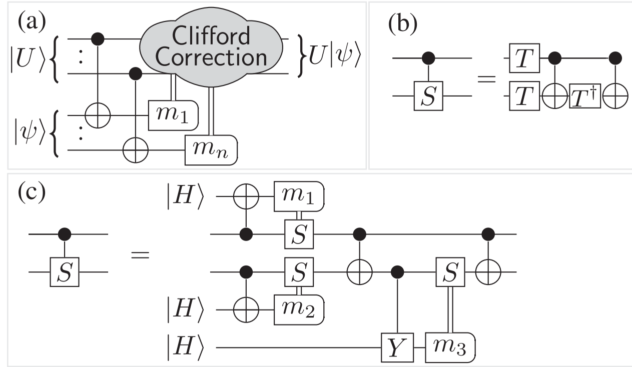

FIG. 1. (a) Half-teleportation gadget for implementing diagonal U in the third level of the Clifford hierarchy. The circuit uses a resource state $|U\rangle = U|+\rangle^{\otimes n}$. (b) Exact synthesis of CS gate using T and CNOT gates. (c) The CS circuit as a gadgetized circuit using three $|H\rangle$ magic states. Our techniques show this synthesis is provably optimal.

Numerical results.—We performed substantial numerical investigations up to five-qubit systems, for which there are over two million stabilizer states and over one thousand Pauli operators. For ascending $m \leq 5$, we calculated $\mathcal{R}(|H^{\otimes m}\rangle)^{1/m}$ as $\{1.414, 1.322, 1.304, 1.301, 1.298\}$. The decrease with m shows a strongly submultiplicative behavior and reduces simulation overheads (though analytic lower bounds derived in the Supplemental Material [32] show going to higher m cannot reduce this value below 1.207). Specifically, $\mathcal{R}(|H^{\otimes 5}\rangle)^{2t/5} \approx 1.298^{2t} = 1.685^t$ characterizes the complexity of our Clifford plus t T gate simulation. This gives exponential improvement over the method used in [19] with complexity scaling as 1.9185^t. A more efficient use of the stabilizer Schmidt decomposition was subsequently established in [18], leading to complexity scaling as 1.385^t, although the restriction to pure states, as in [19], also holds here. Even better scaling with t can be obtained by using approximate states [18], but at the price of δ^{-5} overhead in the precision δ. A quantum computation using z CCZ gates, can be implemented using $4z$ T gates [31], implying our simulation complexity scales as 8.067^z. However, as discussed earlier, we can use $|\mathrm{CCZ}\rangle$ resource states. We found $\mathcal{R}(|\mathrm{CCZ}\rangle) = 2.555$, and so z CCZs can be simulated with an overhead dominated by $\mathcal{R}(|\mathrm{CCZ}\rangle)^{2z} = 6.531^z$. This gives exponential improvement over using the four-T-gate gadgetization. Alternative gadgetizations using $|U\rangle = U|+\rangle$ with third-level diagonal U follow naturally, and the simulation overhead is given by $\mathcal{R}(|U\rangle)$, see Supplemental Material [32] for many possible examples.

Lower bounds on gate synthesis.—Beyond classical simulation, robustness can help us investigate gate synthesis. Above, we noted that four T gates can exactly synthesize a CCZ gate. How can we be sure that a more complicated or clever scheme does not use even fewer T gates, not just for CCZ gates, but more generally? This is an important instance where our resource theory of magic applies; a potential resource state $|U\rangle = U|+\rangle$ for a non-Clifford unitary U cannot be made using t T gates if $\mathcal{R}(|U\rangle) > \mathcal{R}(|H^{\otimes t}\rangle)$. The resource state $|\mathrm{CCZ}\rangle$, Clifford equivalent to a Toffoli resource, has $\mathcal{R}(|\mathrm{CCZ}\rangle) = 2.555$ implying

$$\mathcal{R}(|H^{\otimes 3}\rangle) < \mathcal{R}(|\mathrm{CCZ}\rangle) < \mathcal{R}(|H^{\otimes 4}\rangle). \tag{6}$$

Therefore, we establish that the four T gate synthesis of CCZ is optimal. Furthermore, the standard decomposition of CS into 3 T gates (see Fig. 1) is provably optimal since $\mathcal{R}(|H^{\otimes 2}\rangle) < \mathcal{R}(|\mathrm{CS}\rangle) < \mathcal{R}(|H^{\otimes 3}\rangle)$. The numerical results in the Supplemental Material [32] allow for the inference of optimality for many more unitaries.

Improved gate synthesis.— Circuit synthesis can be purely unitary over the Clifford + T gate set or, more generally, can make use of stabilizer ancillas and measurement, thereby using even fewer T gates. The scale of the potential savings is exemplified by the four T gate realization of CCZ [31], which is assisted by ancillas and measurement. Purely unitary synthesis of CCZ over Clifford + T is known to need at least seven T gates [24]. We shed new light on this phenomena by showing that it emerges from Clifford equivalence of magic states, and give new examples of improved synthesis. The interesting examples arise when $C|U\rangle = |V\rangle$, and yet, unitary synthesis of U uses fewer T gates than V. For these remarkable examples, despite Clifford equivalence of states $|U\rangle$ and $|V\rangle$, there do not exist Cliffords C_1 and C_2 such that $C_1 U C_2 = V$. One explicit example, comparable to Jones' construction [31], starts with the Toff* gate corresponding to $\mathrm{CCZ}_{123}\mathrm{CS}_{12}$, which is known to be unitarily synthesizable using four T gates [29]. Because the "square-root-of-NOT" $\sqrt{X}$ is a Clifford gate, we also have that

$$(II\sqrt{X})|\mathrm{Toff}^*\rangle = |\mathrm{CCZ}\rangle. \tag{7}$$

Clifford equivalence of magic states provides an alternative proof that CCZ can be performed with four T gates.

We have found a number of similar examples using the following method: (i) Identify U and V with different T counts, but whose states have the same robustness $\mathcal{R}(|U\rangle) = \mathcal{R}(|V\rangle)$; (ii) Search for the Clifford C that takes $|U\rangle$ to $|V\rangle$. Existence of such a Clifford is not guaranteed by virtue of $\mathcal{R}(|U\rangle) = \mathcal{R}(|V\rangle)$, but we found a Clifford in every instance investigated. Note, also, that our T count is over the CNOT + T basis, which is less general than the Clifford + T basis, but existing techniques for the latter [24] are impractical for more than three qubits. The two methods (Clifford + T and CNOT + T) give the same T count for CCZ, and it is an interesting open question whether they always agree on the T cost of synthesizing a third-level gate. We list, in compact notation, a few of the new synthesis results and the T savings (more are provided in the Supplemental Material [32]). For example, the CCZ construction discussed above would be represented as

$$\mathrm{CCZ}_{123} \xrightarrow{7\to 4} \mathrm{CS}_{12}\mathrm{CCZ}_{123}, \tag{8}$$

where the subscripts denote the qubits on which a third-level gate acts and the numbers above the arrow denote the T cost. Other examples include

$$\mathrm{CCZ}_{123} \xrightarrow{7\to 4} \mathrm{CS}_{12}\mathrm{CS}_{13}, \tag{9}$$

$$\mathrm{CCZ}_{123,145} \xrightarrow{11\to 8} \mathrm{CS}_{12,13,14,15}, \tag{10}$$

$$T_{1,2,3}\mathrm{CS}_{12,23,13} \xrightarrow{6\to 5} T_{2,3}\mathrm{CS}_{12,23,13}. \tag{11}$$

Discussion.—Reformulating robustness as an optimization in (2) facilitates a comparison with recent related works, see Table I. For qudit-based computation, Veitch *et al.* [7] showed that sum negativity $sn(\rho)$ of a state's discrete Wigner function was a well-defined resource, and Pashayan *et al.* [21] showed how the run time of a Monte Carlo type sampling algorithm was slower by a factor quadratic in the size of the sum negativity. In the qudit setting, the natural choice for the columns of A in Eq. (2) are the vertices of the Wigner polytope (a larger, but more geometrically simple object than the stabilizer polytope), and phase point operators form a natural operator basis. With these choices, b is a vectorized version of the Wigner representation of ρ, and the matrix A becomes the identity matrix. $\mathcal{R}(\rho)$ is simply the sum negativity (equal to the ℓ_1 norm) of the Wigner quasiprobability distribution associated with ρ. In other words, sum negativity is just robustness relative to the set of operators with non-negative discrete Wigner function. Unlike our approach, the discrete Wigner function is not easily adapted to qubits (although, see [49,50]).

In work by Bravyi, Smith, and Smolin [19] t-fold copies of $|H\rangle$ are decomposed as linear combinations of stabilizer vectors [20]; the number of terms χ—the stabilizer Schmidt rank—in the decomposition quantifies the simulation overhead. Finding the optimal decomposition is an ℓ_0 minimization ($\|x\|_0 = |\{i : x_i \neq 0\}|$), which is nonconvex and Non-deterministic polynomial-time hard, limiting calculations to a small number of qubits. Bravyi and Gossett [18]

TABLE I. Restatement of related work in terms of norm-minimizing solutions of a system of equations $Ax = b$. The amount of resource in an ancillary state $|\psi\rangle$ or ρ quantifies the classical simulation overhead. In the first and third lines, the columns of A are n-qubit stabilizer states (as complex vectors or generalized Bloch vectors, respectively). In the second line, the columns of A are extreme points of the Wigner polytope.

Resource	$\|\psi\rangle \in \mathbb{C}^d$	$\rho \in B(\mathcal{H})$
$\chi(\|\psi\rangle) = \|x\|_0$	Refs. [18,19]	
$sn(\rho) = \|x\|_1$		Refs. [7,21]
$\mathcal{R}(\rho) = \|x\|_1$		This Letter

extended this analysis by efficiently finding approximate decompositions that are still sufficient for the task of simulating the outcome of a quantum algorithm. This approximation precludes the possibility of ordering states by the amount of resource, however. We note that it is well known in the signal processing literature [51] that the solution to ℓ_1 minimization also provides a (qualitatively) good solution for ℓ_0 minimization.

In Bravyi, Smith, and Smolin [19], it is conjectured that $|H^{\otimes t}\rangle$ has the smallest χ of all t-fold copies of a single-qubit nonstabilizer state. In this Letter, we see that $|H^{\otimes t}\rangle$ has relatively large robustness. This is curious and worthy of further investigation but is also strongly reminiscent of [52] where the (entanglement) Schmidt rank is seen to disagree with almost every other continuous entanglement measure. A related open problem is to reconcile the fact that small angle ancillae $(1, e^{i\phi\approx 0})/\sqrt{2}$ are cheap in our framework, yet are harder to synthesize over the Clifford $+ T$ gate [28] set and harder to fault-tolerantly distill [26,53]. Considerations such as this suggest that a combination of both the stabilizer Schmidt rank and robustness pictures of magic could prove useful.

The authors thank Naomi Nickerson for reviving our interest in this problem and acknowledge funding from Engineering and Physical Sciences Research Council (Grant No. EP/M024261/1).

*m.howard@sheffield.ac.uk
†earltcampbell@gmail.com

[1] F. G. S. L. Brandão and G. Gour, Phys. Rev. Lett. **115**, 070503 (2015).
[2] B. Coecke, T. Fritz, and R. W. Spekkens, Inf. Comput. **250**, 59 (2016).
[3] A. Grudka, K. Horodecki, M. Horodecki, P. Horodecki, R. Horodecki, P. Joshi, W. Kłobus, and A. Wójcik, Phys. Rev. Lett. **112**, 120401 (2014).
[4] M. Horodecki and J. Oppenheim, Int. J. Mod. Phys. B **27**, 1345019 (2013).
[5] C. Napoli, T. R. Bromley, M. Cianciaruso, M. Piani, N. Johnston, and G. Adesso, Phys. Rev. Lett. **116**, 150502 (2016).
[6] D. Stahlke, Phys. Rev. A **90**, 022302 (2014).
[7] V. Veitch, S. A. H. Mousavian, D. Gottesman, and J. Emerson, New J. Phys. **16**, 013009 (2014).
[8] G. Vidal and R. Tarrach, Phys. Rev. A **59**, 141 (1999).
[9] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[10] E. Knill, Nature (London) **434**, 39 (2005).
[11] E. T. Campbell, B. M. Terhal, and C. Vuillot, arXiv:1612.07330.
[12] M. Howard, J. Wallman, V. Veitch, and J. Emerson, Nature (London) **510**, 351 (2014).
[13] N. Delfosse, C. Okay, J. Bermejo-Vega, D. E. Browne, and R. Raussendorf, arXiv:1610.07093.
[14] V. Veitch, C. Ferrie, D. Gross, and J. Emerson, New J. Phys. **14**, 113011 (2012).
[15] A. Mari and J. Eisert, Phys. Rev. Lett. **109**, 230503 (2012).
[16] D. Gross, J. Math. Phys. (N.Y.) **47**, 122107 (2006).
[17] S. Aaronson and D. Gottesman, Phys. Rev. A **70**, 052328 (2004).
[18] S. Bravyi and D. Gosset, Phys. Rev. Lett. **116**, 250501 (2016).
[19] S. Bravyi, G. Smith, and J. A. Smolin, Phys. Rev. X **6**, 021043 (2016).
[20] H. J. García, I. L. Markov, and A. W. Cross, Quantum Inf. Comput. **14**, 683 (2014).
[21] H. Pashayan, J. J. Wallman, and S. D. Bartlett, Phys. Rev. Lett. **115**, 070501 (2015).
[22] M. Amy, D. Maslov, and M. Mosca, IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst. **33**, 1476 (2014).
[23] M. Amy and M. Mosca, arXiv:1601.07363.
[24] D. Gosset, V. Kliuchnikov, M. Mosca, and V. Russo, Quantum Inf. Comput. **14**, 1261 (2014).
[25] A. Bocharov, Y. Gurevich, and K. M. Svore, Phys. Rev. A **88**, 012313 (2013).
[26] G. Duclos-Cianci and K. M. Svore, Phys. Rev. A **88**, 042325 (2013).
[27] A. Paetznick and K. M. Svore, Quantum Inf. Comput. **14**, 1277 (2014).
[28] N. J. Ross and P. Selinger, arXiv:1403.2975.
[29] P. Selinger, Phys. Rev. A **87**, 042302 (2013).
[30] N. Wiebe and M. Roetteler, Quantum Inf. Comput. **16**, 134 (2016).
[31] C. Jones, Phys. Rev. A **87**, 022328 (2013).
[32] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.118.090501 for an additional file with proofs, discussion and calculations. It also contains plain text files 3QubitCalculations.txt and 4QubitCalculations.txt, which are the results of numerous robustness calculations. Finally, files Amat2.mat, Amat3.mat, Amat4.mat, and Amat5.mat correspond to the A matrices for the linear system $Ax = b$, which includes Refs. [33–37].
[33] D. Andersson, I. Bengtsson, K. Blanchfield, and H. B. Dang, J. Phys. A **48**, 345301 (2015).
[34] B. W. Reichardt, Quantum Inf. Comput. **9**, 1030 (2009).
[35] E. T. Campbell, Phys. Rev. A **83**, 032317 (2011).
[36] A. W. Harrow and M. A. Nielsen, Phys. Rev. A **68**, 012308 (2003).
[37] W. van Dam and M. Howard, Phys. Rev. A **83**, 032310 (2011).
[38] S. G. Hoggar, Geometriae Dedicata **69**, 287 (1998).
[39] M. Grant and S. Boyd, in *Recent Advances in Learning and Control*, edited by V. Blondel, S. Boyd, and H. Kimura, Lecture Notes in Control and Information Sciences (Springer-Verlag Limited, Berlin, 2008), pp. 95–110.
[40] M. Grant and S. Boyd, CVX: MATLAB Software for Disciplined Convex Programming, Version 2.1 (2014).
[41] This certificate of optimality can be used to obtain a magic witness—an operator whose expectation value with respect to stabilizer states is in the interval $[-1, 1]$ and whose expectation with respect to ρ is $\mathcal{R}(\rho)$. These witnesses can be used to derive exact expressions for robustness as in the Supplemental Material [32].
[42] H. Buhrman, R. Cleve, M. Laurent, N. Linden, A. Schrijver, and F. Unger, in *47th Annual IEEE Symposium on Foundations of Computer Science: FOCS 2006, Berkeley, California* (IEEE Computer Society, Los Alamitos, CA, 2006), pp. 411–419.

[43] S. Virmani, S. F. Huelga, and M. B. Plenio, Phys. Rev. A **71**, 042328 (2005).
[44] M. B. Plenio and S. Virmani, New J. Phys. **12**, 033012 (2010).
[45] D. Gottesman and I. L. Chuang, Nature (London) **402**, 390 (1999).
[46] X. Zhou, D. W. Leung, and I. L. Chuang, Phys. Rev. A **62**, 052316 (2000).
[47] E. T. Campbell and M. Howard, arXiv:1606.01904, [Phys. Rev. A (to be published)].
[48] E. T. Campbell and M. Howard, arXiv:1606.01906, [Phys. Rev. Lett. (to be published)].
[49] N. Delfosse, P. Allard Guerin, J. Bian, and R. Raussendorf, Phys. Rev. X **5**, 021003 (2015).
[50] R. Raussendorf, D. E. Browne, N. Delfosse, C. Okay, and J. Bermejo-Vega, arXiv:1511.08506.
[51] E. Candes and T. Tao, IEEE Trans. Inf. Theory **51**, 4203 (2005).
[52] M. Van den Nest, Phys. Rev. Lett. **110**, 060504 (2013).
[53] E. T. Campbell and J. O'Gorman, Quantum Sci. Technol. **1**, 015007 (2016).