Robert Raussendorf

Statement

and

Readings

Decoherence in quantum computation - foe or friend?

Robert Raussendorf, University of British Columbia

Abstract: Decoherence is detrimental to quantum computation because it makes the computation "noisy". Or is it? Upon closer inspection, it turns out that decoherence can both compromize and help realize quantum computation. Which of the two applies does very much depend on the decoherence model considered.

I will start out by proving the expected, namely that decoherence, for a certain (justifiable) class of decoherence models, does indeed compromise quantum computation. In this regard, I will review a result of Bravyi and Kitaev [1b]/ van Dam and Howard [1a] demonstrating an *upper* bound to the error threshold for fault-tolerant quantum computation. The significance of this upper bound is that no method of error correction, however clever, can put the quantum computation back on track if the decoherence level per elementary gate operation is above the threshold value.

In the second part of my introduction, I will discuss two computational models [2], [3] that *use* decoherent dynamics to realize quantum computation. In the case of [2], the computation is driven by local projective measurements on a highly entangled quantum state. Therein, the entanglement of the initial quantum state is progressively destroyed as the computation proceeds. Thus, entanglement is a resource for this computational model. In the second case, Ref. [3], universal quantum computation is implemented in a dissipative quantum system whose evolution is governed by time-independent and local couplings to the environment. Due to the purely dissipative nature of the process, this way of doing quantum computation exhibits some inherent robustness and defies some of the DiVincenzo criteria for quantum computation.

Suggested Reading:

[1a] Sergey Bravyi and Alexei Kitaev, Phys. Rev. A **71**, 022316 (2005).

[1b] Wim van Dam and Mark Howard, Phys. Rev. Lett. **103**, 170504 (2009).

[2] R. Raussendorf and H.J Briegel, Phys. Rev. Lett. **86**, 5188 (2001).

[3] F. Verstraete, M. Wolf and J.I. Cirac, Nature Physics **5**, 633 - 636 (2009).

Remark: The results of refs. [1a] and [1b] are very closely related. For background reading, I recommend Ref. [1b] over [1a] because it is shorter. In my introduction, I will discuss [1a] (first part only), however, because the result therein is better suited for graphical display.

Universal quantum computation with ideal Clifford gates and noisy ancillas

Sergey Bravyi* and Alexei Kitaev†
Institute for Quantum Information, California Institute of Technology, Pasadena, 91125 California, USA

(Received 6 May 2004; published 22 February 2005)

We consider a model of quantum computation in which the set of elementary operations is limited to Clifford unitaries, the creation of the state $|0\rangle$, and qubit measurement in the computational basis. In addition, we allow the creation of a one-qubit ancilla in a mixed state ρ, which should be regarded as a parameter of the model. Our goal is to determine for which ρ universal quantum computation (UQC) can be efficiently simulated. To answer this question, we construct purification protocols that consume several copies of ρ and produce a single output qubit with higher polarization. The protocols allow one to increase the polarization only along certain "magic" directions. If the polarization of ρ along a magic direction exceeds a threshold value (about 65%), the purification asymptotically yields a pure state, which we call a magic state. We show that the Clifford group operations combined with magic states preparation are sufficient for UQC. The connection of our results with the Gottesman-Knill theorem is discussed.

DOI: 10.1103/PhysRevA.71.022316 PACS number(s): 03.67.Lx, 03.67.Pp

I. INTRODUCTION AND SUMMARY

The theory of fault-tolerant quantum computation defines an important number called the error threshold. If the physical error rate is less than the threshold value δ, it is possible to stabilize computation by transforming the quantum circuit into a fault-tolerant form where errors can be detected and eliminated. However, if the error rate is above the threshold, then errors begin to accumulate, which results in rapid decoherence and renders the output of the computation useless. The actual value of δ depends on the error correction scheme and the error model. Unfortunately, this number seems to be rather small for all known schemes. Estimates vary from 10^{-6} (see Ref. [1]) to 10^{-4} (see Refs. [2–4]), which is hardly achievable with the present technology.

In principle, one can envision a situation in which qubits do not decohere, and a subset of the elementary gates is realized *exactly* due to special properties of the physical system. This scenario could be realized experimentally using spin, electron, or other many-body systems with topologically ordered ground states. Excitations in two-dimensional topologically ordered systems are anyons—quasiparticles with unusual statistics described by nontrivial representations of the braid group. If we have sufficient control of anyons, i.e., are able to move them around each other, fuse them, and distinguish between different particle types, then we can realize some set of unitary operators and measurements exactly. This set may or may not be computationally universal. While the universality can be achieved with sufficiently nontrivial types of anyons [5–8], more realistic systems offer only decoherence protection and an incomplete set of topological gates. (See Refs. [9,10] about non-Abelian anyons in quantum Hall systems and Refs. [11,12] about topological orders in Josephson junction arrays.) Nevertheless, universal computation is possible if we introduce some additional operations (e.g., measurements by Aharonov-Bohm interference [13] or some gates that are not related to topology at all). Of course, these nontopological operations cannot be implemented exactly and thus are prone to errors.

In this situation, the threshold error rate δ may become significantly larger than the values given above because we need to correct only errors of certain special type and we introduce a smaller amount of error in the correction stage. The main purpose of the present paper is to illustrate this statement by a particular computational model.

The model is built upon the *Clifford group*—the group of unitary operators that map the group of Pauli operators to itself under conjugation. The set of elementary operations is divided into two parts: $\mathcal{O}=\mathcal{O}_{\text{ideal}}\cup\mathcal{O}_{\text{faulty}}$. Operations from $\mathcal{O}_{\text{ideal}}$ are assumed to be perfect. We list these operations below:

(i) prepare a qubit in the state $|0\rangle$;

(ii) apply unitary operators from the Clifford group;

(iii) measure an eigenvalue of a Pauli operator (σ^x,σ^y, or σ^z) on any qubit.

Here we mean nondestructive projective measurement. We also assume that no errors occur between the operations.

It is well known that these operations are not sufficient for universal quantum computation (UQC) (unless a quantum computer can be efficiently simulated on a classical computer). More specifically, the Gottesman-Knill theorem states that by operations from $\mathcal{O}_{\text{ideal}}$ one can only obtain quantum states of a very special form called *stabilizer states*. Such a state can be specified as an intersection of eigenspaces of pairwise commuting Pauli operators, which are referred to as *stabilizers*. Using the stabilizer formalism, one can easily simulate the evolution of the state and the statistics of measurements on a classical probabilistic computer (see Ref. [14] or a textbook [15] for more details).

The set $\mathcal{O}_{\text{faulty}}$ describes faulty operations. In our model, it consists of just one operation: prepare an ancillary qubit in a mixed state ρ. The state ρ should be regarded as a parameter of the model. From the physical point of view, ρ is mixed due to imperfections of the preparation procedure (entanglement of the ancilla with the environment, thermal fluctua-

*Email address: serg@cs.caltech.edu

†Email address: kitaev@iqi.caltech.edu

tions, etc.). An essential requirement is that by preparing n qubits we obtain the state $\rho^{\otimes n}$, i.e., all ancillary qubits are independent. The independence assumption is similar to the uncorrelated errors model in the standard fault-tolerant computation theory.

Our motivation for including all Clifford group gates into $\mathcal{O}_{\rm ideal}$ relies mostly on the recent progress in the fault-tolerant implementation of such gates. For instance, using a concatenated stabilizer code with good error correcting properties to encode each qubit and applying gates transversally (so that errors do not propagate inside code blocks) one can implement Clifford gates with an arbitrary high precision, see Ref. [16]. However, these nearly perfect gates act on *encoded* qubits. To establish a correspondence with our model, one needs to prepare an *encoded* ancilla in the state ρ. It can be done using the schemes for fault-tolerant encoding of an arbitrary *known* one-qubit state described by Knill in Ref. [17]. In the more recent paper [18] Knill constructed a scheme of fault-tolerant quantum computation which combines (i) the teleported computing and error correction technique by Gottesman and Chuang [19]; (ii) the method of purification of CSS states by Dür and Briegel [20]; and (iii) the magic states distillation algorithms described in the present paper. As was argued in Ref. [18], this scheme is likely to yield a much higher value for the threshold δ (it may be up to 1%).

Unfortunately, ideal implementation of the Clifford group cannot be currently achieved in any realistic physical system with a topological order. What universality classes of anyons allow one to implement all Clifford group gates (but do not allow one to simulate UQC) is an interesting open problem.

To fully utilize the potential of our model, we allow *adaptive* computation. It means that a description of an operation to be performed at step t may be a function of all measurement outcomes at steps $1,\ldots,t-1$. (For even greater generality, the dependence may be probabilistic. This assumption does not actually strengthen the model since tossing a fair coin can be simulated using $\mathcal{O}_{\rm ideal}$) At this point, we need to be careful because the proper choice of operations should not only be defined mathematically—it should be computed by some *efficient algorithm*. In all protocols described below, the algorithms will actually be very simple. (Let us point out that dropping the computational complexity restriction still leaves a nontrivial problem: can we prepare an arbitrary multiqubit pure state with any given fidelity using only operations from the basis $\mathcal{O}$?)

The main question that we address in this paper is as follows: For which density matrices ρ can one efficiently simulate universal quantum computation by adaptive computation in the basis $\mathcal{O}$?

It will be convenient to use the Bloch sphere representation of one-qubit states:

$$\rho = \tfrac{1}{2}(I + \rho_x\sigma^x + \rho_y\sigma^y + \rho_z\sigma^z).$$

The vector (ρ_x,ρ_y,ρ_z) will be referred to as the *polarization vector* of ρ. Let us first consider the subset of states satisfying

$$|\rho_x| + |\rho_y| + |\rho_z| \leqslant 1.$$

This inequality says that the vector (ρ_x,ρ_y,ρ_z) lies inside the octahedron O with vertices (±1, 0, 0), (0, ±1, 0), (0, 0, ±1),

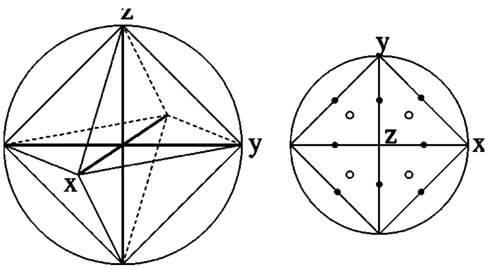

FIG. 1. Left: the Bloch sphere and the octahedron O. Right: the octahedron O projected on the $x-y$ plane. The magic states correspond to the intersections of the symmetry axes of O with the Bloch sphere. The empty and filled circles represent T-type and H-type magic states, respectively.

see Fig. 1. The six vertices of O represent the six eigenstates of the Pauli operators σ^x,σ^y, and σ^z. We can prepare these states by operations from $O_{\rm ideal}$ only. Since ρ is a convex linear combination (probabilistic mixture) of these states, we can prepare ρ by operations from $\mathcal{O}_{\rm ideal}$ and by tossing a coin with suitable weights. Thus we can rephrase the Gottesman-Knill theorem in the following way.

Theorem 1. Suppose the polarization vector (ρ_x,ρ_y,ρ_z) of the state ρ belongs to the convex hull of (±1, 0, 0), (0, ±1, 0), (0, 0, ±1). Then any adaptive computation in the basis $\mathcal{O}$ can be efficiently simulated on a classical probabilistic computer.

This observation leads naturally to the following question: is it true that UQC can be efficiently simulated whenever ρ lies in the exterior of the octahedron O? In an attempt to provide at least a partial answer, we prove the universality for a large set of states. Specifically, we construct two particular schemes of UQC simulation based on a method which we call *magic states distillation*. Let us start by defining the magic states.

Definition 1. Consider pure states $|H\rangle, |T\rangle \in \mathbb{C}^2$ such that

$$|T\rangle\langle T| = \frac{1}{2}\left[I + \frac{1}{\sqrt{3}}(\sigma^x + \sigma^y + \sigma^z)\right],$$

and

$$|H\rangle\langle H| = \frac{1}{2}\left[I + \frac{1}{\sqrt{2}}(\sigma^x + \sigma^z)\right].$$

The images of $|T\rangle$ and $|H\rangle$ under the action of one-qubit Clifford operators are called magic states of T type and H type, respectively.

[This notation is chosen since $|H\rangle$ and $|T\rangle$ are eigenvectors of certain Clifford group operators: the Hadamard gate H and the operator usually denoted T, see Eq. (7).] Denote the one-qubit Clifford group by $\mathcal{C}_1$. Overall, there are 8 magic states of T type, $\{U|T\rangle, U \in \mathcal{C}_1\}$ (up to a phase) and 12 states of H type, $\{U|H\rangle, U \in \mathcal{C}_1\}$, see Fig. 1. Clearly, the polarization vectors of magic states are in one-to-one correspondence with rotational symmetry axes of the octahedron O (H-type states correspond to 180° rotations and T-type states correspond to 120° rotations). The role of magic states in our construction is twofold. First, adaptive computation in the basis $\mathcal{O}_{\rm ideal}$ together with the preparation of magic states (of either type) allows one to simulate UQC (see Sec. III). Second, by adaptive computation in the basis $\mathcal{O}_{\text{ideal}}$ one can "purify" imperfect magic states. It is a rather surprising coincidence that one and the same state can comprise both of these properties, and that is the reason why we call them magic states.

More exactly, a magic state distillation procedure yields one copy of a magic state (with any desired fidelity) from several copies of the state ρ, provided that the initial fidelity between ρ and the magic state to be distilled is large enough. In the course of distillation, we use only operations from the set $\mathcal{O}_{\text{ideal}}$. By constructing two particular distillation schemes, for T-type and H-type magic states, respectively, we prove the following theorems.

Theorem 2. Let $F_T(\rho)$ be the maximum fidelity between ρ and a T-type magic state, i.e.,

$$F_T(\rho) = \max_{U\in\mathcal{C}_1} \sqrt{\langle T|U^\dagger \rho U|T\rangle}.$$

Adaptive computation in the basis $\mathcal{O}=\mathcal{O}_{\text{ideal}}\cup\{\rho\}$ allows one to simulate universal quantum computation whenever

$$F_T(\rho) > F_T = \left[\frac{1}{2}\left(1+\sqrt{\frac{3}{7}}\right)\right]^{1/2} \approx 0.910.$$

Theorem 3. Let $F_H(\rho)$ be the maximum fidelity between ρ and an H-type magic state,

$$F_H(\rho) = \max_{U\in\mathcal{C}_1} \sqrt{\langle H|U^\dagger \rho U|H\rangle}.$$

Adaptive computation in the basis $\mathcal{O}=\mathcal{O}_{\text{ideal}}\cup\{\rho\}$ allows one to simulate universal quantum computation whenever

$$F_H(\rho) > F_H \approx 0.927.$$

The quantities F_T and F_H have the meaning of threshold fidelity since our distillation schemes increase the polarization of ρ, converging to a magic state as long as the inequalities $F_T(\rho)>F_T$ or $F_H(\rho)>F_H$ are fulfilled. If they are not fulfilled, the process converges to the maximally mixed state. The conditions stated in the theorems can also be understood in terms of the polarization vector (ρ_x,ρ_y,ρ_z). Indeed, let us associate a "magic direction" with each of the magic states. Then Theorems 2 and 3 say that the distillation is possible if there is a T direction such that the projection of the vector (ρ_x,ρ_y,ρ_z) onto that T direction exceeds the threshold value of $2F_T^2-1\approx 0.655$, or if the projection on some of the H directions is greater than $2F_H^2-1\approx 0.718$.

Let us remark that, although the proposed distillation schemes are probably not optimal, the threshold fidelities F_T and F_H cannot be improved significantly. Indeed, it is easy to check that the octahedron O corresponding to probabilistic mixtures of stabilizer states can be defined as

$$\mathcal{O} = \{\rho : F_T(\rho) \leq F_T^*\},$$

where

$$F_T^* = \left[\frac{1}{2}\left(1+\sqrt{\frac{1}{3}}\right)\right]^{1/2} \approx 0.888.$$

It means that F_T^* is a lower bound on the threshold fidelity F_T for any protocol distilling T-type magic states. Thus any potential improvement to Theorem 2 may only decrease F_T from 0.910 down to $F_T^*=0.888$. From a practical perspective, the difference between these two numbers is not important.

On the other hand, such an improvement would be of great theoretical interest. Indeed, if Theorem 2 with F_T replaced by F_T^* is true, it would imply that the Gottesman-Knill theorem provides necessary and sufficient conditions for the classical simulation, and that a transition from classical to universal quantum behavior occurs at the boundary of the octahedron O. This kind of transition has been discussed in context of a general error model [21]. Our model is simpler, which gives hope for sharper results.

By the same argument, one can show that the quantity

$$F_H^* \overset{\text{def}}{=} \max_{\rho\in O} \sqrt{\langle H|\rho|H\rangle} = \left[\frac{1}{2}\left(1+\sqrt{\frac{1}{2}}\right)\right]^{1/2} \approx 0.924$$

is a lower bound on the threshold fidelity F_H for any protocol distilling H-type magic states.

A similar approach to UQC simulation was suggested in Ref. [22], where Clifford group operations were used to distill the entangled three-qubit state $|000\rangle+|001\rangle+|010\rangle+|100\rangle$, which is necessary for the realization of the Toffoli gate.

The rest of the paper is organized as follows. Section II contains some well-known facts about the Clifford group and stabilizer formalism, which will be used throughout the paper. In Sec. III we prove that magic states together with operations from $\mathcal{O}_{\text{ideal}}$ are sufficient for UQC. In Sec. IV ideal magic are substituted by faulty ones and the error rate that our simulation algorithm can tolerate is estimated. In Sec. V we describe a distillation protocol for T-type magic states. This protocol is based on the well-known five-qubit quantum code. In Sec. VI a distillation protocol for H-type magic states is constructed. It is based on a certain CSS stabilizer code that encodes one qubit into 15 and admits a nontrivial automorphism [23]. Specifically, the bitwise application of a certain *non-Clifford* unitary operator preserves the code subspace and effects the same operator on the encoded qubit. We conclude with a brief summary and a discussion of open problems.

II. CLIFFORD GROUP, STABILIZERS, AND SYNDROME MEASUREMENTS

Let $\mathcal{C}_n$ denote the n-qubit *Clifford group*. Recall that it is a finite subgroup of $\mathrm{U}(2^n)$ generated by the Hadamard gate H (applied to any qubit), the phase-shift gate K (applied to any qubit), and the controlled-not gate $\Lambda(\sigma^x)$ (which may be applied to any pair qubits),

$$H=\frac{1}{\sqrt{2}}\begin{pmatrix}1 & 1\\ 1 & -1\end{pmatrix},\quad K=\begin{pmatrix}1 & 0\\ 0 & i\end{pmatrix},\quad \Lambda(\sigma^x)=\begin{pmatrix}I & 0\\ 0 & \sigma^x\end{pmatrix}. \tag{1}$$

The Pauli operators $\sigma^x,\sigma^y,\sigma^z$ belong to $\mathcal{C}_1$, for instance, $\sigma^z=K^2$ and $\sigma^x=HK^2H$. The *Pauli group* $P(n)\subset\mathcal{C}_n$ is generated by the Pauli operators acting on n qubits. It is known [24] that the Clifford group $\mathcal{C}_n$ augmented by scalar unitary operators $e^{i\varphi}I$ coincides with the normalizer of $P(n)$ in the unitary group $\mathrm{U}(2^n)$. Hermitian elements of the Pauli group are of particular importance for quantum error correction theory; they are referred to as *stabilizers*. These are operators of the form

$$\pm\sigma^{\alpha_1}\otimes\cdots\otimes\sigma^{\alpha_n},\quad \alpha_j\in\{0,x,y,z\},$$

where $\sigma^0=I$. Let us denote by $S(n)$ the set of all n-qubit stabilizers:

$$S(n)=\{S\in P(n)\ :\ S^\dagger=S\}.$$

For any two stabilizers S_1,S_2 we have $S_1S_2=\pm S_2S_1$ and $S_1^2=S_2^2=I$. It is known that for any set of pairwise commuting stabilizers $S_1,\ldots,S_k\in S(n)$ there exists a unitary operator $V\in\mathcal{C}_n$ such that

$$VS_jV^\dagger=\sigma^z[j],\quad j=1,\ldots,k,$$

where $\sigma^z[j]$ denotes the operator σ^z applied to the jth qubit, e.g., $\sigma^z[1]=\sigma^z\otimes I\otimes\cdots\otimes I$.

These properties of the Clifford group allow us to introduce a very useful computational procedure which can be realized by operations from $\mathcal{O}_{\mathrm{ideal}}$. Specifically, we can perform a joint nondestructive eigenvalue measurement for any set of pairwise commuting stabilizers $S_1,\ldots,S_k\in S(n)$. The outcome of such a measurement is a sequence of eigenvalues $\lambda=(\lambda_1,\ldots,\lambda_k)$, $\lambda_j=\pm1$, which is usually called a *syndrome*. For any given outcome, the quantum state is acted upon by the projector

$$\Pi_\lambda=\prod_{j=1}^{k}\frac{1}{2}(I+\lambda_jS_j).$$

Now, let us consider a computation that begins with an arbitrary state and consists of operations from $\mathcal{O}_{\mathrm{ideal}}$. It is clear that we can defer all Clifford operations until the very end if we replace the Pauli measurements by general syndrome measurements. Thus the most general transformation that can be realized by $\mathcal{O}_{\mathrm{ideal}}$ is an *adaptive syndrome measurement*, meaning that the choice of the stabilizer S_j to be measured next depends on the previously measured values of $\lambda_1,\ldots,\lambda_{j-1}$. In general, this dependence may involve coin tossing. Without loss of generality one can assume that S_j commutes with all previously measured stabilizers $S_1,\ldots,S_{j-1}$ (for all possible values of $\lambda_1,\ldots,\lambda_{j-1}$ and coin tossing outcomes). Adaptive syndrome measurement has been used in Ref. [25] to distill entangled states of a bipartite system by local operations.

III. UNIVERSAL QUANTUM COMPUTATION WITH MAGIC STATES

In this section, we show that operations from $\mathcal{O}_{\mathrm{ideal}}$ are sufficient for universal quantum computation if a supply of *ideal* magic states is also available. First, consider a one-qubit state

$$|A_\theta\rangle=2^{-1/2}(|0\rangle+e^{i\theta}|1\rangle)\tag{2}$$

and suppose that θ is not a multiple of $\pi/2$. We now describe a procedure that implements the phase shift gate

$$\Lambda(e^{i\theta})=\begin{pmatrix}1&0\\0&e^{i\theta}\end{pmatrix}$$

by consuming several copies of $|A_\theta\rangle$ and using only operations from $\mathcal{O}_{\mathrm{ideal}}$.

Let $|\psi\rangle=a|0\rangle+b|1\rangle$ be the unknown initial state which should be acted on by $\Lambda(e^{i\theta})$. Prepare the state $|\Psi_0\rangle=|\psi\rangle\otimes|A_\theta\rangle$ and measure the stabilizer $S_1=\sigma^z\otimes\sigma^z$. Note that both outcomes of this measurement appear with probability $1/2$. If the outcome is "+1", we are left with the state

$$|\Psi_1^+\rangle=(a|0,0\rangle+be^{i\theta}|1,1\rangle).$$

In the case of "−1" outcome, the resulting state is

$$|\Psi_1^-\rangle=(ae^{i\theta}|0,1\rangle+b|1,0\rangle).$$

Let us apply the gate $\Lambda(\sigma^x)[1,2]$ (the first qubit is the control one). The above two states are mapped to

$$|\Psi_2^+\rangle=\Lambda(\sigma^x)[1,2]|\Psi_1^+\rangle=(a|0\rangle+be^{i\theta}|1\rangle)\otimes|0\rangle,$$

$$|\Psi_2^-\rangle=\Lambda(\sigma^x)[1,2]|\Psi_1^-\rangle=(ae^{i\theta}|0\rangle+b|1\rangle)\otimes|1\rangle.$$

Now the second qubit can be discarded, and we are left with the state $a|0\rangle+be^{\pm i\theta}|1\rangle$, depending upon the measured eigenvalue. Thus the net effect of this circuit is the application of a unitary operator that is chosen randomly between $\Lambda(e^{i\theta})$ and $\Lambda(e^{-i\theta})$ (and we know which of the two possibilities has occurred).

Applying the circuit repeatedly, we effect the transformations $\Lambda(e^{ip_1\theta})$, $\Lambda(e^{ip_2\theta}),\ldots$ for some integers $p_1,p_2,\ldots$ which obey the random-walk statistics. It is well known that such a random walk visits each integer with the probability 1. It means that sooner or later we will get $p_k=1$ and thus realize the desired operator $\Lambda(e^{i\theta})$. The probability that we will need more than N steps to succeed can be estimated as $cN^{-1/2}$ for some constant $c>0$. Note also that if θ is a rational multiple of 2π, we actually have a random walk on a cyclic group $\mathbb{Z}_q$. In this case, the probability that we will need more than N steps decreases exponentially with N.

The magic state $|H\rangle$ can be explicitly written in the standard basis as

$$|H\rangle=\cos\left(\frac{\pi}{8}\right)|0\rangle+\sin\left(\frac{\pi}{8}\right)|1\rangle.\tag{3}$$

Note that $HK|H\rangle=e^{i\pi/8}|A_{-\pi/4}\rangle$. So if we are able to prepare the state $|H\rangle$, we can realize the operator $\Lambda(e^{-i\pi/4})$. It does not belong to the Clifford group. Moreover, the subgroup of U(2) generated by $\Lambda(e^{-i\pi/4})$ and $\mathcal{C}_1$ is dense in U(2). [1] Thus the operators from $\mathcal{C}_1$ and $\mathcal{C}_2$ together with $\Lambda(e^{-i\pi/4})$ constitute a universal basis for quantum computation.

The magic state $|T\rangle$ can be explicitly written in the standard basis:

[1] Recall that the action of the Clifford group $\mathcal{C}_1$ on the set of operators $\pm\sigma^x$, $\pm\sigma^y$, $\pm\sigma^z$ coincides with the action of rotational symmetry group of a cube on the set of unit vectors $\pm e_x$, $\pm e_y$, $\pm e_z$, respectively.

$$|T\rangle = \cos\beta|0\rangle + e^{i(\pi/4)}\sin\beta|1\rangle, \quad \cos(2\beta) = \frac{1}{\sqrt{3}}. \qquad (4)$$

Let us prepare an initial state $|\Psi_0\rangle = |T\rangle \otimes |T\rangle$ and measure the stabilizer $S_1 = \sigma^z \otimes \sigma^z$. The outcome +1 appears with probability $p_+ = \cos^4\beta + \sin^4\beta = 2/3$. If the outcome is −1, we discard the reduced state and try again, using a fresh pair of magic states. (On average, we need three copies of the $|T\rangle$ state to get the outcome +1.) The reduced state corresponding to the outcome +1 is

$$|\Psi_1\rangle = \cos\gamma|0,0\rangle + i\sin\gamma|1,1\rangle, \quad \gamma = \frac{\pi}{12}.$$

Let us apply the gate $\Lambda(\sigma^x)[1,2]$ and discard the second qubit. We arrive at the state

$$|\Psi_2\rangle = \cos\gamma|0\rangle + i\sin\gamma|1\rangle.$$

Next apply the Hadamard gate H:

$$|\Psi_3\rangle = H|\Psi_2\rangle = 2^{-1/2}e^{i\gamma}(|0\rangle + e^{-2i\gamma}|1\rangle) = |A_{-\pi/6}\rangle.$$

We can use this state as described above to realize the operator $\Lambda(e^{-i\pi/6})$. It is easy to check that Clifford operators together with $\Lambda(e^{-i\pi/6})$ constitute a universal set of unitary gates.

Thus we have proved that the sets of operations $\mathcal{O}_{\text{ideal}} \cup \{|H\rangle\}$ and $\mathcal{O}_{\text{ideal}} \cup \{|T\rangle\}$ are sufficient for universal quantum computation.

IV. ERROR ANALYSIS

To establish a connection between the simulation algorithms described in Sec. III and the universality theorems stated in the introduction we have to substitute *ideal* magic states by *faulty* ones. Before doing that let us discuss the ideal case in more detail. Suppose that a quantum circuit to be simulated uses a gate basis in which the only non-Clifford gate is the phase shift $\Lambda(e^{-i\pi/4})$ or $\Lambda(e^{-i\pi/6})$. One can apply the algorithm of Sec. III to simulate each non-Clifford gate independently. To avoid fluctuations in the number of magic states consumed at each round, let us set a limit of K magic states per round, where K is a parameter to be chosen later. As was pointed out in Sec. III, the probability for some particular simulation round to "run out of budget" scales as $\exp(-\alpha K)$ for some constant $\alpha > 0$. If at least one simulation round runs out of budget, we declare a failure and the whole simulation must be aborted. Denote the total number of non-Clifford gates in the circuit by L. The probability p_a for the whole simulation to be aborted can be estimated as

$$p_a \sim 1 - [1 - \exp(-\alpha K)]^L \sim L\exp(-\alpha K) \ll 1,$$

provided that $L\exp(-\alpha K) \ll 1$. We will assume

$$K \gtrsim \alpha^{-1}\ln L,$$

so the abort probability can be neglected.

Each time the algorithm requests an ideal magic state, it actually receives a slightly nonideal one. Such nearly perfect magic states must be prepared using the distillation methods described in Secs. V and VI. Let us estimate an affordable error rate ϵ_{out} for *distilled* magic states. Since there are L non-Clifford gates in the circuit, one can tolerate an error rate of the order $1/L$ in implementation of these gates.[2] Each non-Clifford gate requires $K \sim \ln L$ magic states. Thus the whole simulation is reliable enough if one chooses

$$\epsilon_{\text{out}} \sim 1/(L\ln L). \qquad (5)$$

What are the resources needed to distill one copy of a magic state with the error rate ϵ_{out}? To be more specific, let us talk about H-type states. It will be shown in Sec. VI that the number n of raw (undistilled) ancillas needed to distill one copy of the $|H\rangle$ magic state with an error rate not exceeding ϵ_{out} scales as

$$n \sim [\ln(1/\epsilon_{\text{out}})]^\gamma, \quad \gamma = \log_3 15 \approx 2.5,$$

see Eq. (39). Taking ϵ_{out} from Eq. (5), one gets

$$n \sim (\ln L)^\gamma.$$

Since the whole simulation requires $KL \sim L\ln L$ copies of the distilled $|H\rangle$ state, we need

$$N \sim L(\ln L)^{\gamma+1}$$

raw ancillas overall.

Summarizing, the simulation theorems stated in the introduction follow from the following results (the last one will be proved later):

(i) the circuits described in Sec. III allow one to simulate UQC with the sets of operations $\mathcal{O}_{\text{ideal}} \cup \{|H\rangle\}$ and $\mathcal{O}_{\text{ideal}} \cup \{|T\rangle\}$;

(ii) these circuits work reliably enough if the states $|H\rangle$ and $|T\rangle$ are slightly noisy, provided that the error rate does not exceed $\epsilon_{\text{out}} \sim 1/(L\ln L)$;

(iii) a magic state having an error rate ϵ_{out} can be prepared from copies of the raw ancillary state ρ using the distillation schemes provided that $F_T(\rho) > F_T$ or $F_H(\rho) > F_H$. The distillation requires resources that are polynomial in $\ln L$.

V. DISTILLATION OF *T*-TYPE MAGIC STATES

Suppose we are given n copies of a state ρ, and our goal is to distill one copy of the magic state $|T\rangle$. The polarization vector of ρ can be brought into the positive octant of the Bloch space by a Clifford group operator, so we can assume that

$$\rho_x, \rho_y, \rho_z \geq 0.$$

In this case, the fidelity between ρ and $|T\rangle$ is the largest one among all T-type magic states, i.e.,

$$F_T(\rho) = \sqrt{\langle T|\rho|T\rangle}.$$

A related quantity,

[2] This fault tolerance does not require any redundancy in the implementation of the circuit (e.g., the use of concatenated codes). It is achived automatically because in the worst case the error probability accumulates linearly in the number of gates. In our model only non-Clifford gates are faulty.

$$\epsilon = 1 - \langle T|\rho|T\rangle = \frac{1}{2}\left[1 - \frac{1}{\sqrt{3}}(\rho_x + \rho_y + \rho_z)\right],$$

will be called the *initial error probability*. By definition, $0 \leqslant \epsilon \leqslant 1/2$.

The output of the distillation algorithm will be some one-qubit mixed state ρ_{out}. To quantify the proximity between ρ_{out} and $|T\rangle$, let us define a *final error probability*:

$$\epsilon_{\mathrm{out}} = 1 - \langle T|\rho_{\mathrm{out}}|T\rangle.$$

It will be certain function of n and ϵ. The asymptotic behavior of this function for $n \to \infty$ reveals the existence of a *threshold error probability*,

$$\epsilon_0 = \frac{1}{2}\left(1 - \sqrt{\frac{3}{7}}\right) \approx 0.173,$$

such that for $\epsilon < \epsilon_0$ the function $\epsilon_{\mathrm{out}}(n,\epsilon)$ converges to zero. We will see that for small ϵ,

$$\epsilon_{\mathrm{out}}(n,\epsilon) \sim (5\epsilon)^{n^{\xi}}, \quad \xi = 1/\log_2 30 \approx 0.2. \tag{6}$$

On the other hand, if $\epsilon > \epsilon_0$, the output state converges to the maximally mixed state, i.e., $\lim_{n\to\infty}\epsilon_{\mathrm{out}}(n,\epsilon) = 1/2$.

Before coming to a detailed description of the distillation algorithm, let us outline the basic ideas involved in its construction. The algorithm recursively iterates an elementary distillation subroutine that transforms five copies of an imperfect magic state into one copy having a smaller error probability. This elementary subroutine involves a syndrome measurement for certain commuting stabilizers $S_1, S_2, S_3, S_4 \in \mathcal{S}(5)$. If the measured syndrome $(\lambda_1,\lambda_2,\lambda_3,\lambda_4)$ is nontrivial ($\lambda_j = -1$ for some j), the distillation attempt fails and the reduced state is discarded. If the measured syndrome is trivial ($\lambda_j = 1$ for all j), the distillation attempt is successful. Applying a decoding transformation (a certain Clifford operator) to the reduced state, we transform it to a single-qubit state. This qubit is the output of the subroutine.

Our construction is similar to concatenated codes used in many fault-tolerant quantum computation techniques, but it differs from them in two respects. First, we do not need to *correct* errors—it suffices only to *detect* them. Once an error has been detected, we simply discard the reduced state, since it does not contain any valuable information. This allows us to achieve higher threshold error probability. Second, we do not use quantum codes in the way for which they were originally designed: in our scheme, the syndrome is measured on a product state.

The state $|T\rangle$ is an eigenstate for the unitary operator

$$T = e^{i\pi/4}KH = \frac{e^{i\pi/4}}{\sqrt{2}}\begin{pmatrix} 1 & 1 \\ i & -i \end{pmatrix} \in \mathcal{C}_1. \tag{7}$$

Note that T acts on the Pauli operators as follows:[3]

$$T\sigma^x T^\dagger = \sigma^z, \quad T\sigma^z T^\dagger = \sigma^y, \quad T\sigma^y T^\dagger = \sigma^x. \tag{8}$$

We will denote its eigenstates by $|T_0\rangle$ and $|T_1\rangle$, so that

$$T|T_0\rangle = e^{+i\pi/3}|T_0\rangle, \quad T|T_1\rangle = e^{-i\pi/3}|T_1\rangle,$$

$$|T_{0,1}\rangle\langle T_{0,1}| = \frac{1}{2}\left[I \pm \frac{1}{\sqrt{3}}(\sigma^x + \sigma^y + \sigma^z)\right].$$

Note that $|T_0\rangle \overset{\mathrm{def}}{=} |T\rangle$ and $|T_1\rangle = \sigma^y H|T_0\rangle$ are T-type magic states.

Let us apply a dephasing transformation,

$$D(\eta) = \frac{1}{3}(\eta + T\eta T^\dagger + T^\dagger \eta T), \tag{9}$$

to each copy of the state ρ. The transformation D can be realized by applying one of the operators I, T, T^{-1} chosen with probability $1/3$ each. Since

$$D(|T_0\rangle\langle T_1|) = D(|T_1\rangle\langle T_0|) = 0,$$

we have

$$D(\rho) = (1-\epsilon)|T_0\rangle\langle T_0| + \epsilon|T_1\rangle\langle T_1|. \tag{10}$$

We will assume that the dephasing transformation is applied at the very first step of the distillation, so ρ has the form (10). Thus the initial state for the elementary distillation subroutine is

$$\rho_{\mathrm{in}} = \rho^{\otimes 5} = \sum_{x \in \{0,1\}^5} \epsilon^{|x|}(1-\epsilon)^{5-|x|}|T_x\rangle\langle T_x|, \tag{11}$$

where $x = (x_1, \ldots, x_5)$ is a binary string, $|x|$ is the number of 1's in x, and

$$|T_x\rangle \overset{\mathrm{def}}{=} |T_{x_1}\rangle \otimes \cdots \otimes |T_{x_5}\rangle.$$

The stabilizers $S_1, \ldots, S_4$ to be measured on the state ρ_{in} correspond to the famous five-qubit code, see Refs. [26,27]. They are defined as follows:

$$S_1 = \sigma^x \otimes \sigma^z \otimes \sigma^z \otimes \sigma^x \otimes I,$$

$$S_2 = I \otimes \sigma^x \otimes \sigma^z \otimes \sigma^z \otimes \sigma^x,$$

$$S_3 = \sigma^x \otimes I \otimes \sigma^x \otimes \sigma^z \otimes \sigma^z,$$

$$S_4 = \sigma^z \otimes \sigma^x \otimes I \otimes \sigma^x \otimes \sigma^z. \tag{12}$$

This code has a cyclic symmetry, which becomes explicit if we introduce an auxiliary stabilizer, $S_5 = S_1 S_2 S_3 S_4 = \sigma^z \otimes \sigma^z \otimes \sigma^x \otimes I \otimes \sigma^x$. Let $\mathcal{L}$ be the two-dimensional code subspace specified by the conditions $S_j|\Psi\rangle = |\Psi\rangle$, $j = 1, \ldots, 4$, and Π be the orthogonal projector onto $\mathcal{L}$:

$$\Pi = \frac{1}{16}\prod_{j=1}^{4}(I + S_j). \tag{13}$$

It was pointed out in Ref. [16] that the operators

$$\hat{X} = (\sigma^x)^{\otimes 5}, \quad \hat{Y} = (\sigma^y)^{\otimes 5}, \quad \hat{Z} = (\sigma^z)^{\otimes 5},$$

and

[3] The operator denoted by T in Ref. [16] does not coincide with our T. They are related by the substitution $T \to e^{-i\pi/4}T^\dagger$ though.

$$\hat{T}=(T)^{\otimes 5} \tag{14}$$

commute with Π, thus preserving the code subspace. Moreover, $\hat{X},\hat{Y},\hat{Z}$ obey the same algebraic relations as one-qubit Pauli operators, e.g., $\hat{X}\hat{Y}=i\hat{Z}$. Let us choose a basis in $\mathcal{L}$ such that $\hat{X},\hat{Y}$, and $\hat{Z}$ become logical Pauli operators σ^x,σ^y, and σ^z, respectively. How does the operator $\hat{T}$ act in this basis? From Eq. (8) we immediately get

$$\hat{T}\hat{X}\hat{T}^\dagger=\hat{Z},\quad \hat{T}\hat{Z}\hat{T}^\dagger=\hat{Y},\quad \hat{T}\hat{Y}\hat{T}^\dagger=\hat{X}.$$

Therefore $\hat{T}$ coincides with the logical operator T up to an overall phase factor. This factor is fixed by the condition that the logical T has eigenvalues $e^{\pm i(\pi/3)}$.

Let us find the eigenvectors of $\hat{T}$ that belong to $\mathcal{L}$. Consider two particular states from $\mathcal{L}$, namely

$$|T_1^L\rangle=\sqrt{6}\Pi|T_{00000}\rangle,\quad \text{and } |T_0^L\rangle=\sqrt{6}\Pi|T_{11111}\rangle.$$

In the Appendix we show that

$$\langle T_{00000}|\Pi|T_{00000}\rangle=\langle T_{11111}|\Pi|T_{11111}\rangle=\frac{1}{6}, \tag{15}$$

so that the states $|T_0^L\rangle$ and $|T_1^L\rangle$ are normalized. Taking into account that $[\hat{T},\Pi]=0$ and that

$$\hat{T}|T_x\rangle=e^{i(\pi/3)(5-2|x|)}|T_x\rangle \text{ for all } x\in\{0,1\}^5, \tag{16}$$

we get

$$\hat{T}|T_1^L\rangle=\sqrt{6}\hat{T}\Pi|T_{00000}\rangle=\sqrt{6}\Pi\hat{T}|T_{00000}\rangle=e^{-i\pi/3}|T_1^L\rangle.$$

Analogously, one can check that

$$\hat{T}|T_0^L\rangle=e^{+i\pi/3}|T_0^L\rangle.$$

It follows that $\hat{T}$ is exactly the logical operator T, including the overall phase, and $|T_0^L\rangle$ and $|T_1^L\rangle$ are the logical states $|T_0\rangle$ and $|T_1\rangle$ (up to some phase factors, which are not important for us). Therefore we have

$$|T_{0,1}^L\rangle\langle T_{0,1}^L|=\Pi\frac{1}{2}\left[I\pm\frac{1}{\sqrt{3}}(\hat{X}+\hat{Y}+\hat{Z})\right]. \tag{17}$$

Now we are in a position to describe the syndrome measurement performed on the state ρ_{in}. The unnormalized reduced state corresponding to the trivial syndrome is as follows:

$$\rho_s=\Pi\rho_{\text{in}}\Pi=\sum_{x\in\{0,1\}^5}\epsilon^{|x|}(1-\epsilon)^{5-|x|}\Pi|T_x\rangle\langle T_x|\Pi, \tag{18}$$

see Eq. (11). The probability for the trivial syndrome to be observed is

$$p_s=\text{Tr}\,\rho_s.$$

Note that the state $\Pi|T_x\rangle$ is an eigenvector of $\hat{T}$ for any $x\in\{0,1\}^5$. But we know that the restriction of $\hat{T}$ on $\mathcal{L}$ has eigenvalues $e^{\pm i\pi/3}$. At the same time, Eq. (16) implies that

$$\hat{T}\Pi|T_x\rangle=-\Pi|T_x\rangle$$

whenever $|x|=1$ or $|x|=4$. This eigenvalue equation is not a contradiction only if

$$\Pi|T_x\rangle=0 \text{ for } |x|=1,4.$$

This equality can be interpreted as an error correction property. Indeed, the initial state ρ_{in} is a mixture of the desired state $|T_{00000}\rangle$ and unwanted states $|T_x\rangle$ with $|x|>0$. We can interpret the number of "1" components in x as a number of errors. Once the trivial syndrome has been measured, we can be sure that either no errors or at least two errors have occurred. Such error correction, however, is not directly related to the minimal distance of the code.

It follows from Eq. (16) that for $|x|=2$, 3 one has $\hat{T}\Pi|T_x\rangle=e^{\pm i\pi/3}\Pi|T_x\rangle$, so that $\Pi|T_x\rangle$ must be proportional to one of the states $|T_0^L\rangle$, $|T_1^L\rangle$. Our observations can be summarized as follows:

$$\Pi|T_x\rangle=\begin{cases}6^{-1/2}|T_1^L\rangle, & \text{if } |x|=0,\\ 0, & \text{if } |x|=1,\\ a_x|T_0^L\rangle, & \text{if } |x|=2,\\ b_x|T_1^L\rangle, & \text{if } |x|=3,\\ 0, & \text{if } |x|=4,\\ 6^{-1/2}|T_0^L\rangle, & \text{if } |x|=5.\end{cases} \tag{19}$$

Here the coefficients a_x,b_x depend upon x in some way. The output state (18) can now be written as

$$\rho_s=\left[\frac{1}{6}\epsilon^5+\epsilon^2(1-\epsilon)^3\sum_{x:|x|=2}|a_x|^2\right]|T_0^L\rangle\langle T_0^L| + \left[\frac{1}{6}(1-\epsilon)^5+\epsilon^3(1-\epsilon)^2\sum_{x:|x|=3}|b_x|^2\right]|T_1^L\rangle\langle T_1^L|. \tag{20}$$

To exclude the unknown coefficients a_x and b_x, we can use the identity

$$|T_0^L\rangle\langle T_0^L|+|T_1^L\rangle\langle T_1^L|=\Pi=\sum_{x\in\{0,1\}^5}\Pi|T_x\rangle\langle T_x|\Pi.$$

Substituting Eq. (19) into this identity, we get

$$\sum_{x:|x|=2}|a_x|^2=\sum_{x:|x|=3}|b_x|^2=\frac{5}{6}.$$

So the final expression for the output state ρ_s is as follows:

$$\rho_s=\left[\frac{\epsilon^5+5\epsilon^2(1-\epsilon)^3}{6}\right]|T_0^L\rangle\langle T_0^L|+\left[\frac{(1-\epsilon)^5+5\epsilon^3(1-\epsilon)^2}{6}\right]\times|T_1^L\rangle\langle T_1^L|. \tag{21}$$

Accordingly, the probability to observe the trivial syndrome is

$$p_s=\frac{\epsilon^5+5\epsilon^2(1-\epsilon)^3+5\epsilon^3(1-\epsilon)^2+(1-\epsilon)^5}{6}. \tag{22}$$

A decoding transformaion for the five-qubit code is a unitary operator $V\in\mathcal{C}_5$ such that

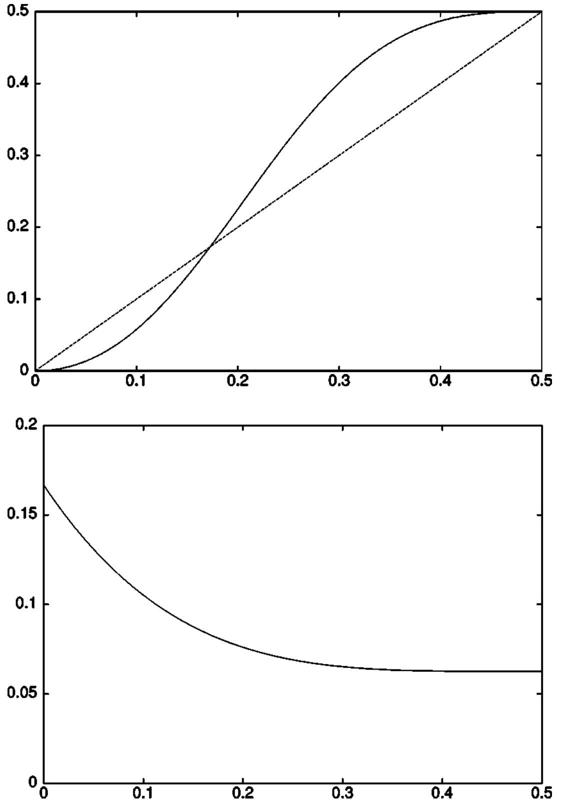

FIG. 2. The final error probability ϵ_{out} and the probability p_s to measure the trivial syndrome as functions of the initial error probability ϵ for the T-type states distillation.

$$V\mathcal{L} = \mathbb{C}^2 \otimes |0,0,0,0\rangle.$$

In other words, V maps the stabilizers S_j, $j=2$, 3, 4, 5 to $\sigma^z[j]$. The logical operators $\hat{X},\hat{Y},\hat{Z}$ are mapped to the Pauli operators $\sigma^x,\sigma^y,\sigma^z$ acting on the first qubit. From Eq. (17) we infer that

$$V|T^L_{0,1}\rangle = |T_{0,1}\rangle \otimes |0,0,0,0\rangle$$

(maybe up to some phase). The decoding should be followed by an additional operator $A=\sigma^y H \in \mathcal{C}_1$, which swaps the states $|T_0\rangle$ and $|T_1\rangle$ (note that for small ϵ the state ρ_s is close to $|T^L_1\rangle$, while our goal is to distill $|T_0\rangle$). After that we get a normalized output state

$$\rho_{\text{out}} = (1-\epsilon_{\text{out}})|T_0\rangle\langle T_0| + \epsilon_{\text{out}}|T_1\rangle\langle T_1|,$$

where

$$\epsilon_{\text{out}} = \frac{t^5+5t^2}{1+5t^2+5t^3+t^5}, \quad t = \frac{\epsilon}{1-\epsilon}. \tag{23}$$

The plot of the function $\epsilon_{\text{out}}(\epsilon)$ is shown on Fig. 2. It indicates that the equation $\epsilon_{\text{out}}(\epsilon)=\epsilon$ has only one nontrivial solution, $\epsilon=\epsilon_0\approx 0.173$. The exact value is

$$\epsilon_0 = \frac{1}{2}\left(1-\sqrt{\frac{3}{7}}\right).$$

If $\epsilon<\epsilon_0$, we can recursively iterate the elementary distillation subroutine to produce as good an approximation to the state $|T_0\rangle$ as we wish. On the other hand, if $\epsilon>\epsilon_0$, the distillation subroutine increases the error probability and iterations converge to the maximally mixed state. Thus ϵ_0 is a threshold error probability for our scheme. The corresponding threshold polarization is $1-2\epsilon_0=\sqrt{3/7}\approx 0.655$. For a sufficiently small ϵ, one can use the approximation $\epsilon_{\text{out}}(\epsilon)\approx 5\epsilon^2$.

The probability $p_s=p_s(\epsilon)$ to measure the trivial syndrome decreases monotonically from $1/6$ for $\epsilon=0$ to $1/16$ for $\epsilon=1/2$, see Fig. 2. In the asymptotic regime where ϵ is small, we can use the approximation $p_s\approx p_s(0)=1/6$.

Now the construction of the whole distillation scheme is straightforward. We start from $n\gg 1$ copies of the state $\rho=(1-\epsilon)|T_0\rangle\langle T_0|+\epsilon|T_1\rangle\langle T_1|$. Let us split these states into groups containing five states each and apply the elementary distillation subroutine described above to each group independently. In some of these groups the distillation attempt fails, and the outputs of such groups must be discarded. The average number of "successful" groups is obviously $p_s(\epsilon)\times(n/5)\approx n/30$ if ϵ is small. Neglecting the fluctuations of this quantity, we can say that our scheme provides a constant *yield* $r=1/30$ of output states that are characterized by the error probability $\epsilon_{\text{out}}(\epsilon)\approx 5\epsilon^2$. Therefore we can obtain r^2n states with $\epsilon_{\text{out}}\approx 5^3\epsilon^4$, r^3n states with $\epsilon_{\text{out}}\approx 5^7\epsilon^8$, and so on. We have created a hierarchy of states with n states on the first level and four or fewer states on the last level. Let k be the number of levels in this hierarchy and ϵ_{out} the error probability characterizing the states on the last level. Up to small fluctuations, the numbers $n,k,\epsilon_{\text{out}}$, and ϵ are related by the following obvious equations:

$$\epsilon_{\text{out}} \approx \tfrac{1}{5}(5\epsilon)^{2^k}, \quad r^k n \approx 1. \tag{24}$$

Their solution yields Eq. (6).

VI. DISTILLATION OF *H*-TYPE MAGIC STATES

A distillation scheme for H-type magic states also works by recursive iteration of a certain elementary distillation subroutine based on a syndrome measurement for a suitable stabilizer code. Let us start with introducing some relevant coding theory constructions, which reveal an unusual symmetry of this code and explain why it is particularly useful for H-type magic states distillation.

Let $\mathbb{F}_2^n$ be the n-dimensional binary linear space and A be a one-qubit operator such that $A^2=I$. With any binary vector $u=(u_1,\ldots,u_n)\in\mathbb{F}_2^n$ we associate the n-qubit operator

$$A(u) = A^{u_1}\otimes A^{u_2}\otimes \cdots \otimes A^{u_n}.$$

Let $(u,v)=\sum_{i=1}^n u_i v_i \bmod 2$ denote the standard binary inner product. If $\mathcal{L}\subseteq\mathbb{F}_2^n$ is a linear subspace, we denote by $\mathcal{L}^\perp$ the set of vectors which are orthogonal to $\mathcal{L}$. The Hamming weight of a binary vector u is denoted by $|u|$. Finally, $u\cdot v\in\mathbb{F}_2^n$ designates the bitwise product of u and v, i.e., $(u\cdot v)_i=u_i v_i$.

A systematic way of constructing stabilizer codes was suggested by Calderbank, Shor, and Steane, see Refs. [28,29]. Codes that can be described in this way will be referred to as *standard CSS codes*. In addition, we consider their images under an arbitrary unitary transformation $V \in \mathrm{U}(2)$ applied to every qubit. Such "rotated" codes will be called *CSS codes*.

Definition 2. Consider a pair of one-qubit Hermitian operators A,B such that

$$A^2 = B^2 = I, \quad AB = -BA,$$

and a pair of binary vector spaces $\mathcal{L}_A,\mathcal{L}_B \subseteq \mathbb{F}_2^n$, such that

$$(u,v) = 0 \text{ for all } u \in \mathcal{L}_A, v \in \mathcal{L}_B.$$

A quantum code $\mathrm{CSS}(A,\mathcal{L}_A;B,\mathcal{L}_B)$ is a decomposition

$$(\mathbb{C}^2)^{\otimes n} = \bigoplus_{\mu\in\mathcal{L}_A^*} \bigoplus_{\eta\in\mathcal{L}_B^*} \mathcal{H}(\mu,\eta), \tag{25}$$

where the subspace $\mathcal{H}(\mu,\eta)$ is defined by the conditions

$$A(u)|\Psi\rangle = (-1)^{\mu(u)}|\Psi\rangle, \quad B(v)|\Psi\rangle = (-1)^{\eta(v)}|\Psi\rangle$$

for all $u \in \mathcal{L}_A$ and $v \in \mathcal{L}_B$. The linear functionals μ and η are referred to as A syndrome and B syndrome, respectively. The subspace $\mathcal{H}(0,0)$ corresponding to the trivial syndromes $\mu = \eta = 0$ is called the code subspace.

The subspaces $\mathcal{H}(\mu,\eta)$ are well defined since the operators $A(u)$ and $B(v)$ commute for any $u \in \mathcal{L}_A$ and $v \in \mathcal{L}_B$:

$$A(u)B(v) = (-1)^{(u,v)}B(v)A(u) = B(v)A(u).$$

The number of logical qubits in a CSS code is

$$k = \log_2[\dim \mathcal{H}(0,0)] = n - \dim \mathcal{L}_A - \dim \mathcal{L}_B.$$

Logical operators preserving the subspaces $\mathcal{H}(\mu,\eta)$ can be chosen as

$$\{A(u)\ :\ u \in \mathcal{L}_B^{\perp}/\mathcal{L}_A\} \text{ and } \{B(v)\ :\ v \in \mathcal{L}_A^{\perp}/\mathcal{L}_B\}.$$

(By definition, $\mathcal{L}_A \subseteq \mathcal{L}_B^{\perp}$ and $\mathcal{L}_B \subseteq \mathcal{L}_A^{\perp}$, so the factor spaces are well defined.) In the case where A and B are Pauli operators, we get a standard CSS code. Generally, $A = V\sigma^z V^{\dagger}$ and $B = V\sigma^x V^{\dagger}$ for some unitary operator $V \in \mathrm{SU}(2)$, so an arbitrary CSS code can be mapped to a standard one by a suitable bitwise rotation. By a syndrome measurement for a CSS code we mean a projective measurement associated with the decomposition (25).

Consider a CSS code such that some of the operators $A(u)$, $B(v)$ do not belong to the Pauli group $P(n)$. Let us pose this question: can one perform a syndrome measurement for this code by operations from $\mathcal{O}_{\mathrm{ideal}}$ only? It may seem that the answer is no, because by definition of $\mathcal{O}_{\mathrm{ideal}}$ one cannot measure an eigenvalue of an operator unless it belongs to the Pauli group. Surprisingly, this naive answer is wrong. Indeed, imagine that we have measured part of the operators $A(u)$, $B(v)$ (namely, those that belong to the Pauli group). Now we may restrict the remaining operators to the subspace corresponding to the obtained measurement outcomes. It may happen that the restriction of some unmeasured operator $A(u)$, which does not belong to the Pauli group, coincides with the restriction of some other operator $\widetilde{A}(\widetilde{u}) \in P(n)$. If this is the case, we can safely measure $\widetilde{A}(\widetilde{u})$ instead of $A(u)$. The 15-qubit code that we use for the distillation is actually the simplest (to our knowledge) CSS code exhibiting this strange behavior. We now come to an explicit description of this code.

Consider a function f of four Boolean variables. Denote by $[f] \in \mathbb{F}_2^{15}$ the table of all values of f except $f(0000)$. The table is considered as a binary vector, i.e.,

$$[f] = (f(0001), f(0010), f(0011), \ldots, f(1111)).$$

Let $\mathcal{L}_1$ be the set of all vectors $[f]$, where f is a linear function satisfying $f(0)=0$. In other words, $\mathcal{L}_1$ is the linear subspace spanned by the four vectors $[x_j]$, $j=1, 2, 3, 4$ (where x_j is an indicator function for the jth input bit):

$$\mathcal{L}_1 = \text{linear span}([x_1],[x_2],[x_3],[x_4]).$$

Let also $\mathcal{L}_2$ be the set of all vectors $[f]$, where f is a polynomial of degree at most 2 satisfying $f(0)=0$. In other words, $\mathcal{L}_2$ is the linear subspace spanned by the four vectors $[x_j]$ and the six vectors $[x_ix_j]$:

$$\begin{aligned}\mathcal{L}_2 = \text{linear span}([x_1],[x_2],[x_3],[x_4],[x_1x_2],[x_1x_3],\\ [x_1x_4],[x_2x_3],[x_2x_4],[x_3x_4]).\end{aligned} \tag{26}$$

The definition of $\mathcal{L}_1$ and $\mathcal{L}_2$ resembles the definition of punctured Reed-Muller codes of order 1 and 2, respectively, see Ref. [30]. Note also that $\mathcal{L}_1$ is the dual space for the 15-bit Hamming code. The relevant properties of the subspaces $\mathcal{L}_j$ are stated in the following lemma.

Lemma 1.

(1) For any $u \in \mathcal{L}_1$ one has $|u| \equiv 0 \pmod 8$.

(2) For any $v \in \mathcal{L}_2$ one has $|v| \equiv 0 \pmod 2$.

(3) Let $[1]$ be the unit vector $(1, 1, \ldots, 1, 1)$. Then $\mathcal{L}_1^{\perp} = \mathcal{L}_2 \oplus [1]$ and $\mathcal{L}_2^{\perp} = \mathcal{L}_1 \oplus [1]$.

(4) For any vectors $u,v \in \mathcal{L}_1$ one has $|u\cdot v| \equiv 0 \pmod 4$.

(5) For any vectors $u \in \mathcal{L}_1$ and $v \in \mathcal{L}_2^{\perp}$ one has $|u\cdot v| \equiv 0 \pmod 4$.

Proof.

(1) Any linear function f on $\mathbb{F}_2^4$ satisfying $f(0)=0$ takes value 1 exactly eight times (if $f\neq 0$) or zero times (if $f=0$).

(2) All basis vectors of $\mathcal{L}_2$ have weight equal to 8 (the vectors $[x_i]$) or 4 (the vectors $[x_ix_j]$). By linearity, all elements of $\mathcal{L}_2$ have even weight.

(3) One can easily check that all basis vectors of $\mathcal{L}_1$ are orthogonal to all basis vectors of $\mathcal{L}_2$, therefore $\mathcal{L}_1 \subseteq \mathcal{L}_2^{\perp}$, $\mathcal{L}_2 \subseteq \mathcal{L}_1^{\perp}$. Besides, we have already proved that $[1] \in \mathcal{L}_1^{\perp}$ and $[1] \in \mathcal{L}_2^{\perp}$. Now the statement follows from dimension counting, since $\dim \mathcal{L}_1 = 4$ and $\dim \mathcal{L}_2 = 10$.

(4) Without loss of generality we may assume that $u \neq 0$ and $v \neq 0$. If $u=v$, the statement has been already proved, see property 1. If $u \neq v$, then $u=[f]$, $v=[g]$ for some linearly independent linear functions f and g. We can introduce new coordinates (y_1,y_2,y_3,y_4) on $\mathbb{F}_2^4$ such that $y_1=f(x)$ and $y_2=g(x)$. Now $|u\cdot v| = |[y_1y_2]| = 4$.

(5) Let $u \in \mathcal{L}_1$ and $v \in \mathcal{L}_2^{\perp}$. Since $\mathcal{L}_2^{\perp} = \mathcal{L}_1 \oplus [1]$, there are two possibilities: $v \in \mathcal{L}_1$ and $v = [1] + w$ for some $w \in \mathcal{L}_1$. The first case has been already considered. In the second case we have

$$|u\cdot v| = \sum_{j=1}^{15} u_j(1-w_j) = |u| - |u\cdot w|.$$

It follows from properties 1 and 4 that $|u\cdot v| \equiv 0 (\mathrm{mod}\, 4)$. □

Now consider the one-qubit Hermitian operator

$$A = \frac{1}{\sqrt{2}}(\sigma^x + \sigma^y) = \begin{pmatrix} 0 & e^{-i(\pi/4)} \\ e^{+i(\pi/4)} & 0 \end{pmatrix} = e^{-1(\pi/4)} K\sigma^x,$$

where K is the phase shift gate, see Eq. (1). By definition, A belongs to the Clifford group $\mathcal{C}_1$. One can easily check that $A^2 = I$ and $A\sigma^z = -\sigma^z A$, so the code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1)$ is well defined. We claim that its code subspace coincides with the code subspace of a certain stabilizer code.

Lemma 2. Consider the decomposition

$$(\mathbb{C}^2)^{\otimes 15} = \bigoplus_{\mu\in\mathcal{L}_2^*} \bigoplus_{\eta\in\mathcal{L}_1^*} \mathcal{H}(\mu,\eta),$$

associated with the code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1)$ and the decomposition

$$(\mathbb{C}^2)^{\otimes 15} = \bigoplus_{\mu\in\mathcal{L}_2^*} \bigoplus_{\eta\in\mathcal{L}_1^*} \mathcal{G}(\mu,\eta),$$

associated with the stabilizer code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; \sigma^x, \mathcal{L}_1)$. For any syndrome $\eta \in \mathcal{L}_1^*$ one has

$$\mathcal{H}(0,\eta) = \mathcal{G}(0,\eta).$$

Moreover, for any $\mu \in \mathcal{L}_2^*$ there exists some $w \in \mathbb{F}_2^{15}$ such that for any $\eta \in \mathcal{L}_1^*$

$$\mathcal{H}(\mu,\eta) = A(w)\mathcal{G}(0,\eta). \tag{27}$$

This Lemma provides a strategy to measure a syndrome of the code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1)$ by operations from $\mathcal{O}_{\mathrm{ideal}}$. Specifically, we measure μ (i.e., the σ^z part of the syndrome) first, compute $w = w(\mu)$, apply $A(w)^\dagger$, measure η using the stabilizers $\sigma^x([x_j])$, and apply $A(w)$.

Proof of the lemma. Consider an auxiliary subspace,

$$\mathcal{H} = \bigoplus_{\eta\in\mathcal{L}_1^*} \mathcal{H}(0,\eta) = \bigoplus_{\eta\in\mathcal{L}_1^*} \mathcal{G}(0,\eta),$$

corresponding to the trivial σ^z syndrome for both CSS codes. Each state $|\Psi\rangle \in \mathcal{H}(0)$ can be represented as

$$|\Psi\rangle = \sum_{v\in\mathcal{L}_2^\perp} c_v |v\rangle,$$

where c_v are some complex amplitudes and $|v\rangle = |v_1, \ldots, v_{15}\rangle$ are vectors of the standard basis. Let us show that

$$A(u)|\Psi\rangle = \sigma^x(u)|\Psi\rangle \text{ for any } |\Psi\rangle \in \mathcal{H}, \quad u \in \mathcal{L}_1.$$

To this end, we represent A as $\sigma^x e^{i\pi/4} K^\dagger$. For any $u \in \mathcal{L}_1$ and $v \in \mathcal{L}_2^\perp$ we have

$$A(u)|v\rangle = \sigma^x(u) e^{i(\pi/4)|u| - i(\pi/2)|u\cdot v|} |v\rangle = \sigma^x(u)|v\rangle,$$

because $|u| \equiv 0 (\mathrm{mod}\, 8)$ and $|u\cdot v| \equiv 0 (\mathrm{mod}\, 4)$ (see Lemma 1, parts 1 and 5).

Since for any $u \in \mathcal{L}_1$ the operators $A(u)$ and $\sigma^x(u)$ act on $\mathcal{H}$ in the same way, their eigenspaces must coincide, i.e., $\mathcal{H}(0,\eta) = \mathcal{G}(0,\eta)$ for any $\eta \in \mathcal{L}_1^*$.

Let us now consider the subspace $\mathcal{H}(\mu,\eta)$ for arbitrary $\mu \in \mathcal{L}_2^*$, $\eta \in \mathcal{L}_1^*$. By definition, μ is a linear functional on $\mathcal{L}_2 \subseteq \mathbb{F}_2^{15}$; we can extend it to a linear functional on $\mathbb{F}_2^{15}$, i.e., represent it in the form $\mu(v) = (w,v)$ for some $w \in \mathbb{F}_2^{15}$. Then for any $|\Psi\rangle \in \mathcal{H}(\mu,\eta)$, $v \in \mathcal{L}_2$, and $u \in \mathcal{L}_1$ we have

$$\sigma^z(v)A(w)^\dagger|\Psi\rangle = (-1)^{(w,v)} A(w)^\dagger \sigma^z(v)|\Psi\rangle = A(w)^\dagger|\Psi\rangle,$$

$$A(u)A(w)^\dagger|\Psi\rangle = A(w)^\dagger A(u)|\Psi\rangle = (-1)^{\eta(v)} A(w)^\dagger|\Psi\rangle$$

(as σ^z and A anticommute), hence $A(w)^\dagger|\Psi\rangle \in \mathcal{H}(0,\eta)$. Thus

$$\mathcal{H}(\mu,\eta) = A(w)\mathcal{H}(0,\eta) = A(w)\mathcal{G}(0,\eta).$$

□

Lemma 2 is closely related to an interesting property of the stabilizer code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; \sigma^x, \mathcal{L}_1)$, namely the existence of a non-Clifford automorphism [23]. Consider a one-qubit unitary operator W such that

$$W\sigma^z W^\dagger = \sigma^z \text{ and } W\sigma^x W^\dagger = A.$$

It is defined up to an overall phase and obviously does not belong to the Clifford group $\mathcal{C}_1$. However, the bitwise application of W, i.e., the operator $W^{\otimes 15}$, preserves the code subspace $\mathcal{G}(0,0)$. Indeed, $W^{\otimes 15}\mathcal{G}(0,0)$ corresponds to the trivial syndrome of the code

$$\mathrm{CSS}(W\sigma^z W^\dagger, \mathcal{L}_2; W\sigma^x W^\dagger, \mathcal{L}_1) = \mathrm{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1).$$

Thus $W^{\otimes 15}\mathcal{G}(0,0) = \mathcal{H}(0,0)$. But $\mathcal{H}(0,0) = \mathcal{G}(0,0)$ due to the lemma.

Now we are in a position to describe the distillation scheme and to estimate its threshold and yield. Suppose we are given 15 copies of the state ρ, and our goal is to distill one copy of an H-type magic state. We will actually distill the state,

$$|A_0\rangle = \frac{1}{\sqrt{2}}(|0\rangle + e^{i\frac{\pi}{4}}|1\rangle) = e^{i\frac{\pi}{8}} HK^\dagger |H\rangle.$$

Note that $|A_0\rangle$ is an eigenstate of the operator A; specifically, $A|A_0\rangle = |A_0\rangle$. Let us also introduce the state

$$|A_1\rangle = \sigma^z|A_0\rangle,$$

which satisfies $A|A_1\rangle = -|A_1\rangle$. Since the Clifford group $\mathcal{C}_1$ acts transitively on the set of H-type magic states, we can assume that the fidelity between ρ and $|A_0\rangle$ is the maximum one among all H-type magic states, so that

$$F_H(\rho) = \sqrt{\langle A_0|\rho|A_0\rangle}.$$

As in Sec. V we define the initial error probability

$$\epsilon = 1 - [F_H(\rho)]^2 = \langle A_1|\rho|A_1\rangle.$$

Applying the dephasing transformation

$$D(\eta) = \frac{1}{2}(\eta + A\eta A^\dagger)$$

to each copy of ρ, we can guarantee that ρ is diagonal in the $\{A_0, A_1\}$ basis, i.e.,

$$\rho = D(\rho) = (1-\epsilon)|A_0\rangle\langle A_0| + \epsilon|A_1\rangle\langle A_1|.$$

Since $A \in \mathcal{C}_1$, the dephasing transformation can be realized by operations from $\mathcal{O}_{\rm ideal}$. Thus our initial state is

$$\rho_{\rm in} = \rho^{\otimes 15} = \sum_{u \in \mathbb{F}_2^{15}} \epsilon^{|u|}(1-\epsilon)^{15-|u|}|A_u\rangle\langle A_u|, \qquad (28)$$

where $|A_u\rangle = |A_{u_0}\rangle \otimes \cdots \otimes |A_{u_{15}}\rangle$.

According to the remark following the formulation of Lemma 2, we can measure the syndrome (μ, η) of the code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; A, \mathcal{L}_1)$ by operations from $\mathcal{O}_{\rm ideal}$ only. Let us follow this scheme, omitting the very last step. So, we begin with the state $\rho_{\rm in}$, measure μ, compute $w = w(\mu)$, apply $A(w)^\dagger$, and measure η. We consider the distillation attempt successful if $\eta = 0$. The measured value of μ is not important at this stage. In fact, for any $\mu \in \mathcal{L}_2^*$ the unnormalized post-measurement state is

$$\rho_s = \Pi A(w)^\dagger \rho_{\rm in} A(w)\Pi = \Pi\rho_{\rm in}\Pi.$$

In this equation Π is the projector onto the code subspace $\mathcal{H}(0,0) = \mathcal{G}(0,0)$, i.e., $\Pi = \Pi_z\Pi_A$ for

$$\Pi_z = \frac{1}{|\mathcal{L}_2|}\sum_{v \in \mathcal{L}_2} \sigma^z(v), \quad \Pi_A = \frac{1}{|\mathcal{L}_1|}\sum_{u \in \mathcal{L}_1} A(u). \qquad (29)$$

Let us compute the state $\rho_s = \Pi\rho_{\rm in}\Pi$. Since

$$A(u)|A_w\rangle = (-1)^{(u,w)}|A_w\rangle, \quad \sigma^z(v)|A_w\rangle = |A_{w+v}\rangle,$$

one can easily see that $\Pi_A|A_w\rangle = |A_w\rangle$ if $w \in \mathcal{L}_1^\perp$, otherwise $\Pi_A|A_w\rangle = 0$. On the other hand, $\Pi_z|A_w\rangle$ does not vanish and depends only on the coset of $\mathcal{L}_2$ that contains w. There are only two such cosets in $\mathcal{L}_1^\perp$ (because $\mathcal{L}_1^\perp = \mathcal{L}_2 \oplus [1]$, see Lemma 1), and the corresponding projected states are

$$|A_0^L\rangle = \sqrt{|\mathcal{L}_2|}\Pi_z|A_{0\cdots 0}\rangle = \frac{1}{\sqrt{|\mathcal{L}_2|}}\sum_{v \in \mathcal{L}_2}|A_v\rangle,$$

$$|A_1^L\rangle = \sqrt{|\mathcal{L}_2|}\Pi_z|A_{1\cdots 1}\rangle = \frac{1}{\sqrt{|\mathcal{L}_2|}}\sum_{v \in \mathcal{L}_2}|A_{v+[1]}\rangle. \qquad (30)$$

The states $|A_{0,1}^L\rangle$ form an orthonormal basis of the code subspace. The projections of $|A_w\rangle$ for $w \in \mathcal{L}_1^\perp$ onto the code subspace are given by these formulas:

$$\Pi|A_w\rangle = \frac{1}{\sqrt{|\mathcal{L}_2|}}|A_0^L\rangle \text{ if } w \in \mathcal{L}_2,$$

$$\Pi|A_w\rangle = \frac{1}{\sqrt{|\mathcal{L}_2|}}|A_1^L\rangle \text{ if } w \in \mathcal{L}_2 + [1].$$

Now the unnormalized final state $\rho_s = \Pi\rho_{\rm in}\Pi$ can be expanded as

$$\rho_s \frac{1}{|\mathcal{L}_2|}\sum_{v \in \mathcal{L}_2}(1-\epsilon)^{15-|v|}\epsilon^{|v|}|A_0^L\rangle\langle A_0^L|$$

$$\times + \frac{1}{|\mathcal{L}_2|}\sum_{v \in \mathcal{L}_2}\epsilon^{15-|v|}(1-\epsilon)^{|v|}|A_1^L\rangle\langle A_1^L|.$$

The distillation succeeds with probability

$$p_s = |\mathcal{L}_2|\mathrm{Tr}\,\rho_s = \sum_{v \in \mathcal{L}_1^\perp}\epsilon^{15-|v|}(1-\epsilon)^{|v|}.$$

(The factor $|\mathcal{L}_2|$ reflects the number of possible values of μ, which all give rise to the same state ρ_s.)

To complete the distillation procedure, we need to apply a decoding transformation that would map the two-dimensional subspace $\mathcal{H}(0,0) \subset (\mathbb{C}^2)^{\otimes 15}$ onto the Hilbert space of one qubit. Recall that $\mathcal{H}(0,0) = \mathcal{G}(0,0)$ is the code subspace of the stabilizer code $\mathrm{CSS}(\sigma^z, \mathcal{L}_2; \sigma^x, \mathcal{L}_1)$. Its logical Pauli operators can be chosen as

$$\hat{X} = (\sigma^x)^{\otimes 15}, \quad \hat{Y} = (\sigma^y)^{\otimes 15}, \quad \hat{Z} = -(\sigma^z)^{\otimes 15}.$$

It is easy to see that $\hat{X}, \hat{Y}, \hat{Z}$ obey the correct algebraic relations and preserve the code subspace. The decoding can be realized as a Clifford operator $V \in \mathcal{C}_{15}$ that maps $\hat{X}, \hat{Y}, \hat{Z}$ to the Pauli operators $\sigma^x, \sigma^y, \sigma^z$ acting on the first qubit. (The remaining 14 qubits become unentangled with the first one, so we can safely disregard them.) Let us show that the logical state $|A_0^L\rangle$ is transformed into $|A_0\rangle$ (up to some phase). For this, it suffices to check that $\langle A_0^L|\hat{X}|A_0^L\rangle = \langle A_0|\sigma^x|A_0\rangle$, $\langle A_0^L|\hat{Y}|A_0^L\rangle = \langle A_0|\sigma^y|A_0\rangle$, and $\langle A_0^L|\hat{Z}|A_0^L\rangle = \langle A_0|\sigma^z|A_0\rangle$. Verifying these identities becomes a straightforward task if we represent $|A_0^L\rangle$ in the standard basis:

$$|A_0^L\rangle = |\mathcal{L}_2|^{1/2}2^{-15/2}\sum_{u \in \mathcal{L}_2^\perp} e^{i(\pi/4)|u|}|u\rangle$$

$$= 2^{-5/2}\sum_{u \in \mathcal{L}_1}(|u\rangle + e^{-i(\pi/4)}|u+[1]\rangle).$$

To summarize, the distillation subroutine consists of the following steps.

(1) Measure eigenvalues of the Pauli operators $\sigma^z([x_j])$, $\sigma^z([x_jx_k])$ (for $j,k = 1,2,3,4$). The outcomes determine the σ^z syndrome, $\mu \in \mathcal{L}_2^*$.

(2) Find $w = w(\mu) \in \mathbb{F}_2^{15}$ such that $(w,v) = \mu(v)$ for any $v \in \mathcal{L}_2$.

(3) Apply the correcting operator $A(w)^\dagger$.

(4) Measure eigenvalues of the operators $\sigma^x([x_j])$. The outcomes determine the A syndrome, $\eta \in \mathcal{L}_1^*$.

(5) Declare failure if $\eta \neq 0$, otherwise proceed to the next step.

(6) Apply the decoding transformation, which takes the

code subspace to the Hilbert space of one qubit.

The subroutine succeeds with probability

$$p_s = \sum_{v \in \mathcal{L}_1^\perp} \epsilon^{15-|v|}(1-\epsilon)^{|v|}. \tag{31}$$

In the case of success, it produces the normalized output state

$$\rho_{\text{out}} = (1-\epsilon_{\text{out}})|A_0\rangle\langle A_0| + \epsilon_{\text{out}}|A_1\rangle\langle A_1| \tag{32}$$

characterized by the error probability

$$\epsilon_{\text{out}} = p_s^{-1} \sum_{v \in \mathcal{L}_2} \epsilon^{15-|v|}(1-\epsilon)^{|v|}. \tag{33}$$

The sums in Eqs. (31) and (33) are special forms of so-called weight enumerators. The *weight enumerator* of a subspace $\mathcal{L} \subseteq \mathbb{F}_2^n$ is a homogeneous polynomial of degree n in two variables, namely

$$W_{\mathcal{L}}(x,y) = \sum_{u \in \mathcal{L}} x^{n-|u|} y^{|u|}.$$

In this notation,

$$p_s = W_{\mathcal{L}_1^\perp}(\epsilon, 1-\epsilon), \quad \epsilon_{\text{out}} = \frac{W_{\mathcal{L}_2}(\epsilon,1-\epsilon)}{W_{\mathcal{L}_1^\perp}(\epsilon,1-\epsilon)}.$$

The MacWilliams identity [30] relates the weight enumerator of $\mathcal{L}$ to that of $\mathcal{L}^\perp$:

$$W_{\mathcal{L}}(x,y) = \frac{1}{|\mathcal{L}^\perp|} W_{\mathcal{L}^\perp}(x+y, x-y).$$

Applying this identity and taking into account that $\mathcal{L}_2^\perp = \mathcal{L}_1 \oplus [1]$ and that $|u| \equiv 0 (\text{mod } 2)$ for any $u \in \mathcal{L}_1$ (see Lemma 1), we get

$$p_s = \frac{1}{16} W_{\mathcal{L}_1}(1, 1-2\epsilon), \quad \epsilon_{\text{out}} = \frac{1}{2}\left(1 - \frac{W_{\mathcal{L}_1}(1-2\epsilon,1)}{W_{\mathcal{L}_1}(1,1-2\epsilon)}\right). \tag{34}$$

The weight enumerator of the subspace $\mathcal{L}_1$ is particularly simple:

$$W_{\mathcal{L}_1}(x,y) = x^{15} + 15x^7y^8.$$

Substituting this expression into Eq. (34), we arrive at the following formulas:

$$p_s = \frac{1 + 15(1-2\epsilon)^8}{16}, \tag{35}$$

$$\epsilon_{\text{out}} = \frac{1 - 15(1-2\epsilon)^7 + 15(1-2\epsilon)^8 - (1-2\epsilon)^{15}}{2[1 + 15(1-2\epsilon)^8]}. \tag{36}$$

The function $\epsilon_{\text{out}}(\epsilon)$ is plotted in Fig. 3. Solving the equation $\epsilon_{\text{out}}(\epsilon) = \epsilon$ numerically, we find the threshold error probability:

$$\epsilon_0 \approx 0.141. \tag{37}$$

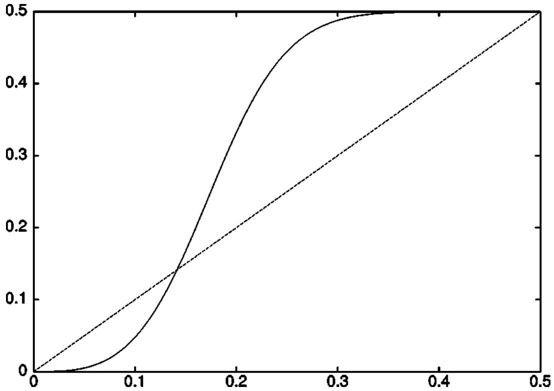

FIG. 3. The final error probability $\epsilon_{\text{out}}(\epsilon)$ for the H-type states distillation.

Let us examine the asymptotic properties of this scheme. For small ϵ the distillation subroutine succeeds with probability close to 1, therefore the yield is close to 1/15. The output error probability is

$$\epsilon_{\text{out}} \approx 35\epsilon^3. \tag{38}$$

Now suppose that the subroutine is applied recursively. From n copies of the state ρ with a given ϵ, we distill one copy of the magic state $|A_0\rangle$ with the final error probability

$$\epsilon_{\text{out}}(n,\epsilon) \approx \frac{1}{\sqrt{35}}(\sqrt{35}\epsilon)^{3^k}, \quad 15^k \approx n,$$

where k is the number of recursion levels (here we neglect the fluctuations in the number of successful distillation attempts). Solving these equation, we obtain the relation

$$\epsilon_{\text{out}}(n,\epsilon) \sim (\sqrt{35}\epsilon)^{n^\xi}, \quad \xi = 1/\log_3 15 \approx 0.4. \tag{39}$$

It characterizes the efficiency of the distillation scheme.

VII. CONCLUSION AND SOME OPEN PROBLEMS

We have studied a simplified model of fault-tolerant quantum computation in which operations from the Clifford group are realized exactly, whereas decoherence occurs only during the preparation of nontrivial ancillary states. The model is fully characterized by a one-qubit density matrix ρ describing these states. It is shown that a good strategy for simulating universal quantum computation in this model is "magic states distillation." By constructing two particular distillation schemes we find a threshold polarization of ρ above which the simulation is possible.

The most exciting open problem is to understand the computational power of the model in the region of parameters $1 < |\rho_x| + |\rho_y| + |\rho_z| \leq 3/\sqrt{7}$ (which corresponds to $F_T^* < F_T(\rho) \leq F_T$, see Sec. I). In this region, the distillation scheme based on the five-quit code does not work, while the Gottesman-Knill theorem does not yet allow the classical simulation. One possibility is that a transition from classical to universal quantum behavior occurs on the octahedron boundary, $|\rho_x| + |\rho_y| + |\rho_z| = 1$.

To prove the existence of such a transition, one it suffices to construct a T-type states distillation scheme having the threshold fidelity F_T^*. A systematic way of constructing such schemes is to replace the five-qubit by a $GF(4)$-linear stabilizer code. A nice property of these codes is that the bitwise application of the operator T preserves the code subspace and acts on the encoded qubit as T, see Ref. [31] for more details. One can check that the error-correcting effect described in Sec. V takes place for an arbitrary GF(4)-linear stabilizer code, provided that the number of qubits is $n=6k-1$ for any integer k. Unfortunately, numerical simulations we performed for some codes with $n=11$ and $n=17$ indicate that the threshold fidelity increases as the number of qubits increases. So it may well be the case that the five-qubit code is the best GF(4)-linear code as far as the distillation is concerned.

From the experimental point of view, an exciting open problem is to design a physical system in which reliable storage of quantum information and its processing by Clifford group operations is possible. Since our simulation scheme tolerates strong decoherence on the ancilla preparation stage, such a system would be a good candidate for a practical quantum computer.

ACKNOWLEDGMENTS

We thank Mikhail Vyalyi for bringing to our attention many useful facts about the Clifford group. This work has been supported in part by the National Science Foundation under Grant No. EIA-0086038.

APPENDIX

The purpose of this section is to prove Eq. (15). Let us introduce this notation:

$$|\hat{T}_0\rangle = |T_{00000}\rangle \text{ and } |\hat{T}_1\rangle = |T_{11111}\rangle.$$

Consider the set $S_+(5) \subset S(5)$ consisting of all possible tensor products of the Pauli operators $\sigma^x, \sigma^y, \sigma^z$ on five qubits (clearly, $|S_+(5)| = 4^5 = |S(5)|/2$ since elements of $S(5)$ may have a plus or minus sign). For each $g \in S_+(5)$ let $|g| \in [0,5]$ be the number of qubits on which g acts nontrivially (e.g., $|\sigma^x \otimes \sigma^x \otimes \sigma^y \otimes I \otimes I| = 3$). We have

$$|\hat{T}_0\rangle\langle\hat{T}_0| = \frac{1}{2^5} \sum_{g \in S_+(5)} \left(\frac{1}{\sqrt{3}}\right)^{|g|} g.$$

Now let us expand the formula (13) for the projector Π. Denote by $G \subset P(5)$ the Abelian group generated by the stabilizers S_1, S_2, S_3, S_4. It consists of 16 elements. Repeatedly conjugating the stabilizer S_1 by the operator $\hat{T} = T^{\otimes 5}$, we get three elements of G:

$$S_1 = \sigma^x \otimes \sigma^z \otimes \sigma^z \otimes \sigma^x \otimes I,$$

$$S_1 S_3 S_4 = \sigma^z \otimes \sigma^y \otimes \sigma^y \otimes \sigma^z \otimes I,$$

$$S_3 S_4 = \sigma^y \otimes \sigma^x \otimes \sigma^x \otimes \sigma^y \otimes I.$$

Due to the cyclic symmetry mentioned in Sec. V, the 15 cyclic permutations of these elements also belong to G; together with the identity operator they exhaust the group G. Thus $G \subset S_+(5)$, and we have

$$\Pi = \frac{1}{16} \sum_{h \in G} h.$$

Taking into account that $\mathrm{Tr}(gh) = 2^5 \delta_{g,h}$ for any $g, h \in S_+(5)$, we get

$$\langle\hat{T}_0|\Pi|\hat{T}_0\rangle = \frac{1}{2^9} \sum_{h \in G} \sum_{g \in S_+(5)} 3^{-|g|/2} \mathrm{Tr}(gh) = \frac{1}{16} \sum_{g \in G} 3^{-|g|/2} = \frac{1}{6}.$$

Similar calculations show that $\langle\hat{T}_1|\Pi|\hat{T}_1\rangle = \frac{1}{6}$.

[1] E. Knill, R. Laflamme, and W. Zurek, Science **279**, 342 (1998).

[2] C. Zalka, e-print quant-ph/9612028.

[3] A. Steane, Phys. Rev. Lett. **78**, 2252 (1997).

[4] E. Dennis, A. Kitaev, A. Landahl, and J. Preskill, J. Math. Phys. **43**, 4452 (2002).

[5] A. Kitaev, Ann. Phys. (N.Y.) **303**, 2 (2003).

[6] M. Freedman, M. Larsen, and Z. Wang, e-print quant-ph/0001108.

[7] M. Freedman, A. Kitaev, M. Larsen, and Z. Wang, Bull., New Ser., Am. Math. Soc. **40**, 31 (2002).

[8] C. Mochon, Phys. Rev. A **69**, 032306 (2004).

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

[10] C. Nayak and F. Wilczek, Nucl. Phys. B **479**, 529 (1996).

[11] B. Doucot and J. Vidal, Phys. Rev. Lett. **88**, 227005 (2001).

[12] M. Feigel'man and L. Ioffe, Phys. Rev. B **66**, 224503 (2002).

[13] J. Preskill, e-print quant-ph/9712048.

[14] D. Gottesman, Ph.D. thesis, Caltech, Pasadena, 1997, URL http://arxiv.org/abs/quant-ph/9705052.

[15] M. Nielsen and I. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).

[16] D. Gottesman, Phys. Rev. A **57**, 127 (1998).

[17] E. Knill, e-print quant-ph/0402171.

[18] E. Knill, e-print quant-ph/0404104.

[19] D. Gottesman and I. Chuang, Nature (London) **402**, 390 (1999).

[20] W. Dur and H. Briegel, Phys. Rev. Lett. **90**, 067901 (2003).

[21] D. Aharonov, e-print quant-ph/9602019.

[22] E. Dennis, Phys. Rev. A **63**, 052314 (2001).

[23] E. Knill, R. Laflamme, and W. Zurek, e-print quant-ph/9610011.

[24] A. Calderbank, E. Rains, P. Shor, and N. Sloane, Phys. Rev. Lett. **78**, 405 (1997).

[25] A. Ambainis and D. Gottesman, e-print quant-ph/0310097.

[26] C. Bennett, D. DiVincenzo, J. Smolin, and W. Wootters, Phys. Rev. A **54**, 3824 (1996).

[27] R. Laflamme, C. Miquel, J. Paz, and W. Zurek, Phys. Rev. Lett. **77**, 198 (1996).

[28] A. Calderbank and P. Shor, Phys. Rev. A **54**, 1098 (1996).

[29] A. Steane, Proc. R. Soc. London, Ser. A **452**, 2551 (1996).

[30] F. MacWilliams and N. Sloane, *The Theory of Error-Correcting Codes* (North-Holland, Amsterdam, 1981).

[31] A. Calderbank, E. Rains, P. Shor, and N. Sloane, e-print quant-ph/9608006.

Tight Noise Thresholds for Quantum Computation with Perfect Stabilizer Operations

Wim van Dam*
Department of Computer Science, University of California, Santa Barbara, California 93106, USA
and Department of Physics, University of California, Santa Barbara, California 93106, USA

Mark Howard†
Department of Physics, University of California, Santa Barbara, California 93106, USA

(Received 21 July 2009; published 23 October 2009)

We study how much noise can be tolerated by a universal gate set before it loses its quantum-computational power. Specifically we look at circuits with perfect stabilizer operations in addition to imperfect nonstabilizer gates. We prove that for all unitary single-qubit gates there exists a tight depolarizing noise threshold that determines whether the gate enables universal quantum computation or if the gate can be simulated by a mixture of Clifford gates. This exact threshold is determined by the Clifford polytope spanned by the 24 single-qubit Clifford gates. The result is in contrast to the situation wherein nonstabilizer qubit states are used; the thresholds in that case are not currently known to be tight.

DOI: 10.1103/PhysRevLett.103.170504 PACS numbers: 03.67.Lx, 03.67.Pp

Introduction.—A way to study the resources needed for universal quantum computation (UQC) is to analyze the transition from a system that can provide UQC to one that is classically efficiently simulable. A particularly useful example of a classically simulable system is given by the stabilizer operations, which are made by a combination of preparation of $|0\rangle$ states, unitary Clifford gates, measurements in the $\{|0\rangle, |1\rangle\}$ basis, and classical control determined by the measurement outcomes. The Gottesman-Knill theorem tells us that stabilizer operations can be efficiently simulated classically (see, for example, [1], Theorem 10.7), while it also known that the addition of any other one-qubit gate outside the Clifford group will enable the system to perform UQC. This fact provides us with a framework for testing tolerance to noise—one can examine how noisy this additional non-Clifford gate can be before it becomes classically simulable itself. If the non-Clifford operation has become a probabilistic combination of Clifford gates due to the noise, then we know that we are unequivocally in the classical computational regime. The noise rate where the extra gate becomes simulable (where it enters the "Clifford polytope" [2]) is thus an upper bound for fault tolerance. If the converse is true—i.e., if any operation outside the Clifford polytope enables UQC, then the threshold is tight. In this Letter we show that for single-qubit gates undergoing depolarizing such a tight noise threshold does indeed apply. We will do so by proving that any depolarized gate that lies outside the Clifford polytope of single-qubit operations, in combination with noiseless stabilizing operations, allows for UQC. This result should be contrasted to the situation for nonstabilizer qubit *states* where the thresholds in that case are not currently known to be tight. In fact, a recent result by Campbell and Browne [3] states that achieving tight thresholds for all nonstabilizer qubit states is impossible if the number of copies of the resource state must be finite.

We will consistently assume that Clifford gates can be implemented perfectly, motivated by the fact that these gates can be implemented fault tolerantly by applying them transversally and to encoded states [4–7]. The fault-tolerant implementation of Clifford gates naturally carries with it a threshold of its own, independent of the kind we discuss in this Letter. The current model is particularly relevant to the so-called Pfaffian quantum Hall state in topological quantum computation [8,9], the two-qubit Clifford group (but only the Clifford group) can be implemented using braiding making these operations naturally fault tolerant. The additional resource required to perform UQC will likely be highly noisy, and so we can see the parallels with our model.

We will begin by listing a couple of previously known results in this area. Next, we will discuss the connection between the geometry of the Clifford polytope and stabilizer measurements, and show that tightness of a magic-state distillation procedure for single-qubit states automatically ensures tight thresholds for non-Clifford gates undergoing any kind of unital noise. Finally, we show that currently known magic-state distillation techniques are sufficient to prove tight thresholds for a non-Clifford gate undergoing depolarizing noise.

Previously known results.—The idea of using perfect stabilizer operations in conjunction with imperfect nonstabilizer states to perform UQC originates with Knill [7]. Shortly after, Bravyi and Kitaev [10] showed that most nonstabilizer qubit states (when sufficiently many copies are available) can be purified ("distilled"), using only stabilizer operations, towards a pure nonstabilizer state (a "magic state"). Since a universal gate set can be created from perfect stabilizer operations and a supply of magic states [10], we see that allowing access to a supply of appropriate (possibly impure) nonstabilizer qubit ancillas promotes the power of stabilizer operations from classically simulable to UQC. The conditions on the ancillary qubits to enable UQC is that they are sufficiently close to being one of the 20 so-called magic states that lie on the surface of the Bloch sphere. The two classes of magic state (see Fig. 1) are the $|H\rangle$ type and $|T\rangle$ type, where all $|H\rangle$ type states can be derived by applying a Clifford operation to some canonical representative $|H\rangle = (|0\rangle + e^{i\pi/4}|1\rangle)/\sqrt{2}$, and similarly for $|T\rangle$-type states like $|T\rangle = \cos(\vartheta)|0\rangle + e^{i\pi/4}\sin(\vartheta)|1\rangle$, where $\cos(2\vartheta) = 1/\sqrt{3}$. The routines used in [10] were unable to distill qubit states just outside the edges and faces of the octahedron of Fig. 1. Reichardt [11] subsequently proposed an improved routine that closed the gap in the $|H\rangle$ direction (along the edges of the octahedron). Virmani *et al.* [12] suggested using the convex hull of Clifford operations in order to find gates' robustness to various types of noise. In particular, they considered gates that are diagonal in the computational basis. Plenio and Virmani [13] subsequently extended this idea by analyzing cases where noise was allowed to affect the stabilizer operations too. Buhrman *et al.* [2] used a similar idea (that noise causes non-Clifford gates to eventually become able to be implemented via Clifford gates only) to find the non-Clifford gate that is most resistant to depolarizing noise—a $\pi/8$ rotation about the Z axis (or the same gate modulo some Clifford operation). Reichardt [14] showed that this particular gate enabled UQC right up to its threshold noise rate (about 45%), as well as considering in detail the process of reducing multiqubit states to single-qubit states using postselected stabilizer operations. Our current result here generalizes this tightness result to all possible single-qubit gates.

Preliminaries and notation.—Let us parameterize an arbitrary single-qubit SU(2) gate as follows

$$U(\theta,\gamma,\delta) = \begin{pmatrix} e^{i\gamma}\cos(\theta) & -e^{i\delta}\sin(\theta) \\ e^{-i\delta}\sin(\theta) & e^{-i\gamma}\cos(\theta) \end{pmatrix}. \qquad (1)$$

The representation of this rotation in SO(3) is denoted by $R(\theta,\gamma,\delta)$. Implementing a rotation R while suffering depolarizing noise (with noise rate p), means that this noisy operation is represented by the rescaling $M = (1-p)R$, a fact that we will need later.

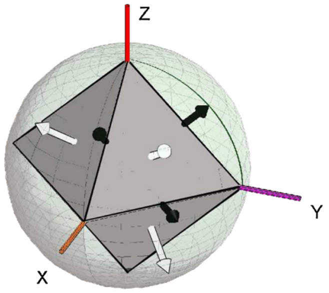

FIG. 1 (color online). Magic states and the octahedron: Some of the single-qubit magic states: $|H\rangle$ type states are designated with black arrows, $|T\rangle$ type states with white arrows. The octahedron defined by $|x| + |y| + |z| \leq 1$ depicts the single-qubit states that can be created by stabilizer operations. Reichardt [11] showed that distillation techniques work right up to the edges of the octahedron (i.e., tight in the $|H\rangle$ direction). Current distillation techniques are unable to distill states just outside the faces of the octahedron (i.e., not tight in the $|T\rangle$ direction).

Often we will apply the unitary $U(\theta,\gamma,\delta)$ to one half of an entangled Bell pair, $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$, yielding

$$\rho = (I \otimes U)|\Phi\rangle\langle\Phi|(I\otimes U)^\dagger. \qquad (2)$$

If we use the two-qubit Pauli operators as a basis for the density matrix ρ then we can find the 16 real coefficients $c_{ij} = \mathrm{Tr}(\rho(\sigma_i \otimes \sigma_j))$ so that

$$\rho = \frac{1}{4}\sum c_{ij}(\sigma_i \otimes \sigma_j), \qquad i,j \in \{I,X,Y,Z\}. \qquad (3)$$

Since we have applied a local unitary to a maximally entangled state, the coefficients (c_{IX}, c_{IY}, c_{IZ}, c_{XI}, c_{YI}, c_{ZI}) are always zero. Comparing the 9 coefficients $\{c_{XX}, c_{XY}, \ldots, c_{ZZ}\}$ one can see that these are the same as the entries of the SO(3) matrix $R(\theta,\gamma,\delta)$. More precisely,

$$R(\theta,\gamma,\delta) = \begin{pmatrix} c_{XX} & -c_{YX} & c_{ZX} \\ c_{XY} & -c_{YY} & c_{ZY} \\ c_{XZ} & -c_{YZ} & c_{ZZ} \end{pmatrix}, \qquad (4)$$

where the c_{ij} are obviously also functions of (θ,γ,δ).

If we represent the 24 single-qubit Clifford operations as SO(3) matrices, then they are simply signed permutation matrices with unit determinant (they are a matrix representation of the elements of the chiral octahedral symmetry group or, equivalently, the symmetry group S_4). We label these operations C_i and so the convex hull of the C_i (the so-called Clifford polytope) is given by

$$\mathcal{P} = \left\{\sum_{i=1}^{24} p_i C_i \;\middle|\; \text{with } p_i \geq 0 \text{ and } \sum_{i=1}^{24} p_i = 1\right\}. \qquad (5)$$

Geometrically, the Clifford polytope is a closed polyhedron in $\mathbb{R}^9$ that has 24 vertices (each vertex representing one of the C_i). This polytope can also be defined by the bounding inequalities of its 120 facets. The concise description of these facets used by Buhrman *et al.* [2] is given by the set

$$\mathcal{F} = \{C_i F C_j | i,j \in \{1,\ldots,24\}, F \in \{A, A^T, B\}\}, \qquad (6)$$

where

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{pmatrix}. \qquad (7)$$

At times we will have reason to refer to different subsets of the set of facets $\mathcal{F}$ so we use the obvious notation $\mathcal{F} = \mathcal{F}_A \cup \mathcal{F}_{A^T} \cup \mathcal{F}_B$. It is useful to note that all the facets derived from A comprise a single column with ± 1 entries and zeros elsewhere, and similarly for the row facets derived from A^T; hence, $|\mathcal{F}_A| = |\mathcal{F}_{A^T}| = 3 \times 2^3 = 24$. There are $|\mathcal{F}_B| = 72$ "B-type" facets, which can be constructed as follows: (i) Pick one position in a 3×3 matrix F, e.g., row i and column j and put ± 1 there ($9 \times 2 = 18$

choices), (ii) Fill the remaining entries not in row i or column j with ± 1 such that $\det(F) = -2$ (4 choices).

To determine whether or not an operation M is inside the Clifford polytope $\mathcal{P}$ we take the elementwise inner product (or Frobenius inner product) between M and the facets $F \in \mathcal{F}$ of the polytope

$$M \cdot F = \sum_{i,j=1}^{3} M_{i,j} F_{i,j} = \mathrm{Tr}(M^T F). \qquad (8)$$

Using the above notation, a 3×3 matrix M is inside the polytope $\mathcal{P}$ if and only if for all $F \in \mathcal{F}$ we have $M \cdot F \leq 1$.

Interpreting the facets of the Clifford polytope.—Our proof will involve applying some non-Clifford gate to one half of a Bell Pair [as in Eq. (2)] and then postselecting on the outcomes of various stabilizer measurements. After some further stabilizer operations, this measurement ultimately has the effect of taking our two-qubit state ρ to a single-qubit state ρ' (times some stabilizer state that we do not care about), which we then distill using magic-state distillation (see [14] for a more general discussion of these kinds of techniques). For example, performing a YX measurement on ρ and postselecting on a "+1" outcome (i.e., projecting with $\Pi = \frac{1}{2}(II + YX)$) leads to a single-qubit state ρ' with a Bloch vector given by

$$\vec{r}(\rho') = \left(0, \frac{c_{XZ} - c_{ZY}}{c_{II} + c_{YX}}, -\frac{c_{XY} + c_{ZZ}}{c_{II} + c_{YX}}\right). \qquad (9)$$

The form of this vector means that it lies in the YZ plane (see Fig. 1), where we know that distillation techniques work right up to the edge $|y| + |z| = 1$ of the octahedron. We can check if $\vec{r}$ is outside the octahedron by simply comparing the L^1 norm of $\vec{r}$ with 1. Rearranged, the condition $\|\vec{r}\|_1 > 1$ for being outside the octahedron is

$$|c_{XZ} - c_{ZY}| + |-(c_{XY} + c_{ZZ})| > |c_{II} + c_{YX}|. \qquad (10)$$

Given the correspondence between the coefficients c_{ij} and the elements of R [see Eq. (4)] we can rewrite the above condition (dropping the absolute value operators) as a facet inequality

$$R \cdot F > 1 \quad \text{where} \quad F = \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}. \qquad (11)$$

This facet is a legitimate "B-type" facet and a little thought shows that, had we applied the single-qubit Pauli operations [as SO(3) rotations] X, Y or Z to $\vec{r}(\rho')$ above, we would arrive at three other "B-type" facets

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ -1 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & -1 \\ 1 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ -1 & 0 & 1 \\ -1 & 0 & -1 \end{pmatrix}, \qquad (12)$$

respectively. Note that all four facet inequalities combined could be simplified to the form Eq. (10) above. We omit the details, but it is straightforward to show that all 72 "B-type" facets correspond to postselecting on some (weight two) Pauli operator, and possibly performing a single-qubit Pauli rotation on the resulting ρ'.

It is somewhat more straightforward to see the geometrical interpretation of the "$A^{(T)}$-type" facets. For example, the canonical A given in Eq. (7), merely returns the sum of the elements of the Bloch vector $\vec{r}$, arising from a rotation applied to the X "+1" eigenstate.

$$R \cdot A = \sum_{i=1}^{3} r_i \quad \text{where} \quad \vec{r} = R \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}. \qquad (13)$$

In general, an operation M having an inner product greater than one with some "A-type" facet simply means that M, applied to some initial vector corresponding to a stabilizer state, brings that vector to a final position outside the octahedron.

The preceding discussion shows us that if magic-state distillation were possible everywhere outside the octahedron, then every unital operation outside the Clifford polytope would be distillable—either straightforwardly or by using postselection, depending on which facet it violated. Using current (not tight) distillation routines however, we would be unable to deal with some operations violating an "A-type" facet by a fairly small amount. In the next section we show that, for depolarizing noise, any noisy rotation violating an "A-type" facet also violates a "B-type" facet. Since "B-type" facets correspond to $|H\rangle$ state distillation, the results we obtain are tight.

Tight threshold for depolarizing noise.—The claim we shall prove is that, anytime a matrix $M = (1 - p)R$, representing a depolarized rotation, is outside some "A-type" facet then there exists a "B-type" facet that M also lies outside. In fact, we will prove the slightly stronger statement that for all $R \in \mathrm{SO}(3)$

$$\forall A \in \mathcal{F}_A \cup \mathcal{F}_{A^T}, \quad \exists B \in \mathcal{F}_B \ \text{ such that } R \cdot (B - A) \geq 0. \qquad (14)$$

To simplify the proof we will repeatedly make use of the symmetries of the problem [see Eq. (6)]. We will pick a canonical "A-type" facet and assume that this gives the largest inner product with R of all the $F \in \mathcal{F}_A$. If there was a larger inner product with some $F \in \mathcal{F}_{A^T}$, then we could just relabel R^T as R. We can assume that the facet with ones in the first column [the A in Eq. (7)] gives the biggest inner product since $\mathrm{Tr}[R^T(C_i A C_j)] = \mathrm{Tr}(C_j R^T C_i A) = \mathrm{Tr}[(C_k R C_l)^T A]$, which shows that the inner product of R with a different $F \in \mathcal{F}_A$ is the same as the inner product between a different (but related via Clifford operations) rotation and the canonical "A-type" facet.

The proof will hinge on an entry of R, outside of the first column, being larger than the rest of the elements outside the first column. As such, let us define 12 matrices closely related to A and call them A'_i

$$A'_1 = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad A'_2 = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix} \quad \cdots \quad A'_{12} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} \qquad (15)$$

such that the index i of the largest inner product $R \cdot A_i'$ tells us the sign and location of the largest magnitude element outside the first column. Once again, symmetry allows us to assume that A_1' yields the largest inner product because the rest of the A_i' can be derived from A_1' via Clifford rotations

$$\{A_i'\} = \left\{ \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}^j A_1' \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}^k \middle| \begin{matrix} j \in \{1,2,3\} \\ k \in \{1,2,3,4\} \end{matrix} \right\}. \tag{16}$$

For a matrix R to be an SO(3) rotation there are constraints on the signs of the elements $R_{i,j}$; i.e., there are eight choices for the first column, 6 choices for the second column and 2 for the third. Given that A is the maximum facet for R, we have fixed the signs positively in the first column, reducing the number of types of rotation to $6 \times 2 = 12$. Since A_1' gives the maximum inner product with R of all A_i' we have that $R_{1,2} < 0$, which reduces the number of rotation types further to $3 \times 2 = 6$. It can be shown that $R_{1,2}$ having larger magnitude than the rest of the elements $R_{i,j} (i \in \{1, 2, 3\}, j \in \{2, 3\})$ restricts the type of rotation further to one the following four types

$$R \in \left\{ \begin{pmatrix} + & - & + \\ + & + & - \\ + & + & + \end{pmatrix}, \begin{pmatrix} + & - & - \\ + & + & - \\ + & + & + \end{pmatrix}, \begin{pmatrix} + & - & - \\ + & + & - \\ + & - & + \end{pmatrix}, \begin{pmatrix} + & - & + \\ + & - & - \\ + & + & + \end{pmatrix} \right\}. \tag{17}$$

This should not be surprising if one considers that $R_{1,2} = -(R_{2,1}R_{3,3} - R_{2,3}R_{3,1})$ because of the structure of SO(3) matrices, and the sign patterns listed above ensure $|R_{1,2}|$ is as large as possible.

We claim that the $B \in \mathcal{F}_B$ of Eq. (9) will suffice to prove the desired inequality $R \cdot (B - A) \geq 0$, which reads in matrix form

$$\begin{pmatrix} + & - & \cdot \\ + & \cdot & - \\ + & \cdot & + \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix} \geq 0. \tag{18}$$

Using the relevant entries of R we define a pair of 2 vectors $\vec{u}$ and $\vec{v}$ as $\vec{u} = (R_{1,1}, R_{1,2})$, $\vec{v} = (R_{2,3}, R_{3,3})$ so that we can rewrite the above inequality Eq. (17) as

$$\|\vec{v}\|_1 - \|\vec{u}\|_1 \geq 0. \tag{19}$$

The L^2 normalization of all the rows and columns of the rotation matrix R means that $\vec{u}$ and $\vec{v}$ have the same L^2 norm. With reference to Fig. 2, it should be clear that because $\vec{u}$ has an L^∞ norm at least as big as that of $\vec{v}$ (because $|R_{1,2}| \geq |R_{2,3}|, |R_{3,3}|$), it holds that the L^1 norm of $\vec{v}$ is automatically at least as large as the L^1 norm of $\vec{u}$, as desired.

Summary.—We showed that for any unitary one-qubit gate undergoing depolarizing noise with rate p it holds that if its SO(3) representation $M = (1 - p)R$ lies outside the Clifford polytope $\mathcal{P}$, then it must be the case that there is a facet $B \in \mathcal{F}_B$ such that $M \cdot B > 1$. In turn, this means that if this noisy gate is applied to a Bell pair $|\Phi\rangle = \frac{1}{\sqrt{2}}(|00\rangle + |11\rangle)$ and an appropriate stabilizer measurement is performed, then, conditionally on the outcome of the measurement, one obtains a state that can be transformed using Clifford gates into a single-qubit state with $|y| + |z| > 1$ in the Bloch ball representation. By the result of Reichardt [11] such states enable stabilizing operations to perform universal quantum computation.

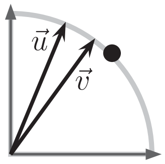

FIG. 2. Proof of Eq. (19): For any pair of two vectors $\vec{u}$ and $\vec{v}$ with the same L^2 norm, the vector with greater L^∞ norm has smaller L^1 norm. A vector pointing towards the black dot has simultaneously minimal L^∞ norm and maximal L^1 norm.

This material is based upon work supported by the National Science Foundation under Grant No. 0917244.

*vandam@cs.ucsb.edu
†mhoward@physics.ucsb.edu

[1] Michasel A. Nielsen and Isaac L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, England, 2000).
[2] H. Buhrman, R. Cleve, M. Laurent, N. Linden, A. Schrijver, and F. Unger, *Annual IEEE Symposium on Foundations of Computer Science* (2006), p 411.
[3] E. T. Campbell and D. E. Browne, arXiv:0908.0836v2.
[4] A. R. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).
[5] A. Steane, Proc. R. Soc. A **452**, 2551 (1996).
[6] D. Gottesman, Phys. Rev. A **57**, 127 (1998).
[7] E. H. Knill, arXiv:quant-ph/0402171v1.
[8] M. Freedman, C. Nayak, and K. Walker, Phys. Rev. B **73**, 245307 (2006).
[9] L. S. Georgiev, Phys. Rev. B **74**, 235112 (2006).
[10] S. Bravyi and A. Kitaev, Phys. Rev. A **71**, 022316 (2005).
[11] Ben W. Reichardt, Quant. Info. Proc. **4**, 251 (2005).
[12] S. Virmani, S. F. Huelga, and M. B. Plenio, Phys. Rev. A **71**, 042328 (2005).
[13] M. B. Plenio and S. Virmani, arXiv:0810.4340.
[14] B. W. Reichardt, arXiv:quant-ph/0608085v1.

A One-Way Quantum Computer

Robert Raussendorf and Hans J. Briegel
Theoretische Physik, Ludwig-Maximilians-Universität München, Germany

(Received 25 October 2000)

We present a scheme of quantum computation that consists entirely of one-qubit measurements on a particular class of entangled states, the cluster states. The measurements are used to imprint a quantum logic circuit on the state, thereby destroying its entanglement at the same time. Cluster states are thus one-way quantum computers and the measurements form the program.

DOI: 10.1103/PhysRevLett.86.5188 PACS numbers: 03.67.Lx, 03.65.Ud

A quantum computer promises efficient processing of certain computational tasks that are intractable with classical computer technology [1]. While basic principles of a quantum computer have been demonstrated in the laboratory [2], scalability of these systems to a large number of qubits [3], essential for practical applications such as the Shor algorithm, represents a formidable challenge. Most of the current experiments are designed to implement sequences of highly controlled interactions between selected particles (qubits), thereby following models of a quantum computer as a (sequential) network of quantum logic gates [4,5].

Here we propose a different model of a scalable quantum computer. In our model, the entire resource for the quantum computation is provided initially in the form of a specific entangled state (a so-called cluster state [6]) of a large number of qubits. Information is then written onto the cluster, processed, and read out from the cluster by one-particle measurements only. The entangled state of the cluster thereby serves as a universal "substrate" for any quantum computation. Cluster states can be created efficiently in any system with a quantum Ising-type interaction (at very low temperatures) between two-state particles in a lattice configuration.

We consider two- and three-dimensional arrays of qubits that interact via an Ising-type next-neighbor interaction [6] described by a Hamiltonian $H_{\mathrm{int}} = g(t) \times \sum_{\langle a,a'\rangle} \frac{1+\sigma_z^{(a)}}{2}\frac{1-\sigma_z^{(a')}}{2} \cong -\frac{1}{4}g(t)\sum_{\langle a,a'\rangle}\sigma_z^{(a)}\sigma_z^{(a')}$ [7] whose strength $g(t)$ can be controlled externally. A possible realization of such a system is discussed below. A qubit at site a can be in two states $|0\rangle_a \equiv |0\rangle_{z,a}$ or $|1\rangle_a \equiv |1\rangle_{z,a}$, the eigenstates of the Pauli phase flip operator $\sigma_z^{(a)}$ $[\sigma_z^{(a)}|i\rangle_a = (-1)^i|i\rangle_a]$. These two states form the computational basis. Each qubit can equally be in an arbitrary superposition state $\alpha|0\rangle + \beta|1\rangle$, $|\alpha|^2 + |\beta|^2 = 1$. For our purpose, we initially prepare all qubits in the superposition $|+\rangle = (|0\rangle + |1\rangle)/\sqrt{2}$, an eigenstate of the Pauli spin flip operator σ_x $[\sigma_x|\pm\rangle = \pm|\pm\rangle]$. H_{int} is then switched on for an appropriately chosen finite time interval T, where $\int_0^T dt\, g(t) = \pi$, by which a unitary transformation S is realized. Since H_{int} acts uniformly on the lattice, entire clusters of neighboring particles become entangled in one single step. The quantum state $|\Phi\rangle_C$, the state of a cluster (C) of neighboring qubits, which is thereby created provides in advance all entanglement that is involved in the subsequent quantum computation. It has been shown [6] that the cluster state $|\Phi\rangle_C$ is characterized by a set of eigenvalue equations

$$\sigma_x^{(a)} \bigotimes_{a'\in ngbh(a)} \sigma_z^{(a')}|\Phi\rangle_C = \pm|\Phi\rangle_C\,, \tag{1}$$

where $ngbh(a)$ specifies the sites of all qubits that interact with the qubit at site $a \in C$. The eigenvalues are determined by the distribution of the qubits on the lattice. The equations (1) are central for the proposed computation scheme. As an example, a measurement on an individual qubit of a cluster has a random outcome. On the other hand, Eqs. (1) imply that any two qubits at sites $a, a' \in C$ can be projected into a Bell state by measuring a subset of the other qubits in the cluster. This property will be used to define quantum channels that allow us to propagate quantum information through a cluster.

We show that a cluster state $|\Phi\rangle_C$ can be used as a substrate on which any quantum circuit can be imprinted by one-qubit measurements. In Fig. 1 this scheme is illustrated. For simplicity, we assume that in a certain region of the lattice each site is occupied by a qubit. This requirement is not essential as will be explained below [see (d)]. In the first step of the computation, a subset of qubits is measured in the basis of σ_z which effectively removes them. In Fig. 1 these qubits are denoted by " ⊙."

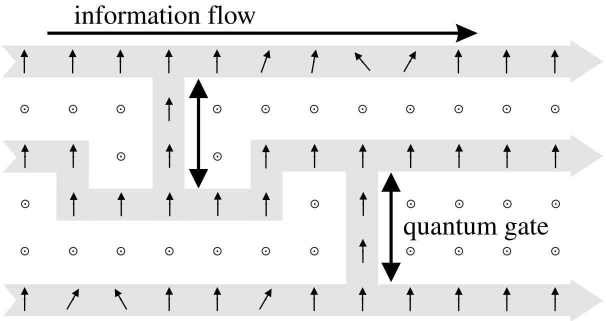

FIG. 1. Quantum computation by measuring two-state particles on a lattice. Before the measurements the qubits are in the cluster state $|\Phi\rangle_C$ of (1). Circles ⊙ symbolize measurements of σ_z, vertical arrows are measurements of σ_x, while tilted arrows refer to measurements in the x-y plane.

0031-9007/01/86(22)/5188(4)$15.00

The state $|\Phi\rangle_C$ is thereby projected into a tensor product $|\mu\rangle_{C\backslash\mathcal{N}} \otimes |\tilde{\Phi}\rangle_{\mathcal{N}}$ consisting of the state $|\mu\rangle_{C\backslash\mathcal{N}}$ of all measured particles (subset $C\backslash\mathcal{N}$) on one side and an entangled state $|\tilde{\Phi}\rangle_{\mathcal{N}}$ of yet unmeasured particles (subset $\mathcal{N} \subset C$), on the other side. These unmeasured particles define a "network" $\mathcal{N}$ corresponding to the shaded structure in Fig. 1. The state $|\tilde{\Phi}\rangle_{\mathcal{N}}$ of the network is related to a cluster state $|\Phi\rangle_{\mathcal{N}}$ on $\mathcal{N}$ by a local unitary transformation which depends on the set of measurement results μ. More specifically, $|\tilde{\Phi}\rangle_{\mathcal{N}}$ satisfies Eqs. (1)—with C replaced by the subcluster $\mathcal{N}$—except for a possible difference in the sign factors, which are determined by the measurement results μ.

To process quantum information with this network, it suffices to measure its particles in a certain order and in a certain basis. Quantum information is thereby propagated horizontally through the cluster by measuring the qubits on the wire while qubits on vertical connections are used to realize two-bit quantum gates. The basis in which a certain qubit is measured depends in general on the results of preceding measurements. The processing is finished once all qubits except the last one on each wire have been measured. At this point, the results of previous measurements determine in which basis these "output" qubits need to be measured for the final readout. We note that, in the entire process, only one-qubit measurements are required. The amount of entanglement therefore decreases with every measurement [8] and all entanglement involved in the process is provided by the initial resource, the cluster state. This is different from the scheme of Ref. [11], which uses Bell measurements (capable of producing entanglement) to realize quantum gates.

In the following, we show that any quantum logic circuit can be implemented on a cluster state. The purpose of this is twofold. First, it serves as an illustration of how to implement a particular quantum circuit in practice. Second, in showing that any quantum circuit can be implemented on a sufficiently large cluster we demonstrate the universality of the proposed scheme. For pedagogical reasons we first explain a scheme with one essential modification with respect to the proposed scheme: before the entanglement operation S, certain qubits are selected as input qubits and the input information is written onto them, while the remaining qubits are prepared in $|+\rangle$. This step weakens the scheme since it affects the character of the cluster state as a genuine resource. It can, however, be avoided [see (e)]. Points (a) to (c) are concerned with the basic elements of a quantum circuit, quantum gates, and wires, point (d) with the composition of gates to circuits.

(a) Information propagation in a wire for qubits. A qubit can be teleported from one site of a cluster to any other site. In particular, consider a chain of an odd number of qubits 1 to n prepared in the state $|\psi_{\text{in}}\rangle_1 \otimes |+\rangle_2 \otimes \cdots \otimes |+\rangle_n$ and subsequently entangled by S. The state that was originally encoded in qubit 1, $|\psi_{\text{in}}\rangle$, is now delocalized and can be transferred to site n by performing σ_x measurements (basis $\{|+\rangle_j = |0\rangle_{x,j}, |-\rangle_j = |1\rangle_{x,j}\}$) at qubit sites $j = 1, \ldots, n-1$ with measurement outcomes $s_j \in \{0, 1\}$. The resulting state is $|s_1\rangle_{x,1} \otimes \cdots \otimes |s_{n-1}\rangle_{x,n-1} \otimes |\psi_{\text{out}}\rangle_n$. The output state $|\psi_{\text{out}}\rangle$ is related to the input state $|\psi_{\text{in}}\rangle$ by a unitary transformation $U_\Sigma \in \{1, \sigma_x, \sigma_z, \sigma_x\sigma_z\}$ which depends on the outcomes of the σ_x measurements at sites 1 to $n-1$. A similar argument can be given for an even number of qubits. The effect of U_Σ can be accounted for at the end of a computation as shown below [see (d)]. It is noteworthy that not all classical information gained by the σ_x measurements needs to be stored to identify the transformation U_Σ. Instead, U_Σ is determined by the values of only two classical bits which are updated with every measurement.

(b) An arbitrary rotation $U_R \in \mathrm{SU}(2)$ can be achieved in a chain of five qubits. Consider a rotation in its Euler representation $U_R(\xi, \eta, \zeta) = U_x(\zeta)U_z(\eta)U_x(\xi)$, where $U_x(\alpha) = \exp(-i\alpha\frac{\sigma_x}{2}), U_z(\alpha) = \exp(-i\alpha\frac{\sigma_z}{2})$. Initially, the first qubit is in some state $|\psi_{\text{in}}\rangle$, which is to be rotated, and the other qubits are in $|+\rangle$; i.e., their common state reads $|\Psi\rangle_{1,\ldots,5} = |\psi_{\text{in}}\rangle_1 \otimes |+\rangle_2 \otimes |+\rangle_3 \otimes |+\rangle_4 \otimes |+\rangle_5$. After the five qubits are entangled by S they are in the state $S|\Psi\rangle_{1,\ldots,5} = 1/2|\psi_{\text{in}}\rangle_1|0\rangle_2|-\rangle_3|0\rangle_4|-\rangle_5 - 1/2|\psi_{\text{in}}\rangle_1|0\rangle_2|+\rangle_3|1\rangle_4|+\rangle_5 - 1/2|\psi_{\text{in}}^*\rangle_1|1\rangle_2|+\rangle_3|0\rangle_4|-\rangle_5 + 1/2|\psi_{\text{in}}^*\rangle_1|1\rangle_2|-\rangle_3|1\rangle_4|+\rangle_5$, where $|\psi_{\text{in}}^*\rangle = \sigma_z|\psi_{\text{in}}\rangle$. Now, the state $|\psi_{\text{in}}\rangle$ can be rotated by measuring qubits 1 to 4, while it is teleported to site 5 at the same time. The qubits $1, \ldots, 4$ are measured in appropriately chosen bases $\mathcal{B}_j(\alpha_j) = \{\frac{|0\rangle_j + e^{i\alpha_j}|1\rangle_j}{\sqrt{2}}, \frac{|0\rangle_j - e^{i\alpha_j}|1\rangle_j}{\sqrt{2}}\}$ whereby the measurement outcomes $s_j \in \{0, 1\}$ for $j = 1, \ldots, 4$ are obtained. Here, $s_j = 0$ means that qubit j is projected into the first state of $\mathcal{B}_j(\alpha_j)$. The resulting state is $|s_1\rangle_{\alpha_1,1} \otimes |s_2\rangle_{\alpha_2,2} \otimes |s_3\rangle_{\alpha_3,3} \otimes |s_4\rangle_{\alpha_4,4} \otimes |\psi_{\text{out}}\rangle_5$ with $|\psi_{\text{out}}\rangle = U|\psi_{\text{in}}\rangle$. For the choice $\alpha_1 = 0$ (measuring σ_x of qubit 1) the rotation U has the form $U = \sigma_x^{s_2+s_4}\sigma_z^{s_1+s_3}U_R[(-1)^{s_1+1}\alpha_2, (-1)^{s_2}\alpha_3, (-1)^{s_1+s_3}\alpha_4]$. In summary, the procedure to implement an arbitrary rotation $U_R(\xi, \eta, \zeta)$, specified by its Euler angles ξ, η, ζ is (i) measure qubit 1 in $\mathcal{B}_1(0)$; (ii) measure qubit 2 in $\mathcal{B}_2((-1)^{s_1+1}\xi)$; (iii) measure qubit 3 in $\mathcal{B}_3((-1)^{s_2}\eta)$; (iv) measure qubit 4 in $\mathcal{B}_4((-1)^{s_1+s_3}\zeta)$. In this way the rotation U_R' is realized: $U_R'(\xi, \eta, \zeta) = \sigma_x^{s_2+s_4}\sigma_z^{s_1+s_3}U_R(\xi, \eta, \zeta)$. The extra rotation $U_\Sigma = \sigma_x^{s_2+s_4}\sigma_z^{s_1+s_3}$ can be accounted for at the end of the computation, as is described below in (d).

(c) To perform the gate $\text{CNOT}(c, t_{\text{in}} \rightarrow t_{\text{out}}) = |0\rangle_{cc}\langle 0| \otimes 1^{(t_{\text{in}}\rightarrow t_{\text{out}})} + |1\rangle_{cc}\langle 1| \otimes \sigma_x^{(t_{\text{in}}\rightarrow t_{\text{out}})}$ between a control qubit c and a target qubit t, four qubits, arranged as depicted Fig. 2a, are required. During the action of the gate, the target qubit t is transferred from t_{in} to t_{out}. The following procedure has to be implemented. Let qubit 4 be the control qubit. First, the state $|i_1\rangle_{z,1} \otimes |i_4\rangle_{z,4} \otimes |+\rangle_2 \otimes |+\rangle_3$ is prepared and then the entanglement operation S is performed. Second, σ_x of qubits 1 and 2 is measured. The measurement results

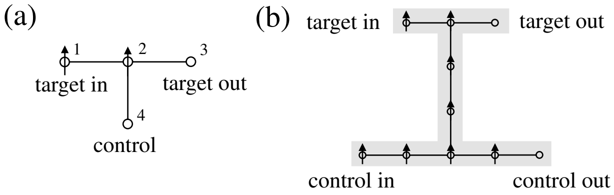

FIG. 2. Realization of a CNOT gate by one-particle measurements. See text.

$s_j \in \{0,1\}$ correspond to projections of the qubits j into $|s_j\rangle_{x,j}$, $j = 1, 2$. The quantum state created by this procedure is $|s_1\rangle_{x,1} \otimes |s_2\rangle_{x,2} \otimes U_\Sigma^{(34)}|i_4\rangle_{z,4} \otimes |i_1 + i_4 \,\mathrm{mod} 2\rangle_{z,3}$, where $U_\Sigma^{(34)} = \sigma_z^{(3)^{s_1+1}} \sigma_x^{(3)^{s_2}} \sigma_z^{(4)^{s_1}}$. The input state is thus acted upon by the CNOT and successive σ_x and σ_z rotations $U_\Sigma^{(34)}$, depending on the measurement results s_1, s_2. These unwanted extra rotations can again be accounted for as described in (d). For practical purposes it is more convenient if the control qubit is, as the target qubit, transferred to another site during the action of the gate. When a CNOT is combined with other gates to form a quantum circuit it will be used in the form shown in Fig. 2b.

To explain the working principle of the CNOT gate we, for simplicity, refer to the minimal implementation with four qubits. The minimal CNOT can be viewed as a wire from qubit 1 to qubit 3 with an additional qubit, No. 4, attached. From the eigenvalue equations (1) it can now be derived that, if qubit 4 is in an eigenstate $|i_4\rangle_{z,4}$ of σ_z, then the value of $i_4 \in \{0,1\}$ determines whether a unit wire or a spin flip σ_x (modulo the same correction $U_\Sigma^{(3)}$ for both values of i_4) is being implemented. In other words, once σ_x of qubits 1 and 2 have been measured, the value i_4 of qubit 4 controls whether the target qubit is flipped or not.

(d) Quantum circuits. The gates described—the CNOT and arbitrary one-qubit rotations—form a universal set [5]. In the implementation of a quantum circuit on a cluster state the site of every output qubit of a gate overlaps with the site of an input qubit of a subsequent gate. Because of this, the entire entanglement operation can be performed at the beginning. To see this, compare the following two strategies. Given a quantum circuit implemented on a network $\mathcal{N}$ of qubits which is divided into two consecutive circuits, circuit 1 is implemented on network $\mathcal{N}_1$ and circuit 2 is implemented on network $\mathcal{N}_2$, and $\mathcal{N} = \mathcal{N}_1 \cup \mathcal{N}_2$. There is an overlap $\mathcal{O} = \mathcal{N}_1 \cap \mathcal{N}_2$ which contains the sites of the output qubits of circuit 1 (these are identical to the sites of the input qubits of circuit 2). The sites of the readout qubits form a set $\mathcal{R} \subset \mathcal{N}_2$. Strategy (i) consists of the following steps: (1) write input and entangle all qubits on $\mathcal{N}$; (2) measure qubits $\in \mathcal{N} \backslash \mathcal{R}$ to implement the circuit. Strategy (ii) consists of (1) write input and entangle the qubits on $\mathcal{N}_1$, (2) measure the qubits in $\mathcal{N}_1 \backslash \mathcal{O}$. This implements the circuit on $\mathcal{N}_1$ and writes the intermediate output to $\mathcal{O}$; (3) entangle the qubits on $\mathcal{N}_2$; (4) measure all qubits in $\mathcal{N}_2 \backslash \mathcal{R}$. Steps 3 and 4 implement the circuit 2 on $\mathcal{N}_2$. The measurements on $\mathcal{N}_1 \backslash \mathcal{O}$ commute with the entanglement operation restricted to $\mathcal{N}_2$, since they act on different subsets of particles. Therefore the two strategies are mathematically equivalent and yield the same results. It is therefore consistent to entangle in a single step at the beginning and perform all measurements afterwards.

Two further points should be addressed in connection with circuits. First, the randomness of the measurement results does not jeopardize the function of the circuit. Depending on the measurement results, extra rotations σ_x and σ_z act on the output qubit of every implemented gate. By use of the relations $U_R(\xi, \eta, \zeta)\sigma_z^s \sigma_x^{s'} = \sigma_z^s \sigma_x^{s'} U_R\big((-1)^s \xi, (-1)^{s'} \eta, (-1)^s \zeta\big)$, and $\mathrm{CNOT}(c,t)\sigma_z^{(t)^{s_t}} \sigma_z^{(c)^{s_c}} \sigma_x^{(t)^{s'_t}} \sigma_x^{(c)^{s'_c}} = \sigma_z^{(t)^{s_t}} \sigma_z^{(c)^{s_c+s_t}} \sigma_x^{(t)^{s'_c+s'_t}} \sigma_x^{(c)^{s'_c}} \mathrm{CNOT}(c,t)$, these extra rotations can be pulled through the network to act upon the output state. There they can be accounted for by adjusting the measurement basis for the final readout. The above relations imply that for a rotation $U_R(\xi, \eta, \zeta)$—different from the CNOT gate—the accumulated extra rotations U_Σ at the input side of U_R need to be determined before the measurement bases that realize U_R can be specified. This introduces a partial temporal ordering of the measurements on the whole cluster. Second, quantum circuits can also be implemented on irregular clusters. In that case, qubits may be missing which are required for the standard implementation of the circuit. This can be compensated by a large flexibility in shape of the gates and wires. The components can be bent and stretched to fit to the cluster structure as long as the topology of the circuit implementation does not change. Irregular clusters are found in lattices with a finite site occupation probability $0 < p < 1$. In such a situation, the possibility of *universal* quantum computation is closely linked to the phenomenon of percolation. For p above a certain critical value p_c, which depends on the dimension of the lattice, an infinitely extended cluster exists that may be used as the carrier of the quantum circuit. In two dimensions, for example, exactly one such cluster C exists. Suppose this cluster is divided into two subclusters C_1 and C_2 by a one-dimensional cut $\mathcal{O} = C_1 \cap C_2$. It can be shown, e.g., by using Russo's formula [12] from percolation theory that, for any cut $\mathcal{O}$, $|\mathcal{O}| = \infty$. Therefore there is no upper bound, in principle, to the "capacity" of the cluster, i.e., to the number of qubits that can be processed across such a cut.

(e) Full scheme. It is important to note that the step of writing the input information onto the qubits before the cluster is entangled was introduced only for pedagogical reasons. For illustration of this point consider a chain of five qubits in the state $S|+\rangle_1 \otimes |+\rangle_2 \otimes \cdots \otimes |+\rangle_5$. Clearly, there is no local information on any of the qubits. However, by measuring qubits 1 to 4 along suitable directions, qubit 5 can be projected into any desired state (modulo U_Σ). What is used here is the knowledge that the resource has been prepared with qubit 1 in the state $|+\rangle_1$ before the entanglement operation. By the four measurements, this qubit is then rotated as described in (b). In order to use qubit 5 for further processing, the five-qubit chain considered here should, of course, be part of a larger cluster such that particle 5 is still entangled with the remaining network, after particles 1 to 4 have been measured. The method of preparing the input state remains the same, in this case, as explained in (d). In a similar manner any desired input state can be prepared if the rotations are replaced by a circuit preceding the proper circuit for computation. In summary, no input information needs to be written to the qubits before they are entangled. Cluster states are thus a genuine resource for quantum computation via measurements only.

For a cluster of a given *finite* size, the number of computational steps may be too large to fit on the cluster. In this case, the computation can be split into consecutive parts, for each of which there is sufficient space on the cluster. The modified procedure consists then of repeatedly (re)entangling the cluster and imprinting the actual part of the circuit—by measuring all of the lattice qubits except the ones carrying the intermediate quantum output—until the whole calculation is performed. This procedure has also the virtue that qubits involved in the later part of a calculation need not be protected from decoherence for a long time while the calculation is still being performed at a remote place of the cluster. Standard error-correction techniques [13,14] may then be used on each part of the circuit to stabilize the computation against decoherence.

A possible implementation of such a quantum computer uses neutral atoms stored in periodic micropotentials [15–18] where Ising-type interactions can be realized by controlled collisions between atoms in neighboring potential wells [16,18]. This system combines small decoherence rates with a high scalability. The question of scalability is linked to the percolation phenomenon, as mentioned earlier. For a site occupation probability above the percolation threshold, there exists a cluster which is bounded in size only by the trap dimensions. For optical lattices in three dimensions, single-atom site occupation with a filling factor of 0.44 has been reported [19] which is significantly above the percolation threshold of 0.31 [20]. As in other proposed implementations for quantum computing, the addressability of single qubits in the lattice is, however, still a problem. (For recent progress, see Ref. [21]). Recently, it has also been shown that implementations based on arrays of capacitively coupled quantum dots may be used to realize an Ising-type interaction [22].

In conclusion, we have described a new scheme of quantum computation that consists entirely of one-qubit measurements on a particular class of entangled states, the cluster states. The measurements are used to imprint a quantum circuit on the state, thereby destroying its entanglement at the same time. Cluster states are thus one-way quantum computers and the measurements form the program.

We thank D. E. Browne, D. P. DiVincenzo, A. Schenzle, and H. Wagner for helpful discussions. This work was supported by the Deutsche Forschungsgemeinschaft.

[1] C. H. Bennett and D. P. DiVincenzo, Nature (London) **404**, 247 (2000).
[2] See Ref. [1] for a recent review.
[3] J. I. Cirac and P. Zoller, Nature (London) **404**, 579 (2000).
[4] D. Deutsch, Proc. R. Soc. London **425**, 73 (1989).
[5] A. Barenco *et al.,* Phys. Rev. A **52**, 3457 (1995).
[6] H.-J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001).
[7] The second Hamiltonian is of the standard Ising form. The symbol "≅" means that the states generated from a given initial state, under the action of these Hamiltonians, are identical up to a local rotation on certain qubits. We use the first Hamiltonian to make the computational scheme more transparent. The conclusions drawn in the paper are, however, the same for both Hamiltonians.
[8] By the "amount of entanglement" contained in the resource, we mean any measure that satisfies the criteria of an entanglement monotone [9]. For cluster states, the entanglement can be calculated, e.g., in terms of the Schmidt measure of Ref. [10].
[9] G. Vidal, J. Mod. Opt. **47**, 355 (2000).
[10] J. Eisert and H.-J. Briegel, quant-ph/0007081.
[11] D. Gottesman and I. L. Chuang, Nature (London) **402**, 390 (1999).
[12] See, e.g., G. Grimmett, *Percolation* (Springer-Verlag, New York, 1989).
[13] A. Calderbank and P. W. Shor, Phys. Rev. A **54**, 1098 (1996).
[14] A. Steane, Phys. Rev. Lett. **77**, 793 (1996).
[15] G. K. Brennen *et al.,* Phys. Rev. Lett. **82**, 1060 (1999).
[16] D. Jaksch *et al.,* Phys. Rev. Lett. **82**, 1975 (1999).
[17] T. Calarco *et al.,* Phys. Rev. A **61**, 022304 (2000).
[18] H.-J. Briegel *et al.,* J. Mod. Opt. **47**, 415 (2000).
[19] M. T. DePue *et al.,* Phys. Rev. Lett. **82**, 2262 (1999).
[20] J. M. Ziman, *Models of Disorder* (Cambridge University Press, Cambridge, United Kingdom, 1979).
[21] R. Scheunemann *et al.,* Phys. Rev. A **62**, 051801(R) (2000).
[22] T. Tanamoto, quant-ph/0009030.

nature physics

LETTERS

PUBLISHED ONLINE: 20 JULY 2009 | DOI: 10.1038/NPHYS1342

Quantum computation and quantum-state engineering driven by dissipation

Frank Verstraete[1]*, Michael M. Wolf[2] and J. Ignacio Cirac[3]*

The strongest adversary in quantum information science is decoherence, which arises owing to the coupling of a system with its environment[1]. The induced dissipation tends to destroy and wash out the interesting quantum effects that give rise to the power of quantum computation[2], cryptography[2] and simulation[3]. Whereas such a statement is true for many forms of dissipation, we show here that dissipation can also have exactly the opposite effect: it can be a fully fledged resource for universal quantum computation without any coherent dynamics needed to complement it. The coupling to the environment drives the system to a steady state where the outcome of the computation is encoded. In a similar vein, we show that dissipation can be used to engineer a large variety of strongly correlated states in steady state, including all stabilizer codes, matrix product states[4], and their generalization to higher dimensions[5].

The situation we have in mind is shown in Fig. 1. A quantum system composed of N particles (such as qubits) is organized in space according to a particular geometry (in the figure, a one-dimensional lattice). Neighbouring systems are coupled to some local environments, which are dissipative in nature and tend to drive the system to a steady state. Our idea is to engineer those couplings, so that the environments drive the system to a desired final state. The coupling to the environment will be static, so that the desired state is obtained after some time without having to actively control the system. Note that the role of the environments is to dissipate (or, more precisely, evacuate) the entropy of the system, and by choosing the couplings appropriately we can use this effect to drive our system.

We will show first how to design the interactions with the environment to implement universal quantum computation. This new method, which we refer to as dissipative quantum computation (DQC), defies some of the standard criteria for quantum computation because it requires neither state preparation, nor unitary dynamics[6]. However, it is nevertheless as powerful as standard quantum computation. Then we will show that dissipation can be engineered[7] to prepare ground states of frustration-free Hamiltonians. Those include matrix product states[4,8,9] (MPSs) and projected entangled pair states[5,9] (PEPSs), such as graph states[10] and Kitaev[11] and Levin–Wen[12] topological codes. Both DQC and dissipative state engineering (DSE) are robust in the sense that, given the dissipative nature of the process, the system is driven towards its steady state independent of the initial state and hence of eventual perturbations along the way.

Here, we will concentrate first on DQC, showing how given any quantum circuit one can construct a locally acting master equation for which the steady state is unique, encodes the outcome of the circuit and is reached in polynomial time (with respect to the one corresponding to the circuit). Then we will show how to construct dissipative processes that drive the system to the ground state of any frustration-free Hamiltonian. In the Methods section, we will prove that MPS (ref. 9) and certain kinds of PEPS (ref. 9) can be efficiently prepared using this method, and in Supplementary Information we will give details of the proofs. In this letter we will not consider specific physical set-ups where our ideas can be implemented. Nevertheless, the Methods section will provide a universal way of engineering the master equations required for DQC and DSE, which can be easily adapted to current experiments[13] based on, for example, atoms in optical lattices[14] or trapped ions[15]. Thus, we expect that our predictions may be experimentally tested in the near future.

Let us start with DQC by considering N qubits in a line and a quantum circuit specified by a sequence of nearest-neighbour qubit operations $\{U_t\}_{t=1}^{\mathrm{T}}$. We define $|\psi_t\rangle := U_t U_{t-1}\ldots U_1|0\rangle_1\otimes\ldots|0\rangle_{\mathrm{N}}$, so that $|\psi_T\rangle$ is the final state after the computation. Our goal is to find a master equation $\dot\rho=\mathcal{L}(\rho)$ with a Liouvillian in Lindblad form[16]

$$\mathcal{L}(\rho)=\sum_k L_k\rho L_k^\dagger-\frac{1}{2}\left\{L_k^\dagger L_k,\rho\right\}_+ \qquad (1)$$

where the L_k acts locally and has a steady state, ρ_0: (1) that is unique; (2) that can be reached in a time poly(T); (3) such that ψ_T can be extracted from it in a time poly(T). As in Feynman's construction of a quantum simulator[3], we consider another auxiliary register with states $\{|t\rangle\}_{t=0}^{\mathrm{T}}$, which will represent the time. We choose the Lindblad operators

$$L_i=|0\rangle_i\langle 1|\otimes|0\rangle_t\langle 0|$$

$$L_t=U_t\otimes|t+1\rangle\langle t|+U_t^\dagger\otimes|t\rangle\langle t+1|$$

where $i=1,\ldots,N$ and $t=0,\ldots,T$. It is clear that the L terms act locally except for the interaction with the extra register, which can be made local as well. Furthermore,

$$\rho_0=\frac{1}{T+1}\sum_t|\psi_t\rangle\langle\psi_t|\otimes|t\rangle\langle t|$$

is a steady state, that is, $\mathcal{L}(\rho_0)=0$. Given such a state, the result of the actual quantum computation can be read out with probability $1/T$ by measuring the time register. In Supplementary Information, we show that ρ_0 is the unique steady state and that the Liouvillian has a spectral gap $\Delta=\pi^2/(2T+3)^2$. This means indeed that the steady state will be reached in polynomial time in T. Note that this gap is independent of N as well as of the actual quantum computation that is carried out (that is, independent of the U_t). It is also shown that the same gap is retained if the clock register is encoded in the unary

[1]Fakultät für Physik, Universität Wien, 1090 Wien, Austria, [2]Niels Bohr Institute, 2100 Copenhagen, Denmark, [3]Max-Planck-Institut für Quantenoptik, 85748 Garching, Germany. *e-mail: fverstraete@gmail.com; ignacio.cirac@mpq.mpg.de.

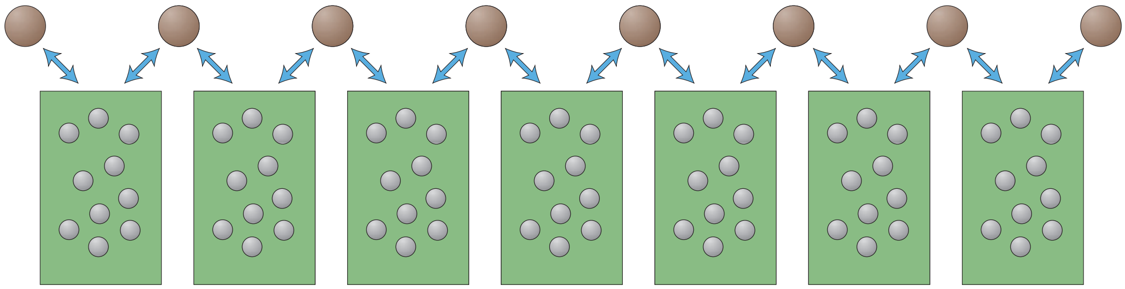

Figure 1 | Schematic representation of the set-up. We consider a collection of N quantum particles, locally coupled to a set of environments. The couplings are engineered in such a way that the system reaches the desired state in the long-time limit.

way proposed by Kitaev and co-workers[17], making the Lindblad operators strictly local. A sketch of the proof is as follows. First, we do a similarity transformation on $\mathcal{L}$ that replaces all gates U_i with the identity gates, showing that its spectrum is independent of the actual quantum computation. Second, another similarity transformation is done that makes $\mathcal{L}$ Hermitian and block-diagonal. Each block can then be diagonalized exactly leading to the claimed gap.

In some sense, the present formalism can be seen as a robust way of doing adiabatic quantum computation[18] (errors do not accumulate and the path does not have to be engineered carefully) and implementing quantum random walks[19], and it might therefore be easier to tackle interesting open questions, such as the quantum probabilistically-checkable-proofs theorem, in this setting[20]. In addition, it seems that the dissipative way of preparing ground states is more natural than to use adiabatic time evolution, as nature itself prepares them by cooling.

Let us now turn to DSE and consider again a quantum system with N particles on a lattice in any dimension. We are interested in ground states Ψ, of Hamiltonians

$$H = \sum_{\lambda} H_{\lambda}$$

that are frustration-free, meaning that Ψ minimizes the energy of each H_λ individually, and local in the sense that H_λ acts non-trivially only on a small set $\lambda \subset \{1,\ldots,N\}$ of sites (for example, nearest neighbours). We can assume the terms H_λ to be projectors and we will denote the orthogonal projectors by $P_\lambda = \mathbf{1} - H_\lambda$. States Ψ of the considered form are, for example, all PEPS (including MPS and stabilizer states[21]).

We will consider discrete time evolution generated by a trace-preserving completely positive map instead of a master equation. These two approaches are basically equivalent[22] as every local completely positive map $\mathcal{T}$ can be associated with a local Liouvillian through $\mathcal{L}(\rho) = N[\mathcal{T}(\rho) - \rho]$, which leads to the same fixed points and spectrum. We choose completely positive maps of the form

$$\mathcal{T}(\rho) = \sum_{\lambda} p_\lambda \left[P_\lambda \rho P_\lambda + \frac{1}{m} \sum_{i=1}^{m} U_{\lambda,i} H_\lambda \rho H_\lambda U_{\lambda,i}^{\dagger} \right] \qquad (2)$$

where the p_λ terms are probabilities and $U_{\lambda,1},\ldots,U_{\lambda,m}$ is a set of unitaries acting non-trivially only within region λ. They effectively rotate part of the high-energy space (with support of H_λ) to the zero-energy space, so that $\mathrm{tr}[\mathcal{T}(\rho)\Psi] \geq \mathrm{tr}[\rho\Psi]$ increases. As for Liouvillians (1), we could similarly take $L_{\lambda,i} = U_i H_\lambda$, or the ones associated with the completely positive map.

We show now that for every frustration-free Hamiltonian, the completely positive map in equation (2) converges to the ground-state space if we choose the unitaries $U_{\lambda,i}$ to be completely depolarizing, that is, $\mathcal{T}(\rho) \propto \sum_\lambda P_\lambda \rho P_\lambda + \mathbf{1}_\lambda \otimes \mathrm{tr}_\lambda[H_\lambda \rho]/\mathrm{tr}[\mathbf{1}_\lambda]$. For ease of notation, we will explain the proof for the case of a one-dimensional ring with nearest-neighbour interactions labelled by the first site $\lambda = 1,\ldots,N$. Assume ρ is such that its expectation value with respect to the projector Ψ onto the ground-state space of H is non-increasing under applications of $\mathcal{T}$, that is, in particular $\mathrm{tr}[\rho\Psi] = \mathrm{tr}[\mathcal{T}^N(\rho)\Psi]$. Expressing this in the Heisenberg picture in which $\mathcal{T}^*(\Psi) = \Psi + \sum_\lambda H_\lambda \mathrm{tr}_\lambda(\Psi)/(d^2N)$, we get

$$\begin{aligned}\mathrm{tr}[\rho\Psi] &\geq \mathrm{tr}[\rho\Psi] + \frac{1}{(d^2N)^N}\mathrm{tr}\left[\rho \sum_{\mu=1}^{N}\prod_{\lambda=1}^{N}(H_{\lambda+\mu}\mathrm{tr}_{\lambda+\mu})(\Psi)\right] \\ &\geq \mathrm{tr}[\rho\Psi] + \frac{\nu^N}{(d^2N)^N}\mathrm{tr}[\rho H]\end{aligned}$$

where the first inequality comes from discarding (positive) terms in the sum and the second one is due to bounding all partial traces of H_λ from below by the respective smallest eigenvalue ν. Note that the latter is strictly positive unless H has a product state as the ground state (in which case the statement becomes trivial). Hence, we must have $\mathrm{tr}[\rho H] = 0$; that is, ρ is a ground state of H. It is easily seen that the same argument applies for more general interactions on arbitrary lattices.

Once we have shown that the steady state after the application of the completely positive map lies within the desired subspace (the ground-state space of the frustration-free Hamiltonian), the next question to be addressed is how efficient the process is. This depends on the spectral gap, δ, of the completely positive map (or, equivalently, of the corresponding Liouvillian), as the time to reach the steady state, $\tau = \mathcal{O}(1/\delta)$. Thus, the above procedure will be efficient as long as the gap vanishes only polynomially with the number of systems, N. Similarly to what occurs with many-body Hamiltonians, the determination of such a gap is, in general, very complicated. For a wide range of interesting models, however, it can be proved that this gap scales favourably. This is the case for all MPS as well as for a rich subfamily of PEPS that includes all stabilizer states (such as Kitaev's toric code[11] and the Levin–Wen states[12]). In the Methods section, we characterize such a subfamily of states, and in Supplementary Information we give the technical proofs of our statements. Here, we will qualitatively explain how our method works efficiently for some families of states. For that we note that the action of the completely positive map (2) can be interpreted as randomly choosing a region λ (according to p_λ, which we may set equal to $1/N$), then measuring P_λ and applying a correction according to the unitaries if the outcome was negative. We denote by R_n the set of regions λ where φ satisfies the condition $H_\lambda|\varphi\rangle = 0$. If we measure now in one of those regions, we will obviously obtain a positive result, and thus R_n will remain the same. If we measure in another region, we may have a positive or negative result, something that may change the set R_n. By imposing certain conditions on the operators H_λ and $U_{\lambda,i}$, we can make sure that in each step R_n cannot be reduced and that the probability of being enlarged is non-vanishing. This automatically ensures that the τ scales only polynomially with the number of systems. In one dimension, however, one can get rid of all those restrictions and show that any MPS can be prepared in a time that also scales favourably with N. The fact that all MPS states can be prepared with our method, together with the results reported in refs 23, 24, automatically implies the existence of phase transitions driven by dissipation in the following sense. By changing the parameters of the operators H_λ appearing in the completely positive map (2), we change the steady state of that map. It is possible to choose models for which that state changes abruptly at some particular value of that parameter in such a way that the correlation length diverges and an order parameter appears (an example can be found in the Supplementary Information).

We have investigated the computational power of purely dissipative processes, and proved that it is equivalent to that of the quantum circuit model of quantum computation. We have also shown that dissipative dynamics can be used to create ground states (such as MPS or PEPS) of frustration-free Hamiltonians of strongly correlated quantum spin systems. We believe that these new methods can be experimentally tested using atoms or ions with current set-ups (see the Methods section).

Let us stress that we have been concerned here with a proof-of-principle demonstration that dissipation provides us with an alternative way of carrying out quantum computations or state engineering. We believe, however, that much more efficient and practical schemes can be developed and adapted to specific implementations. We also think that these results open up some interesting questions that deserve further investigation: for example, how the use of fault-tolerant computations can make our scheme more robust, or how one can design translationally invariant completely positive maps that prepare MPS more efficiently, or the importance and generality of the set of commuting Hamiltonians (see the Methods section), which is intimately connected to the fixed points of the renormalization group transformations on PEPS (as it happens with MPS; ref. 25). Furthermore, the model of DQC might well lead to the construction of new quantum algorithms, as, for example, quantum random walks can more easily be formulated within this context. Finally, other ideas related to this work can be easily addressed using the methods introduced; for example, thermal states of commuting Hamiltonians can be engineered using DSE because the Metropolis way of sampling over classical spin configurations can be adopted to the case of commuting operators. Similar techniques could be applied to free fermionic and bosonic systems, and, more generally, it should be possible to devise DSE schemes converging to the ground or thermal states of frustrated Hamiltonians by combining unitary and dissipative dynamics.

Note added. Concurrently with the submission of this paper, refs 26 and 27 appeared in which a similar quantum-reservoir engineering was used to prepare many-body states and non-equilibrium quantum phases.

Methods

Engineering dissipation. Here we show how to engineer the local dissipation that gives rise to the master equations (1) and completely positive maps (2). They are composed of local terms, involving few particles (typically two), so that we just have to show how to implement those. To simplify the exposition, we will treat those particles as a single one and assume that one has full control over its dynamics (for example, one can apply arbitrary gates).

Let us start with the completely positive maps. It is clear that by applying a quantum gate to the particle and a 'fresh' ancilla and then tracing the ancilla one can generate any physical action (that is, completely positive map) on the system. Furthermore, by repeating the same process with short time intervals one can subject the system to an arbitrary time-independent master equation. This last process may not be efficient. An alternative way works as follows. Let us assume that the ancilla is a qubit interacting with a reservoir such that it fulfils a master equation with Liouville operator $L_a = \sqrt{\Gamma}\sigma_-$, where $\sigma_- = |0\rangle\langle 1|$. Now, we couple the ancilla to the system with a Hamiltonian $H = \Omega(\sigma_- L^\dagger + \sigma_-^\dagger L)$. In the limit $\Gamma \gg \Omega$, one can adiabatically eliminate the level $|1\rangle$ of the ancilla[28] by applying second-order perturbation theory to the Liouvillian (albeit for non-Hermitian operators). In this way we obtain an effective master equation for ρ describing the system alone, with Liouville operator $\Omega/\sqrt{\Gamma}L$. By using several ancillas with Hamiltonians $H = \Omega(\sigma_- L_i + \sigma_-^\dagger L_i^\dagger)$ and following the same procedure we obtain the desired master equation. Although we have not specified here a physical system, one could use atoms. In that case, the ancilla could be an atom itself with $|0\rangle$ and $|1\rangle$ an electronic ground and excited level, respectively, so that spontaneous emission gives rise to the dissipation. The coupling to the system (other atoms) could be achieved using standard ideas used in the implementation of quantum computation using those systems[13].

Efficient state preparation. We have shown that it is possible to engineer dissipative processes that prepare ground states of frustration-free Hamiltonians in steady state. In the proof, the time for this preparation scales as N^N, which may be an issue for experiments with large number of particles. Here we give much more efficient methods for certain classes of frustration-free Hamiltonians.

We consider first frustration-free Hamiltonians for which $[H_\lambda, H_\mu] = 0$ and show that, under certain conditions, the corresponding ground states can be prepared in a time that scales only polynomially with the number of particles. The corresponding set of ground states contains important families, such as stabilizer states (for example, cluster states and topological codes), or certain kinds of PEPS, namely, those that have (commuting) parent Hamiltonians with the injectivity condition (as defined in refs 8, 29). Note that there was no known way of efficient preparation for the latter.

Loosely speaking, we will consider two classes of Hamiltonians. (1) Hamiltonians for which all excitations can be locally annihilated. In this case the time of convergence scales as $\tau = \mathcal{O}(\log N)$. (2) Interactions where excitations have to be moved along the lattice before they can annihilate and $\tau = \mathcal{O}(N \log N)$.

To see how the first case can occur notice that, when iterating $\mathcal{T}$, the correction on λ does not change the outcome of previous measurements on neighbouring regions because

$$\forall \lambda \neq \lambda' \colon [U_{\lambda,i}, H_{\lambda'}] = 0 \tag{3}$$

In fact, this can always be achieved by regrouping the regions into larger ones having an interior $I(\lambda) \subset \lambda$ on which only H_λ acts non-trivially and letting the $U_{\lambda,i}$ solely act on $I(\lambda)$. Denote by q the largest probability for obtaining twice a negative measurement outcome on the same region λ. The energy $\mathrm{tr}[H\mathcal{T}^M(\rho)]$ after M applications of $\mathcal{T}$ decreases then as $N(1-(1-q)/N)^M$ such that it takes $\mathcal{O}((N\log N)/(1-q))$ steps to converge to a ground state. The relaxation time of the corresponding Liouvillian is thus $\tau = \mathcal{O}(\log N^{1/1-q})$. Clearly, this is a reasonable bound only if $q < 1$, a condition possibly incompatible with equation (3).

Note that for all stabilizer states we can achieve $q = 0$, because there exists always a local unitary (acting on a single qubit) so that $H_\lambda U_\lambda H_\lambda = 0$. A class of stabilizer states where this is compatible with equation (3) are the so-called graph states[10]. In this case, λ labels (with some abuse of notation) a vertex of a graph and $H_\lambda = (\mathbf{1} - \sigma_x^{(\lambda)} \prod_{(\lambda,\mu)\in\mathcal{E}} \sigma_z^{(\mu)})/2$, where $\sigma^{(\lambda)}$ is a Pauli operator acting on site λ and $\mathcal{E}$ is the set of edges of the graph. Obviously, $U_\lambda = \sigma_z^{(\lambda)}$ does the job. In this special case, we can get even faster convergence when using the Liouvillian

$$\mathcal{L}(\rho) = \left(\sum_\lambda U_\lambda H_\lambda \rho H_\lambda U_\lambda^\dagger\right) - \frac{1}{2}\left\{H, \rho\right\}_+$$

The corresponding relaxation time can be determined exactly by realizing that the spectrum of $\mathcal{L}$ equals that of $-(H \otimes \mathbf{1} + \mathbf{1} \otimes H)/2$ so that $\tau = 1$ (see Supplementary Information).

For the second type of commuting Hamiltonians, equation (3) and $q < 1$ are incompatible. However, we can still prove fast convergence by relaxing equation (3) such that within each region λ the U_λ acts on a site closest to a predetermined site (say the origin) on the lattice and thus commutes with all terms H_λ that are further away (see Supplementary Information for details). In this way excitations are moved over the lattice before they can annihilate. As this requires extra time proportional to the system's size, we get $\tau = \mathcal{O}(N \log N)$.

We turn now to another family of ground states of frustration-free Hamiltonians, namely MPS (ref. 9). For the sake of clearness, we will consider here translationally invariant Hamiltonians, although the analysis can be straightforwardly extended to systems without that symmetry. We will specify a completely positive map to prepare states of the form

$$|\Psi\rangle = \sum_{i=1}^{d} \mathrm{tr}(A_{i_1} \ldots A_{i_N}) |i_1 \ldots i_N\rangle$$

where the A terms are $D \times D$ matrices. We assume the injectivity property[29], which implies that Ψ is the unique ground state of a nearest-neighbour frustration-free 'parent' Hamiltonian that has a gap. Denoting by ρ the reduced density operator corresponding to particles k and $k+1$, H_k and $P_k = 1 - H_k$ will denote the projectors onto its kernel and range, respectively. Note that $\mathrm{tr}(P_k) = D^2$. We take $N = 2^n$ for simplicity, but this is clearly not necessary. We construct the channel $\mathcal{T}$ in several steps. We first define a channel acting on two neighbouring particles $k, k+1$, as follows

$$\mathcal{R}_{r,c}(X) := P_k X P_k + \frac{P_k}{D^2}\mathrm{tr}(H_k X)$$

Here, $k = 2^{r-1}(2c-1)$, where $r = 1, \ldots, n$ and $c = 1, \ldots, 2^{n-r}$. The action of these maps has a tree structure, where the index r indicates the row in the tree, whereas c does it for the column. Now we define recursively,

$$\mathcal{S}_{r,c} := \frac{(1-\epsilon_r)}{2}(\mathcal{S}_{r-1,2c} + \mathcal{S}_{r-1,2c+1}) + \epsilon_r \mathcal{R}_{r,c}$$

Here, $r = 2, \ldots, n$, $c = 1, \ldots, 2^{n-r}$, $\mathcal{S}_{1,c} := \mathcal{R}_{1,c}$ and $\epsilon_{r+1} = 1/M^r$, where $M = CN^2$ and $C \gg 1$ (see Supplementary Information). Note that $\mathcal{S}_{r,1}$ acts on the first 2^r particles, $\mathcal{S}_{r,2}$ on the next 2^r and so on. We finally define

$$\mathcal{T} := (1-\epsilon_{n+1})\mathcal{S}_{n,1} + \epsilon_{n+1}\mathcal{R}_{n,2} \qquad (4)$$

In the Supplementary Information, we show that this map achieves the fixed point (up to an exponentially small error in C) in a time $\mathcal{O}(N^{\log_2(N)})$. The intuition behind the completely positive map (4) is that the channels $\mathcal{S}_{1,c}$, which are the ones that most often applied, project the state of every second nearest neighbour onto the right subspace. Then $\mathcal{S}_{2,c}$ do the same with half of the pairs that have not been projected. Then $\mathcal{S}_{3,c}$ does the same on half of the rest, and so on.

Received 11 March 2008; accepted 18 June 2009; published online 20 July 2009

References

1. Aliferis, P., Gottesman, D. & Preskill, J. Quantum accuracy threshold for concatenated distance-3 codes. *Quant. Inf. Comput.* **6,** 97–165 (2006).
2. Nielsen, M. A. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge Univ. Press, 2000).
3. Feynman, R. P. Simulating physics with computers. *Int. J. Theor. Phys.* **21,** 467–488 (1982).
4. Fannes, M., Nachtergaele, B. & Werner, R. F. Finitely correlated states on quantum spin chains. *Commun. Math. Phys.* **144,** 443–490 (1992).
5. Verstraete, F. & Cirac, J. I. Renormalization algorithms for quantum-many body systems in two and higher dimensions. Preprint at <http://arxiv.org/abs/cond-mat/0407066> (2004).
6. DiVincenzo, D. P. The physical implementation of quantum computation. *Fortschr. Phys.* **48,** 771–783 (2000).
7. Poyatos, J. F., Cirac, J. I. & Zoller, P. Quantum Reservoir Engineering with laser cooled trapped ions. *Phys. Rev. Lett.* **77,** 4728–4731 (1996).
8. Perez-Garcia, D., Verstraete, F., Wolf, M. M. & Cirac, J. I. Matrix product state representations. *Quant. Inf. Comput.* **7,** 401–430 (2007).
9. Verstraete, F., Murg, V. & Cirac, J. I. Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems. *Adv. Phys.* **57,** 143–224 (2008).
10. Briegel, H. J. & Raussendorf, R. Persistent entanglement in arrays of interacting qubits. *Phys. Rev. Lett.* **86,** 910–913 (2001).
11. Kitaev, A. Y. Fault-tolerant quantum computation by anyons. *Ann. Phys.* **303,** 2–30 (2003).
12. Levin, M. A. & Wen, X. G. String-net condensation: A physical mechanism for topological phases. *Phys. Rev. B* **71,** 045110 (2005).
13. Cirac, J. I. & Zoller, P. New frontiers in quantum information with atoms and ions. *Phys. Today* **57,** 38–44 (2004).
14. Bloch, I., Dalibard, J. & Zwerger, W. Many-body physics with ultracold gases. *Rev. Mod. Phys.* **80,** 885–964 (2008).
15. Leibfried, D., Blatt, R., Monroe, C. & Wineland, D. Quantum dynamics of single trapped ions. *Rev. Mod. Phys.* **75,** 281–324 (2003).
16. Lindblad, G. On the generators of quantum dynamical semigroups. *Commun. Math. Phys.* **48,** 119–130 (1976).
17. Kempe, J., Kitaev, A. Y. & Regev, O. The complexity of the local Hamiltonian problem. *SIAM J. Comput.* **35,** 1070–1097 (2004).
18. Aharonov, D. *et al.* Adiabatic quantum computation is equivalent to standard quantum computation. *SIAM J. Comput.* **37,** 166–194 (2007).
19. Kempe, J. Quantum random walks—an introductory overview. *Contemp. Phys.* **44,** 307–327 (2003).
20. Arora, S. & Safra, S. Probabilistic checking of proofs: A new characterization of NP. *J. ACM* **45,** 70–122 (1998).
21. Gottesman, D. A theory of fault-tolerant quantum computation. *Phys. Rev. A* **57,** 127–137 (1998).
22. Wolf, M. M. & Cirac, J. I. Dividing quantum channels. *Commun. Math. Phys.* **279,** 147–168 (2008).
23. Wolf, M. M., Ortiz, G., Verstraete, F. & Cirac, J. I. Quantum phase transitions in matrix product systems. *Phys. Rev. Lett.* **97,** 110403 (2006).
24. Verstraete, F., Wolf, M. M., Perez-Garcia, D. & Cirac, J. I. Criticality, the area law, and the computational power of PEPS. *Phys. Rev. Lett.* **96,** 220601 (2006).
25. Verstraete, F., Cirac, J. I., Latorre, J. I., Rico, E. & Wolf, M. M. Renormalization-group transformations on quantum states. *Phys. Rev. Lett.* **94,** 140601 (2005).
26. Diehl, S. *et al.* Quantum states and phases in driven open quantum systems with cold atoms. *Nature Phys.* **4,** 878–883 (2008).
27. Kraus, B. *et al.* Preparation of entangled states by quantum Markov processes. *Phys. Rev. A* **78,** 042307 (2008).
28. Cohen-Tannoudji, C., Dupont-Roc, J. & Grynberg, G. *Atom-Photon Interactions* (Wiley, 1992).
29. Perez-Garcia, D., Verstraete, F., Cirac, J. I. & Wolf, M. M. PEPS as unique ground states of local Hamiltonians. *Quant. Inf. Comput.* **8,** 0650–0663 (2008).

Acknowledgements

We thank D. Perez-Garcia for discussions and acknowledge financial support by the EU projects QUEVADIS, SCALA, the FWF, QUANTOP, FNU, SFB FoQuS, the DFG, Forschungsgruppe 635, the Munich Center for Advanced Photonics (MAP) and Caixa Manresa.

Author contributions

All authors have contributed equally to this paper.

Additional information

Supplementary information accompanies this paper on www.nature.com/naturephysics. Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions. Correspondence and requests for materials should be addressed to F.V. or J.I.C.