Quantum macroscopicity versus distillation of macroscopic superpositions

Benjamin Yadin[1] and Vlatko Vedral[1,2]

[1]*Atomic and Laser Physics, Clarendon Laboratory, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*
[2]*Centre for Quantum Technologies, National University of Singapore, Singapore 117543*
(Received 10 July 2014; published 28 August 2015)

We suggest a way to quantify a type of macroscopic entanglement via distillation of Greenberger-Horne-Zeilinger states by local operations and classical communication. We analyze how this relates to an existing measure of quantum macroscopicity based on the quantum Fisher information in several examples. Both cluster states and Kitaev surface code states are found to not be macroscopically quantum but can be distilled into macroscopic superpositions. We look at these distillation protocols in more detail and ask whether they are robust to perturbations. One key result is that one-dimensional cluster states are not distilled robustly but higher-dimensional cluster states are.

I. INTRODUCTION

Despite the overwhelming successes of quantum mechanics, one of its greatest remaining problems is to explain why it appears to break down at the macroscopic scale. In particular, macroscopic objects are never seen in quantum superpositions. The well-known thought experiment of Schrödinger's cat highlights the absurdity of a cat existing in a superposition of alive and dead states, yet in principle this is possible within quantum theory. It is therefore important to attempt to create macroscopic quantum states in experiments, in order to probe the boundary between quantum and classical mechanics—to decide if a fundamental size limit exists or if the challenge is purely a matter of isolating a system from its noisy environment.

Some recent experiments have sought cat states in photonic systems or similar macroscopic superpositions in superconducting circuits, molecular interferometers, and mechanical resonators [1–5]. Due to this great variety, one needs a general measure of quantum macroscopicity to compare experiments in which qualitatively different states are produced. Such a measure may also help us better understand the transition to macroscopic classical behavior.

There is no single measure generally agreed to quantify macroscopicity; typically, proposed measures are motivated along the lines of the working definition 1 given by Fröwis and Dür [6]: a quantum state is macroscopic if it is able to display nonclassical behavior at a large scale that is not simply an accumulation of microscopic quantum effects. The need to rule out accumulated phenomena was originally appreciated by Leggett [7,8]. These include, for example, bulk properties of condensed matter systems that are explained only by quantum physics, yet which are built up from effects extending over the atomic scale. In other words, one expects that a macroscopic quantum state necessarily has many-body or long-range quantum correlations.

An appropriate measure should then describe the largest scale to which quantum effects extend in a given state—this is often referred to as an effective size, denoted here by N^*. We do not impose any cutoff above which a value of N^* counts as macroscopic, but instead consider families of states parametrized by some obvious size quantity N (e.g., the number of qubits). Then the relevant property of the family is the scaling of N^* with N. The case $N^* = O(N)$ is maximally macroscopic [9].

In this work, we explore the consequences of viewing macroscopicity as a statement about quantum correlations. Thus we propose an effective size based on distilling macroscopic superpositions, as a way of quantifying a kind of macroscopic entanglement. Statements about entanglement are easiest for finite-dimensional systems, so our work is currently restricted to these (we focus on systems of qubits here), although characterizations of macroscopicity exist for continuous-variable systems [10–12]. We compare this quantity against an existing widely studied measure of macroscopicity based on the quantum Fisher information.

In Sec. II, we first introduce the quantum Fisher information measure of macroscopicity; we then propose a measure of pure-state macroscopic entanglement via distillation of Greenberger-Horne-Zeilinger (GHZ) states [13] in Sec. III and investigate how these relate in specific examples. We find that cluster states and Kitaev surface code ground states have macroscopic entanglement, but this is not detected by the Fisher information measure. In Sect. IV, we ask whether these distillation protocols are sensitive to imperfections, finding answers via mappings onto statistical spin models. We conclude in Sec. V.

II. FISHER INFORMATION MEASURE OF MACROSCOPICITY

Given a state, ρ, and an observable, A, the quantum Fisher information can be defined by

$$\mathcal{F}(\rho, A) = 2 \sum_{a,b} \frac{(p_a - p_b)^2}{p_a + p_b} |\langle \psi_a | A | \psi_b \rangle|^2, \qquad (1)$$

where p_a and $|\psi_a\rangle$ are the eigenvalues and eigenstates of ρ. We consider the class $\mathcal{A}$ of observables which can be written as $A = \sum_{i=1}^{N} A_i$ over local A_i, each acting nontrivially on a single qubit i and with fixed norm $\|A_i\| = 1$. The effective size proposed by Fröwis and Dür [6] is

$$N_{\mathcal{F}}^*(\rho) := \max_{A \in \mathcal{A}} \frac{\mathcal{F}(\rho, A)}{4N} \qquad (2)$$

and lies in the range $[1,N]$. ($\mathcal{A}$ may be extended to k-local A with A_i acting on groups of k qubits, with k bounded independent of N, in which case the denominator of equation (2) contains the number of groups instead of N.)

Observables in $\mathcal{A}$ are supposed to model the kinds of quantities that are easily measured at the macroscopic scale with coarse-grained, noisy classical detectors. For pure states, $\frac{1}{4}\mathcal{F}$ equals the variance, and $N_{\mathcal{F}}^*$ is seen to quantify the largest quantum fluctuations of any macroscopic observable—originally identified in Refs. [14,15]. It has been shown [6] that $N_{\mathcal{F}}^*$ is more inclusive than a variety of measures [16–18] looking at macroscopic superpositions of two states: maximal macroscopicity according to any of these measures implies $N_{\mathcal{F}}^* = O(N)$.

In general, $N_{\mathcal{F}}^*$ describes the usefulness of a state for quantum metrology. Consider the family of states $\rho_\theta = e^{-i\theta A}\rho e^{i\theta A}$ encoding the parameter $\theta \in \mathbb{R}$. From n independent copies of ρ, the quantum Cramér-Rao bound sets a lower limit on the uncertainty with which θ can be estimated: $\delta\theta \geqslant 1/\sqrt{n\mathcal{F}(\rho, A)}$ [19]. A macroscopic quantum state with $N_{\mathcal{F}}^* = O(N)$ makes $\delta\theta \propto 1/N$ possible, a qualitative improvement over the classical $\delta\theta \propto 1/\sqrt{N}$.

The authors have recently provided a motivation for this measure as a quantifier of macroscopic coherence in a precise sense [20]. (That work also notes some similarities between this measure and other approaches motivated from very different starting points [10,21].)

Furthermore, it has been demonstrated that large $N_{\mathcal{F}}^*$ is a witness of macroscopic entanglement in the following ways. For pure k-producible states in which blocks of up to k sites may be entangled, $\frac{1}{4}\mathcal{F}(|\psi\rangle\langle\psi|, A) \leqslant kN$ for 1-local A (similar bounds exist for mixtures) [22,23]. Also, $N_{\mathcal{F}}^* = O(N)$ for a pure state implies that macroscopically many [i.e., $O(N^2)$] pairs of sites have a nonvanishing $O(1)$ amount of localizable entanglement [24,25].

However, the converse is false: there are highly entangled states with small $N_{\mathcal{F}}^*$. To put this precisely, we introduce a measure aiming to quantify macroscopic entanglement.

III. MEASURE OF MACROSCOPIC ENTANGLEMENT

For a pure quantum state, one can quantify the amount of bipartite entanglement by counting the number of maximally entangled states that can be distilled by local operations and classical communication (LOCC) from many copies of the given state [26].

In the multipartite case there is no unique maximally entangled state [27]. To give a reasonable notion of macroscopic entanglement, we suggest using the GHZ states $|\mathrm{GHZ}_n\rangle :=$ $(|0\rangle^{\otimes n} + |1\rangle^{\otimes n})/\sqrt{2}$ as the target for distillation. $|\mathrm{GHZ}_n\rangle$ is often described as the typical qubit model of a macroscopic superposition and is the maximally macroscopic state of n qubits in the sense that $N_{\mathcal{F}}^* = n$. Furthermore, it is easy to motivate assigning an effective size of n to such a state.

To define our measure, take any pure state $|\psi\rangle$ of N qubits and consider acting on $|\psi\rangle$ with stochastic LOCC (SLOCC), described by measurement operators $\{M_a\}$ corresponding to the outcomes $\sqrt{p_a}|\phi_a\rangle := M_a|\psi\rangle$ with probabilities $p_a = \langle\psi|M_a^\dagger M_a|\psi\rangle$. We denote this kind of transformation

by $|\psi\rangle \to \{|\phi_a\rangle, p_a\}$. Now restrict these operations to the set D_ψ such that every outcome is of the form $|\phi_a\rangle = |\mathrm{GHZ}_{S_a}\rangle|0\rangle^{N-|S_a|}$, where $S_a \subseteq \{1, 2, \ldots, N\}$ is some subset of N qubits of cardinality $|S_a|$. We associate with each $|\phi_a\rangle$ a size $n_a = |S_a|$ unless a trivial GHZ state of size 1 is obtained, in which case $n_a = 0$. Our measure is

$$N_D^*(|\psi\rangle) := \max_{\{M_a\}\in D_\psi} \sum_a p_a n_a. \tag{3}$$

This is supposed to describe the size of GHZ-type entanglement present in the state $|\psi\rangle$. We have restricted each final state to a single GHZ, rather than a general product $\bigotimes_i |\mathrm{GHZ}_{n_i}\rangle$, since we are only interested in the largest GHZ; the remaining parts could be converted deterministically into product states. This prescription, instead of summing the sizes, rules out the accumulated phenomena mentioned earlier. Thus a state like $(|00\rangle + |11\rangle)^{\otimes n}$ has $N_D^* = O(1)$ instead of $O(N)$. In general, $N_D^*(\bigotimes_i |\psi_i\rangle) = \max_i N_D^*(|\psi_i\rangle)$.

It is simple to show that N_D^* is an entanglement monotone—it cannot increase on average under SLOCC. Suppose $|\psi\rangle \to \{|\chi_\mu\rangle, p_\mu\}$ by SLOCC. Then for each μ there exists an optimal ensemble, $\{|\phi_{\mu,a}\rangle, p_{\mu,a}\}$, distilled from $|\chi_\mu\rangle$ such that

$$N_D^*(|\chi_\mu\rangle) = \sum_a p_{\mu,a} n_{\mu,a}, \tag{4}$$

where $|\phi_{\mu,a}\rangle$ contains a GHZ state of size $n_{\mu,a}$. By composing the two SLOCC protocols, it follows that $|\psi\rangle \to \{|\phi_{\mu,a}\rangle, p_\mu p_{\mu,a}\}$ is a valid distillation operation in D_ψ, so

$$N_D^*(|\psi\rangle) \geqslant \sum_{\mu,a} p_\mu p_{\mu,a} n_{\mu,a}$$
$$= \sum_\mu p_\mu N_D^*(|\chi_\mu\rangle), \tag{5}$$

which proves the monotonicity.

The optimization involved in determining N_D^* will generally be intractable—the best we can do is find bounds. A lower bound must come from an explicit construction of distillation operations, and this will be difficult except for some particular cases. By extending a method in Ref. [28], we present a simple upper bound for states that are symmetric under exchange of any two sites (see Appendix A). This comes from the geometric entanglement [29] of a single site with the rest: $E_G(|\psi\rangle) = 1 - \lambda_{\max}(|\psi\rangle)$, where $\lambda_{\max}$ is the largest eigenvalue of $\rho_1 := \mathrm{Tr}_{2,3,\ldots,N}|\psi\rangle\langle\psi|$. Using the monotonicity of E_G under SLOCC, we find

$$N_D^*(|\psi\rangle) \leqslant 2N[1 - \lambda_{\max}(|\psi\rangle)]. \tag{6}$$

A. Generalized GHZ states

Generalized GHZ states [28] were our initial motivation for considering GHZ distillation. These depend on N and the parameter $\epsilon \in \mathbb{R}$ and were suggested to be a reasonable description of the macroscopic current superpositions in superconducting qubits. They can be written as $|\psi_N^\epsilon\rangle \propto |\epsilon\rangle^{\otimes n} + |-\epsilon\rangle^{\otimes n}$, where $|\pm\epsilon\rangle := \cos(\epsilon/2)|0\rangle \pm \sin(\epsilon/2)|1\rangle$. For $\epsilon = \frac{\pi}{2}$ we recover $|\mathrm{GHZ}_N\rangle$, while $\epsilon = 0$ gives $|0\rangle^{\otimes n}$. Thus we expect ϵ to vary the macroscopicity smoothly between the minimal and maximal values. Indeed, for $\epsilon \ll 1 \ll N\epsilon^2$, $N_{\mathcal{F}}^* \approx N\epsilon^2$ [6].
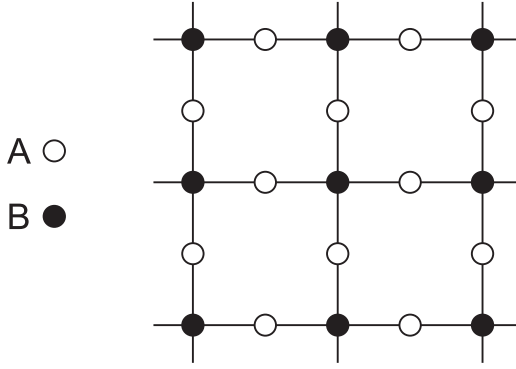
FIG. 1. The graph used to define the cluster state $|C_N\rangle$. Measurements are performed on the A sites in order to distill the B sites into a GHZ state.

The distillation protocol constructed in Ref. [28] produces an average distilled size of approximately $N\epsilon^2/2$ in the above limit. Moreover, $\lambda_{\max} \approx \cos^2(\epsilon/2)$, giving $N_D^* \leqslant N\epsilon^2/2$—so the protocol is exactly optimal [30], and $N_D^*(|\psi_N^\epsilon\rangle) \approx N\epsilon^2/2$. Here we have $N_D^* \approx N_{\mathcal{F}}^*/2$.

B. Cluster states

It has already been noted for cluster states that $N_{\mathcal{F}}^* = O(1)$, but it is nevertheless possible to deterministically distill from them GHZ states of size $O(N)$ [6,12]. Recall that a d-dimensional cluster state is defined for qubits associated with the vertices of (a subset of) a d-dimensional square lattice. It can be constructed from $(|0\rangle + |1\rangle)^{\otimes n}$ by applying a controlled-σ^z gate between each neighboring pair [31].

To see the scaling of $N_{\mathcal{F}}^*$, note that the variance quantifies the total amount of two-point correlations [24]. In cluster states, it can be shown that two regions are uncorrelated unless they share a boundary. Since each region of bounded $O(1)$ size has a bounded number of neighbors, it follows that $N_{\mathcal{F}}^* = O(1)$.

For simplicity, we focus on the family of N-qubit cluster states $|C_N\rangle$ defined by graphs of the type shown in Fig. 1— these are two-dimensional cluster states with a fraction of sites removed. A measurement of each site of set A in the σ^x eigenbasis projects B into $|\text{GHZ}_{N_B}\rangle$ up to local σ^x gates depending on the outcomes, where N_B is the number of B sites. Classical communication can remove these errors, making the final state deterministically $|\text{GHZ}_{N_B}\rangle$. Therefore $N_D^*(|C_N\rangle) \geqslant N_B = O(N)$. It is simple to see that this generalizes to cluster states of any dimension.

C. Kitaev surface code states

Kitaev's surface code model is the simplest lattice system displaying topological order, which is of great interest for condensed matter physics and topological quantum computation [32–34]. Its ground states are sensitive to *global* topological properties of their lattice. Since they could exist on a macroscopic lattice, this makes them interesting candidates for macroscopicity.

The two-dimensional version is defined by a square lattice where each edge represents a qubit, with Hamiltonian $H = -\sum_s z_s - \sum_p x_p$. Here, z_s is the product of σ^z operators over a star and x_p is the product of σ^x operators over a plaquette
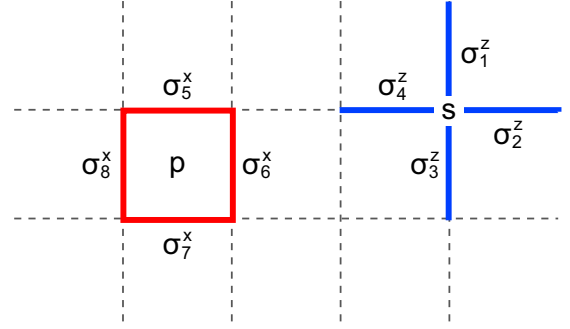

FIG. 2. (Color online) A star s and plaquette p in Kitaev's model and their corresponding operators: $z_s = \sigma_1^z \sigma_2^z \sigma_3^z \sigma_4^z$ and $x_p = \sigma_5^x \sigma_6^x \sigma_7^x \sigma_8^x$.

(see Fig. 2), and the sums are over all stars and plaquettes in the lattice. There may be a number of degenerate ground states, depending on the topology of the lattice; however, these have equivalent entanglement properties [35] so we concentrate on the simplest ground state.

It is possible to describe the structure of this state explicitly. We define a configuration of the lattice to be a product state where the qubits lying on a certain set of curves (open or closed) are all in the state $|1\rangle$ while the rest are $|0\rangle$. The ground state $|K_N\rangle$ is an equal superposition of all configurations containing only closed curves which are topologically trivial (contractible to a point).

$|K_N\rangle$, like $|C_N\rangle$, has no correlations between non-neighboring regions [36], so $N_{\mathcal{F}}^* = O(1)$. As a further parallel, one can deterministically distill macroscopic GHZ states from $|K_N\rangle$ by SLOCC: choose any topologically trivial non-self-intersecting loop B of $N_B = O(N)$ qubits and perform local projective measurements on the remainder of the lattice in the z basis. Consider the case where the outcome of each measurement is $|0\rangle$—then B ends up in $|\text{GHZ}_{N_B}\rangle$ containing the only two consistent closed-curve configurations. For different outcomes, the final state differs by local σ^x operations, so we can obtain $|\text{GHZ}_{N_B}\rangle$ deterministically. Hence $N_D^*(|K_N\rangle) = O(N)$.

D. Dicke states

Dicke states [37] have recently been studied for their interesting multipartite entanglement properties [6,22,38–40]. The N-qubit versions $|N,k\rangle$ are defined by symmetrizing $|0\rangle^{\otimes(N-k)}|1\rangle^{\otimes k}$. For any observable of the form $A = \sum_i \cos(\theta)\sigma_i^x + \sin(\theta)\sigma_i^y$ we obtain the maximal variance giving $N_{\mathcal{F}}^* = 1 + 2f(1-f)N$, where the quantity $f = k/N$ controls the macroscopicity in the same way as ϵ for $|\psi_N^\epsilon\rangle$.

Our upper bound gives $N_D^*(|N,k\rangle) \leqslant 2\min\{k, N-k\}$. We are not aware of any SLOCC protocols to distill GHZ states from $|N,k\rangle$, and so cannot provide a lower bound. We can only state that $N_D^* \leqslant 2(N_{\mathcal{F}}^* - 1)$, using $\min\{f, 1-f\} \leqslant 2f(1-f)$. Thus $N_{\mathcal{F}}^* < O(N) \Rightarrow N_D^* < O(N)$. However we do not know if $N_{\mathcal{F}}^* = O(N) \Rightarrow N_D^* = O(N)$.

IV. ROBUSTNESS OF DISTILLATION

In the above cases of cluster states and Kitaev surface code states, the distillation accomplishes something of practical

value: it extracts states which are useful for quantum metrology from initial states which are not, without increasing entanglement. Equivalently, it amplifies the macroscopicity of the states according to the measure $N_{\mathcal{F}}^*$. If this is to be regarded as something with practical significance, then the final states must be robust with respect to imperfections in the protocol. Indeed, one might doubt the physical meaning of the distillation if it is not robust in this sense.

To be specific, we neglect environmental noise since the GHZ states produced are maximally sensitive to decoherence [28], so any experiment taking advantage of them must have tolerably low noise. Instead, we suppose that the measurement device operating on the individual qubits may couple to them imprecisely. Therefore we say that the distillation of GHZ states is robust if the average Fisher information of the final state retains the same scaling with N for any small perturbation to the measurement operators.

The definition of the perturbation will be clear from the following examples.

A. Cluster states

For cluster states, we perturb the projective measurements to generalized measurement operators $\{E, \bar{E}\}$, satisfying $E^\dagger E + \bar{E}^\dagger \bar{E} = I$. Via the action of controlled-σ^z gates, the initial state can be written as $|C_N\rangle \propto \sum_{\mathbf{b}} |\mathbf{a}(\mathbf{b})_A^x\rangle |\text{GHZ}(\mathbf{b})_B^z\rangle$, where A and B are expressed in the x and z bases, respectively, $\mathbf{a} \in \{0,1\}^{N_A}$, $\mathbf{b} \in \{0,1\}^{N_B}$, and $|\text{GHZ}(\mathbf{b})\rangle := (|\mathbf{b}\rangle + |\bar{\mathbf{b}}\rangle)/\sqrt{2}$, with $\bar{b}_i := 1 - b_i$. Each value of $\mathbf{a}$ is determined by $\mathbf{b}$, and without loss of generality we fix $b_1 = 0$ (see Appendix B for details).

In the unperturbed case, projective measurements on A give full GHZ states on B with equal probability for each outcome. From this symmetry it is clear that we need examine only a single branch, say the outcome $E^{\otimes N_A}$. The final state is $|\psi\rangle \propto \sum_{\mathbf{b}} [E^{\otimes N_A} |\mathbf{a}(\mathbf{b})_A^x\rangle] |\text{GHZ}(\mathbf{b})_B^z\rangle$.

The most general form (up to normalization) for $E^\dagger E$ in the x basis is

$$E^\dagger E = \begin{pmatrix} 1 & \delta^* \\ \delta & \epsilon \end{pmatrix}, \quad \delta \in \mathbb{C}, \ \epsilon \in \mathbb{R}. \qquad (7)$$

For the observable $Z_B = \sum_{i \in B} \sigma_i^z$, $\mathcal{F}$ depends on ϵ but not δ:

$$\frac{1}{4}\mathcal{F}(|\psi\rangle\langle\psi|, Z_B) = \frac{1}{\mathcal{Z}} \sum_{\mathbf{b}} \epsilon^{|\mathbf{a}(\mathbf{b})|}(N_B - 2|\mathbf{b}|)^2, \qquad (8)$$

where $\mathcal{Z} = \sum_{\mathbf{b}} \epsilon^{|\mathbf{a}(\mathbf{b})|}$ and $|\mathbf{a}| := \sum_i a_i$. We interpret this by mapping the problem onto a two-dimensional square-lattice ferromagnetic classical Ising model. Each $\mathbf{b}$ maps onto a spin configuration with individual magnetic moments of ± 1 and a total magnetization $M = N_B - 2|\mathbf{b}|$, and $\mathbf{a}(\mathbf{b})$ is the corresponding bond configuration. There is an energy cost of $2J$ for each $a_i = 1$, so we associate ϵ with $e^{-2\beta J}$ and $\mathcal{Z}$ with the partition function. Therefore $\frac{1}{4}\mathcal{F}(|\psi\rangle\langle\psi|, Z_B) = \langle M^2 \rangle$.

This model has a low-temperature ferromagnetic phase; from a bound by Griffiths [41] we obtain $\frac{1}{4}\mathcal{F}(|\psi\rangle\langle\psi|, Z_B) \geqslant N_B^2(1 - O(\epsilon^4))$—so the distillation is robust.

This fails for one-dimensional cluster states, because the Ising model has no $T > 0$ phase transition in one dimension

[42]. So distillation from one-dimensional cluster states is not robust.

B. Kitaev surface code states

$|K_N\rangle$ works similarly, and the closed-loop restriction lets us interpret the configurations as domain walls of an Ising model with a spin on each plaquette. The robustness of distillation again relies on $T > 0$ ferromagnetic order. However, the relevant statistical model is more complex and depends on δ. As discussed in Appendix C, the disorder associated with δ maps onto probabilistically turning off the bonds, giving a dilute Ising model. This has a ferromagnetic phase [43], letting us conclude $\langle Z_B^2 \rangle = O(N_B^2)$ for sufficiently small δ and ϵ.

An important caveat is that this only works when the loop B divides the lattice into two-dimensional regions—for the same reason as the failure of one-dimensional cluster states. Therefore it seems we can only conclude robustness when $N_B = O(\sqrt{N})$. The distilled states, while robust, have $\mathcal{F} = O(N)$ and do not offer a metrological advantage over the initial state.

V. CONCLUSIONS

In summary, we have proposed a notion of macroscopic entanglement which measures the size of GHZ states that can be extracted by SLOCC and compared this with an existing measure of macroscopicity. We find that cluster states and Kitaev model ground states are macroscopically entangled by our definition but are not macroscopically quantum by the Fisher information. However, we lack statements examining the converse: whether $N_{\mathcal{F}}^* = O(N) \Rightarrow N_D^* = O(N)$. Extensions to our measure beyond qubits are also important—for finite-dimensional systems, one could distill $|\text{GHZ}_n^k\rangle = \frac{1}{\sqrt{k}} \sum_{i=0}^{k-1} |i\rangle^{\otimes n}$. We suggest ascribing the same size to this as to $|\text{GHZ}_n^2\rangle$, as $|\text{GHZ}_n^2\rangle^{\otimes m}$ looks like $|\text{GHZ}_n^{2^m}\rangle$ at the m-qubit level.

The distillation can also be interpreted as increasing the advantage of these states in high-precision parameter estimation. A criterion requiring the distillation to be of practical utility, by being robust against errors, tells us that two-dimensional, but not one-dimensional, cluster states are useful. This dependence on dimensionality also appears in the statement that two-dimensional cluster states are universal for measurement-based quantum computation, while the one-dimensional versions are not [44]. Kitaev model ground states are robust, but this can be proved only up to a nonmacroscopic $O(\sqrt{N})$ distilled size. It is noteworthy that our analysis for these states is very close to the proof that surface codes can be error corrected [45]. Thus we speculate that our results might be related to the ability of these states to encode quantum information and do so robustly.

As suggested by others [12], it would be helpful to develop a resource theory for macroscopicity, requiring an understanding of the operations unable to increase the effective size. The SLOCC distillation operations used here are not included in the set of free operations for macroscopic coherence defined in Ref. [20], under which the Fisher information cannot increase. The crucial point is that the final σ^x gates, conditioned on measurement outcomes, are not free in that framework and

instead entail manipulation of coherence between eigenstates of Z_B. This is where the present notion of macroscopic entanglement departs from macroscopic coherence.

APPENDIX A: UPPER BOUND ON N_D^* FOR SYMMETRIC STATES

First note that for symmetric states, we may always construct a symmetric optimal distillation protocol: if an asymmetric optimal protocol is found, one just needs to probabilistically perform all its permutations with equal weighting. As a result, the probability p_a for each $|\text{GHZ}_{S_a}\rangle$ is a function of n_a only. We denote by q_n the total probability of obtaining any GHZ state of size n—it is clear that $q_n = \binom{N}{n} p_a$ for any a such that $|S_a| = n$.

For the distilled state $|\phi_a\rangle = |\text{GHZ}_{S_a}\rangle |0\rangle^{\otimes N - |S_a|}$ the geometric entanglement E_G of a site i with the rest of the system is $\frac{1}{2}$ if $i \in S_a$ and 0 otherwise. The number of subsets of size n containing i is $\binom{N-1}{n-1}$.

The average E_G for the final ensemble is therefore

$$\sum_a p_a E_G(|\phi_a\rangle) = \sum_{n=2}^N q_n \frac{\binom{N-1}{n-1}}{\binom{N}{n}} \frac{1}{2}$$

$$= \frac{1}{2} \sum_{n=2}^N q_n \frac{n}{N}$$

$$= \frac{1}{2N} N_D^*(|\psi\rangle). \tag{A1}$$

By the monotonicity of E_G under SLOCC, we have

$$N_D^*(|\psi\rangle) \leqslant 2N E_G(|\psi\rangle). \tag{A2}$$

APPENDIX B: DISTILLATION FROM CLUSTER STATES $|C_N\rangle$

To construct $|C_N\rangle$, we define $\{|0^x\rangle, |1^x\rangle\} := \{|+\rangle, |-\rangle\}$ and write the state $|+\rangle^{\otimes n}$ before applying the U_{CZ} gates as

$$\bigotimes_{i \in A} |0_i^x\rangle \otimes \bigotimes_{j \in B} \left(|0_j^z\rangle + |1_j^z\rangle\right) = \sum_{\mathbf{b}} |\mathbf{0}_A^x\rangle \otimes |\mathbf{b}_B^z\rangle. \tag{B1}$$

The action of U_{CZ} on two qubits, i and j, is determined by

$$U_{CZ} |a_i^x\rangle |b_j^z\rangle = |(a_i \oplus b_j)^x\rangle |b_j^z\rangle, \tag{B2}$$

where $\oplus$ denotes addition modulo 2. Therefore we have that

$$|C_N\rangle \propto \sum_{\mathbf{b}} |\mathbf{a}(\mathbf{b})_A^x\rangle |\mathbf{b}_B^z\rangle, \tag{B3}$$

where each $\mathbf{a}(\mathbf{b})$ is determined by the following rule: $a_i = 0$ when the neighboring b_i are equal, and $a_i = 1$ when they

are different. One can see that $\mathbf{a}(\mathbf{b}^{(1)}) = \mathbf{a}(\mathbf{b}^{(2)})$ if and only if $\mathbf{b}^{(2)} = \mathbf{b}^{(1)}$ or $\overline{\mathbf{b}^{(1)}}$. This leads to

$$|C_N\rangle \propto \sum_{\mathbf{b}} |\mathbf{a}(\mathbf{b})_A^x\rangle |\text{GHZ}(\mathbf{b})_B^z\rangle, \tag{B4}$$

where we fix $b_1 = 0$ without loss of generality—any single site in B could be fixed. It is clear from this expression that a measurement of every A site in the x basis will project B into a GHZ state; there are $2^{N_B - 1}$ different outcomes.

With imperfect measurements, as discussed in the main text, we only need to consider a single branch with the outcome E for each measurement, so we use the final state

$$|\psi\rangle = \frac{1}{\sqrt{\mathcal{Z}}} \sum_{\mathbf{b}} \left[E^{\otimes N_A} |\mathbf{a}(\mathbf{b})_A^x\rangle \right] |\text{GHZ}(\mathbf{b})_B^z\rangle. \tag{B5}$$

To calculate the Fisher information in the variable $Z_B = \sum_{i \in B} \sigma_i^z$, we just need the variance $\langle\psi|Z_B^2|\psi\rangle - \langle\psi|Z_B|\psi\rangle^2$. Now it is easy to see from symmetry that $\langle\psi|Z_B|\psi\rangle = 0$, while

$$\sum_i \sigma_i^z(|\mathbf{b}^z\rangle \pm |\overline{\mathbf{b}}^z\rangle) = \sum_i (-1)^{b_i}(|\mathbf{b}^z\rangle \mp |\overline{\mathbf{b}}^z\rangle)$$

$$= (N_B - 2|\mathbf{b}|)(|\mathbf{b}^z\rangle \mp |\overline{\mathbf{b}}^z\rangle) \tag{B6}$$

gives $Z_B^2 |\text{GHZ}(\mathbf{b})_B^z\rangle = (N_B - 2|\mathbf{b}|)^2 |\text{GHZ}(\mathbf{b})_B^z\rangle$. Therefore

$$\langle\psi|Z_B^2|\psi\rangle = \frac{1}{\mathcal{Z}} \sum_{\mathbf{b}, \mathbf{b}'} \langle \mathbf{a}(\mathbf{b})|(E^\dagger E)^{\otimes N_A}|\mathbf{a}'(\mathbf{b}')\rangle$$

$$\times \langle \text{GHZ}(\mathbf{b})_B^z | Z_B^2 | \text{GHZ}(\mathbf{b}')_B^z \rangle$$

$$= \frac{1}{\mathcal{Z}} \sum_{\mathbf{b}, \mathbf{b}'} \langle \mathbf{a}(\mathbf{b})|(E^\dagger E)^{\otimes N_A}|\mathbf{a}'(\mathbf{b}')\rangle$$

$$\times \delta_{\mathbf{b}, \mathbf{b}'}(N_B - 2|\mathbf{b}|)^2$$

$$= \frac{1}{\mathcal{Z}} \sum_{\mathbf{b}} \epsilon^{|\mathbf{a}(\mathbf{b})|}(N_B - 2|\mathbf{b}|)^2, \tag{B7}$$

given that $\langle 0|E^\dagger E|0\rangle = 1$, $\langle 1|E^\dagger E|1\rangle = \epsilon$. Similarly, by setting $\langle\psi|\psi\rangle = 1$ we obtain $\mathcal{Z} = \sum_{\mathbf{b}} \epsilon^{|\mathbf{a}(\mathbf{b})|}$.

As described in the main text, $\frac{1}{4}\mathcal{F}(|\psi\rangle\langle\psi|, Z_B)$ can be interpreted as $\langle M^2\rangle$ for a classical Ising model with $\epsilon = e^{-2\beta J}$. Griffiths [41] establishes a bound of the form $\langle|M|\rangle \geqslant N_B[1 - f(T)]$, where f is independent of N_B and $\lim_{T \to 0} f(T) = 0$. Using $0 \leqslant \langle(N_B - |M|)^2\rangle = N_B^2 - 2N_B\langle|M|\rangle + \langle M^2\rangle$, it follows that $\langle M^2\rangle \geqslant 2N_B\langle|M|\rangle - N_B^2 \geqslant N_B^2[1 - 2f(T)]$. To leading order, $f(T) \approx e^{-8\beta J}$, which gives the bound

$$\frac{1}{4}\mathcal{F}(|\psi\rangle\langle\psi|, Z_B) \geqslant N_B^2[1 - O(\epsilon^4)]. \tag{B8}$$

APPENDIX C: DISTILLATION FROM KITAEV MODEL GROUND STATES $|K_N\rangle$

As above, we need to calculate $\langle Z_B^2\rangle$ and verify that it remains $O(N_B^2)$ for sufficiently small perturbations. Our approach again involves a map onto a statistical model where a spin σ_i is placed at the center of each plaquette. B is chosen to be a rectangular loop of size $N_B = O(\sqrt{N})$ which cuts the remainder A of the lattice into two independent two-dimensional regions. If we label each edge in B by $b_i = 0$ or 1 with neighboring inside and outside Ising spins $\sigma_i^{(i)}$ and
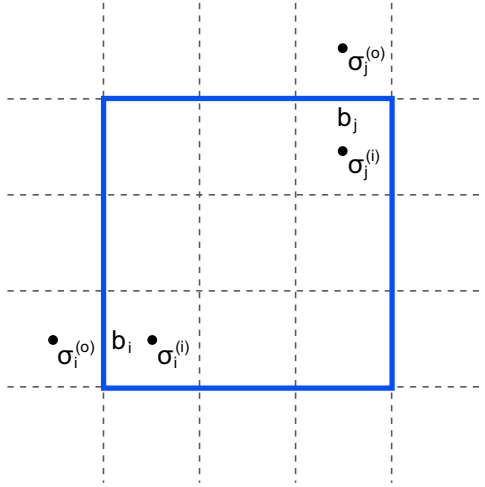
FIG. 3. (Color online) The highlighted loop B is the set of qubits to be distilled into a GHZ state. Each segment b_i has neighboring Ising spins $\sigma_i^{(i)}$ and $\sigma_i^{(o)}$.

$\sigma_i^{(o)} = \pm 1$ (see Fig. 3), then $\sigma_i^{(i)}\sigma_i^{(o)} = 1 - 2b_i$ and $\langle Z_B^2 \rangle$ will be given by

$$
\begin{aligned}
\langle (N_B - 2|\mathbf{b}|)^2 \rangle &= \left\langle \left(\sum_{i \in B} \sigma_i^{(i)}\sigma_i^{(o)} \right)^2 \right\rangle \\
&= N_B + 2 \sum_{i<j} \langle \sigma_i^{(i)}\sigma_j^{(i)} \rangle \langle \sigma_i^{(o)}\sigma_j^{(o)} \rangle,
\end{aligned} \tag{C1}
$$

where the expectation value is taken with respect to the statistical model to be defined below. Hence it will follow that $\langle Z_B^2 \rangle = O(N_B^2)$ if our model results in an ordered phase where the two-point correlation functions are bounded strictly above zero independent of N.

To define the required mapping, we recall that the premeasurement state is a sum of (topologically trivial) closed-loop configurations in the z basis:

$$
|K_N\rangle \propto \sum_{c:\, \partial c = 0} |c_A\rangle |c_B\rangle, \tag{C2}
$$

where ∂c is the boundary of c and c_A and c_B are the parts of c existing on A and B, respectively. (For the remainder of this section, a quantity such as c or c_A will be understood to be a vector, while c_i is a single component.) Since B is a closed loop, every distinct c_A in this sum corresponds to exactly two values of c_B, which are the opposites of each other. Therefore, by choosing to fix a single site of B, say $b_1 = 0$ (which we do implicitly from now on), we can write the state as

$$
|K_N\rangle \propto \sum_{c:\, \partial c = 0} |c_A\rangle |\mathrm{GHZ}(c_B)\rangle. \tag{C3}
$$

Furthermore, for each c_B we can generate all the corresponding c_A by choosing a particular representative $c_A(c_B)$ and adding to this all the possible closed loops z_A lying strictly within A. (Note that adding z_A to c_A creates a configuration $|c_A \oplus z_A\rangle$ in which c_A is deformed through the region bounded by z_A.)

This provides an additional representation:

$$
|K_N\rangle \propto \sum_{c_B} \left(\sum_{z_A:\, \partial z_A = 0} |c_A(c_B) \oplus z_A\rangle \right) |\mathrm{GHZ}(c_B)\rangle. \tag{C4}
$$

Hence we can write the postmeasurement state $E^{\otimes N_A} |K_N\rangle$ as

$$
|\psi\rangle = \frac{1}{\sqrt{\mathcal{Z}}} \sum_{c:\, \partial c = 0} E^{\otimes N_A} |c_A\rangle |\mathrm{GHZ}(c_B)\rangle \tag{C5}
$$

$$
= \frac{1}{\sqrt{\mathcal{Z}}} \sum_{c_B} \left(\sum_{z_A:\, \partial z_A = 0} E^{\otimes N_A} |c_A(c_B) \oplus z_A\rangle \right) |\mathrm{GHZ}(c_B)\rangle \tag{C6}
$$

such that $\mathcal{Z}$ gives the normalization $\langle \psi | \psi \rangle = 1$. We can use these two forms simultaneously to determine $\mathcal{Z}$:

$$
\begin{aligned}
\mathcal{Z} &= \sum_{c:\, \partial c = 0} \sum_{c_B'} \sum_{z_A:\, \partial z_A = 0} \langle c_A'(c_B') \oplus z_A | \mathcal{E} | c_A \rangle \\
&\quad \times \langle \mathrm{GHZ}(c_B') | \mathrm{GHZ}(c_B) \rangle \\
&= \sum_{c:\, \partial c = 0} \sum_{z_A:\, \partial z_A = 0} \langle c_A \oplus z_A | \mathcal{E} | c_A \rangle, \tag{C7}
\end{aligned}
$$

where $\mathcal{E} := (E^\dagger E)^{\otimes N_A}$. In the second line, we have used the fact that $c_B' = c_B$ to choose the representative $c_A'(c_B')$ to equal c_A. Similarly we find

$$
\langle Z_B^2 \rangle = \frac{1}{\mathcal{Z}} \sum_{c:\, \partial c = 0} \sum_{z_A:\, \partial z_A = 0} \langle c_A \oplus z_A | \mathcal{E} | c_A \rangle (N_B - |c_B|)^2. \tag{C8}
$$

From the general form of $E^\dagger E$ described in the main text, we have [46]

$$
\langle c_A \oplus z_A | \mathcal{E} | c_A \rangle = \prod_{i \in A} \delta^{z_i} \epsilon^{c_i(1 - z_i)}. \tag{C9}
$$

We shall see that this calculation can be mapped onto a statistical spin model with the following Hamiltonian:

$$
H = \sum_{\langle i,j \rangle} J_{ij} \sigma_i \sigma_j, \tag{C10}
$$

where the sum is over neighboring pairs of spins, and $J_{ij} = 0$ with probability p and $J_{ij} = -J$ $(J > 0)$ with probability $1 - p$. Such a model describes an Ising ferromagnet with random disorder caused by removing some of the bonds; we take this disorder to be annealed (as opposed to quenched), meaning that the J_{ij} are considered dynamical variables. It is known that this model is in a ferromagnetic phase for sufficiently small T and p [43].

Let us first rewrite the Hamiltonian in terms of the variables $f_{ij} = 0$ or 1 with probabilities $(1 - p)$ and p, respectively, and $b_{ij} = \frac{1}{2}(1 - \sigma_i \sigma_j)$:

$$
H(f,b) = -J \sum_e (1 - f_e)(1 - 2b_e), \tag{C11}
$$

where we replace (ij) with e labeling the edges of the lattice dual to the spins. Up to normalization, the probability of the

configuration (f,b) is

$$P(f,b) \propto e^{-\beta H(f,b)} \prod_e \left(\frac{p}{1-p}\right)^{f_e}$$

$$\propto \prod_e e^{-\beta J(f_e + 2b_e(1-f_e))} \left(\frac{p}{1-p}\right)^{f_e}$$

$$= \prod_e \left[\left(\frac{p}{1-p}\right) e^{-\beta J}\right]^{f_e} (e^{-2\beta J})^{b_e(1-f_e)}. \quad (C12)$$

Therefore we can make the identification $\langle c_A \oplus z_A | \mathcal{E} | c_A \rangle \to P(f,b)$ as long as we map

$$\delta \to \left(\frac{p}{1-p}\right) e^{-\beta J}, \quad \epsilon \to e^{-2\beta J},$$
$$z_i \to f_e, \quad c_i \to b_e. \quad (C13)$$

The last of these is compatible with the restriction $\partial c = 0$ because each bond configuration b for the spins must be formed of closed loops. However, our condition that $\partial z_A = 0$ means that we must similarly restrict the patterns of f in the spin model; it is nevertheless clear that this only reduces the disorder and so must preserve the ferromagnetic phase. There are also no bonds lying on the curve B, explaining our earlier claim that B cuts the remainder A into two noninteracting sections.

Therefore, to sum up, each term $\langle c_A \oplus z_A | \mathcal{E} | c_A \rangle$ can be interpreted as the probability of a configuration in this disordered
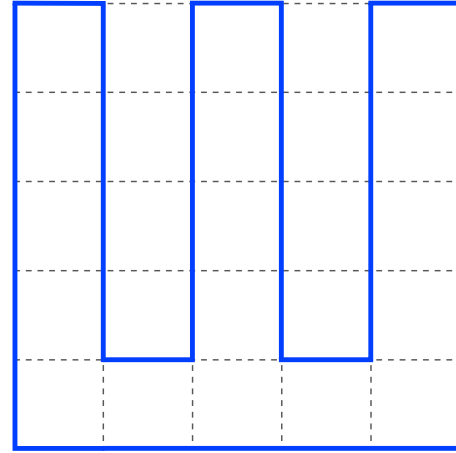

FIG. 4. (Color online) An example of a loop to be distilled into a size $O(N)$ GHZ state.

spin model, so that $\mathcal{Z}$ becomes the partition function and $\langle Z_B^2 \rangle$ is related to the two-point correlators as described earlier. The existence of a ferromagnetic phase then lets us conclude that $\langle Z_B^2 \rangle = O(N_B^2)$ for sufficiently small δ and ϵ.

This argument assumed B to be a rectangular loop of size $O(\sqrt{N})$, such that the distilled GHZ states have Fisher information of only $O(N)$—the same as the initial state $|K_N\rangle$. If we want an improvement for metrology, we would need to choose a loop of size $O(N)$. As depicted in Fig. 4, it seems to us that any such loop will divide the lattice into one-dimensional rather than two-dimensional regions, in which case the above argument does not apply. By analogy with the cluster state example, we might conjecture that no $O(N)$ loop is robust to perturbations.

[1] W.-B. Gao *et al.*, Nat. Phys. **6**, 331 (2010).
[2] C. H. Van der Wal *et al.*, Science **290**, 773 (2000).
[3] O. Nairz, M. Arndt, and A. Zeilinger, Am. J. Phys. **71**, 319 (2003).
[4] T. Farrow and V. Vedral, Opt. Commun. **337**, 22 (2015).
[5] M. Arndt and K. Hornberger, Nat. Phys. **10**, 271 (2014).
[6] F. Fröwis and W. Dür, New J. Phys. **14**, 093039 (2012).
[7] A. Leggett, Prog. Theor. Phys. Suppl. **69**, 80 (1980).
[8] A. J. Leggett, J. Phys.: Condens. Matter **14**, R415 (2002).
[9] We use the convention $f(N) = O[g(N)] \Leftrightarrow \lim_{N \to \infty} f(N)/g(N) \in (0, \infty)$.
[10] C.-W. Lee and H. Jeong, Phys. Rev. Lett. **106**, 220401 (2011).
[11] P. Sekatski, N. Sangouard, and N. Gisin, Phys. Rev. A **89**, 012116 (2014).
[12] F. Fröwis, N. Sangouard, and N. Gisin, Opt. Commun. **337**, 2 (2015).
[13] D. M. Greenberger, M. A. Horne, and A. Zeilinger, Going beyond Bell's theorem, in *Bell's Theorem, Quantum Theory and Conceptions of the Universe* (Springer, New York, 1989), pp. 69–72.
[14] A. Shimizu and T. Miyadera, Phys. Rev. Lett. **89**, 270403 (2002).
[15] A. Shimizu and T. Morimae, Phys. Rev. Lett. **95**, 090401 (2005).
[16] J. I. Korsbakken, K. B. Whaley, J. Dubois, and J. I. Cirac, Phys. Rev. A **75**, 042106 (2007).
[17] F. Marquardt, B. Abel, and J. von Delft, Phys. Rev. A **78**, 012109 (2008).
[18] G. Björk and P. G. L. Mana, J. Optics. B: Quantum Semiclassical Opt. **6**, 429 (2004).
[19] S. L. Braunstein and C. M. Caves, Phys. Rev. Lett. **72**, 3439 (1994).
[20] B. Yadin and V. Vedral, arXiv:1505.03792.
[21] S. Nimmrichter and K. Hornberger, Phys. Rev. Lett. **110**, 160403 (2013).
[22] G. Tóth, Phys. Rev. A **85**, 022322 (2012).
[23] P. Hyllus, W. Laskowski, R. Krischek, C. Schwemmer, W. Wieczorek, H. Weinfurter, L. Pezze, and A. Smerzi, Phys. Rev. A **85**, 022321 (2012).
[24] T. Morimae, Phys. Rev. A **81**, 010101 (2010).
[25] M. Popp, F. Verstraete, M. A. Martin-Delgado, and J. I. Cirac, Phys. Rev. A **71**, 042306 (2005).
[26] C. H. Bennett, H. J. Bernstein, S. Popescu, and B. Schumacher, Phys. Rev. A **53**, 2046 (1996).
[27] W. Dür, G. Vidal, and J. I. Cirac, Phys. Rev. A **62**, 062314 (2000).

[28] W. Dür, C. Simon, and J. I. Cirac, Phys. Rev. Lett. **89**, 210402 (2002).

[29] T.-C. Wei and P. M. Goldbart, Phys. Rev. A **68**, 042307 (2003).

[30] This is an improvement over the bound $N\epsilon^2 \log_2(1/\epsilon)/2$ in Ref. [28].

[31] H. J. Briegel and R. Raussendorf, Phys. Rev. Lett. **86**, 910 (2001).

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

[33] X.-G. Wen, *Quantum Field Theory of Many-Body Systems from the Origin of Sound to an Origin of Light and Electrons* (Oxford University Press, New York, 2004).

[34] H. Bombin, in *Quantum Error Correction*, edited by Daniel A. Lidar and Todd A. Brun (Cambridge University Press, New York, 2013).

[35] R. Orús, T.-C. Wei, O. Buerschaper, and M. Van den Nest, New J. Phys. **16**, 013015 (2014).

[36] Y. A. Lee and G. Vidal, Phys. Rev. A **88**, 042318 (2013).

[37] R. Dicke, Phys. Rev. **93**, 99 (1954).

[38] F. Fröwis and W. Dür, Phys. Rev. Lett. **109**, 170401 (2012).

[39] M. Bergmann and O. Gühne, J. Phys. A: Math. Theor. **46**, 385304 (2013).

[40] L. Novo, T. Moroder, and O. Gühne, Phys. Rev. A **88**, 012305 (2013).

[41] R. B. Griffiths, Phys. Rev. **136**, A437 (1964).

[42] J. M. Yeomans, *Statistical Mechanics of Phase Transitions* (Oxford University Press, London, 1992).

[43] M. Thorpe and D. Beeman, Phys. Rev. B **14**, 188 (1976).

[44] M. A. Nielsen, Rep. Math. Phys. **57**, 147 (2006).

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

[46] We take δ to be real without loss of generality, as the sums for $\mathcal{Z}$ and $\langle Z_B^2 \rangle$ will select the real part in any case.