Magic state distillation from entangled states

Ning Bao, ChunJun Cao, and Vincent Paul Su

Magic State Distillation from Entangled States

Ning Bao,[1, *] ChunJun Cao,[2, †] and Vincent Paul Su[3, ‡]

[1] *Computational Science Initiative, Brookhaven National Laboratory, Upton, New York, 11973*
[2] *Joint Center for Quantum Information and Computer Science,*
University of Maryland, College Park, MD, 20742, USA
[3] *Center for Theoretical Physics and Department of Physics, University of California, Berkeley, CA 94720*

Magic can be distributed non-locally in many-body entangled states, such as the low energy states of condensed matter systems. Using the Bravyi-Kitaev magic state distillation protocol, we find that non-local magic is distillable and can improve the distillation outcome. We analyze a few explicit examples and show that spin squeezing can be used to convert non-distillable states into distillable ones. Our analysis also suggests that the conventional product input states assumed by magic distillation protocols are extremely atypical among general states with distillable magic. It further justifies the need for studying a diverse range of entangled inputs that yield magic states with high probability.

I. INTRODUCTION

Fault tolerance is an absolute necessity for leveraging the full power of quantum computing. It is known that one cannot implement a universal set of gates transversally using a single quantum error correction code [1–5]. For many popular choices of codes, Clifford gates are often transversal while the non-Clifford gates are not. Therefore, the challenge is to perform the remaining non-Clifford element, such as a T gate, fault-tolerantly. The difficulty of implementing these T gates forms a major bottleneck for the development of universal quantum computation. Various solutions for implementing non-Clifford gates have been proposed [6–14].

One such proposal is to prepare so-called "magic states" and perform T-gates through measurement-based quantum computation schemes [6, 15–17] using only fault-tolerant Clifford operations. To this end, magic state distillation (MSD) protocols have been developed in order to convert product states with large overlap with the magic states into the magic states of a particular form; for the seminal work in this area, please see [6, 7]. However, these magic states can be costly to prepare with arbitrarily high accuracy. Such is the prevalent assumption behind the program of Clifford+T circuit decompositions and related compilation schemes [18–22]. Because of the inefficiencies in distilling these magic states, the ability to develop a "magic state factory" that is capable of effectively generating such states in large quantities would be a great advance to the scalability of quantum computation. For an alternate approach to fault tolerant computation using code switching, see [23–28].

The conventional setup for magic state distillation takes a number of noisy ancillae, which are often represented as product of identical single qubit mixed states, and projects them onto the code subspace of a carefully

chosen quantum error correction code. One then decodes using a fault-tolerant Clifford circuit and obtains a number of magic states that have higher fidelity provided the inputs are above a distillation threshold. A great number of procedures and variants have been devised to various degrees of efficacy [6, 7, 29–33].

While the product of identical mixed states is a natural input state, its form is rather restrictive considering the vast pool of many-body quantum states that occur naturally in physical systems that can function as magic reservoirs [34–38]. In particular, recent progress in quantum many-body magic indicates that, magic, like entanglement, can take on various forms in such states. Indeed, while some of the magic is found locally, not unlike the conventional input states in magic distillation protocols, a large portion remains in the correlations and is distributed non-locally in entangled states. Therefore, it is natural to ask whether such "non-local" magic from these systems can be a) distilled and b) used to improve the outcome of distillations. More generally, it is beneficial, not only theoretically but also practically, to understand how MSD performs on a much wider class of inputs, where magic is not concentrated strictly locally on each ancilla.

In this work, we take the first step in characterizing how more generic input states, such as entangled states, can alter the outcomes of distillation. In particular, we answer both of the above questions in the affirmative — it is possible to distill magic that is distributed non-locally, as opposed to locally in the traditional input states. Non-local magic (NLM) can also improve the success probabilities of the distillation protocol. A wider range of input states can thus help reduce the overall cost for generating a much needed quantum resource. From a practical perspective, we show that well-known procedures used in spin squeezing that generate entanglement and non-local magic can render undistillable magic states distillable. More concretely, we conduct our explicit analysis using the Bravyi-Kitaev (BK) magic distillation protocol [6] and study how it performs over different forms of entangled 5-qubit states.

We will review some basic aspects of magic states and

* ningbao75@gmail.com
† ccj991@gmail.com
‡ vipasu@berkeley.edu

their distillation in Section II. Then in section III, we construct different entangled states with varying magic distributions and perform the Bravyi-Kitaev protocol for T state distillation using these as input states. We show that although the conventional form of input states explicitly have only single qubit magic, states with purely non-local magic such that individual qubit reduced density matrices look non-magical can remain fully distillable. We then consider both conceptually simple and practically relevant examples where non-local magic and entanglement improve distillation outcome. This is made particularly apparent with the example of spin squeezing, where non-distillable states are rendered distillable with a slight addition of non-local magic and entanglement. Then in Section IV, we further analyze the properties that make certain states more distillable from the point of view of the quantum error correction code in question. By analyzing random distillable states, we conclude that the conventional MSD inputs are extremely atypical whereas typical distillable states have high entanglement and zero local magic. We then provide some intuitions for the patterns behind the distribution of the distillable states. Finally, in Section V, we summarize our findings, speculations, and future directions.

II. BACKGROUND

A. Magic

Stabilizer states are generated by Clifford operations acting on the all zeros state. They have a wide range of applications in quantum information, condensed matter physics and quantum gravity [39–41]. However, it is well known that they are not dense in the Hilbert space, and thus the Clifford operations, which can be implemented transversally for certain encoding schemes, are not sufficient for universal quantum computation. In fact, the Gottesman-Knill theorem [42] gives us a constructive way to efficiently simulate Clifford circuits purely classically.

However, they can be supplemented with a non-Clifford gate using measurement based protocols that only involves Clifford operations [6, 7] and noisy ancilla states, thus rendering the computation universal. These procedures require distilling "non-stabilizer-ness", or magic, into a useable form. Therefore, magic, being a resource [43] in such fault-tolerant computation schemes, aims to quantify the degree to which states fail to be captured by stabilizer states. There are many magic measures. For example, the most straightforward definition is the minimum distance between a state of interest ρ and the set of stabilizer states

$$D(\rho) = \min_{\sigma \in \text{STAB}} ||\rho - \sigma||, \qquad \text{(II.1)}$$

where STAB denotes the set of stabilizer states. However, this measure is difficult to compute practically. A number of computable measures have also been proposed [43–46].

In this work, we will use one such measure designed for qubit systems called the robustness of magic (ROM) [44], defined for an arbitrary density matrix ρ

$$\mathcal{R}(\rho) = \min \left\{ |x|_1 : \rho = \sum_i x_i s_i, \, x_i \in \mathbb{R} \right\} \qquad \text{(II.2)}$$

where the s_i correspond to density operators of stabilizer states. Note that the coefficients can be negative, allowing us to express arbitrary states in this way. This quantity is lower bounded by 1 in order to have a normalized density matrix. If ρ is a stabilizer state, then there will simply be a single non-zero coefficient.

As ROM is unity for all stabilizer states, here we use $\ln(\mathcal{R})$ which we call LROM to avoid the offset for stabilizer states and for similarities with definitions like mana [43]. However, unlike mana which is additive for product of magic states, LROM is, in general, subadditive

$$\mathcal{R}(\rho_1 \otimes \rho_2) \leq \mathcal{R}(\rho_1)\mathcal{R}(\rho_2)$$
$$\implies \ln(\mathcal{R}(\rho_1 \otimes \rho_2)) \leq \ln \mathcal{R}(\rho_1) + \ln \mathcal{R}(\rho_2). \qquad \text{(II.3)}$$

To distinguish local from non-local magic, we will compute magic for different subsystems in a single state. If, for instance,

$$\ln \mathcal{R}(\rho_{12}) > \ln \mathcal{R}(\rho_1) + \ln \mathcal{R}(\rho_2) \qquad \text{(II.4)}$$

we will then conclude that there is magic stored non-locally in the joint system ρ_{12} which is not found in each of the individual subsystems.

B. Bravyi-Kitaev Magic Distillation

Here we review the Bravyi-Kitaev(BK) MSD protocol for distilling a particular magic state, the T state. Like in the original work, we assume that stabilizer operations are easy. Therefore, T-type magic can be supplied by any state which obtained by a stabilizer operation acting on a specific T state, which we call $|T_0\rangle$. Those familiar with the Bravyi-Kitaev protocol can mostly skim this section, though we will establish some notation that will be used later in future sections.

The $|T_0\rangle$ state is the following qubit state.

$$|T_0\rangle = \cos(\theta_T) |0\rangle + e^{i\pi/4} \sin(\theta_T) |1\rangle, \qquad \text{(II.5)}$$

where

$$\theta_T = \frac{1}{2} \cos^{-1} \left(\frac{1}{\sqrt{3}} \right).$$

It can also be advantageous to write $|T_0\rangle$ in terms of its density matrix.

$$|T_0\rangle \langle T_0| = \frac{1}{2} \left(I + \frac{1}{\sqrt{3}}(X + Y + Z) \right). \qquad \text{(II.6)}$$

where X, Y and Z are the single qubit Pauli matrices. The density matrix makes manifest that one can also describe the state by its polarization vector $\vec{r} = \frac{1}{\sqrt{3}}(1, 1, 1)$ on the Bloch sphere. Similarly, one can write the $|T_1\rangle$ state as

$$|T_1\rangle \langle T_1| = \frac{1}{2}\left(I - \frac{1}{\sqrt{3}}(X + Y + Z)\right). \qquad \text{(II.7)}$$

From the polarization vector, one can see that $|T_0\rangle$ points in the direction normal to one of the triangular faces of the stabilizer octahedron. By the symmetries of the stabilizer octahedron, there are eight states which serve the same magical purpose, all related by single qubit Clifford operations. Thus successful T distillation occurs whenever any of the eight states are produced of which $|T_0\rangle$ and $|T_1\rangle$ are two. Likewise, a T-state may refer to any of these states, and we define the T-fidelity to be the maximum fidelity of a state with any of the T states,

$$F_T(\rho) = \max_{U \in \mathcal{C}^{(2)}} \{F(\rho, U|T_0\rangle\langle T_0|U^\dagger)\}. \qquad \text{(II.8)}$$

Here we use $\mathcal{C}^{(2)}$ to denote the Clifford group over a single qubit.

The Bravyi-Kitaev protocol for T-state distillation involves a recursive subroutine based on the $[[5, 1, 3]]$ code [47] whose stabilizer group is

$$\mathcal{S} = \langle XZZXI, IXZZX, XIXZZ, ZXIXZ \rangle. \qquad \text{(II.9)}$$

It distills T states because the code admits a transversal implementation of the Clifford gate $B = \exp(i\pi/4)SH$ for which $|T\rangle$ is an eigenstate. As usual, S, H are the single qubit $\pi/2$-phase gate and the Hadamard gate respectively.

Given five copies of an approximate T state of the following form

$$\rho_{in} = (1 - \epsilon)|T_0\rangle\langle T_0| + \epsilon|T_1\rangle\langle T_1|, \qquad \text{(II.10)}$$

the MSD protocol involves measuring the generators of $\mathcal{S}$ and post-selecting the trivial syndrome, effectively projecting the input state onto the code subspace. If a non-trivial syndrome is measured, the state is simply discarded and the protocol is restarted. The decoding circuit is then applied to return a single qubit state, which iteratively approaches a pure T state, provided the fidelity of ρ_{in} with any T-state is sufficiently high[1]. For every iteration of this procedure, one obtains noisy ancilla with a different fidelity. We plot the input and output fidelities as a result of this procedure in Figure 1.

It is worth noting that the input state, which is a tensor product of one qubit states, is certainly not in the code

―――――――――

[1] Equivalently, it is the same to first apply the decoding circuit then measure the disentangled syndrome qubits separately [48].
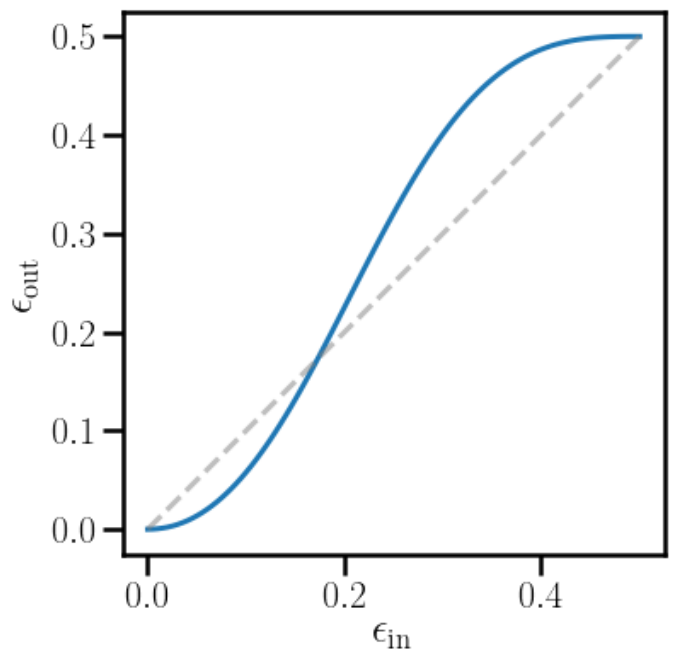

FIG. 1: The output fidelity (solid blue) as a function of input fidelity after a single round of Bravyi Kitaev magic distillation. For reference, the dashed gray line indicates the identity channel $\epsilon_{\text{in}} = \epsilon_{\text{out}}$. The cross over point is $\epsilon^* \approx 0.173$. If the input states are not sufficiently close to a T state, then they will limit to the maximally mixed state.

subspace. Because there is only a single logical qubit, one can decompose the projected state in the basis of the encoded magic states $|T_0^L\rangle$ and $|T_1^L\rangle$. The crux of the analysis lies in showing that for sufficiently small ϵ, the T-fidelity increases. If for any $\epsilon < \epsilon^*$ such that the protocol always increases the T-fidelity of the input state, then $1 - \epsilon^*$ is known as the distillation threshold.

When it comes to the efficiency of the procedure in creating magic states, there are two factors that determine the resource estimate. Assuming there is a target T-fidelity, one can use the curve in Figure 1 to estimate how many rounds of distillation are required starting from an initial pool of noisy ancilla. The second quantity of practical importance is the success probability for measuring trivial syndrome. By success probability we mean the probability of measuring the trivial syndrome corresponding to a single iteration of the distillation, as opposed to producing an approximate T state with the desired final fidelity after multiple rounds.

For input states of the form (II.10), the success probability depends on ϵ as well. As seen in Figure 3, this is upper bounded by $1/6$, meaning every round of distillation in expectation requires at least 30 approximately magical qubits. Since it is assumed that all 5 input qubits are the same, one actually needs access to an exponential number of noisy ancilla due to the recursive nature of the

protocol.

The utility of T states was also discussed in the same paper by Bravyi and Kitaev. They showed explicitly how to enact a non-Clifford gate with access to T states and only Clifford operations. The fact that the same state was the fixed point of the distillation procedure and enables Clifford operations to enact universal quantum computation was the reason for calling these states "magic states".

III. MAGIC DISTILLATION BEYOND PRODUCT STATES

In this section, we investigate whether existing MSD procedure can distill magic from non-product states. Indeed, the conventional input for such protocols $\rho_{\rm in}^{\otimes n}$ is a product state. Magic is localized to each bit and there is clearly no magic hidden in the quantum correlations. However, one might ask whether other forms of multi-qubit magic states can also be used to distill desired magic states as long as certain conditions are satisfied. In particular, can the distillation procedure distill magic from states where the state has entanglement or correlation such that magic is non-locally distributed, i.e. magic of any single qubit is small whereas that of a larger subsystem is non-trivial? For example, such states can exist in the ground state of quantum many-body systems [34, 35]. To this aim, we construct a few informative examples below each with distinct but carefully designed magic distributions and answer the above question in the affirmative by testing them using the BK protocol. For all of the distillation procedures below, we only use the entangled states as inputs for the first iteration. The ensuing iterations works under the conventional assumptions where multiple tensor copies of the outputs from earlier iterations are used.

A. Phase-GHZ and GHZ-T states

We begin with a few structurally simple states to build intuition. Consider GHZ states in the computational basis with a phase.

$$|PG(\alpha)\rangle_n = \alpha |0\rangle^{\otimes n} + e^{i\phi}\beta |1\rangle^{\otimes n}. \tag{III.1}$$

For instance, one can choose $\alpha \in [0, 1]$ which sets the relative weighting of the superposition. By normalization, we choose $\beta = \sqrt{1-\alpha^2} \in \mathbb{R}$. For extreme values of α/β, this becomes arbitrarily close to a simple product state, allowing us to "tune" the presence of magic. The specific magic distribution is shown in Figure 2a.

This state provides a steep contrast with the conventional input states of MSD — no proper subsystem is magical as they are all density matrices diagonal in the computational basis. However, the overall state clearly is magical because one can turn this state into the product
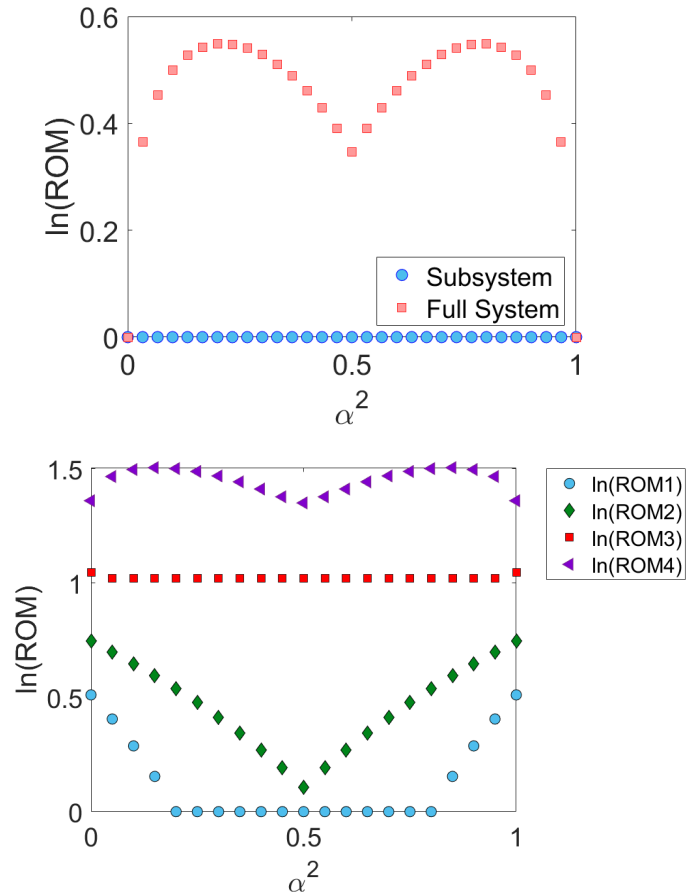

FIG. 2: (a) Magic distribution of the Phase-GHZ state with phase $\phi = \pi/4$. For PG, all subsystems have the same form of reduced density matrix. (b) Magic of various subsystems for the 4-qubit GHZ-T state as a function of α^2.

of $|0\rangle$s and a single qubit magic state $\alpha|0\rangle + e^{i\phi}\beta|1\rangle$ by applying a series of Clifford gates.

This state is distillable using BK (Figure 3) for most values of α because the output is above the distillation threshold (dashed line). Although BK is suboptimal for distilling magic from such kind of states, it clearly demonstrates that non-local magic remains distillable using existing protocols.

Other than the conventional product input states which have magic strictly localized to each qubit and the Phase-GHZ state above where magic is strictly global, one can also construct states that have magic at different scales. Consider a state that takes the following form on n qubits [2].

$$|GT(\alpha)\rangle_n = \alpha |T_0\rangle^{\otimes n} + e^{i\phi}\beta |T_1\rangle^{\otimes n}. \tag{III.2}$$

———

[2] While we do not use the degree of freedom of the relative phase in this work, it is one that's available to us and others for future work.
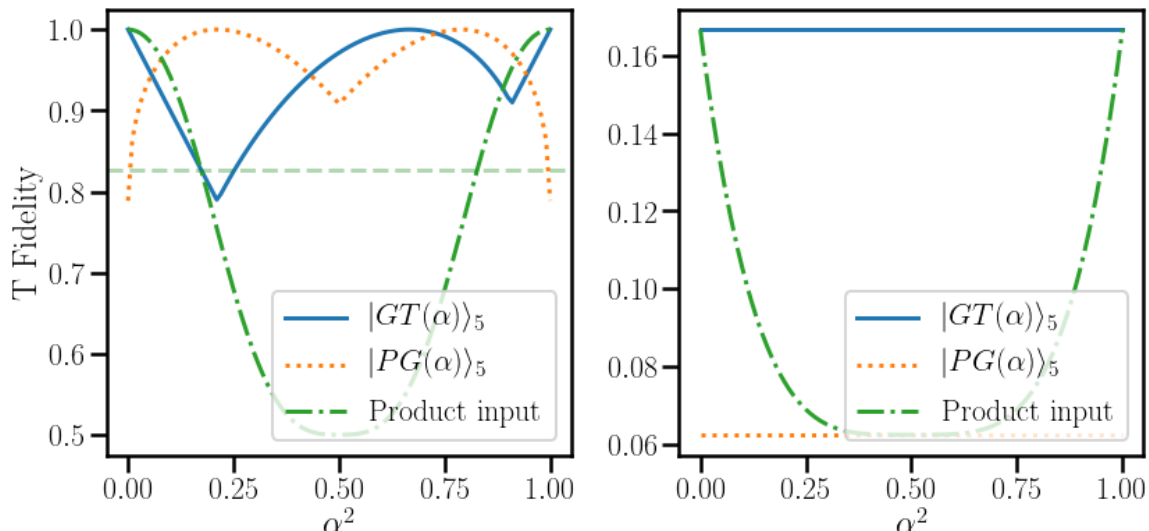
FIG. 3: With our choice of simple entangled states, we apply the BK protocol and measure the magic fidelity of the resulting qubit. For comparison, we also compute it for the product states considered in the original work. Notably, the output fidelity is above threshold across a wide range of $\alpha^2 \sim 1 - \epsilon$. The output of this round can then be recursively iterated. The green line shows the distillable threshold for states of the form (II.10). For the other states, they reach a pure T state except a few specific values of α. This more lenient threshold is not a contradiction since the output density matrix of this round is not of the same form. On the right, we show the probability to measure the trivial syndrome. To actually produce a T state, one needs the procedure to both converge to a T state and measure the trivial syndrome.

Such a state is chosen precisely because it is a combination of the noisy ancilla input and the Phase-GHZ state above. While each one-qubit reduced density matrix mimics the conventional noisy ancilla input that shows up in the protocol (e.g. $\epsilon = \beta^2$), the overall state is also similar to the Phase-GHZ state where there is additional magic globally. Thus for α or β small, magic is locally distributed. However, for other choices of α, β the state can pick up a non-trivial contribution from non-local magic in the n-qubit GHZ state. One can see that magic (Figure 2b) on small subsystems of one or two qubits has a distribution similar to that of the conventional input states. However, as we go to larger subsystems, there is also more non-local magic which adds to the total global magic available in the state.

Magic in such states are mostly distillable. See Figure 3. However, the dips in the distillability plot now no longer coincide with the locations where there is less total magic, which was the case for Phase-GHZ states.

We can ask whether how global magic helps with distillation in these examples. Other than the clear improvement in distillability in the region where $\beta^2 = \epsilon \in [\epsilon^*, 1 - \epsilon^*]$, where the conventional inputs are completely undistillable, there is also other improvements in the form of success probability, i.e., the probability of measuring the trivial syndrome (Figure 3). Indeed, the GHZ-T state, which has single qubit states identical to conventional noisy ancilla input but have non-zero non-local magic retains a relatively high success probability for all
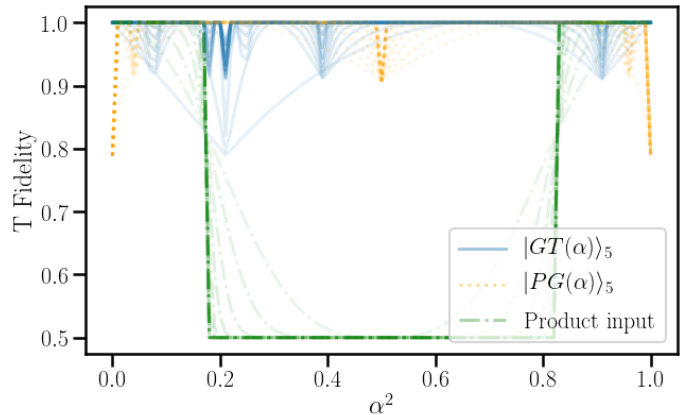

FIG. 4: Multiple iterations of the BK protocol for simple input states. Over a wide range, our choice of toy entangled states are distillable. Original product input states are limited to the tail ends where $\alpha^2 < \epsilon_*$ or $\alpha^2 > 1 - \epsilon_*$.

values of α, which is distinct from the product state counterpart.

B. Squeezed Spin States from Noisy Ancillae

The above examples are conceptually clean and easy to understand, but may be hard to prepare in the labo-

ratory. For entangled input states that are slightly more realistic, we consider initially states that are product stabilizer states that suffer from single qubit coherent noise. We then perform a global entangling unitary usually used in spin squeezing procedures [49] on such states and examine how it improves distillation outcome. Spin squeezed states have been implemented in atomic cavities going back a couple of decades, primarily with application to quantum metrology and detecting quantum entanglement [50–55].

The initial ancilla states before squeezing are spins that are parallel to, say, the x-axis, but have small misalignments across the spins. More explicitly,

$$|\psi_{\mathrm{ini}}\rangle = \bigotimes_{i=1}^{5} |\theta_i\rangle, \qquad |\theta_i\rangle = \cos\theta_i |0\rangle + \sin\theta_i |1\rangle, \quad \text{(III.3)}$$

where the orientation of each spin θ_i can fluctuate to simulate inhomogeneity of the initial spin alignment. This is modelled by choosing each $\theta_i \in [\bar{\theta}_i - \theta_{\max}, \bar{\theta}_i + \theta_{\max}]$ randomly from a uniform distribution. We take the means $\bar{\theta}_i$ to be equal for all i. For example, $\bar{\theta}_i = \pi/4$ creates a product state that is approximately $|+++++\rangle$ but with small misalignment across the spins. Despite the small amount of local magic introduced to $|\psi_{\mathrm{ini}}\rangle$ through the spin misalignments, such states remain mostly undistillable, especially for $\theta_{\max}$ small.

To define the unitary squeezing process, consider a total spin-5/2 system consisting of 5 qubits. It is known that a pure product state of qubits is not a squeezed state [49, 55] and squeezing introduces entanglement among the qubits. For example,

$$|+++++\rangle = 2^{5/2} \sum_{k=0}^{5} \sqrt{\binom{5}{k}} |S = \frac{5}{2}, S_z = \frac{5}{2} - k\rangle \quad \text{(III.4)}$$

Squeezing can then be performed by considering some unitary rotation on the $S = 5/2$ submanifold. The initial spin misalignments sometimes introduce components that are outside the $S = 5/2$ submanifold, and the unitary mapping over the 5 qubits leave these components unchanged. We can use the one-axis twisting method [49] with unitary $U(t)$ to generate a number of squeezed states $|\psi_{\mathrm{fi}}\rangle$

$$U(t) = \exp\left[-itS_z^2\right] \oplus I, \qquad |\psi_{\mathrm{fi}}(t)\rangle = U(t)|\psi_{\mathrm{ini}}\rangle. \quad \text{(III.5)}$$

Note that although $U(t)$ is not a Clifford unitary, there is also no need for it to be fault-tolerant or extremely precise as a range of different t's can all yield similar outcomes.

For small t, the action of $U(t)$ introduces a small amount of mostly non-local magic and entanglement to the initial state. See Figure 5b for results when $\langle \theta_i \rangle = \pi/4$. Surprisingly, the previously undistillable initial states are now mostly distillable by applying $U(t)$ (Figure 5a). Here cost is defined as the average number of copies of the same state needed to distill a single copy of
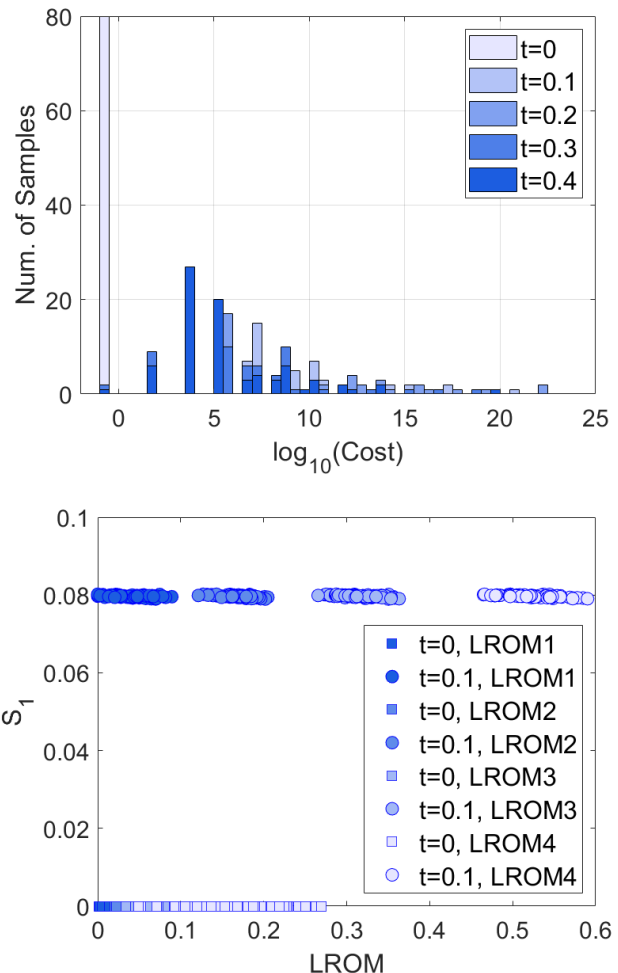

FIG. 5: (a) Figure of distillability of states with $(t \neq 0)$ vs without $(t = 0)$ squeezing $(\theta_{\max} = 0.05)$. States without squeezing (blue) are undistillable and are assigned cost -1. (b) Single site entanglement and magic of various subsystems of squeezed (colored disks) vs unsqueezed state (colored squares) with $\bar{\theta}_i = \pi/4$.

the T state with a reasonably high fidelity $(F_T(\rho) \geq 0.97$ for the plots). If the resulting output does not exceed the prescribed fidelity cut off after 15 iterations, we deem the state undistillable and it is assigned cost -1 to differentiate from the rest of the samples. Heuristically, we can also see spin squeezing as reducing what is otherwise an infinite distillation cost to a finite number as none of the intial input states were distillable.

This is another clear demonstration that entangled states generated in a more practical laboratory setting [50–53] can help enhance the outcome of magic state distillation. While this is encouraging, squeezing does not always improve distillation outcomes. This is particularly visible when the input states are already T-states with relatively high fidelity because they act like a source of correlated noise. We further discuss the role played by spin squeezing in Appendix B. We wish to point out that

it is difficult to directly compare the cost of a procedure with spin squeezing with the cost from conventional distillation procedures. This is because their inputs are in totally different regimes — the cost in the original BK protocol is computed assuming product of noisy ancilla states above distillation threshold, whereas the advantage of spin squeezing is to take states that are below the distillation threshold, e.g. very close to stabilizer states, and make them distillable. For this reason, the starting point for the procedures is hard to standardize between the two approaches. Taking a step back, we would like to emphasize that the lesson learned from spin squeezing is that a totally different class of states, even without magic, from the original noisy ancilla proposed can be made distillable. For those interested in making magic factories, it broadens the horizon of possible inputs and encourages us to look for other experimentally viable entangled states.

IV. UNDERSTANDING DISTILLABILITY

In this section, we examine the underlying $[[5, 1, 3]]$ code that powers the BK distillation procedure to understand properties of states that contain distillable magic. In doing so, we present a few key observations. First, using elementary ideas from quantum error correction, we can deduce the form of non-local magic content for the ideal input state. We then investigate these intuitions by generating random BK-distillable states whose success probability is lower bounded by that of product input states. Studying their entanglement and magic distribution properties tells us that the typical product ansatz is highly fine-tuned. Second, we can explain explicitly why the total magic in a state does not correlate with the distillable magic for this protocol. We expect the lessons here to generalize appropriately for any magic distillation procedure which relies on a quantum error correcting code.

As mentioned previously, the Bravyi-Kitaev distillation procedure (decoding followed by postselecting the trivial syndrome) can be represented as a linear map $M : \mathcal{H}_{\text{in}} \rightarrow \mathcal{H}_{\text{out}}$, where $\mathcal{H}_{\text{in}}, \mathcal{H}_{\text{out}}$ are 32 and 2 dimensional Hilbert spaces, respectively. Then for each normalized 5-qubit input $|\psi_{\text{in}}\rangle \in \mathcal{H}_{\text{in}}$, we can obtain a single-qubit output $|\psi_{\text{out}}\rangle \in \mathcal{H}_{\text{out}}$ such that

$$\lambda |\psi_{\text{out}}\rangle = M |\psi_{\text{in}}\rangle. \qquad (\text{IV.1})$$

As a portion of $|\psi_{\text{in}}\rangle$ may not lie in the code subspace, the probability of success $|\lambda|^2$, i.e. measuring a trivial syndrome, satisfies $0 \leq |\lambda| \leq 1$. It is clear that states with more overlap with the code subspace have higher success probability.

Working in the basis of the code subspace, any input state with unit success probability can be written as

$$|\psi_{\text{in}}\rangle = a|T_0^L\rangle + b|T_1^L\rangle \qquad (\text{IV.2})$$

where the size of a will determine the output state fidelity. This works because the perfect code encodes a single qubit.

It is immediately clear that states close to the encoded T states will have both high success probability and good T-fidelity. Furthermore, because the code corrects two qubit erasure errors, any encoded information must be found in 3 or more qubit subsystems. Therefore, for states with high success rate, i.e., large overlap with the code subspace, the amount of magic found in 3 or more qubit subsystems should positively correlate with the output fidelity. On the other hand, such states will have very little magic in one or two qubit subsystems.

However, when a state does not overlap significantly with the code subspace, namely $|\lambda|$ small, the connection between the amount of magic, its distribution and output fidelity is no longer clear.

To get an idea of a typical distillable state, we generate random states uniformly in the space of states where their output fidelity is above the BK distillation threshold and $|\lambda| > 0.16$ is reasonably large. As a reference, the success probability for product input states close to T states is ~ 0.167 (see Figure 3).

We see that the typical distillable states are not only entangled, but also have non-local magic (Figure 6). As we see, any subsystem is close to being maximally mixed while magic localized to any single qubit is virtually nonexistent. This is in sharp contrast with the conventional input states which have neither. Therefore, this observation indicates that having many-body quantum states beyond the typical product input states can improve the magic distillation outcomes.

However, unlike the conventional input states, there is no obvious correlation between the total amount of magic or entanglement a state has and the distillation outcome. In fact, if anything, too much magic can negatively impact distillability in terms of the success probability. We can try to understand this observation from two limits.

In the limit where the state is perturbatively close the an encoded T state, then having a state whose magic or entanglement pattern that deviates from the state that has the single best distillable outcome will obviously result in a worse distillation outcome. The maximum amount of magic for a state in the code subspace is upper bounded by that of a single T state (LROM ~ 0.55). If the input state has more magic than this, then it is necessarily out of the code subspace thus resulting in a lower success probability (Figure 7). If, on the other hand, the total magic is less than the upper bound, then it can have a large overlap with the code subspace. In that case, the success probability is close to 1, but its output fidelity must suffer for being farther from the encoded T state. There is a direct correlation between the amount of magic and output fidelity. This can be seen in Figure 8, but is less pronounced in Figure 7. Interestingly, it is also difficult to find a distillable state that has less total magic than that of a T state but also small overlap with the code subspace. This is likely related to
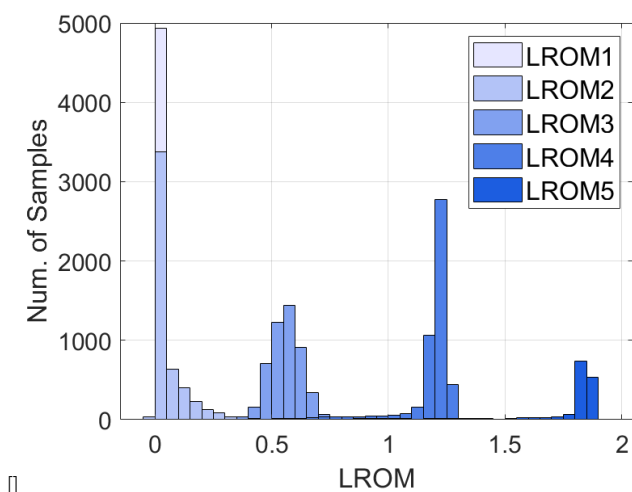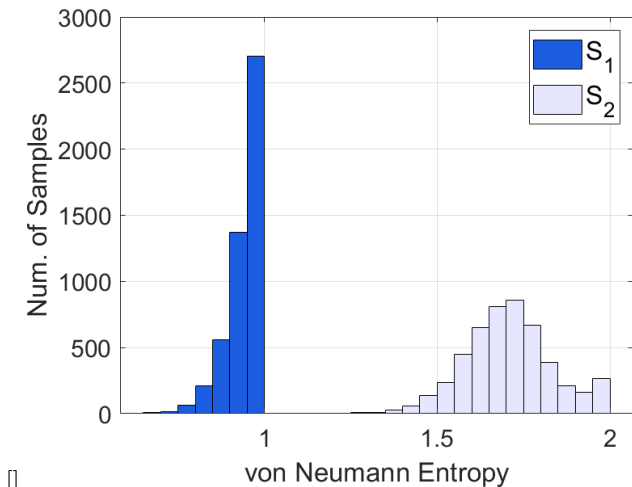
FIG. 6: (a) Distribution of von Neumann entropies for single and two qubit subsystems. (b) LROM distribution in the random BK-distillable states.
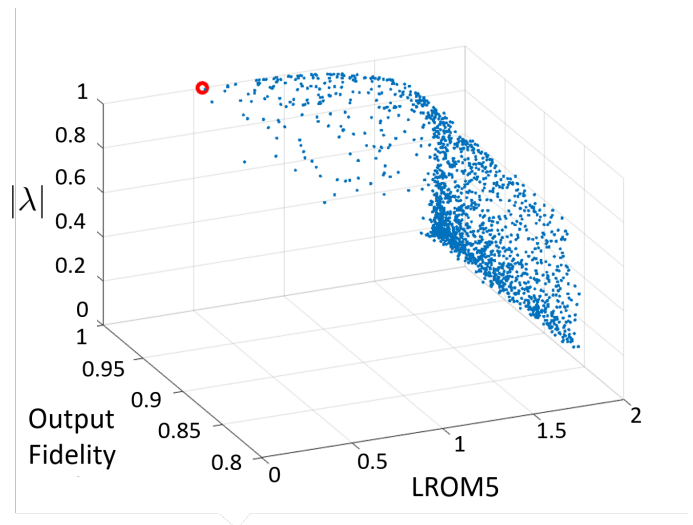

FIG. 7: LROM5 of random states (blue) against output fidelity and success probability. Red circle denotes the location of the encoded T state.
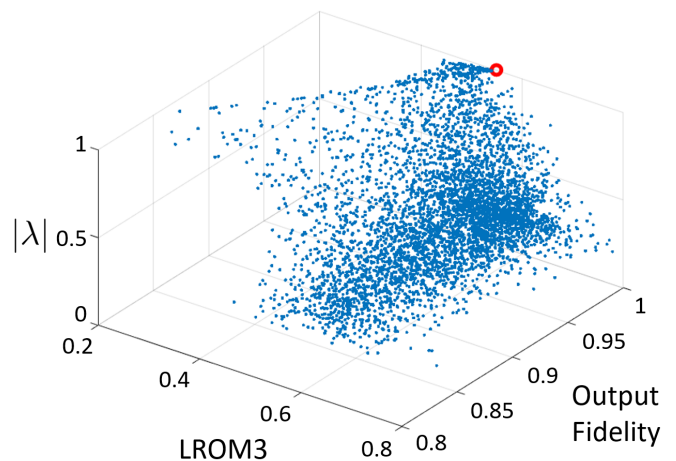
the fact that (Haar) random states are magical [38][3].

In the limit where a state has small overlap with the code subspace ($|\lambda| \lesssim 0.4$), the previous arguments no longer apply. Instead, there is no clear correlation between output fidelity, success probability, and the amount of magic.

One can also compute the average cost of distilling magic from these typical states. Again, there is no obvious correlation between the total magic of a state and the total cost for generating a magic state with desired T-fidelity $F_T(\rho) > 0.999$.

---

[3] Interestingly, LROM of a larger subsystem has such a smaller variance compared to that of a smaller system. This is also qualitatively consistent with results from Haar random states [38] where the variance of magic scales roughly as $\sim e^{-L}$ where L is the number of sites.

FIG. 8: LROM3 of random states (blue) against output fidelity and success probability. Red circle denotes the location of the encoded T state. On the plane $|\lambda| = 1$, there is clear correlation between LROM and output fidelity.

However, there are states with magic lower than that of a typical state that have slightly lower cost, as shown in the thin narrow bands stretching to the left of the cluster (Figure 9). These are precisely the states that have higher success probabilities which are closer to the encoded T states. Curiously, these states concentrate on the ridges but are not found anywhere in-between, unlike the typical states that contain more magic.
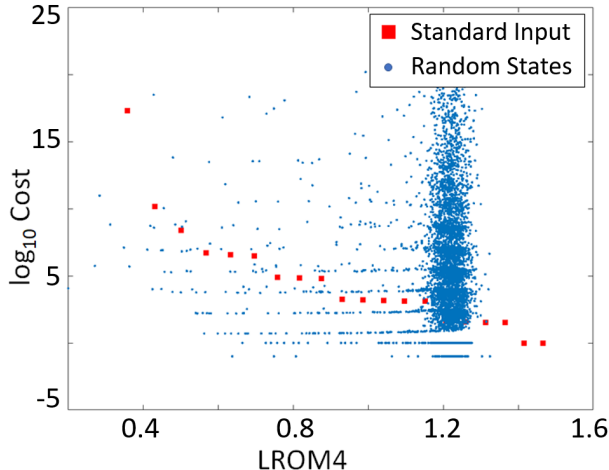
FIG. 9: Plotting LROM4 and cost for random states (blue) and conventional inputs (red). Cost is set to -1 if the state does not exceed the prescribed fidelity $F_T(\rho) > 0.999$ within 15 iterations of the protocol.

V. DISCUSSION

In this work, we have examined how entanglement and non-local magic can impact the outcome of magic state distillation. Using the BK protocol on 5 qubits, we establish that non-local magic is indeed distillable. Furthermore, numerical experiments indicate that non-local magic in these entangled states can also "assist" the usual magic distillation process by lowering the cost of overall distillation. This is outlined conceptually by studying the GHZ T and GHZ phase states. More practically, we have also constructed a spin-squeezing-assisted procedure for noisy ancilla states. It is shown that certain squeezing operations indeed render non-distillable states distillable, which may be more relevant in an experimental setting.

Although the overall impact of many-body magic is still largely unclear in the context of distillation, we build up some generalizable intuition using the 5-qubit code and argue that for the existing distillation protocols based on quantum error correction codes, a typical distillable state should be entangled.

Because of the common assumptions behind magic state distillation, one might have suspected that entanglement generally introduces a source of correlated noise, thereby reducing the output fidelity and success rate of magic distillations. However, this expectation is likely a streetlight effect for only searching perturbatively near the conventional product input states. While it is a valid expectation in certain cases (e.g. Figure 13), it is abundantly clear that by analyzing the entanglement structure and magic distribution of the encoded T state, a typical distillable state is anything but a product state with local magic.

One might also suspect that a larger amount of to-

tal magic would generally lead to better distillation outcomes. This is somewhat suggestive from Figures 2a and 4 for Phase GHZ states. However, this correlation is only manifest when the input lies predominantly in the code subspace. Analysis in Section IV indicates that this intuition has limited applicability when the overlap between the input and the code subspace is small. There it is unclear what properties would positively correlate with distillability.

Since our work suggests that entangled states may assist magic distillations, it opens up a number of directions in the area of many-body magic [34–36, 38]. Thus far we have only focused on examples and numerical analysis, however a more systematic approach is needed to understand many-body magic distillability that is not protocol specific. Generally, we want to understand the features in the distillation protocol and those of the many-body states that will lead to better distillation outcomes. In particular, it opens up the question of what a "typical" distillable magic state should look like if one looks beyond product of identical states. It is also unclear whether existing protocols are optimal for distilling many-body magic, as the input states are drastically different. For example, while it is well-known that bound magic states exist [56] assuming identical inputs, it is not clear to what extent the same conclusion applies when one also considers more general input states. We would also like to understand what kind of protocols are near-optimal in distilling magic from condensed matter systems such as the low energy states of a spin-chain where magic is abundant.

It is possible that holographic quantum error correction codes may also be useful in this regard for distilling multi-scale non-local magic from CFTs due to its multi-scale, self-similar structure. It is pointed out recently [57] that holographic codes with good complementary recovery properties have difficulties implementing transversal non-Clifford gates. Nevertheless, it is still entirely possible to construct holographic codes in general that do have transversal T gates. For instance, this can be done by using codes with transversal non-Clifford gates [58, 59] such as the $[[15, 1, 3]]$ CSS code [60, 61] as the base tensor, which supports a transversal logical $\bar{T}$ gate. Using the general methodology outlined in [62] (and more comprehensively in [63]) for building tensor networks on 2d hyperbolic space with regular tessellations, one can produce a holographic code constructed using the $\{15, q \geq 4\}$ tiling. This results in an $[[n, k]]$ CSS code that admits a transversal $\bar{T}^{\otimes k}$ gate[4]. The same methodology also applies to more general $[[2^{n-1}, 1, 3]]$ punctured Reed-Muller codes [64] in a $\{2^{n-1}, q \geq 4\}$ tessellation. The resulting stabilizer code would again admit a transversal non-Clifford gate. It should also be noted that transversal

—————

[4] Note that the tensors or codes used to build the holographic code can be the same up to local Clifford equivalences.

non-Clifford gates are not always necessary for distillation. Indeed, the 5 qubit code does not admit a transversal T gate, however, the transversal B gate enables distillation of T states. A holographic code that serves a similar purpose is the $[[n, k]]$ stabilizer code known as the HaPPY pentagon code [65]. There the logical $\bar{B}^{\otimes k}$ is transversal and a protocol similar to BK may be sufficient for T state distillation.

ACKNOWLEDGEMENTS

[1] Bryan Eastin and Emanuel Knill. Restrictions on transversal encoded quantum gate sets. *Physical Review Letters*, 102(11):110502, 2009.

[2] Bei Zeng, Andrew Cross, and Isaac L. Chuang. Transversality versus universality for additive quantum codes. *IEEE Transactions on Information Theory*, 57(9):6272–6284, 2011.

[3] Sergey Bravyi and Robert König. Classification of topologically protected gates for local stabilizer codes. *Phys. Rev. Lett.*, 110:170503, Apr 2013.

[4] Fernando Pastawski and Beni Yoshida. Fault-tolerant logical gates in quantum error-correcting codes. *Phys. Rev. A*, 91:012305, Jan 2015.

[5] Tomas Jochym-O'Connor, Aleksander Kubica, and Theodore J. Yoder. Disjointness of stabilizer codes and limitations on fault-tolerant logical gates. *Phys. Rev. X*, 8:021047, May 2018.

[6] Sergey Bravyi and Alexei Kitaev. Universal quantum computation with ideal clifford gates and noisy ancillas. *Physical Review A*, 71(2), Feb 2005.

[7] E. Knill. Fault-Tolerant Postselected Quantum Computation: Schemes. *arXiv e-prints*, pages quant–ph/0402171, February 2004.

[8] Adam Paetznick and Ben W. Reichardt. Universal Fault-Tolerant Quantum Computation with Only Transversal Gates and Error Correction. *Phys. Rev. Lett.*, 111(9):090505, August 2013.

[9] Theodore J. Yoder. Universal fault-tolerant quantum computation with Bacon-Shor codes. *arXiv e-prints*, page arXiv:1705.01686, May 2017.

[10] Theodore J. Yoder, Ryuji Takagi, and Isaac L. Chuang. Universal Fault-Tolerant Gates on Concatenated Stabilizer Codes. *Physical Review X*, 6(3):031039, July 2016.

[11] Tomas Jochym-O'Connor and Stephen D. Bartlett. Stacked codes: Universal fault-tolerant quantum computation in a two-dimensional layout. *Phys. Rev. A*, 93:022323, Feb 2016.

[12] Sergey Bravyi and Andrew Cross. Doubled Color Codes. *arXiv e-prints*, page arXiv:1509.03239, September 2015.

[13] Cody Jones, Peter Brooks, and Jim Harrington. Gauge color codes in two dimensions. *Phys. Rev. A*, 93:052332, May 2016.

[14] Benjamin J. Brown. A fault-tolerant non-Clifford gate for the surface code in two dimensions. *Science Advances*, 6(21):eaay4929, May 2020.

[15] Daniel Gottesman and Isaac L. Chuang. Demonstrating the viability of universal quantum computation using teleportation and single-qubit operations. *Nature (London)*, 402(6760):390–393, November 1999.

[16] Michael A. Nielsen. Quantum computation by measurement and quantum memory. *Physics Letters A*, 308(2-3):96–100, February 2003.

[17] D. W. Leung. Two-qubit Projective Measurements are Universal for Quantum Computation. *arXiv e-prints*, pages quant–ph/0111122, November 2001.

[18] David Gosset, Vadym Kliuchnikov, Michele Mosca, and Vincent Russo. An algorithm for the T-count. *arXiv e-prints*, page arXiv:1308.4134, August 2013.

[19] Matthew Amy, Dmitri Maslov, Michele Mosca, and Martin Roetteler. A meet-in-the-middle algorithm for fast synthesis of depth-optimal quantum circuits. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 32(6):818–830, 2013.

[20] Yunseong Nam, Neil J. Ross, Yuan Su, Andrew M. Childs, and Dmitri Maslov. Automated optimization of large quantum circuits with continuous parameters. *npj Quantum Information*, 4:23, May 2018.

[21] Luke Heyfron and Earl T. Campbell. An Efficient Quantum Compiler that reduces T count. *arXiv e-prints*, page arXiv:1712.01557, December 2017.

[22] Qingfeng Wang, Ming Li, Christopher Monroe, and Yunseong Nam. Resource-Optimized Fermionic Local-Hamiltonian Simulation on Quantum Computer for Quantum Chemistry. *arXiv e-prints*, page arXiv:2004.04151, April 2020.

[23] Héctor Bombín. Gauge color codes: optimal transversal gates and gauge fixing in topological stabilizer codes. *New Journal of Physics*, 17(8):083002, aug 2015.

[24] Aleksander Kubica and Michael E. Beverland. Universal transversal gates with color codes: A simplified approach. *Phys. Rev. A*, 91:032330, Mar 2015.

[25] Michael E. Beverland, Aleksander Kubica, and Krysta M. Svore. Cost of universality: A comparative study of the overhead of state distillation and code

switching with color codes. *PRX Quantum*, 2:020341, Jun 2021.

[26] Aleksander Kubica, Beni Yoshida, and Fernando Pastawski. Unfolding the color code. *New Journal of Physics*, 17(8):083026, aug 2015.

[27] Hector Bombin. Transversal gates and error propagation in 3d topological codes, 2018.

[28] Hector Bombin. 2d quantum computation with 3d topological codes, 2018.

[29] Adam M. Meier, Bryan Eastin, and Emanuel Knill. Magic-state distillation with the four-qubit code. *arXiv e-prints*, page arXiv:1204.4221, April 2012.

[30] Sergey Bravyi and Jeongwan Haah. Magic-state distillation with low overhead. *Phys. Rev. A*, 86(5):052329, November 2012.

[31] Cody Jones. Low-overhead constructions for the fault-tolerant toffoli gate. *Phys. Rev. A*, 87:022328, Feb 2013.

[32] Cody Jones. Multilevel distillation of magic states for quantum computing. *Phys. Rev. A*, 87:042305, Apr 2013.

[33] Jeongwan Haah, Matthew B. Hastings, D. Poulin, and D. Wecker. Magic State Distillation with Low Space Overhead and Optimal Asymptotic Input Count. *arXiv e-prints*, page arXiv:1703.07847, March 2017.

[34] S. Sarkar, C. Mukhopadhyay, and A. Bayat. Characterization of an operational quantum resource in a critical many-body system. *New Journal of Physics*, 22(8):083077, August 2020.

[35] Christopher David White, ChunJun Cao, and Brian Swingle. Conformal field theories are magical. *Phys. Rev. B*, 103(7):075145, 2021.

[36] Zi-Wen Liu and Andreas Winter. Many-body quantum magic. *arXiv e-prints*, page arXiv:2010.13817, October 2020.

[37] Tyler D. Ellison, Kohtaro Kato, Zi-Wen Liu, and Timothy H. Hsieh. Symmetry-protected sign problem and magic in quantum phases of matter. *arXiv e-prints*, page arXiv:2010.13803, October 2020.

[38] Christopher David White and Justin H. Wilson. Mana in Haar-random states. *arXiv e-prints*, page arXiv:2011.13937, November 2020.

[39] Daniel Gottesman. The Heisenberg Representation of Quantum Computers. *arXiv e-prints*, pages quant–ph/9807006, July 1998.

[40] A. Yu. Kitaev. Fault-tolerant quantum computation by anyons. *Annals of Physics*, 303(1):2–30, January 2003.

[41] Fernando Pastawski, Beni Yoshida, Daniel Harlow, and John Preskill. Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence. *Journal of High Energy Physics*, 2015:149, June 2015.

[42] Daniel Gottesman. Theory of fault-tolerant quantum computation. *Phys. Rev. A*, 57:127–137, Jan 1998.

[43] Victor Veitch, S A Hamed Mousavian, Daniel Gottesman, and Joseph Emerson. The resource theory of stabilizer quantum computation. *New Journal of Physics*, 16(1):013009, 2014.

[44] Mark Howard and Earl Campbell. Application of a resource theory for magic states to fault-tolerant quantum computing. *Physical Review Letters*, 118(9):090501, 2017.

[45] Xin Wang, Mark M. Wilde, and Yuan Su. Efficiently computable bounds for magic state distillation. *arXiv e-prints*, page arXiv:1812.10145, December 2018.

[46] Michael Beverland, Earl Campbell, Mark Howard, and Vadym Kliuchnikov. Lower bounds on the non-Clifford resources for quantum computations. *Quantum Science and Technology*, 5(3):035009, July 2020.

[47] Raymond Laflamme, Cesar Miquel, Juan Pablo Paz, and Wojciech Hubert Zurek. Perfect quantum error correcting code. *Phys. Rev. Lett.*, 77:198–201, Jul 1996.

[48] Daniel Gottesman. *Stabilizer codes and quantum error correction*. PhD thesis, California Institute of Technology, January 1997.

[49] Masahiro Kitagawa and Masahito Ueda. Squeezed spin states. *Phys. Rev. A*, 47:5138–5143, Jun 1993.

[50] Monika H. Schleier-Smith, Ian D. Leroux, and Vladan Vuletić. Squeezing the collective spin of a dilute atomic ensemble by cavity feedback. *Phys. Rev. A*, 81(2):021804, February 2010.

[51] Ian D. Leroux, Monika H. Schleier-Smith, and Vladan Vuleti. Implementation of cavity squeezing of a collective atomic spin. *Physical Review Letters*, 104(7), Feb 2010.

[52] Takafumi Ono, Javier Sabines-Chesterking, Hugo Cable, Jeremy L O'Brien, and Jonathan C F Matthews. Optical implementation of spin squeezing. *New Journal of Physics*, 19(5):053005, may 2017.

[53] J. Hald, J. L. Sørensen, C. Schori, and E. S. Polzik. Spin squeezed atoms: A macroscopic entangled ensemble created by light. *Phys. Rev. Lett.*, 83:1319–1322, Aug 1999.

[54] R. J. Sewell, M. Koschorreck, M. Napolitano, B. Dubost, N. Behbood, and M. W. Mitchell. Magnetic sensitivity beyond the projection noise limit by spin squeezing. *Phys. Rev. Lett.*, 109:253605, Dec 2012.

[55] Géza Tóth, Christian Knapp, Otfried Gühne, and Hans J. Briegel. Spin squeezing and entanglement. *Phys. Rev. A*, 79(4):042334, April 2009.

[56] Earl T. Campbell and Dan E. Browne. Bound States for Magic State Distillation in Fault-Tolerant Quantum Computation. *Phys. Rev. Lett.*, 104(3):030503, January 2010.

[57] Sam Cree, Kfir Dolev, Vladimir Calvera, and Dominic J. Williamson. Fault-tolerant logical gates in holographic stabilizer codes are severely restricted. *arXiv e-prints*, page arXiv:2103.13404, March 2021.

[58] Narayanan Rengaswamy, Robert Calderbank, Michael Newman, and Henry D. Pfister. On Optimality of CSS Codes for Transversal T. *arXiv e-prints*, page arXiv:1910.09333, October 2019.

[59] Bei Zeng, Andrew Cross, and Isaac L. Chuang. Transversality versus Universality for Additive Quantum Codes. *arXiv e-prints*, page arXiv:0706.1382, June 2007.

[60] E. Knill, R. Laflamme, and W. Zurek. Threshold Accuracy for Quantum Computation. *arXiv e-prints*, pages quant–ph/9610011, October 1996.

[61] Jonas T. Anderson, Guillaume Duclos-Cianci, and David Poulin. Fault-Tolerant Conversion between the Steane and Reed-Muller Quantum Codes. *Phys. Rev. Lett.*, 113(8):080501, August 2014.

[62] ChunJun Cao, Jason Pollack, and Yixu Wang. Hyper-Invariant MERA: Approximate Holographic Error Correction Codes with Power-Law Correlations. *arXiv e-prints*, page arXiv:2103.08631, March 2021.

[63] ChunJun Cao and Brad Lackey. Building Quantum Error Correction Codes from Tensor Networks. *in prep*, 2021.

[64] Bei Zeng, Hyeyoun Chung, Andrew W. Cross, and Isaac L. Chuang. Local unitary versus local Clifford equivalence of stabilizer and graph states. *Phys. Rev.*

A, 75(3):032325, March 2007.

[65] Fernando Pastawski, Beni Yoshida, Daniel Harlow, and John Preskill. Holographic quantum error-correcting codes: toy models for the bulk/boundary correspondence. *Journal of High Energy Physics*, 2015:149, June 2015.

Appendix A: Distillability of various Phase-GHZ and GHZ-T states

In this appendix we examine the distillability of states similar to those presented in III A, the Phase GHZ and GHZ T states. As a reminder, these were n qubit states that took the following form.

$$|PG(\alpha)\rangle_n = \alpha |0\rangle^{\otimes n} + e^{i\phi}\beta |1\rangle^{\otimes n} \ , \ |GT(\alpha)\rangle_n = \alpha |T_0\rangle^{\otimes n} + e^{i\phi}\beta | \tag{A.1}$$

Because the Bravyi-Kitaev protocol takes as input a 5 qubit state, it was a natural choice to pick $n = 5$. Not only does this make each of the reduced density matrices the same, the input state is symmetric between the five qubits. Here, we now consider a more broad construction where we make a 5 qubit state out of a combination of smaller $|PG\rangle_n$ ($|GT\rangle_n$) states. This may lead to easier experimental implementation if it is much easier to generate, say, $|PG\rangle_2$ and $|PG\rangle_3$ than $|PG\rangle_5$. On the theoretical side, it is interesting to ask whether the presence of more magic yields better distillation outcomes. This is because $|PG\rangle_n$ has the same magic regardless of n, so $|PG\rangle_2 |PG\rangle_3$ would have twice as much global magic as $|PG\rangle_5$ for the same α.

We consider two primary variations. The first is where we comprise the 5 qubit state of multiple $|PG\rangle_k$ states such that $\sum k = 5$. We make a few observations. The first is that adding more total magic does not seem to enhance the ability to distill. This lines up with the expectations laid out in Section IV. Interestingly, all of these states remained distillable for generic α.

The second variation allows for some of the qubits to be in the $|0\rangle$ state in the computational basis, a consideration for experimental setups where the $|0\rangle$ state is effectively just resetting a clean qubit. Clearly, these qubits do not add to the available magic. The results are displayed in Figure 11. With the exception of Phase-GHZ states with even numbers, these states also have a wide range of distillability.

Appendix B: More on Spin Squeezing and Distillation

Although a global entangling unitary like the one defined in Section III B can improve the distillation outcome, this improvement will depend on a number of factors and is not universal. Here we analyze how different initial states modified the same squeezing unitary $U(t)$ may impact the distillation outcome. Then we examine the effect of a different squeezing unitary $V(t)$. We find
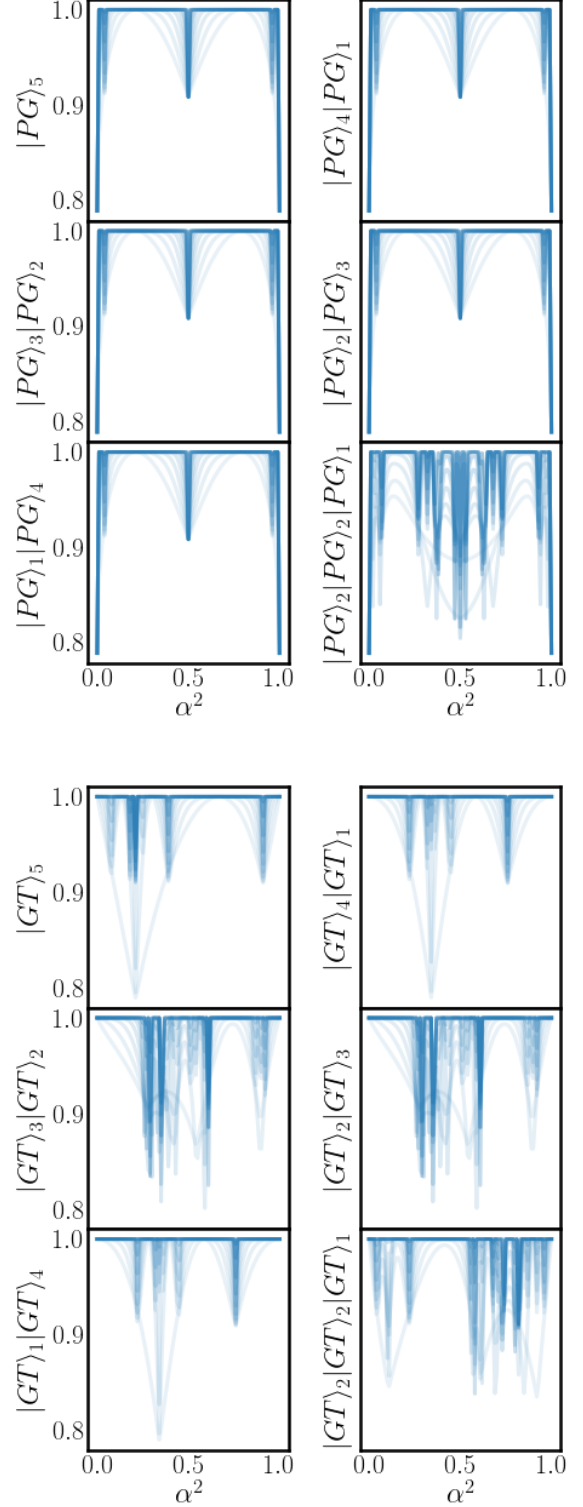

FIG. 10: Distillability of composite $|PG\rangle_k$ and $|GT\rangle_k$ states, where α is taken to uniform across individual components. We plot the magic fidelity of the qubit upon a successful distillation round with darker curves corresponding to later rounds. The darkest line is the resulting T-fidelity after 10 rounds. Because $|PG\rangle_k$ and $|GT\rangle_k$ contain the same amount of magic for all values of k, composite 5 qubit states with more individual states have more total magic in the system. We find that the presence of more total magic does not generally enhance distillability for the BK protocol.
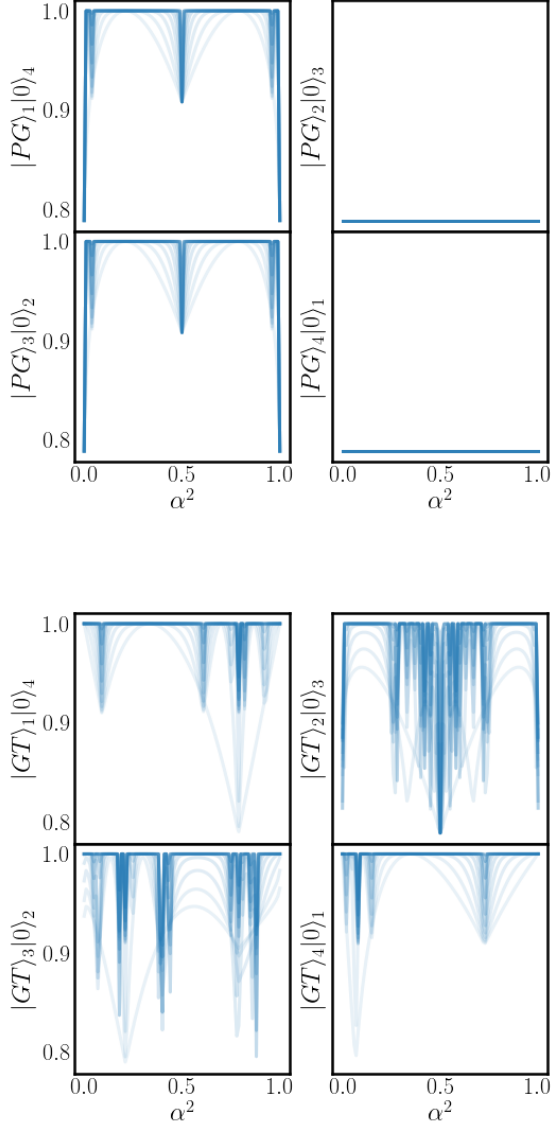
FIG. 11: Distillability of composite $|PG\rangle_k$ and $|GT\rangle_k$ states, with some qubits in the $|0\rangle$ state. A notable finding is that one can get similar distillability without having to entangle the full 5 qubit state.

that improved distillation is unlikely to be connected to squeezing and overall amount of entanglement. However, its success may be related to the non-local magic injected during the squeezing process.

First of all, squeezing alone is not the key to improving the distillation outcome — clearly a state needs to contain enough magic to be distillable at all. Although squeezing in general can add magic to a stabilizer state, a slight squeezing of an initial state without spin misalignment does not improve distillation outcome. Furthermore, states that are unsqueezed under the same unitary

$U(t)$ (eqn III.5) also sees improvement in distillation. For example, distillation outcome is just as good, if not better, for initial states where $\bar{\theta}_i = 0$ (Figure 12). However $U(t)$ does not squeeze $|00000\rangle$, which is approximately the initial state we have.
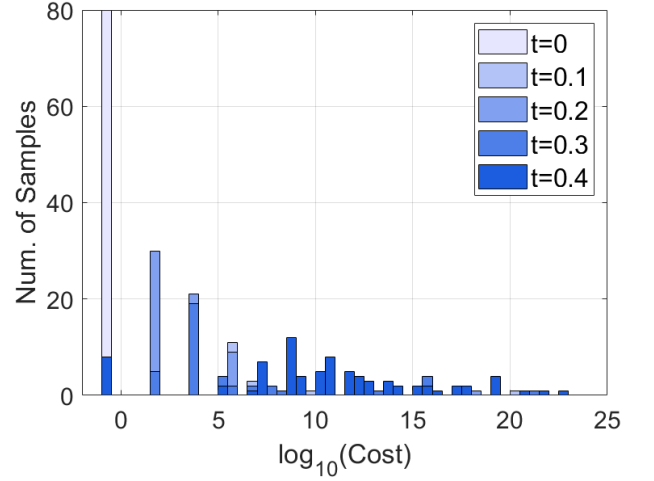

FIG. 12: BK distillation cost applied to states after one-axis twisting $|00000\rangle$ initial states with random misalignment. Here $\theta_{\max} = 0.05$.

The same global unitary action $U(t)$ also need not improve distillation for other types of initial states. Sometimes it acts as coherent noise to the system, which impedes distillation. For example, consider the scenario where the individual qubit input states are sufficiently close to $|T\rangle$ such that $\overline{\beta_i} = \theta_T = \arccos(1/\sqrt{3})/2$, $\phi = \pi/4$

$$|\tau_0^i\rangle = \cos(\beta_i)|0\rangle + \exp(i\phi)\sin(\beta_i)|1\rangle \qquad (B.1)$$

$$|\tau_1^i\rangle = -\sin(\beta_i)|0\rangle + \exp(i\phi)\cos(\beta_i)|1\rangle, \qquad (B.2)$$

and the initial state is given by

$$\rho_{\mathrm{ini}} = \bigotimes_i (p_i|\tau_0^i\rangle\langle\tau_0^i| + (1-p_i)|\tau_1^i\rangle\langle\tau_1^i|). \qquad (B.3)$$

As usual, β_i is drawn uniformly random for an interval $[\overline{\beta_i} - \beta_{\max}, \overline{\beta_i} + \beta_{\max}]$. The probability p_i is also drawn uniformly for an interval $[\overline{p_i} - p_{\max}, \overline{p_i} + p_{\max}]$. These are precisely the target magic T states when $p_{\max}, \beta_{\max} = 0$. Then squeezing now acts as a source of noise that renders distillation less efficient (Figures 13).

On the other hand, if the initial state is close to $|++++ +\rangle$ or $|00000\rangle$ but mixed, then the improvement in the distillation cost also quickly diminishes in these examples as a one increases $p_{\max}$. For instance, consider states of the form (B.3) but with $\phi = 0, \overline{\beta_i} = \pi/4$, the distillation cost is shown in Figure 14 where the number of undistillable states can increase sharply even when the purity is still relatively high.

Now let us briefly examine the same process but with a different global unitary applied to the $\bar{\theta}_i = \pi/4$ initial
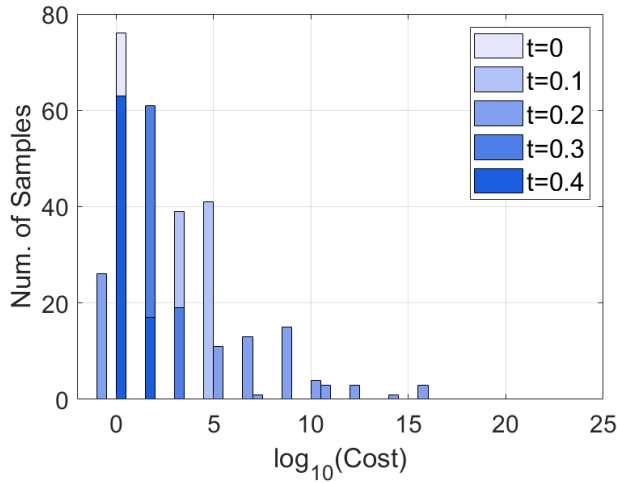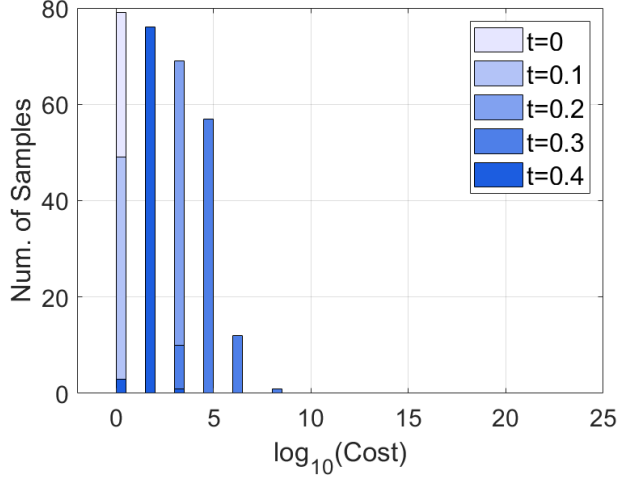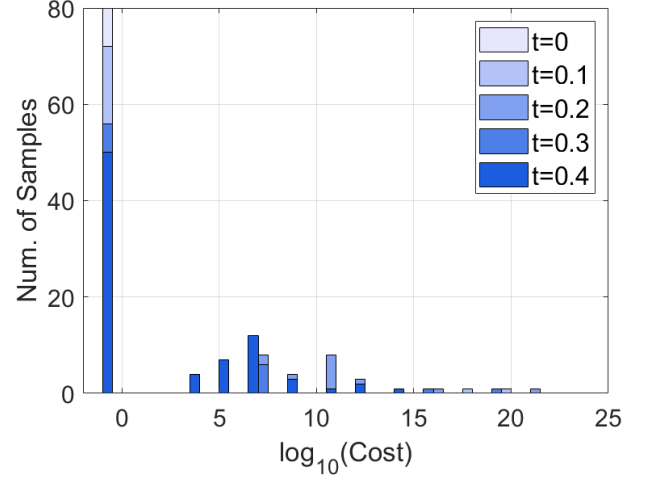
FIG. 14: One-axis twisting applied to initial mixed states that are close to a product stabilizer state. $\theta_{\max} = 0.05$ and $p_{\max} = 0.02$.
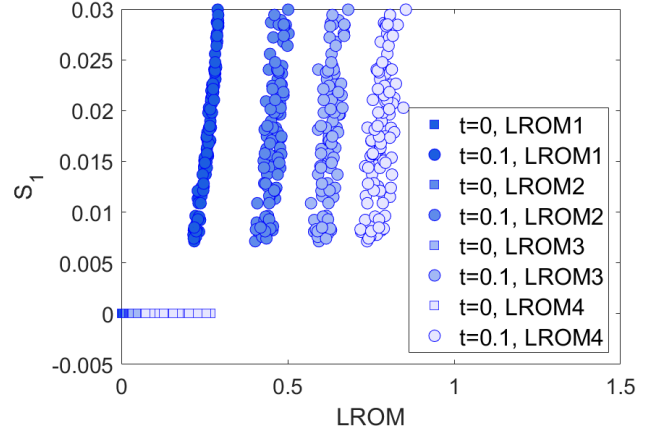

FIG. 13: (a) Squeezing the tensor product of five noisy T states via one-axis twisting. (b) Squeezing the tensor product of give noisy T states via two-axis countertwisting. Here $\beta_{\max} = 0.05, p_{\max} = 0.1$ for both plots. Squeezing increases the overall cost of distillation in both cases.

FIG. 15: LROM and single site entanglement of the state with (colored disks) and without squeezing (colored squares).

state (III.3). Consider a different squeezing procedure using two-axis countertwisting [49] with $V(t) = \exp(-itH)$ where

$$H = \frac{1}{2i}(S_+^2 - S_-^2), \quad S_\pm = S_x \pm iS_y. \tag{B.4}$$

Then squeezing produces a similar level entanglement as the previous one-axis twisting procedure for similar values of t. It also increases the total magic in the state. However, no visible enhancement for the distillation can be found (Figure 15).

The lack of improvements here may be attributed to the amount of non-local magic that is added to the system. Note that a pure state with non-local magic is most certainly entangled, but an entangled state need not contain non-local magic. While there is a slight addition

to non-local magic using the one-axis twisting unitary, there is virtually no change in the two-axis countertwisting scheme (Figure 16). We estimate non-local magic using

$$\mathcal{LROM}(1:2) = \ln(\mathcal{R}(\rho_{12})) - [\ln(\mathcal{R}(\rho_1)) + \ln(\mathcal{R}(\rho_2))], \tag{B.5}$$

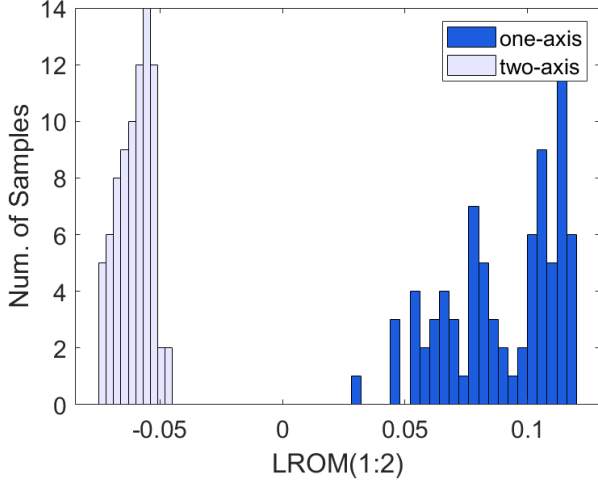where ρ_{12} is an arbitrarily chosen two-site subsystem.

FIG. 16: Histogram contrasting two-body nonlocal magic $\mathcal{LROM}(1:2)$ after the one-axis twisting and the two-axis countertwisting options that result in similar levels of entanglement ($S_1 \sim 0.08$ and $S_1 \sim 0.02$ respectively). Slight negativity can be attributed to the non-additivity of LROM.