

# A sharp phase transition in linear cross-entropy benchmarking

Brayden Ware,<sup>1,2</sup> Abhinav Deshpande,<sup>3,\*</sup> Dominik Hangleiter,<sup>2</sup> Pradeep Niroula,<sup>1,2</sup> Bill Fefferman,<sup>4</sup> Alexey V. Gorshkov,<sup>1,2</sup> and Michael J. Gullans<sup>2</sup>

<sup>1</sup>*Joint Quantum Institute, NIST/University of Maryland, College Park, MD 20742, USA*

<sup>2</sup>*Joint Center for Quantum Information and Computer Science,  
NIST/University of Maryland, College Park, MD 20742, USA*

<sup>3</sup>*Institute for Quantum Information and Matter,  
California Institute of Technology, Pasadena, CA 91125, USA*

<sup>4</sup>*Department of Computer Science, University of Chicago, Chicago, Illinois 60637, USA*

(Dated: May 10, 2023)

Demonstrations of quantum computational advantage and benchmarks of quantum processors via quantum random circuit sampling are based on evaluating the linear cross-entropy benchmark (XEB). A key question in the theory of XEB is whether it approximates the fidelity of the quantum state preparation. Previous works have shown that the XEB generically approximates the fidelity in a regime where the noise rate per qudit  $\varepsilon$  satisfies  $\varepsilon N \ll 1$  for a system of  $N$  qudits and that this approximation breaks down at large noise rates. Here, we show that the breakdown of XEB as a fidelity proxy occurs as a sharp phase transition at a critical value of  $\varepsilon N$  that depends on the circuit architecture and properties of the two-qubit gates, including in particular their entangling power. We study the phase transition using a mapping of average two-copy quantities to statistical mechanics models in random quantum circuit architectures with full or one-dimensional connectivity. We explain the phase transition behavior in terms of spectral properties of the transfer matrix of the statistical mechanics model and identify two-qubit gate sets that exhibit the largest noise robustness.

## I. INTRODUCTION

Near-term quantum devices consist of dozens of imperfect qubits, with error rates far too large to outperform classical computers in tasks like factoring large integers. Instead, several recent attempts to demonstrate the computational power of noisy quantum devices have focused on *quantum random circuit sampling (RCS)* [1–5]. In RCS, the task is to repeatedly produce and measure a single, highly entangled many-body wavefunction  $\psi(x)$ , generating a sequence of outcomes  $\{x_i\}$  that ideally occur with probabilities  $p(x) = |\psi(x)|^2$ . As generating complex entangled wavefunctions is a core capability of quantum computers that cannot be replicated by classical computers (see e.g., [6–9]), RCS is a natural task for testing prototype quantum devices.

Of course, error prone qubits do not allow for precisely reproducing the same wavefunction repeatedly, or for precisely sampling from the correct distribution  $p(x)$ . To compare the performance of a particular quantum device with disparate quantum devices or classical computers at this task, benchmarks—quantitative measures of success—are needed. A straightforward but strict benchmark is the fidelity between the target wavefunction and the output wavefunction. However, measuring fidelity in large quantum systems is intractable on current hardware since the best available schemes require the application of deep noise-free random Clifford circuits for shadow fidelity estimation [10, 11]. Moreover, when comparing with classical algorithms, the fidelity is only meaningful

for certain algorithms that also produce a classical representation of the wave function.

In place of fidelity, the *linear cross-entropy benchmark (XEB)* has been used as a measure of success in RCS [1, 2, 12, 13] and beyond [14–16]. The XEB

$$\chi = 2^N \sum_x \overline{p(x)q(x)} - 1 \quad (1)$$

measures the correlation between the ideal distribution  $p(x)$  and the actual distribution  $q(x)$  of sample outcomes, averaged over the random circuit instances. Importantly, we expect that the XEB can be sample-efficiently estimated just from the produced samples  $x_i$  of a few random quantum circuits.<sup>1</sup> This sample efficiency is crucial for the usefulness of the benchmark. However, the ideal probabilities  $p(x)$  of the measured samples must still be known to estimate the benchmark, and these must be computed via classical algorithms with exponential runtime.

There are two crucial theoretical questions about XEB for understanding how well it serves its purpose as a benchmark for RCS. First, one must understand what scores are, in principle, achievable by noisy quantum devices. Secondly, to use XEB as a benchmark of computational power, one must understand what scores are achievable by competing classical devices, running arbitrary algorithms.

In this work, we address the first question by phrasing it as a comparison between XEB and fidelity: can high XEB scores be achieved without producing states

\* Currently at IBM Quantum, Almaden Research Center, San Jose, California 95120, USA

<sup>1</sup> See [5, Sec. V.B.3] for the argument.

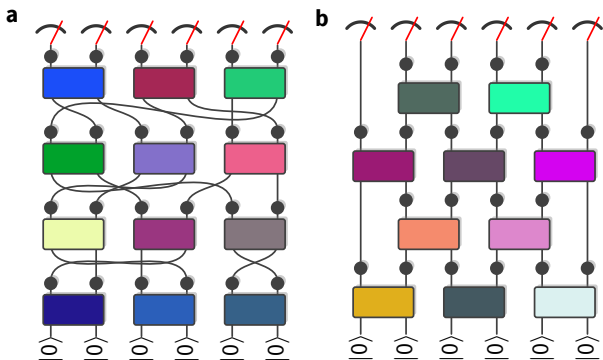


FIG. 1: Circuit geometries considered in this paper. The evolution consists of layers of random two-qudit gates chosen either from a Haar random distribution or according to Eq. (2) and single-qudit noise (shown as black circles) specified by an arbitrary (but fixed) noise channel  $\mathcal{N}$  applied after each gate layer. (a) In the all-to-all geometry, the qubits are permuted randomly between gate layers. (b) In the 1D geometry, alternating layers of gates are applied in a brickwork fashion with open boundary conditions.

with high fidelity to the target? This question has been studied previously in work of Dalzell *et al.* [12] and Gao *et al.* [17], who come to the conclusion that XEB closely approximates fidelity in a regime of significantly weak, spatially homogenous, and temporally independent noise. For an  $N$ -qubit system with local noise rate  $\varepsilon$ , this regime is characterized by  $\varepsilon N \ll 1$ . However, they have also shown that violations of these assumptions break the correlation between XEB and fidelity. Our work addresses the exact dependence of the match between XEB and fidelity as the noise is increased. We show that this match breaks suddenly as noise is increased, occurring as a sharp phase transition in the behavior of XEB at a critical value of  $\varepsilon N$ . This remains true for variants of the XEB [12] which, in the low-noise limit, track the fidelity in a wider parameter regime. In addition to explaining the mechanism behind the transition, we investigate how the critical noise value depends on the gate set and circuit geometry used.

In order to analyze the noise dependence of the XEB, we make use of a mapping between average properties of two copies of the quantum wavefunction—which includes both the fidelity and the XEB—to observables of a statistical model [12, 17–23]. In this model, the average dynamics of the quantum system can be analyzed in terms of a transfer matrix on a  $2^N$ -dimensional configuration space, and the behavior of the XEB can be explained via spectral properties of this transfer matrix. We analyze the model using simulations in two extremal circuit architectures, namely a fully connected (all-to-all) architecture, and a one-dimensional architecture (see Fig. 1), and find the same qualitative behavior. This suggests that the phase transition in the XEB is in fact

a universal feature of random quantum circuits in any architecture.

In the all-to-all model we exploit the permutation symmetry of the qudits to exactly solve the corresponding statistical mechanical model with large numbers of qudits in terms of an exponentially reduced state space. To analyze one-dimensional circuits, we use matrix-product state techniques for computing the dynamics and spectra of the statistical model.

In the following, we will first introduce the statistical mechanical model for the two architectures (Section II), then show our main results on the sharp phase transition of the XEB in Section III for the paradigmatic gate set of Haar-random gates, and finally discuss the gate-set dependence of the location of the transition in Section IV.

## II. BACKGROUND

### A. Random circuit observables

We study noisy random circuits acting on  $N$  qudits of dimension  $q$ . The circuits consist of two-qudit gates generated in one of two ways: Haar-random gates  $G$  drawn from  $U(q^2)$ , or by using a specific two-qudit gate  $g$  and Haar-random single-qudit gates  $U_1, U_2, U_3, U_4 \in U(q)$ :

$$G = (U_1 \otimes U_2)g (U_3 \otimes U_4). \quad (2)$$

The circuits we study (illustrated in Fig. 1) consist of  $d$  layers of  $N/2$  gates, each of which connects either random disjoint pairs of qudits (the ‘all-to-all architecture’, see Section II C), or neighboring qudits of a chain, alternating in a brickwork pattern (the 1D architecture, see Section II D). After each layer of gates, each qudit experiences noise as specified by an arbitrary single-qudit noise channel  $\mathcal{N}$ . We initialize the circuit with the product state  $(|0\rangle\langle 0|)^{\otimes N}$ .

Our reference quantity is the average fidelity  $F = \mathbb{E}_{\mathbf{U}} \text{Tr}[\rho_{\mathbf{U}}(\mathcal{N})\rho_{\mathbf{U}}(0)]$  between the noisy output state  $\rho_{\mathbf{U}}(\mathcal{N})$  and the corresponding noiseless pure state  $\rho_{\mathbf{U}}(0)$ , where the average is taken over the random circuit instance  $\mathbf{U}$ . The average fidelity  $F$  takes the value 1 in the ideal limit of no noise, and the value  $q^{-N}$  when  $\rho_{\mathbf{U}}(\mathcal{N})$  and  $\rho_{\mathbf{U}}(0)$  are totally uncorrelated. A convenient reformulation of the fidelity,

$$F = \text{Tr}(\mathcal{S}^{\otimes N} \mathbb{E}_{\mathbf{U}} [\rho_{\mathbf{U}}(0) \otimes \rho_{\mathbf{U}}(\mathcal{N})]), \quad (3)$$

considers two copies of the same circuit, one of which is evolved with a noisy evolution and the other with the noiseless evolution, followed by the  $\mathcal{S}^{\otimes N}$  operator that swaps corresponding qudits between the two copies. The second important quantity we will consider in this paper is the *linear cross entropy benchmark* (XEB)

$$\chi = q^N \mathbb{E}_{\mathbf{U}} \sum_{\mathbf{x}} \langle \mathbf{x} | \rho_{\mathbf{U}}(\mathcal{N}) | \mathbf{x} \rangle \langle \mathbf{x} | \rho_{\mathbf{U}}(0) | \mathbf{x} \rangle - 1$$

which has been proposed as a classical estimator of the average fidelity for random quantum circuits. It can be written as a two-copy expectation value in a similar way:

$$\chi = \text{Tr} \left( [q^N \mathcal{P}^{\otimes N} - 1] \mathbb{E}_{\mathbf{U}} [\rho_{\mathbf{U}}(\mathcal{N}) \otimes \rho_{\mathbf{U}}(0)] \right), \quad (4)$$

where

$$\mathcal{P} = \sum_{x=0}^{q-1} |x\rangle \langle x| \otimes |x\rangle \langle x|. \quad (5)$$

Since the values of  $\chi$  for low-depth random circuits can be much larger than one, the variant quantity  $\chi_B = \chi / (q^N Z - 1)$  has been proposed by Dalzell *et al.* [12] as a better estimator of the fidelity<sup>2</sup>. Here,

$$\begin{aligned} Z &= \mathbb{E}_{\mathbf{U}} \sum_{\mathbf{x}} \langle \mathbf{x} | \rho_{\mathbf{U}}(0) | \mathbf{x} \rangle \langle \mathbf{x} | \rho_{\mathbf{U}}(0) | \mathbf{x} \rangle \\ &= \text{Tr}(\mathcal{P}^{\otimes N} \mathbb{E}_{\mathbf{U}} [\rho_{\mathbf{U}}(0) \otimes \rho_{\mathbf{U}}(0)]), \end{aligned} \quad (6)$$

is the collision probability of the noiseless circuit. Like the fidelity,  $\chi_B$  takes values between 0 and 1, and attains its maximal value in the noiseless case. As we will see below, which precise variant of XEB is used will not affect the behavior of the phase transition.

We now introduce the statistical model [12] used to analyze the noise-dependence of these two-copy average quantities.

## B. Statistical model

For an arbitrary pair of density matrices  $\rho_A, \rho_B$  on  $N$  qudits of dimension  $q$ , consider the Haar-averaged two-copy density matrix

$$\overline{\rho^{(2)}} = \mathbb{E}_{\mathbf{V}} (\mathbf{V} \rho_A \mathbf{V}^\dagger) \otimes (\mathbf{V} \rho_B \mathbf{V}^\dagger), \quad (7)$$

where  $\mathbf{V}$  is a product of single-qudit Haar random rotations

$$\mathbf{V} = \prod_{i=1}^N V_i. \quad (8)$$

Such an average has the result of projecting arbitrary two-copy density matrices  $\rho_A \otimes \rho_B$  into the symmetric subspace, i.e., the space invariant under arbitrary rotations  $\mathbf{V} \otimes \mathbf{V}$ . The symmetric subspace is spanned by tensor products of  $I/q^2$  and  $\mathcal{S}/q$ , where  $I$  is the two-qudit identity operator and  $\mathcal{S}$  is the operator that swaps two corresponding qudits in the two copies. Thus, the average can be expressed as

$$\overline{\rho^{(2)}} = \sum_{\sigma \in \{0,1\}^{\otimes N}} p(\sigma) \bigotimes_{i=1}^N \left( \frac{I}{q^2} \right)^{1-\sigma_i} \left( \frac{\mathcal{S}}{q} \right)^{\sigma_i}. \quad (9)$$

We reinterpret  $\overline{\rho^{(2)}}$  as a vector  $|\rho\rangle$  in a many-body state space, using basis states  $|\sigma\rangle$  with  $\sigma \in \{0,1\}^{\otimes N}$  corresponding to the terms of the above sum, so that

$$|\rho\rangle = \sum_{\sigma} p(\sigma) |\sigma\rangle. \quad (10)$$

Using this representation, the fidelity of  $\rho_B$  with respect to  $\rho_A$  and their linear cross-entropy can be expressed as inner products with fixed vectors:  $F = \text{Tr}(\mathbb{S} \overline{\rho^{(2)}}) = \langle \mathbb{S} | \rho \rangle$  and  $X = \langle \mathbb{P} | \rho \rangle$ , where  $\mathbb{S} = \mathcal{S}^{\otimes N}$  and  $\mathbb{P} = \mathcal{P}^{\otimes N}$ . Similarly, the trace of  $\rho_A \otimes \rho_B$  is  $\langle \mathbb{I} | \rho \rangle = 1$ . These vectors are expressed as follows, using the dual basis  $\langle \sigma |$  defined via  $\langle \sigma | \sigma' \rangle = \delta_{\sigma \sigma'}$ :

$$\begin{aligned} \langle \mathbb{I} | &= (\langle 0 | + \langle 1 |)^{\otimes N}, \quad \langle \mathbb{P} | = (q^{-1} \langle 0 | + \langle 1 |)^{\otimes N}, \\ \text{and } \langle \mathbb{S} | &= (q^{-1} \langle 0 | + q \langle 1 |)^{\otimes N}. \end{aligned}$$

In this paper, we compute the circuit-averaged fidelity  $F$  and the XEB  $\chi = q^N X - 1$  by computing the two-copy density matrix

$$\overline{\rho^{(2)}} = \mathbb{E}_{\mathbf{U}} [\rho_{\mathbf{U}}(0) \otimes \rho_{\mathbf{U}}(\mathcal{N})]. \quad (11)$$

If the gates are randomly chosen from a distribution that is invariant under single qudit rotations — such as the Haar distribution or the distribution in Eq. (2) — the averaged two-copy density matrix after each layer of gates has the form of Eq. (9). As the gates are drawn independently at random between layers,  $p(\sigma)$  after a given layer depends only on the distribution  $p$  from the previous layer. The corresponding evolution of  $p$  under the circuit layer  $l$  is thus governed by a transfer matrix  $\mathcal{T}_l(\sigma', \sigma)$ . This transfer matrix can be constructed by composing a matrix  $M$  corresponding to each random two-qudit gate and a matrix  $N$  for each single qudit noise channel as follows:

$$\mathcal{T}_l(\sigma', \sigma) = \prod_i N_{\sigma'_i, \sigma_i} \prod_{\langle ij \rangle} M_{\sigma'_i \sigma'_j, \sigma_i \sigma_j}, \quad (12)$$

where  $\langle ij \rangle$  denotes pairs of qudits  $i, j$  which are acted upon by a random gate in a particular layer  $l$  of the circuit. As  $M$  denotes the update corresponding to the average over a unitary gate, its form is constrained by the requirement that it preserves the trace and the fidelity. We show in Appendix A that the two-qudit update matrix  $M$  must take the following form<sup>3</sup>

$$M = \begin{bmatrix} 1 & \frac{\alpha q^2}{q^2+1} & \frac{\alpha q^2}{q^2+1} & 0 \\ 0 & 1 - \alpha - \beta & \beta & 0 \\ 0 & \beta & 1 - \alpha - \beta & 0 \\ 0 & \frac{\alpha}{q^2+1} & \frac{\alpha}{q^2+1} & 1 \end{bmatrix}, \quad (13)$$

<sup>2</sup> See also Refs. [14, 15, 24, 25] for similar suggestions.

<sup>3</sup> Note that an equivalent but distinct parameterization of the update rule was described in Ref. [17]—translations between our conventions and those of Ref. [17] are given in Appendix A.

where we order the two-site  $\{I/q^2, \mathcal{S}/q\}$  basis (corresponding to  $\sigma \in \{0,1\}$ ) of the statistical model as  $\sigma_i\sigma_j \in \{00,01,10,11\}$  and  $\alpha, \beta \in \mathbb{R}$ . This form follows from the conservation of the trace and fidelity under noiseless operations, as well as an additional principle that guarantees reflection symmetry for  $M$ . For Haar-random two-qudit gates,  $\alpha = 1$  and  $\beta = 0$  [26]; the computation of  $\alpha$  and  $\beta$  corresponding to other gate sets is detailed in Appendix A. The parameter  $\alpha$  can be interpreted as controlling the rate at which  $\mathcal{S}$  factors, corresponding to  $\sigma = 1$ , are duplicated ( $I \otimes \mathcal{S} \rightarrow \mathcal{S} \otimes \mathcal{S}$ ) and destroyed ( $I \otimes \mathcal{S} \rightarrow I \otimes I$ ). The parameter  $\beta$  on the other hand controls the rate at which  $\mathcal{S}$  factors hop between sites  $I \otimes \mathcal{S} \rightarrow \mathcal{S} \otimes I$ .

Similarly, the update rule  $N$  for a single-qudit noise channel  $\mathcal{N}$  can be expressed using the  $\{I/q^2, \mathcal{S}/q\}$  basis as [12]

$$N = \begin{bmatrix} 1 & \gamma \\ 0 & 1 - \gamma \end{bmatrix}. \quad (14)$$

The parameter  $\gamma$  sets a rate at which  $\mathcal{S}$  factors decay  $\mathcal{S} \rightarrow I$ . It is shown in Appendix A that  $\gamma$  is equal to the infidelity of  $\mathcal{N}$  with respect to the identity channel, averaged over pure state inputs, and rescaled by  $q/(q-1)$ .

The corresponding rate for  $I$  factors to decay to  $\mathcal{S}$  is zero, as the maximally mixed two-copy density matrix is a fixed point of  $\mathcal{N}$ . We point out that, for this to be true, no conditions are required on the noise channel  $\mathcal{N}$ —in particular, we do not need to assume it to be unital. While the action of a non-unital noise channel  $\mathcal{N}$  (on one or both copies) will not leave the maximally mixed density matrix  $I \otimes I$  fixed, when  $\mathcal{N}$  acts on only one copy the following average over single-qudit Haar rotations projects  $\mathcal{N}(I) \otimes I$  back onto  $I \otimes I$ . When instead noise acts on both copies of the two-copy density matrix, as occurs when computing the purity or the collision probability of the noisy density matrix, only unital noise channels preserve the maximally mixed density matrix  $I \otimes I$ . We expand upon this in Appendix B.

In summary, the statistical models we consider are defined through a transition matrix  $\mathcal{T}(\alpha, \beta, \gamma)$ , parameterized by three parameters:  $\alpha$  and  $\beta$ , which are derived as a property of the gate set, and  $\gamma$ , which is proportional to the infidelity of the noise channel. The case  $\alpha = 1$  and  $\beta = 0$  corresponds to Haar-random two-qudit gates; we will consider this case primarily (see Section III), and then show that other values of  $\alpha$  and  $\beta$  give rise to similar behavior (see Section IV).

Let us now discuss the concrete implementation of the transfer matrix  $\mathcal{T}$  for the all-to-all geometry (Section II C) and the one-dimensional geometry (Section II D).

### C. All-to-all circuit

We first consider an all-to-all geometry, which is given by a fully connected architecture on which gates act

on random pairs of qudits—without regard to spatial locality—in dense layers that cover all of the qudits. In addition to averaging over the randomness of the gates, we also average over all possible choices of the pairings of qudits; or equivalently, as depicted in Fig. 1(a), after each layer of gates the qudits are randomly permuted. This geometry allows for an efficient solution of the statistical model that exploits the permutation symmetry of the circuit architecture. A similar random circuit architecture, considered for example in Ref. [26], has the gates acting on one pair of qudits at a time rather than in dense layers. We expect that our results would translate to this geometry with minimal changes.

We first note that the transfer matrix in the all-to-all circuit does not depend on the specific gate layer since all layers are constructed identically. It is constructed by averaging over all permutations of the qudits as

$$\mathcal{T}(\sigma', \sigma) = \frac{1}{N!} \sum_{P \in S_N} \mathcal{T}(\sigma', P\sigma), \quad (15)$$

where  $S_N$  denotes the permutation group of  $N$  elements. By construction, the transfer matrix  $\mathcal{T}(\sigma', \sigma)$  only depends on the Hamming weights of  $\sigma$  and  $\sigma'$ , that is the number of sites with  $\sigma_i = 1$ . We denote the Hamming weight of the bitstring  $\sigma$  as  $|\sigma|$ .  $|\sigma|$  just counts the number of tensor factors that are  $\mathcal{S}/q$  in the corresponding term of  $\overline{\rho^{(2)}}$ .

We denote the total probability of all configurations with a fixed  $|\sigma| = S$  as

$$p_S = \sum_{|\sigma|=S} p(\sigma). \quad (16)$$

Since we are averaging over a permutation-symmetric family of circuits, we can restrict our evolution of  $p(\sigma)$  to an evolution of  $p_S$ , with a transfer matrix  $\mathcal{T}_{S',S}$ . As  $S$  ranges from 0 to  $N$ , this is an  $(N+1) \times (N+1)$  sized matrix, which is easy to study numerically for large system sizes.

Furthermore, an explicit formula for  $\mathcal{T}_{S',S}$  can be written down with a bit of combinatorial logic, avoiding the need to explicitly compute the average over permutations. We assume for simplicity that  $N$  is even, so that each qudit encounters a gate in each layer. First, we note that we can construct  $\mathcal{T}_{S',S}(\alpha, \beta, \gamma)$  by composing a transfer matrix  $M_{S',S}(\alpha, \beta)$  for the permutation-averaged action of the two-qudit gates with the action of a layer of noise channels  $N_{S',S}(\gamma)$ :

$$\mathcal{T}_{S',S}(\alpha, \beta, \gamma) = N_{S',S}(\gamma) M_{S',S}(\alpha, \beta). \quad (17)$$

Secondly, we note that  $M_{S',S}(\alpha, \beta)$  in fact does not depend on  $\beta$ , because the only effect of  $\beta$  is to connect bitstrings  $\sigma$  which are the same up to permutations. Finally, we construct the permutation-averaged  $M$  by explicitly counting the number of permutations which yield each

possible value of  $M(\sigma', \sigma)$ : the result is

$$M_{S',S} = \frac{1}{\binom{N}{S}} \sum_{n_0, n_1, n_2} 2^{n_1} \binom{N/2}{n_0, n_1, n_2} \delta_{S, n_1 + 2n_2} \left[ \sum_{a, b, c} \binom{n_1}{a, b, c} \left( \frac{q^2 \alpha}{q^2 + 1} \right)^a (1 - \alpha)^b \left( \frac{\alpha}{q^2 + 1} \right)^c \delta_{S', b + 2c + 2n_2} \right]. \quad (18)$$

In this sum,  $n_0, n_1, n_2$  specify the number of gates which encounter, respectively, 0, 1, or 2 sites with  $\sigma_i = 1$ . Of the  $n_1$  gates that encounter 1 such site,  $a, b, c$  are the number of gates which output 0, 1, or 2 sites with  $\sigma'_i = 1$ , respectively.

For  $N(\sigma', \sigma)$ , we see that the matrix elements only depend on  $\sigma$  and  $\sigma'$  through their Hamming weights  $S, S'$ , and so averaging over permutations yields

$$N_{S',S} = \binom{S}{S'} (1 - \gamma)^{S'} \gamma^{S - S'}. \quad (19)$$

In words, in a layer of noise, each of the  $S$  factors of  $S$  that appear in the input has a probability  $\gamma$  to decay.

The numerical implementation of  $\mathcal{T}_{S',S}$  is straightforward. However, we find it necessary to use high precision numerics to get accurate answers, particularly when  $N$  is increased. Throughout, the all-to-all geometry results use BigFloats with 256 bits of precision, as implemented in Julia [27].

#### D. One-dimensional circuit

Next we consider a one-dimensional brickwork-circuit architecture where all gates are nearest-neighbor, as depicted in Fig. 1(b). We use open boundary conditions throughout, and only consider circuits with an even number of qudits  $N$ ; as such, in every other layer one site on each edge experiences neither a gate nor noise.

For our statistical model simulations, we use matrix-product states (MPS) to represent  $|\rho\rangle$ . We use the time-evolving block decimation (TEBD) algorithm [28–30] to apply the transfer matrix  $\mathcal{T}$  as a sequence of gates, as in Eq. (12). Truncations of small singular values are used to keep the bond dimension of the MPS from growing uncontrollably. All simulation results are checked pointwise for convergence in the total truncated weight  $\epsilon$ , with truncation errors as small as  $\epsilon = 10^{-50}$  used in some parameter regimes to reach convergence. This is achieved by using BigFloats with 256 bits of precision for calculations requiring  $\epsilon \leq 10^{-30}$ .

In addition to TEBD, we use the global MPS-Krylov algorithm [31] to compute several leading eigenvalues of the transfer matrix. This algorithm is identical to the commonly used Arnoldi method [32] with MPS representations of all Krylov vectors.

#### E. Review of noiseless circuit results

We now review some of the important properties of the statistical model considered in this paper that have been derived in previous works. The statistical model has two fixed points in the noiseless ( $\gamma = 0$ ) limit; the important features of the dynamics can be understood in relation to these fixed points [12, 26]. These fixed points are the dominant right-eigenvectors (with eigenvalue 1) of the transfer matrix  $\mathcal{T}$ , and they are represented in our notation as

$$|\mathbb{I}\rangle \equiv |0 \cdots 0\rangle \text{ and } |\mathbb{S}\rangle \equiv |1 \cdots 1\rangle. \quad (20)$$

These eigenvectors represent two-copy density matrices proportional to the maximally mixed density matrix (proportional to  $\mathbb{I}$ ) and the global swap operator  $\mathbb{S}$ , respectively. The corresponding left eigenvectors of  $\mathcal{T}$  are  $\langle \mathbb{I} |$  and  $\langle \mathbb{S} |$ , respectively; these left eigenvectors directly manifest the fact that the noiseless evolution preserves the trace and fidelity, recalling Section II B.

Without noise, all sufficiently deep circuits converge to a linear combination of  $|\mathbb{I}\rangle$  and  $|\mathbb{S}\rangle$ . The specific linear combination is in fact fully determined by the two conserved quantities, trace and fidelity. When the evolution is initialized with a pure state the initial fidelity is given by  $F = 1$ . This constrains the fixed-point state to be

$$|H\rangle = \frac{q^N}{q^N + 1} |\mathbb{I}\rangle + \frac{1}{q^N + 1} |\mathbb{S}\rangle. \quad (21)$$

A two-copy average density matrix with statistical model representation  $|H\rangle$  corresponds to an ensemble of quantum states that are indistinguishable from globally Haar-random states using second moment properties [26]—the states form a so-called spherical 2-design.

It is useful to notice that swapping the two copies of the density matrix is a symmetry of the noiseless evolution. This symmetry interchanges the two fixed points  $|\mathbb{I}\rangle$  and  $|\mathbb{S}\rangle$ . Thus, this statistical model can be interpreted as a classical Ising system in a ferromagnetic phase [19, 22]; this analogy will turn out to be a fruitful way to understand the physics of the circuit with noise.

#### F. Review of noisy circuit results

Having reviewed the noiseless circuit results, we now turn to the noisy case. When noise is added ( $\gamma > 0$ ), fidelity is no longer conserved and output density matrices always converge to the maximally mixed density matrix. If the error rate  $\varepsilon = (1 - q^{-2})\gamma$  of the single-qubit noise channel is sufficiently small, Ref. [12] showed that the global white noise approximation

$$\rho \approx F \rho_{\text{ideal}} + (1 - F) \frac{\mathbb{I}}{q^N} \quad (22)$$

accurately predicts second-moment quantities such as fidelity and linear cross-entropy [2]. As consequences, in

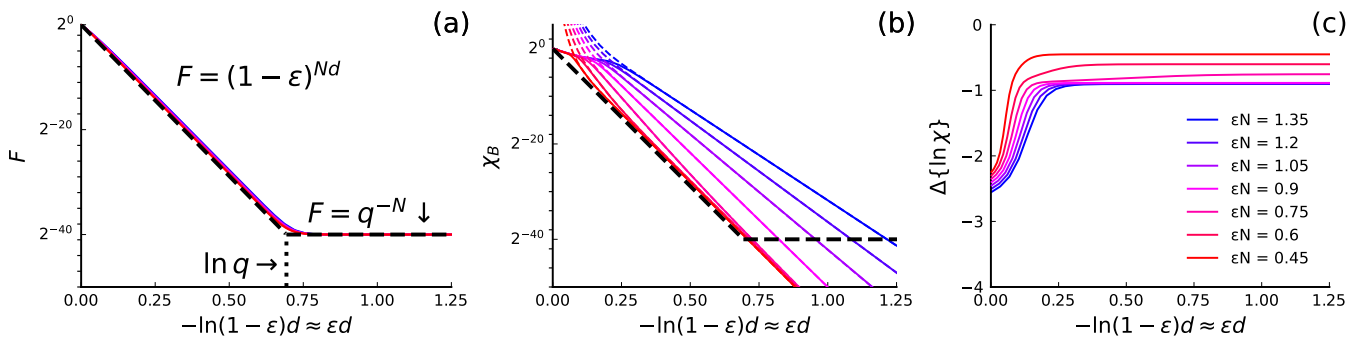


FIG. 2: Fidelity and XEB in the all-to-all circuit model with  $N = 40$  qubits and Haar-random gates, for various values of the noise rate  $\varepsilon$  (see color code in (c)), plotted against rescaled circuit depth. (a) Fidelity  $F$  curves collapse onto the global-white noise prediction  $(1 - \varepsilon)^{Nd}$  until saturating at  $q^{-N}$ . (b) Normalized XEB  $\chi_B$  as a function of rescaled circuit depth (solid lines) and unnormalized XEB  $\chi$  (dashed lines), which are essentially equivalent after an initial equilibration time. For low noise  $\varepsilon N \lesssim 0.92$ , the XEB decay rate matches that of fidelity. (c) Decay rate of the XEB. For  $\varepsilon N \gtrsim 0.92$ , the decay rate saturates to a  $\varepsilon N$ -independent constant,  $\Delta\{\ln \chi\} \approx -0.92$ .

the same regime of noise, XEB approximates fidelity, and both decay exponentially with a rate proportional to the noise rate and the total number of gates. This exponential decay takes the form

$$F(d) \sim (1 - \varepsilon)^{Nd}, \quad (23)$$

where  $\varepsilon$  is the error rate of the single-qudit noise channel and  $d$  is the circuit depth. In fact, this form of the decay is universal for systems with single-qudit noise and two-qudit gate count  $Nd/2$ —it is independent of the circuit architecture or gate set [12].

In the context of the statistical model, the global white noise regime is one in which the dynamics is approximately captured by a linear combination of the two fixed points of the noiseless transfer matrix:

$$|\rho(d)\rangle \approx a(d) |\mathbb{S}\rangle + (1 - a(d)) |\mathbb{I}\rangle, \quad (24)$$

where  $a(d) = (q^N F - 1)/(q^{2N} - 1)$ . While the maximally mixed density matrix continues to be a fixed point of  $\mathcal{T}$  with non-zero noise, the other fixed point  $|\mathbb{S}\rangle$  is converted into a metastable state. To lowest order in the noise, the decay of  $|\mathbb{S}\rangle$  in one layer of the circuit is given by

$$\frac{\langle \mathbb{S} | \mathcal{T} | \mathbb{S} \rangle}{\langle \mathbb{S} | \mathbb{S} \rangle} = (q^{-2}\gamma + (1 - \gamma))^N = (1 - \varepsilon)^N, \quad (25)$$

where  $\varepsilon = \gamma(1 - q^{-2})$ , and thus the prediction for the fidelity  $F(d)$  at depth  $d$  matches Eq. (23).

Intuition for this regime can be gained in a ‘heralded noise’ picture, which models the evolution with error events that occur at fixed locations [12]. If the noise rate is small enough, the noiseless evolution has sufficient time between error events to return the state to the fixed points  $|\mathbb{I}\rangle$  and  $|\mathbb{S}\rangle$ . By expanding the full noisy evolution of  $\mathcal{T}$  into a sum of terms corresponding to different locations and times of the errors, Ref. [12] showed that a sufficient condition on the noise to be in this regime is

$$\varepsilon \ll \frac{1}{N \ln N}, \quad (26)$$

while conjecturing that  $\varepsilon < \frac{c}{N}$  for some constant  $c$  would suffice. Ref. [17] also studied the statistical model with noise quantitatively for several values of  $\varepsilon$  and several choices of gate set, showing that the global white noise approximation fails to describe XEB at larger noise values.

### III. FIDELITY AND XEB TRANSITION

In this section, we will first look empirically at the behavior of fidelity and XEB beyond the global white noise approximation, and then explain the behavior with a spectral analysis of the transfer matrix  $\mathcal{T}$ . Our results confirm the conjecture of Ref. [12], showing that the global white noise approximation is indeed valid when  $\varepsilon N < c$ , with a sharp phase transition at  $\varepsilon N = c$  separating the global white noise regime from a larger noise regime in which the metastable state  $|\mathbb{S}\rangle$  decays too quickly to dominate the contributions to XEB. By contrast, the fidelity continues to track Eq. (23) in the larger noise regime  $\varepsilon N > c$ ; thus, the phase transition also marks the point at which XEB fails to be a proxy for fidelity.

Figure 2 shows the behavior of average fidelity and XEB, computed for the all-to-all circuit geometry and Haar random two-qubit gates. The fidelity [Fig. 2(a)] is close to the global white-noise prediction Eq. (23) until saturation, when fidelity approaches a minimal value,  $F = q^{-N}$ , which is exactly the fidelity of any state with a maximally mixed state. Similar behavior occurs in one-dimensional circuits, as shown in Appendix C (??).

By contrast, the decay rate of the XEB closely matches the global white noise value only for sufficiently small noise rates, as shown in Fig. 2(b). At higher noise, the decay rate of the XEB abruptly stops changing as noise is increased. To sharply distinguish the two regimes, one

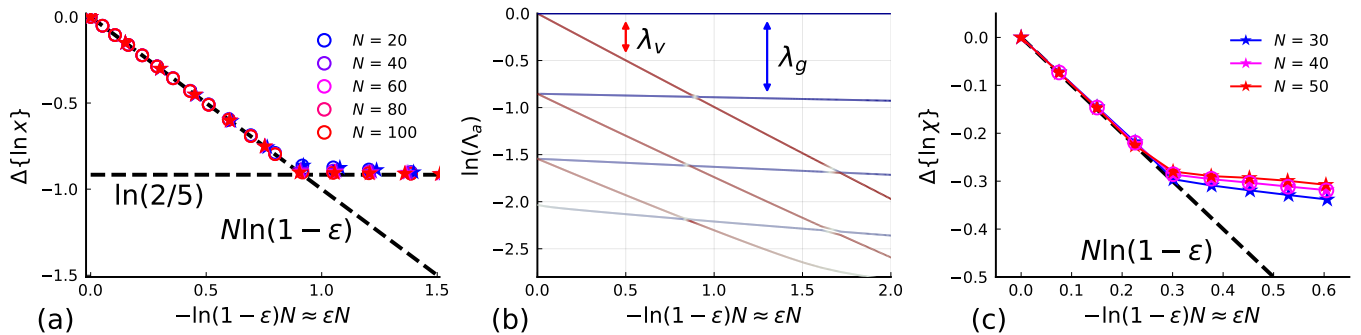


FIG. 3: Decay rate  $\Delta\{\ln \chi\}$  of XEB, when it has reached its asymptotic value (stars) and subleading eigenvalue  $\Lambda_1$  of  $\mathcal{T}$  (circles) for different values of the system size  $N$ , for the all-to-all circuit (a) and the 1D circuit (c), with Haar random gates. The decay rate of the XEB matches the subleading eigenvalue and undergoes a kink at the eigenvalue crossing. In (a), the kink occurs at  $\varepsilon N = \log \frac{5}{2} \approx 0.92$ . (b) Seven largest eigenvalues  $\Lambda_a$  of the transfer matrix  $\mathcal{T}$  in the all-to-all geometry with  $N = 40$  qubits and Haar random gates. States with low Hamming weight (blue tones) are insensitive to  $\varepsilon N$ , while states with extensive Hamming weight (red tones) have decay rates  $\ln(\Lambda_a)$  linear in  $\varepsilon N$ .

should look at the asymptotic decay rate of the XEB,

$$\Delta\{\ln \chi(d)\} := \ln \frac{\chi(d+1)}{\chi(d)}. \quad (27)$$

In all cases,  $\chi(d+1)/\chi(d)$  reaches a plateau as a function of depth  $d$ , as shown in Fig. 2(c). Empirically, we see two regimes: in the low-noise regime,  $\chi(d+1)/\chi(d) \sim (1-\varepsilon)^N$ , while for higher noises  $\chi(d+1)/\chi(d) \sim \text{const.}$ , with a constant that is roughly independent of  $N$  and  $\varepsilon$ .

Figure 3 reveals the nature of the transition in the behavior of XEB. The asymptotic decay rate of XEB, shown in Fig. 3(a) for the all-to-all circuit and in Fig. 3(c) for one-dimensional circuits, experiences a kink at a fixed value of  $\varepsilon N$ . Additionally, this rate is shown to precisely match the value of the largest eigenvalue  $\Lambda_1$  of the transfer matrix, excluding the trivial largest eigenvalue  $\Lambda_0 = 1$  corresponding to the maximally mixed density matrix fixed point. The kink in the value of  $\Lambda_1$  occurs precisely at the location where the leading eigenvalues of  $\mathcal{T}$  cross, as depicted in Fig. 3(b) for the all-to-all model.

This eigenvalue crossing is not accidental, but in fact a generic feature of a wide class of random circuit models. The ubiquity can be explained via an analogy to the Ising model. At zero noise, the leading transfer matrix eigenvalues are doubly degenerate, corresponding to spontaneously symmetry broken states of the copy permutation symmetry; noise plays the role of a field that explicitly breaks the symmetry [22]. In Fig. 3(b), we see that this analogy fits as a description for the leading eigenvectors of  $\mathcal{T}$ , which can be split into two classes: eigenvectors with low Hamming weight, which are analogous to excitations on top of one ferromagnetic vacuum  $|\mathbb{I}\rangle$ , and eigenvectors with extensive Hamming weight, which correspond to excitations on top of the other ferromagnetic vacuum  $|\mathbb{S}\rangle$ .

Such an eigenvalue crossing would be visible not just in the dynamics of XEB, but in nearly all observables that can be computed by expectation values in the av-

erage two-copy density matrix. Suppose the eigenvalues of  $\mathcal{T}$  are  $\Lambda_a$ , with  $a$  being an index, with corresponding left-eigenvectors  $\langle v_a|$  and right-eigenvectors  $|v_a\rangle$ . Then through the eigendecomposition of  $\mathcal{T}$ , any such observable will show dynamics of the form

$$O(d) = \sum_a c_a^O \Lambda_a^d, \quad \text{with } c_a^O = \frac{\langle O|v_a\rangle \langle v_a|\rho_0\rangle}{\langle v_a|v_a\rangle}, \quad (28)$$

where  $|\rho_0\rangle$  is the initial product state and applying  $\langle O|$  corresponds to multiplying with  $O$  and taking the trace. As we observe constant scaling gaps in the spectrum, we expect that this sum will be asymptotically dominated by the largest eigenvalue  $\Lambda_a$  for which the corresponding coefficient  $c_a^O$  is non-zero.

Having answered the question of why XEB experiences a transition in its decay rate, why do we not observe the same transition in the decay rate of the fidelity? To answer this, we analyze in detail the nature of the leading eigenvectors. Let us label the largest non-trivial eigenvalues  $\Lambda_v$  and  $\Lambda_g$  as depicted in Fig. 3(b), so that  $\Lambda_1 = \max\{\Lambda_g, \Lambda_v\}$ ; additionally, we refer to the corresponding eigenvectors as  $|v_v\rangle$  and  $|v_g\rangle$ , respectively.

In the limit of large  $N$  with fixed  $\varepsilon N$ , we find that the eigenvectors  $|v_v\rangle$  and  $|v_g\rangle$  are approximately constant, even as  $\varepsilon N$  is tuned across the transition. We do not observe significant level repulsion or mixing of the eigenvectors at the crossing. This follows our expectations from the Ising analogy, as matrix elements between states that are oppositely polarized should be exponentially suppressed. In Fig. 4, we show the behavior of the coefficients  $c_a^O$  for the observables  $F$  and  $\chi$  in these eigenvectors. We see that while XEB couples significantly to each eigenvector,  $c_g^F$  approaches 0 asymptotically as  $q^{-N}$ . This is a generic behavior, as  $|v_g\rangle$  is an eigenvector with low Hamming weight but  $|\mathbb{S}\rangle$ , the observable vector corresponding to fidelity, is concentrated on strings with extensive Hamming weight.



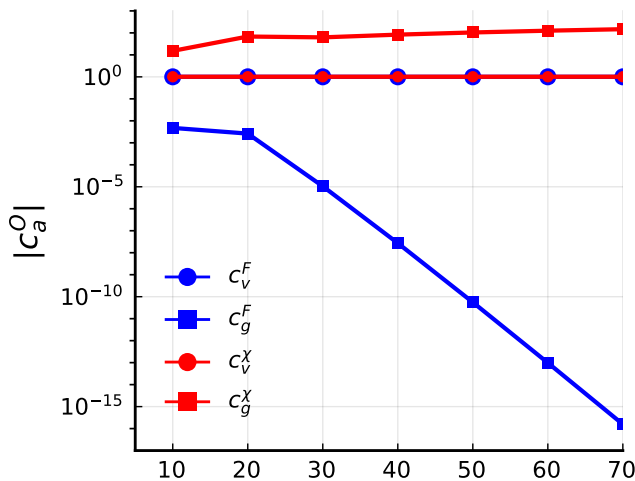


FIG. 4: Coupling constants  $c_a^O$  for fidelity and XEB with the leading non-trivial eigenvectors of the transfer matrix of the all-to-all model with Haar random gates and  $\varepsilon N = 0.01$ .

Thus, the asymptotic behavior of fidelity can be described by truncating the sum in Eq. (28) to the three dominant eigenvalues, resulting in

$$F \approx q^{-N} + (1 - \varepsilon)^{Nd} + \Lambda_g^d \frac{C}{q^N}, \quad (29)$$

for some constant  $C$ . While the third term becomes larger than the second in the large noise phase as  $d \rightarrow \infty$ , the  $q^{-N}$  suppression factor allows the second term to dominate the fidelity decay until the fidelity is close to saturation at  $q^{-N}$  at depths  $d \gtrsim (\ln q)/\varepsilon$ .

In summary, we generically expect a transition in the decay rate of XEB but not fidelity at constant scaling depths. This transition occurs when  $\Lambda_v = \Lambda_g$ .

#### IV. LOCATION OF THE TRANSITION FOR DIFFERENT ENTANGLING GATES

By the arguments of the previous section, the critical value corresponds to the location of the eigenvalue crossing  $\Lambda_v((\varepsilon N)_c) = \Lambda_g((\varepsilon N)_c)$ . In our numerics, these behave roughly as  $\Lambda_g(\varepsilon N) = \Lambda_g(0) \equiv \Lambda_g$  and  $\Lambda_v(\varepsilon N) = (1 - \varepsilon)^N$ . We expect this behavior of the eigenvalues to be exact in the thermodynamic limit based on the analogy to the Ising model explained in the previous section. We confirm this expectation in the all-to-all architecture in Fig. 5, which shows that the finite-size corrections to the eigenvalues scale as  $O(\varepsilon, 1/N)$ , so that these corrections vanish as  $N \rightarrow \infty$  keeping  $\varepsilon N$  fixed. Altogether, this gives us the criterion

$$(1 - \varepsilon)^N = \Lambda_g \implies (\varepsilon N)_c \approx \lambda_g \equiv -\ln \Lambda_g, \quad (30)$$

for the critical value  $(\varepsilon N)_c$  of  $\varepsilon N$ .

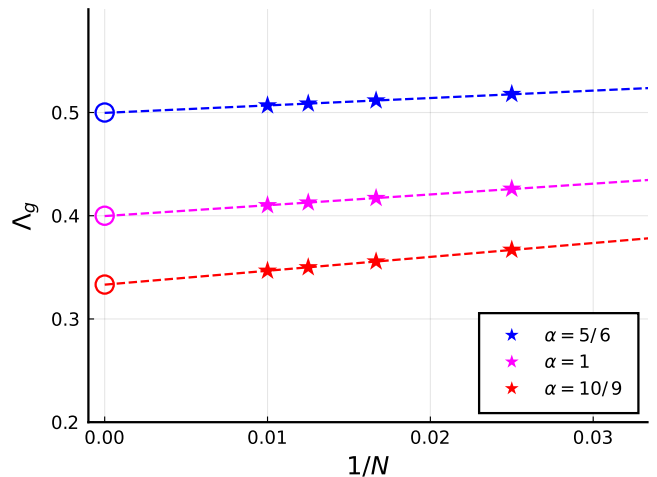


FIG. 5: Gap  $\Lambda_g$  of the all-to-all circuit for several values of  $\alpha$ , computed for various system sizes  $N$  as a function of  $1/N$  (stars) and linear fit (dashed lines). In all cases, the extrapolation to infinite size via the linear fit matches the prediction  $\Lambda_g = 1 - 3/5\alpha$  (circles).

Thus, to solve for the critical value, it suffices to find the gap to the local excitation in the noiseless transfer matrix. In this section, we exploit this to locate the transition for various gate sets and geometries.

Our discussion and results in the previous sections have focused on the case of Haar-random two-qudit gates, for which the statistical model update rule is given by Eq. (13) with  $\alpha = 1$  and  $\beta = 0$ . However, in experimental settings, it is more convenient to use gate sets with a fixed choice of two-qubit gate [2–4]. We consider in this section gate sets with a fixed two-qubit gate dressed with Haar-random single-qubit gates<sup>4</sup>. In Appendix A, we derive the values of  $\alpha$  and  $\beta$  for general two-qubit gates ( $q = 2$ ); it is shown that  $\alpha$  and  $\beta$  are restricted to the region shown in Fig. 6a. We also derive explicit parameterizations for the gates that correspond to the boundaries of this region. Three of the boundaries are given by specific parameterizations of the so-called fSim gate [33],

$$\text{fSim}(\theta, \varphi) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos \theta & -i \sin \theta & 0 \\ 0 & -i \sin \theta & \cos \theta & 0 \\ 0 & 0 & 0 & e^{-i\varphi} \end{pmatrix}, \quad (31)$$

while the last boundary corresponds to so-called *perfect entanglers*,

$$\text{PE}(\phi) = \begin{pmatrix} \sin \phi & 0 & 0 & -i \cos \phi \\ 0 & \cos \phi & -i \sin \phi & 0 \\ 0 & -i \sin \phi & \cos \phi & 0 \\ -i \cos \phi & 0 & 0 & \sin \phi \end{pmatrix}. \quad (32)$$

<sup>4</sup> The statistical model for such gate sets was also derived in Ref. [17].



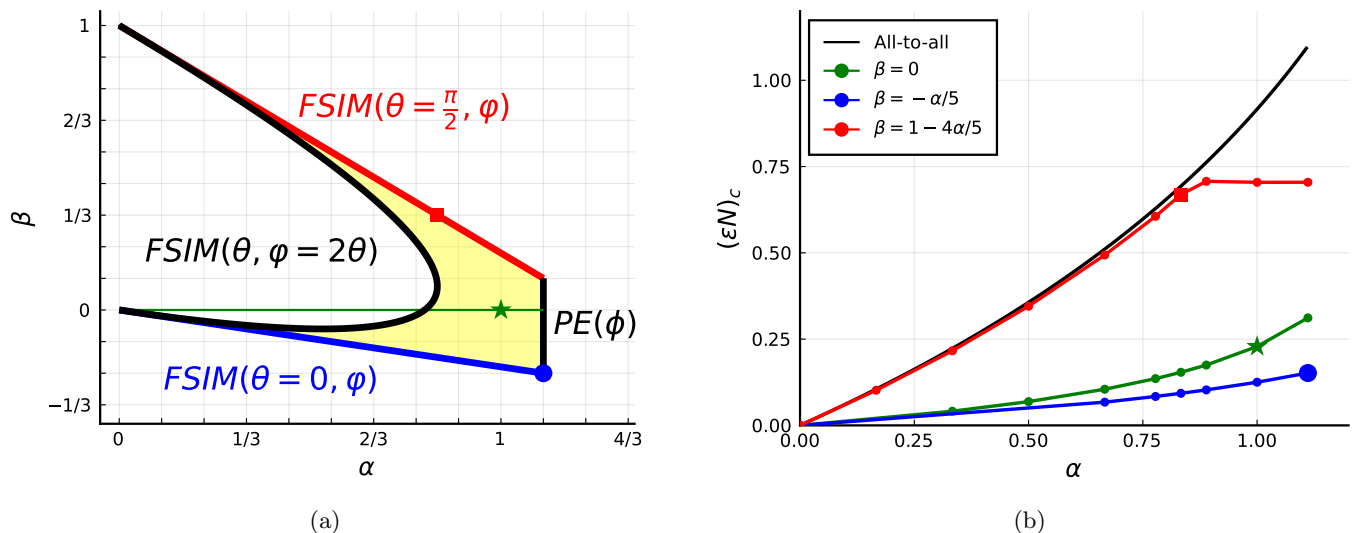


FIG. 6: (a) The valid region of  $\alpha, \beta$  accessible to gate sets with a fixed two qubit gate (shaded yellow), with parameterizations of gates that realize the borders. The red square, green star, and blue circle respectively denote  $fSim(\pi/2, \pi/6)$  used in Ref. [2], Haar random gates, and the control- $Z$  gate. (b) Critical value of  $\epsilon N$  for various geometries computed via Eq. (30). In the all-to-all geometry (black), we use the analytical formula Eq. (33). For 1D circuits (green, blue, and red), we numerically compute  $\Lambda_g$  along the three correspondingly colored lines in (a).

As discussed in Section II C, the transfer matrix for permutation-symmetric circuit architectures has no dependence on  $\beta$ . Thus, for the all-to-all circuit, we only need to compute  $\Lambda_g(\alpha)$ . Empirically, we find that  $\Lambda_g(\alpha) = 1 - 3\alpha/5$ , using finite-size diagonalizations for various  $\alpha$  and  $N$  and extrapolating to  $N \rightarrow \infty$ . This is shown in Fig. 5.

We now present a heuristic argument that in the all-to-all model the exact formula for the gap of the transfer matrix is

$$\Lambda_g = (1 - \alpha) + \alpha \frac{2}{q^2 + 1}, \quad (33)$$

which matches our empirical finding with  $q = 2$ . Consider configurations  $\sigma$  on  $N$  sites with a small fraction  $f$  of  $\mathcal{S}$  factors in random locations, that is, configurations  $\sigma$  with low Hamming weight. When the density of such sites is small, each individual gate encounters either zero or one such site, as the chance of a collision is  $\mathcal{O}(f^2)$ . At each gate that encounters a  $\mathcal{S}$ , there is a chance of  $1 - \alpha$  that the action of the gate leaves one  $\mathcal{S}$ , and a chance of  $\alpha/(q^2 + 1)$  that it produces two  $\mathcal{S}$  factors; thus in total the expected  $\mathcal{S}$  density will be decreased from  $f$  to  $((1 - \alpha) + \alpha \frac{2}{q^2 + 1})f$ . States of the form

$$|\rho\rangle \propto |\mathbb{I}\rangle + a|v_g\rangle \quad (34)$$

satisfy our hypothesis of having only a small expected fraction  $f$  of  $\mathcal{S}$  factors, and after an application of the transfer matrix have a fraction  $\Lambda_g f$  of  $\mathcal{S}$  factors. Thus it seems reasonable to expect that Eq. (33) holds.

The situation becomes more intricate in low-dimensional architectures in which the critical point depends on both  $\alpha$  and  $\beta$ . To analyze the dependence

on the values of  $\alpha$  and  $\beta$ , we numerically compute—and show in Fig. 6b—the critical value  $-\ln \Lambda_g$  in the one-dimensional brickwork architecture [Fig. 1(b)] along the upper and lower boundaries of the allowed region shown in Fig. 6a, as well as the line  $(\alpha, \beta = 0)$ . The computations use the MPS-Krylov algorithm discussed in Section II D, with  $N = 40$  sites and  $\epsilon = 0$ . We find a monotonic dependence of the critical point on both  $\alpha$  and  $\beta$ . In particular, the dependence of the critical value of the  $fSim(\pi/2, \varphi)$ -gate on  $\alpha$  roughly matches that of the all-to-all architecture until it saturates at  $\Lambda_g = 1/2$ , corresponding to  $(\epsilon N)_c = \ln 2$ , while the smaller values of  $\beta$  have a smaller critical value.<sup>5</sup> We leave it as an interesting question to precisely understand this saturation effect at maximal values of  $\beta$ .

Our results show that the critical value of  $\epsilon N$  can be tuned by varying  $\alpha$  and  $\beta$ . In the all-to-all architecture, it can be tuned all the way from 0 to a maximal value of  $\ln 3 \approx 1.1$  by tuning  $\alpha$  from 0 to the maximal value  $10/9$ . In the one-dimensional architecture, it can be tuned from 0 to  $\ln 2 \approx 0.69$  at the maximal value of  $\alpha$  and  $\beta$ . The corresponding gate is the  $i$ SWAP gate [17]. We expect our results to qualitatively carry over to any architecture.

The entangling gate parameter  $\alpha$  is related to a property of the gate known as its *entangling power* [34, 35]. Heuristically, at every value of  $\alpha$ , the gate parameter

<sup>5</sup> Let us also note that, in agreement with Ref. [17], we find that the noise robustness of the  $fSim(\pi/2, \pi/6)$  gate from Ref. [2] exceeds that of Haar-random gates, which in turn exceeds that of controlled- $Z$  gates.

$\beta$  measures the *swapping power* of the corresponding gate—for the minimal value of  $\beta$  given by  $-\alpha/5$ , the gate does not swap the qubits at all, while at its maximal value  $1 - 4\alpha/5$  it corresponds to a full swap. Indeed, multiplying the gates on the blue line (minimum swapping power) with a SWAP gate gives rise to the red line (maximum swapping power); in general multiplying a gate with entangling and swapping power  $\alpha$  and  $\beta$  with a SWAP gate, gives rise to a gate with  $\alpha' = \alpha$ ,  $\beta' = 1 - \alpha - \beta$ . Our results show that computational benchmarking experiments should pick maximally entangling and maximally swapping gates if they wish to maximize the noise level at which the phase transition occurs.

## V. OUTLOOK

We have given strong analytic and numerical evidence for the existence of a phase transition in the dynamics of noisy random circuits as a function of noise. Below the phase transition, the white noise model of Ref. [12] suffices to describe second moment quantities of the density matrix. In particular, in this regime, sample-efficient proxies for the fidelity such as the linear cross-entropy benchmark mimic the fidelity well. Above the phase transition, this correspondence fails to hold. Indeed, in this regime, the XEB and fidelity behave asymptotically differently. Even more drastically, the decay rate of the XEB fails to match that of the fidelity ( $F$ ) or its shifted version ( $F - q^{-n}$ ). Our results hold for any Haar-invariant two-qudit gate set and any single-qudit noise channel (except leakage error). However, as the qualitative results only rely on features that are universal to Ising-like systems with explicit symmetry breaking fields, we expect that they generalize to systems with arbitrary locally-acting and Haar-invariant gate sets and noise channels.

Our results imply that near-term quantum experiments purporting to use the linear cross-entropy benchmark as a proxy for fidelity must pay careful attention to the location of this phase transition. Our results are shown here for the all-to-all geometry and the one-dimensional geometry—we expect the physics of the phase transition to be similar in other experimentally relevant geometries, such as two dimensions, as well. The location of the phase transition depends both on the geometry and the properties of the gate set used. The transfer matrix approach outlined here may be used in

order to estimate the location of the transition for any gate set and geometry.

A separate question that we have not dealt with in this work is the ability to spoof the XEB metric, i.e., the ability to efficiently produce samples that achieve a high XEB score without necessarily producing samples from a state close in fidelity. The noisy output state above the phase transition indeed achieves an XEB score that is higher than the fidelity, but it is unclear whether this output distribution may be classically efficiently simulated. Recent work [36] shows that this is possible for depths logarithmic in system size and larger. Conversely, Gao *et al.* [17] show that, by adversarially selecting which circuit locations experience strong noise, “high” values of the XEB can be achieved classically. It remains an exciting question to what extent the phase transition we have identified in the XEB as a function of noise strength is reflected in the computational complexity of classically simulating the corresponding noisy circuit.

*Note added.*—After completing this work, we became aware of a complementary work studying the same phase transition [37].

## VI. ACKNOWLEDGMENTS

We thank M. Endres, S. Gopalakrishnan, H. Pichler, D. Huse, X. Gao, J. Choi, and C.D. White for helpful discussions. This research was performed while B.W. held an NRC Research Associateship award at the National Institute of Standards and Technology. M.J.G. and P.N. acknowledge support from the NSF QLCI (award No. OMA-2120757) and NSF grant PHY-1748958 through the KITP program on “Quantum Many-Body Dynamics and Noisy Intermediate-Scale Quantum Systems,” where the idea for this work was conceived. A.V.G. and P.N. were supported in part by the DoE Quantum Systems Accelerator, DoE ASCR Accelerated Research in Quantum Computing program (award No. DE-SC0020312), NSF QLCI (award No. OMA-2120757), DoE ASCR Quantum Testbed Pathfinder program (award No. DE-SC0019040), NSF PFCQC program, AFOSR, ARO MURI, AFOSR MURI, and DARPA SAVaNT ADVENT. D.H. acknowledges financial support from the US Department of Defense through a QuICS Hartree fellowship. This research was also supported in part by the National Science Foundation under Grant No. NSF PHY-1748958.

- 
- [1] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, *Characterizing Quantum Supremacy in Near-Term Devices*, *Nature Phys* **14**, 595 (2018), [arxiv:1608.00263](https://arxiv.org/abs/1608.00263).
- [2] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L.

- Brandao, D. A. Buell, B. Burkett, Y. Chen, Z. Chen, B. Chiaro, R. Collins, W. Courtney, A. Dunsworth, E. Farhi, B. Foxen, A. Fowler, C. Gidney, M. Giustina, R. Graff, K. Guerin, S. Habegger, M. P. Harrigan, M. J. Hartmann, A. Ho, M. Hoffmann, T. Huang, T. S. Humble, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, J. Kelly, P. V. Klimov, S. Knysh,

- A. Korotkov, F. Kostritsa, D. Landhuis, M. Lindmark, E. Lucero, D. Lyakh, S. Mandrà, J. R. McClean, M. McEwen, A. Megrant, X. Mi, K. Michielsen, M. Mohseni, J. Mutus, O. Naaman, M. Neeley, C. Neill, M. Y. Niu, E. Ostby, A. Petukhov, J. C. Platt, C. Quintana, E. G. Rieffel, P. Roushan, N. C. Rubin, D. Sank, K. J. Satzinger, V. Smelyanskiy, K. J. Sung, M. D. Trevithick, A. Vainsencher, B. Villalonga, T. White, Z. J. Yao, P. Yeh, A. Zalcman, H. Neven, and J. M. Martinis, *Quantum Supremacy Using a Programmable Superconducting Processor*, *Nature* **574**, 505 (2019).
- [3] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-H. Huo, L. Li, N. Li, S. Li, Y. Li, F. Liang, C. Lin, J. Lin, H. Qian, D. Qiao, H. Rong, H. Su, L. Sun, L. Wang, S. Wang, D. Wu, Y. Xu, K. Yan, W. Yang, Y. Yang, Y. Ye, J. Yin, C. Ying, J. Yu, C. Zha, C. Zhang, H. Zhang, K. Zhang, Y. Zhang, H. Zhao, Y. Zhao, L. Zhou, Q. Zhu, C.-Y. Lu, C.-Z. Peng, X. Zhu, and J.-W. Pan, *Strong Quantum Computational Advantage Using a Superconducting Quantum Processor*, *Phys. Rev. Lett.* **127**, 180501 (2021).
- [4] Q. Zhu, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, Y. Du, D. Fan, M. Gong, C. Guo, C. Guo, S. Guo, L. Han, L. Hong, H.-L. Huang, Y.-H. Huo, L. Li, N. Li, S. Li, Y. Li, F. Liang, C. Lin, J. Lin, H. Qian, D. Qiao, H. Rong, H. Su, L. Sun, L. Wang, S. Wang, D. Wu, Y. Wu, Y. Xu, K. Yan, W. Yang, Y. Yang, Y. Ye, J. Yin, C. Ying, J. Yu, C. Zha, C. Zhang, H. Zhang, K. Zhang, Y. Zhang, H. Zhao, Y. Zhao, L. Zhou, C.-Y. Lu, C.-Z. Peng, X. Zhu, and J.-W. Pan, *Quantum Computational Advantage via 60-Qubit 24-Cycle Random Circuit Sampling*, *Science Bulletin* **67**, 240 (2022), arxiv:2109.03494.
- [5] D. Hangleiter and J. Eisert, *Computational Advantage of Quantum Random Sampling*, (2023), arxiv:2206.04079.
- [6] M. J. Bremner, A. Montanaro, and D. J. Shepherd, *Average-Case Complexity Versus Approximate Simulation of Commuting Quantum Computations*, *Physical Review Letters* **117** (2016), 10.1103/PhysRevLett.117.080501.
- [7] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, *Characterizing Quantum Supremacy in Near-Term Devices*, arXiv:1608.00263 [quant-ph] (2016), arxiv:1608.00263.
- [8] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, *On the Complexity and Verification of Quantum Random Circuit Sampling*, *Nature Phys* **15**, 159 (2019).
- [9] R. Movassagh, *Quantum Supremacy and Random Circuits*, (2020), arxiv:1909.06210.
- [10] H.-Y. Huang, R. Kueng, and J. Preskill, *Efficient Estimation of Pauli Observables by Derandomization*, *Phys. Rev. Lett.* **127**, 030503 (2021), arxiv:2103.07510.
- [11] M. Kliesch and I. Roth, *Theory of Quantum System Certification*, *PRX Quantum* **2**, 010201 (2021), arxiv:2010.05925.
- [12] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, *Random Quantum Circuits Transform Local Noise into Global White Noise*, (2021), arXiv:2111.14907 [quant-ph].
- [13] X. Gao, M. Kalinowski, C.-N. Chou, M. D. Lukin, B. Barak, and S. Choi, *Limitations of Linear Cross-Entropy as a Measure for Quantum Advantage*, (2021), arxiv:2112.01657.
- [14] J. Choi, A. L. Shaw, I. S. Madjarov, X. Xie, R. Finkelstein, J. P. Covey, J. S. Cotler, D. K. Mark, H.-Y. Huang, A. Kale, H. Pichler, F. G. S. L. Brandão, S. Choi, and M. Endres, *Emergent Quantum Randomness and Benchmarking from Hamiltonian Many-body Dynamics*, *Nature* **613**, 468 (2023), arXiv:2103.03535 [cond-mat, physics:physics, physics:quant-ph].
- [15] D. K. Mark, J. Choi, A. L. Shaw, M. Endres, and S. Choi, *Benchmarking Quantum Simulators Using Quantum Chaos*, (2022), arxiv:2205.12211 [cond-mat, physics:physics, physics:quant-ph].
- [16] Y. Li, Y. Zou, P. Glorioso, E. Altman, and M. P. A. Fisher, *Cross Entropy Benchmark for Measurement-Induced Phase Transitions*, (), 2209.00609.
- [17] X. Gao, M. Kalinowski, C.-N. Chou, M. D. Lukin, B. Barak, and S. Choi, *Limitations of Linear Cross-Entropy as a Measure for Quantum Advantage*, (2021), arXiv:2112.01657 [cond-mat, physics:quant-ph].
- [18] T. Zhou and A. Nahum, *Emergent Statistical Mechanics of Entanglement in Random Unitary Circuits*, *Phys. Rev. B* **99**, 174205 (2019).
- [19] R. Vasseur, A. C. Potter, Y.-Z. You, and A. W. W. Ludwig, *Entanglement Transitions from Holographic Random Tensor Networks*, **100**, 134203.
- [20] N. Hunter-Jones, *Unitary Designs from Statistical Mechanics in Random Quantum Circuits*, (2019), arxiv:1905.12053.
- [21] B. Barak, C.-N. Chou, and X. Gao, *Spoofing Linear Cross-Entropy Benchmarking in Shallow Quantum Circuits*, in *12th Innovations in Theoretical Computer Science Conference (ITCS 2021)*, Leibniz International Proceedings in Informatics (LIPIcs), Vol. 185, edited by J. R. Lee (Schloss Dagstuhl–Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2021) pp. 30:1–30:20, arxiv:2005.02421.
- [22] Z. Li, S. Sang, and T. H. Hsieh, *Entanglement Dynamics of Noisy Random Circuits*, **107**, 014307 ().
- [23] A. Deshpande, P. Niroula, O. Shtanko, A. V. Gorshkov, B. Fefferman, and M. J. Gullans, *Tight Bounds on the Convergence of Noisy Random Circuits to the Uniform Distribution*, (2022), arxiv:2112.00716.
- [24] Y. Rinott, T. Shoham, and G. Kalai, *Statistical Aspects of the Quantum Supremacy Demonstration*, *Statistical Science* **37**, 322 (2022), arxiv:2008.05177.
- [25] Y. Liu, M. Otten, R. Bassirianjahromi, L. Jiang, and B. Fefferman, *Benchmarking Near-Term Quantum Computers via Random Circuit Sampling*, arXiv:2105.05232 [quant-ph] (2022), arxiv:2105.05232.
- [26] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, *Random Quantum Circuits Anti-Concentrate in Log Depth*, *PRX Quantum* **3**, 010333 (2022), arXiv:2011.12277 [cond-mat, physics:quant-ph].
- [27] J. Bezanson, A. Edelman, S. Karpinski, and V. B. Shah, *Julia: A fresh approach to numerical computing*, *SIAM review* **59**, 65 (2017).
- [28] U. Schollwöck, *The Density-Matrix Renormalization Group in the Age of Matrix Product States*, *Annals of Physics January 2011 Special Issue*, **326**, 96.
- [29] G. Vidal, *Efficient Classical Simulation of Slightly Entangled Quantum Computations*, *Phys. Rev. Lett.* **91**, 147902 (2003).

- [30] G. Vidal, *Efficient Simulation of One-Dimensional Quantum Many-Body Systems*, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [31] S. Paeckel, T. Köhler, A. Swoboda, S. R. Manmana, U. Schollwöck, and C. Hubig, *Time-Evolution Methods for Matrix-Product States*, *Annals of Physics* **411**, 167998.
- [32] W. E. Arnoldi, *The Principle of Minimized Iterations in the Solution of the Matrix Eigenvalue Problem*, **9**, 17, 43633863.
- [33] B. Foxen, C. Neill, A. Dunsworth, P. Roushan, B. Chiaro, A. Megrant, J. Kelly, Z. Chen, K. Satzinger, R. Barends, F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, S. Boixo, D. Buell, B. Burkett, Y. Chen, R. Collins, E. Farhi, A. Fowler, C. Gidney, M. Giustina, R. Graff, M. Harrigan, T. Huang, S. V. Isakov, E. Jeffrey, Z. Jiang, D. Kafri, K. Kechedzhi, P. Klimov, A. Korotkov, F. Kostritsa, D. Landhuis, E. Lucero, J. McClean, M. McEwen, X. Mi, M. Mohseni, J. Y. Mutus, O. Naaman, M. Neeley, M. Niu, A. Petukhov, C. Quintana, N. Rubin, D. Sank, V. Smelyanskiy, A. Vainsencher, T. C. White, Z. Yao, P. Yeh, A. Zalcman, H. Neven, J. M. Martinis, and Google AI Quantum, *Demonstrating a Continuous Set of Two-qubit Gates for Near-term Quantum Algorithms*, *Phys. Rev. Lett.* **125**, 120504 (2020).
- [34] P. Zanardi, C. Zalka, and L. Faoro, *Entangling Power of Quantum Evolutions*, *Phys. Rev. A* **62**, 030301 (2000).
- [35] X. Wang, B. C. Sanders, and D. W. Berry, *Entangling Power and Operator Entanglement in Qudit Systems*, *Phys. Rev. A* **67**, 042323 (2003).
- [36] D. Aharonov, X. Gao, Z. Landau, Y. Liu, and U. Vazirani, *A Polynomial-Time Classical Algorithm for Noisy Random Circuit Sampling*, (2022), [arxiv:2211.03999](https://arxiv.org/abs/2211.03999) [quant-ph].
- [37] A. Morvan, B. Villalonga, X. Mi, S. Mandrà, A. Bengtsson, P. V. Klimov, Z. Chen, S. Hong, C. Erickson, I. K. Drozdov, J. Chau, G. Laun, R. Movassagh, A. Afaw, L. T. A. N. Brandão, R. Peralta, D. Abanin, R. Acharya, R. Allen, T. I. Andersen, K. Anderson, M. Ansmann, F. Arute, K. Arya, J. Atalaya, J. C. Bardin, A. Bिल्mes, G. Bortoli, A. Bourassa, J. Bovaird, L. Brill, M. Broughton, B. B. Buckley, D. A. Buell, T. Burger, B. Burkett, N. Bushnell, J. Campero, H. S. Chang, B. Chiaro, D. Chik, C. Chou, J. Cogan, R. Collins, P. Conner, W. Courtney, A. L. Crook, B. Curtin, D. M. Debroy, A. D. T. Barba, S. Demura, A. Di Paolo, A. Dunsworth, L. Faoro, E. Farhi, R. Fatemi, V. S. Ferreira, L. F. Burgos, E. Forati, A. G. Fowler, B. Foxen, G. Garcia, E. Genois, W. Giang, C. Gidney, D. Gilboa, M. Giustina, R. Gosula, A. G. Dau, J. A. Gross, S. Habegger, M. C. Hamilton, M. Hansen, M. P. Harrigan, S. D. Harrington, P. Heu, M. R. Hoffmann, T. Huang, A. Huff, W. J. Huggins, L. B. Ioffe, S. V. Isakov, J. Iveland, E. Jeffrey, Z. Jiang, C. Jones, P. Juhas, D. Kafri, T. Khattar, M. Khezri, M. Kieferová, S. Kim, A. Kitaev, A. R. Klots, A. N. Korotkov, F. Kostritsa, J. M. Kreikebaum, D. Landhuis, P. Laptev, K.-M. Lau, L. Laws, J. Lee, K. W. Lee, Y. D. Lensky, B. J. Lester, A. T. Lill, W. Liu, A. Locharla, F. D. Malone, O. Martin, S. Martin, J. R. McClean, M. McEwen, K. C. Miao, A. Mieszala, S. Montazeri, W. Mruczkiewicz, O. Naaman, M. Neeley, C. Neill, A. Nersisyan, M. Newman, J. H. Ng, A. Nguyen, M. Nguyen, M. Y. Niu, T. E. O’Brien, S. Omonije, A. Opremcak, A. Petukhov, R. Potter, L. P. Pryadko, C. Quintana, D. M. Rhodes, C. Rocque, P. Roushan, N. C. Rubin, N. Saei, D. Sank, K. Sankaragomathi, K. J. Satzinger, H. F. Schurkus, C. Schuster, M. J. Shearn, A. Shorter, N. Shutty, V. Shvarts, V. Sivak, J. Skrzynny, W. C. Smith, R. D. Somma, G. Sterling, D. Strain, M. Szalay, D. Thor, A. Torres, G. Vidal, C. V. Heidweiller, T. White, B. W. K. Woo, C. Xing, Z. J. Yao, P. Yeh, J. Yoo, G. Young, A. Zalcman, Y. Zhang, N. Zhu, N. Zobrist, E. G. Rieffel, R. Biswas, R. Babbush, D. Bacon, J. Hilton, E. Lucero, H. Neven, A. Megrant, J. Kelly, I. Aleiner, V. Smelyanskiy, K. Kechedzhi, Y. Chen, and S. Boixo, *Phase transition in Random Circuit Sampling*, (2023), [arXiv:2304.11119](https://arxiv.org/abs/2304.11119) [quant-ph].
- [38] Y. Makhlin, *Nonlocal Properties of Two-Qubit Gates and Mixed States and Optimization of Quantum Computations*, *Quantum Information Processing* **1**, 243 (2002).
- [39] J. Zhang, *Geometric Theory of Nonlocal Two-Qubit Operations*, *Physical Review A* **67** (2003), 10.1103/PhysRevA.67.042313.
- [40] S. Balakrishnan and R. Sankaranarayanan, *Entangling Power and Local Invariants of Two-Qubit Gates*, *Physical Review A* **82**, 034301 (2010).

## Appendix A: Derivations of statistical model rules

In this Appendix, we rederive the rules governing the statistical model following the notation of Ref. [12]. First, we derive the rule corresponding to the application of an arbitrary two-qudit gate followed by an average over single-qudit gates. Next, we derive the rule corresponding to the application of single-qudit [AD: stochastic?] noise.

The initial state  $(|0\rangle\langle 0|)^{\otimes N}$  is first acted upon by single-qudit Haar-random gates, which project the two-copy state into the symmetric subspace spanned by  $\{I, \mathcal{S}\}^{\otimes N}$ . As in the main text,  $I$  is the  $(q^2 \times q^2)$  2-copy identity and  $\mathcal{S}$  the SWAP operation of two qudits acting between the two copies. Upon further applications of arbitrary two-qudit gates followed by Haar-random single-qudit gates, the state remains in the  $\{I, \mathcal{S}\}^{\otimes N}$  space.

We now examine the effect of an arbitrary two-qubit gate  $U$ . The update rule is obtained by performing the following average over Haar-random single-qubit gates  $V$  and  $W$  acting on the two qubits in the two copies of the system:

$$\mathbb{E}_{V,W} \left[ \begin{array}{c} \left[ \begin{array}{cccc} \boxed{V} & \boxed{U} & \boxed{\rho^{(2)}} & \boxed{U^\dagger} & \boxed{V^\dagger} \\ \boxed{W} & \boxed{U} & \boxed{\rho^{(2)}} & \boxed{U^\dagger} & \boxed{W^\dagger} \\ \boxed{V} & \boxed{U} & \boxed{\rho^{(2)}} & \boxed{U^\dagger} & \boxed{V^\dagger} \\ \boxed{W} & \boxed{U} & \boxed{\rho^{(2)}} & \boxed{U^\dagger} & \boxed{W^\dagger} \end{array} \right] \end{array} \right]. \quad (\text{A1})$$

We parameterize the two-qubit unitary as

$$U = \begin{pmatrix} \cos(c_-) & 0 & 0 & -i \sin(c_-) \\ 0 & \cos(c_+)e^{ic_3} & -i \sin(c_+)e^{ic_3} & 0 \\ 0 & -i \sin(c_+)e^{ic_3} & \cos(c_+)e^{ic_3} & 0 \\ -i \sin(c_-) & 0 & 0 & \cos(c_-) \end{pmatrix}, \quad (\text{A2})$$

where  $c_{\pm} = \frac{c_1 \pm c_2}{2}$  with  $c_1, c_2 \in \mathbb{R}$ . Up to local single-qubit rotations, this is the most general two-qubit unitary [38, 39]. Since single-qubit gates are averaged over, all gates that are equivalent up to local single-qubit gates yield the same update rules. The invariants of such two-qubit gates under local rotations have been characterized by Makhlin [38]. Given a gate of the form (A2), these are given by

$$G_1 = \frac{e^{-2ic_3}}{4} (\cos(c_1 - c_2) + e^{2ic_3} \cos(c_1 + c_2))^2, \quad (\text{A3})$$

$$G_2 = \cos 2c_1 + \cos 2c_2 + \cos 2c_3, \quad (\text{A4})$$

and they can be efficiently computed for any gate [38]. The quantity  $|G_1| = \frac{1}{4}(1 + \cos 2c_1 \cos 2c_2 + \cos 2c_2 \cos 2c_3 + \cos 2c_3 \cos 2c_1)$  is related to the entangling power of the two-qubit gate. The relation of the entangling power of a gate to a specific property of the transfer matrix was noted by Gao *et al.* [17].

The two-copy state  $\bar{\rho}^{(2)}$  is a combination of the operators<sup>6</sup>  $II, IS, SI,$  and  $SS$ , where the SWAP operation is between the two copies, i.e., between either the first and third qubits (in the case of  $SI$ ) or the second and fourth qubits (in the case of  $IS$ ) in Eq. (A1). The operators  $II$  and  $SS$  are left invariant under the action of the gates in Eq. (A1). It suffices to consider  $IS$ , as the evolution of  $SI$  may be derived from that of  $IS$  by exchanging the two copies. Denote by  $\overline{IS}$  the evolution of  $IS$  under the update of a single gate layer Eq. (A1). Expanding  $\overline{IS} = aII + bIS + cSI + dSS$  in the basis  $\{I, S\}$ , the coefficients  $a, b, c,$  and  $d$  are determined through the Hilbert-Schmidt inner product with the basis operators. Specifically, we have

$$\text{Tr}[(\overline{IS})(II)] = 2^4 a + 2^3 b + 2^3 c + 2^2 d = 2^3, \quad (\text{A5})$$

$$\text{Tr}[(\overline{IS})(IS)] = 2^3 a + 2^4 b + 2^2 c + 2^3 d = x, \quad (\text{A6})$$

$$\text{Tr}[(\overline{IS})(SI)] = 2^3 a + 2^2 b + 2^4 c + 2^3 d = y, \quad (\text{A7})$$

$$\text{Tr}[(\overline{IS})(SS)] = 2^2 a + 2^3 b + 2^3 c + 2^4 d = 2^3, \quad (\text{A8})$$

with  $x = 6 + 4|G_1| + 2G_2$  and  $y = 6 + 4|G_1| - 2G_2$ . The prefactors can be easily found from the fact that  $\text{tr}[I] = 2^2$ , while  $\text{tr}[S] = 2$ . Moreover, the constraints (A5) and (A8) are determined by the facts that  $\text{tr}[(\overline{IS})(II)] = \text{tr}[(IS)(II)]$  since the trace is conserved

and  $\text{tr}[(\overline{IS})(SS)] = \text{tr}[(IS)(SS)]$  since the fidelity is conserved. Solving these equations gives us

$$a = d = \frac{20 - x - y}{18}, \quad (\text{A9})$$

$$b = \frac{4x + y - 32}{36}, \quad (\text{A10})$$

$$c = \frac{x + 4y - 32}{36}. \quad (\text{A11})$$

The transfer matrix in terms of the normalized operators  $\{I/q^2, S/q\}$  is then

$$M = \begin{pmatrix} 1 & 2a & 2a & 0 \\ 0 & b & c & 0 \\ 0 & c & b & 0 \\ 0 & a/2 & a/2 & 1 \end{pmatrix}, \quad (\text{A12})$$

implying  $\alpha = \frac{5a}{2} = \frac{10(1-|G_1|)}{9}$  and  $\beta = c = \frac{-1}{18} - \frac{G_2}{6} + \frac{5}{9}|G_1|$  from the comparison with Eq. (13). Plugging in the easily-computed values of  $G_1$  and  $G_2$  for any gate, we obtain the transfer matrix

$$M = \begin{pmatrix} 1 & \frac{8}{9}(1 - |G_1|) & \frac{8}{9}(1 - |G_1|) & 0 \\ 0 & \frac{-1}{18} + \frac{G_2}{6} + \frac{5}{9}|G_1| & \frac{-1}{18} - \frac{G_2}{6} + \frac{5}{9}|G_1| & 0 \\ 0 & \frac{-1}{18} - \frac{G_2}{6} + \frac{5}{9}|G_1| & \frac{-1}{18} + \frac{G_2}{6} + \frac{5}{9}|G_1| & 0 \\ 0 & \frac{2}{9}(1 - |G_1|) & \frac{2}{9}(1 - |G_1|) & 1 \end{pmatrix}. \quad (\text{A13})$$

### Notation of Gao *et al.*

We briefly take a detour to compare our notation with the notation of Gao *et al.* [17]. The basis of their statistical mechanical model  $\{I, \Omega\}$  is related to our basis of  $\{I/q^2, S/q\}$  through the linear transformation effected by the matrix  $O$ :

$$O = \begin{pmatrix} 1 & 1 \\ 0 & q^2 - 1 \end{pmatrix}, \text{ with} \quad (\text{A14})$$

$$O^{-1} = \begin{pmatrix} 1 & \frac{-1}{q^2-1} \\ 0 & \frac{1}{q^2-1} \end{pmatrix}. \quad (\text{A15})$$

Performing a similarity transform of transfer matrix  $\mathcal{T}$  using  $O \otimes O$  yields a transfer matrix in the basis  $\{I, \Omega\}$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 - \beta - \frac{\alpha q^2}{q^2+1} & \beta + \frac{\alpha}{q^2+1} & \frac{\alpha}{q^2+1} \\ 0 & \beta + \frac{\alpha}{q^2+1} & 1 - \beta - \frac{\alpha q^2}{q^2+1} & \frac{\alpha}{q^2+1} \\ 0 & \frac{\alpha(q^2-1)}{q^2+1} & \frac{\alpha(q^2-1)}{q^2+1} & 1 - \frac{2\alpha}{q^2+1} \end{pmatrix}, \quad (\text{A16})$$

implying

$$D = \frac{\alpha q^2}{q^2+1} + \beta, \quad (\text{A17})$$

$$R = \frac{\alpha(q^2-1)}{q^2+1}, \quad (\text{A18})$$

$$\eta = q^2 - 1, \quad (\text{A19})$$

in the notation of Gao *et al.*

<sup>6</sup> To ease readability, in the following, we abbreviate  $A \otimes B = AB$  for  $A, B \in \{I, S\}$ .



### Transfer matrix parameters for some gates

We will use the parametrization with  $\alpha$  and  $\beta$  from the main text in order to discuss the possible transfer matrices associated with various gates. The various gates and their parameter values are given in Table I.

We note the constraints on the possible values that  $G_1$  and  $G_2$ , and hence the transfer matrix parameters, can take. First, we use the relation between  $|G_1|$  and the entangling power of the gate [40]:

$$e_p(U) = \frac{2}{9}(1 - |G_1|). \quad (\text{A20})$$

From this and from the fact that the maximum entangling power is between 0 and  $2/9$ , we have the constraints  $0 \leq |G_1| \leq 1$ , giving  $0 \leq \alpha \leq 10/9$ , or  $0 \leq R \leq 2/3$ . The gates that maximize entangling power have been characterized [39, 40]. These take the form  $c_1 = \pi/2, c_2 = 2\phi - \pi/2, c_3 = 0$ , and we denote the gate  $\text{PE}(\phi)$ , defined in the main text.

Reference [17] had also noted that  $0 \leq D \leq 1$ . In addition, they proved that  $D \geq R$ , which has an interesting interpretation in terms of the entangling power of a gate characterized by  $R$  and the apparent entangling power characterized by  $D$ . In terms of the quantities  $\cos 2c_1 \cos 2c_2 + \cos 2c_2 \cos 2c_3 + \cos 2c_3 \cos 2c_1$  and  $\cos 2c_1 + \cos 2c_2 + \cos 2c_3$  that characterize any gate, these two inequalities translate to

$$\cos 2c_1 \cos 2c_2 + \cos 2c_2 \cos 2c_3 + \cos 2c_3 \cos 2c_1 + \frac{1}{2}(\cos 2c_1 + \cos 2c_2 + \cos 2c_3) \geq \frac{-3}{2}, \quad (\text{A21})$$

$$\cos 2c_1 \cos 2c_2 + \cos 2c_2 \cos 2c_3 + \cos 2c_3 \cos 2c_1 - \frac{1}{2}(\cos 2c_1 + \cos 2c_2 + \cos 2c_3) \leq \frac{3}{2}. \quad (\text{A22})$$

These also imply  $\beta + \alpha/(q^2 + 1) \geq 0$  and  $\beta \leq 1 - \alpha q^2/(q^2 + 1)$ . It may be seen that the fSim gate with  $\theta = 0$  and arbitrary  $\varphi$  saturates the  $\beta \geq -\alpha/5$  constraint, with  $D = R = (1 - \cos \varphi)/3$ . Similarly, it can also be seen that the inequality  $\beta \leq 1 - 4\alpha/5$  is saturated with the fSim gate with  $\theta = \pi/2$  and arbitrary  $\varphi$ , with the values  $\alpha = 5(1 + \cos \varphi)/9$  and  $\beta = (5 - 4 \cos \varphi)/9$ .

We find a new constraint using the inequality  $(x + y + z)^2 \geq 3(xy + yz + zx)$  for any  $x, y, z \in \mathbb{R}$ . Applying this to  $x = \cos 2c_1, y = \cos 2c_2, z = \cos 2c_3$ , we obtain

$$G_2^2 - 3(4|G_1| - 1) \geq 0, \text{ or} \quad (\text{A23})$$

$$D - R \leq \left(D - \frac{R}{2}\right)^2, \text{ or} \quad (\text{A24})$$

$$\beta + \frac{\alpha}{5} \leq \left(\beta + \frac{\alpha}{2}\right)^2. \quad (\text{A25})$$

As shown in Fig. 6a, these constraints account for the entire boundary of the region of parameters accessible via a single two-qubit unitary. Furthermore, the equality is satisfied when  $c_1 = c_2 = c_3$ . It may be checked that

the fSIM gate with parameters  $(\theta, \varphi = 2\theta)$  also satisfies these conditions and has

$$\alpha = \frac{5}{12}(1 - \cos 4\theta), \quad (\text{A26})$$

$$\beta = \frac{1}{72}(21 + 15 \cos 4\theta - 36 \cos 2\theta). \quad (\text{A27})$$

As an interesting aside, we also point out that a *fixed* two-qubit gate leads to the same update rules as a Haar-random two-qubit gates. This can be found by solving for  $c_1, c_2$ , and  $c_3$  in our parametrization Eq. (A4): there are several solutions, but one of them is given by

$$c_1 = \frac{\pi}{4}, c_2 = \frac{-1}{2} \tan^{-1} \sqrt{\frac{2}{3}}, c_3 = \frac{\pi}{2} + c_2. \quad (\text{A28})$$

### Statistical model rules for noise channels

Similarly to the update for unitary gates, the update for the noise channel can be obtained by performing an average over a single Haar-random single-qudit gate  $V$  in the two copies of the system. There are two distinct rules, depending on whether the noise acts on one or both copies of the system. If noise acts on one copy, then the necessary average  $\mathcal{N}_1(\rho^{(2)})$  takes the form

$$\sum_a \mathbb{E}_V \left[ \begin{array}{c} \text{---} \text{---} \\ \text{---} \text{---} \\ \text{---} \text{---} \\ \text{---} \text{---} \end{array} \right], \quad (\text{A29})$$

where  $\mathcal{K}_a$  are the Kraus operators associated with the noise channel  $\mathcal{N}$ . If noise acts on both copies, then the average  $\mathcal{N}_2(\rho^{(2)})$  instead takes the form

$$\sum_{a,b} \mathbb{E}_V \left[ \begin{array}{c} \text{---} \text{---} \\ \text{---} \text{---} \\ \text{---} \text{---} \\ \text{---} \text{---} \end{array} \right]. \quad (\text{A30})$$

We begin with the expression for a single-qubit Haar average of an operator  $O$ :

$$\mathbb{E}_V [(V \otimes V)O(V \otimes V)^\dagger] =: \bar{O} = aI + bS, \text{ where} \\ a = \frac{\text{Tr}(O) - \frac{1}{q} \text{Tr}(OS)}{q^2 - 1}, \quad b = \frac{\text{Tr}(OS) - \frac{1}{q} \text{Tr}(O)}{q^2 - 1}. \quad (\text{A31})$$

Since the Kraus operators are trace-preserving, we have  $\text{Tr}(\mathcal{N}_1(O)) = \text{Tr}(\mathcal{N}_2(O)) = \text{Tr}(O)$ . The only parameters that govern the noise dynamics are then  $\text{Tr}(\mathcal{N}_1(O) \cdot S)$  and  $\text{Tr}(\mathcal{N}_2(O) \cdot S)$  for  $O \in \{I, S\}$ .

Gate	Gate invariants		$\{I/q^2, \mathcal{S}/q\}$ basis		$\{I, \Omega\}$ basis	
	$ G_1 $	$G_2$	$\alpha$	$\beta$	$R$	$D$
Controlled- $U$	$G_1$	$2G_1+1$	$\frac{10}{9}(1- G_1 )$	$\frac{-1}{18} - \frac{G_2}{6} + \frac{5}{9} G_1 $	$\frac{2}{3}(1- G_1 )$	$\frac{5}{6} - \frac{G_2}{6} - \frac{1}{3} G_1 $
CNOT	0	1	$\frac{10}{9}$	$\frac{-2}{9}$	$\frac{2}{3}$	$\frac{2}{3}$
SWAP	1	-3	0	1	0	1
iSWAP	0	-1	$\frac{10}{9}$	$\frac{1}{9}$	$\frac{2}{3}$	1
Haar-random gate	-	-	1	0	$\frac{3}{5}$	$\frac{4}{5}$
Single-qubit gate	1	3	0	0	0	0
fSIM( $\theta, \varphi$ )	$\frac{1}{4}(1 + \cos^2 2\theta + 2 \cos 2\theta \cos \varphi)$	$2 \cos 2\theta + \cos \varphi$	$\frac{5}{36}(5 - \cos 4\theta - 4 \cos 2\theta \cos \varphi)$	$\frac{1}{72}(11 + 5 \cos 4\theta - 24 \cos 2\theta + 20 \cos 2\theta \cos \varphi - 12 \cos \varphi)$	$\frac{1}{12}(5 - \cos 4\theta - 4 \cos 2\theta \cos \varphi)$	$\frac{1}{24}(17 - \cos 4\theta - 8 \cos 2\theta - 4 \cos \varphi)$
fSIM( $\theta = \frac{\pi}{2}, \varphi$ )	$\frac{1}{2}(1 - \cos \varphi)$	$-2 + \cos \varphi$	$\frac{5}{9}(1 + \cos \varphi)$	$\frac{1}{9}(5 - 4 \cos \varphi)$	$\frac{1}{3}(1 + \cos \varphi)$	1
fSIM( $\theta = 0, \varphi$ )	$\frac{1}{2}(1 + \cos \varphi)$	$2 + \cos \varphi$	$\frac{5}{9}(1 - \cos \varphi)$	$\frac{-1}{9}(1 - \cos \varphi)$	$\frac{1}{3}(1 - \cos \varphi)$	$\frac{1}{3}(1 - \cos \varphi)$
Special perfect entanglers: $c_1 = \pi/2, c_2 = 2\phi - \frac{\pi}{2}, c_3 = 0$	0	$-\cos 4\phi$	$\frac{10}{9}$	$\frac{1}{18}(-1 + 3 \cos 4\phi)$	$\frac{2}{3}$	$\frac{1}{6}(5 + \cos 4\phi)$

TABLE I: Gate invariants and transfer matrix parameters for some common gates.

Consider  $\text{Tr}(\mathcal{N}_1(I) \cdot \mathcal{S})$ . After noise acts on the state, the operator becomes

$$\mathcal{N}_1(I/q^2) = \mathbb{E}_V \left[ \begin{array}{c} \text{---} \boxed{V} \text{---} \boxed{\rho_1} \text{---} \boxed{V^\dagger} \text{---} \\ \text{---} \boxed{V} \text{---} \boxed{I/q} \text{---} \boxed{V^\dagger} \text{---} \end{array} \right], \quad (\text{A32})$$

where  $\rho_1$  is the state the identity  $I/q$  maps to under the nonunital noise. Therefore  $\text{Tr}(\mathcal{N}_1(I/q^2) \cdot \mathcal{S}) = \text{Tr}(\rho_1 \cdot I/q) = \text{Tr} \rho_1/q = 1/q$ . Thus, this parameter is also independent of the noise strength and type, giving us three remaining parameters.

Let us denote these as

$$\text{Tr}(\mathcal{N}_2(I) \cdot \mathcal{S}) =: q + \mu, \quad (\text{A33})$$

$$\text{Tr}(\mathcal{N}_1(\mathcal{S}) \cdot \mathcal{S}) =: Y_1, \quad (\text{A34})$$

$$\text{Tr}(\mathcal{N}_2(\mathcal{S}) \cdot \mathcal{S}) =: Y_2. \quad (\text{A35})$$

In terms of these parameters, from Eqs. (A31) and (A33) to (A35), we have

$$\mathcal{N}_2 \left[ \frac{I}{q^2} \right] = \left( 1 - \frac{\mu}{q(q^2-1)} \right) \frac{I}{q^2} + \frac{\mu}{q(q^2-1)} \frac{\mathcal{S}}{q}, \quad (\text{A36})$$

$$\mathcal{N}_1 \left[ \frac{\mathcal{S}}{q} \right] = \frac{q^2 - Y_1}{q^2 - 1} \frac{I}{q^2} + \frac{Y_1 - 1}{q^2 - 1} \frac{\mathcal{S}}{q}, \quad (\text{A37})$$

$$\mathcal{N}_2 \left[ \frac{\mathcal{S}}{q} \right] = \frac{q^2 - Y_2}{q^2 - 1} \frac{I}{q^2} + \frac{Y_2 - 1}{q^2 - 1} \frac{\mathcal{S}}{q}. \quad (\text{A38})$$

The parameters are chosen so that in the case of unital noise, we have  $\mu = 0$ . The interpretation of  $\mu$  is that  $\text{Tr}(\mathcal{N}_2(I) \cdot \mathcal{S}) = q^2 \text{Tr}[\rho_1^2] = q + \mu$ . Therefore  $\mu = q^2 \text{Tr}[\rho_1^2] - q$ , which measures how impure the state  $\rho_1$  is. We call  $\mu$  the *nonunitality*.

Dalzell *et al.* [12] have derived the relation between  $Y_1$  and  $Y_2$  and the noise channel's infidelity and nonunitarity for unital noise channels. We follow their derivation and obtain generalized expressions for nonunital noise as well.

Before coming to the evolution of the SWAP operator, let us relate the channel's infidelity and nonunitarity to the quantities  $Y_1$  and  $Y_2$ . From Ref. [12], the average infidelity is given by

$$r = 1 - \mathbb{E}_V \text{Tr} [V |\psi\rangle\langle\psi| V^\dagger \mathcal{N}_1(V |\psi\rangle\langle\psi| V^\dagger)] \quad (\text{A39})$$

$$= 1 - \text{Tr} \left( \mathcal{S} \mathcal{N}_1 \left( \frac{I + \mathcal{S}}{q(q+1)} \right) \right) \quad (\text{A40})$$

$$= 1 - \frac{1}{(q^2-1)(q+1)} \left( q(q^2-1) + Y_1 \frac{q^2-1}{q} \right) \quad (\text{A41})$$

$$= \frac{q - Y_1/q}{q+1}, \text{ implying} \quad (\text{A42})$$

$$Y_1 = q^2 - rq(q+1). \quad (\text{A43})$$

Similarly, the unitarity is given by

$$u = \frac{q}{q-1} \left( \mathbb{E}_V \text{Tr} [(\mathcal{N}_2(V |\psi\rangle\langle\psi| V^\dagger))^2] - \frac{1}{q} \right) \quad (\text{A44})$$

$$= \frac{q}{q-1} \text{Tr} \left[ \mathcal{S} \mathcal{N}_2 \left( \frac{I + \mathcal{S}}{q(q+1)} \right) \right] - \frac{1}{q-1} \quad (\text{A45})$$

$$= \frac{Y_2 + \mu - 1}{q^2 - 1}, \text{ implying} \quad (\text{A46})$$

$$Y_2 = (q^2 - 1)u + 1 - \mu. \quad (\text{A47})$$

Thus, the transfer matrices are as follows:

$$N^{(1)}(\mathcal{N}) = \begin{bmatrix} 1 & \gamma_1 \\ 0 & 1 - \gamma_1 \end{bmatrix}, \quad (\text{A48})$$

where  $\gamma_1 = qr/(q-1)$ .

With two-copy noise, we have instead

$$N^{(2)}(\mathcal{N}) = \begin{bmatrix} 1 - \delta_2 & \gamma_2 \\ \delta_2 & 1 - \gamma_2 \end{bmatrix}, \quad (\text{A49})$$



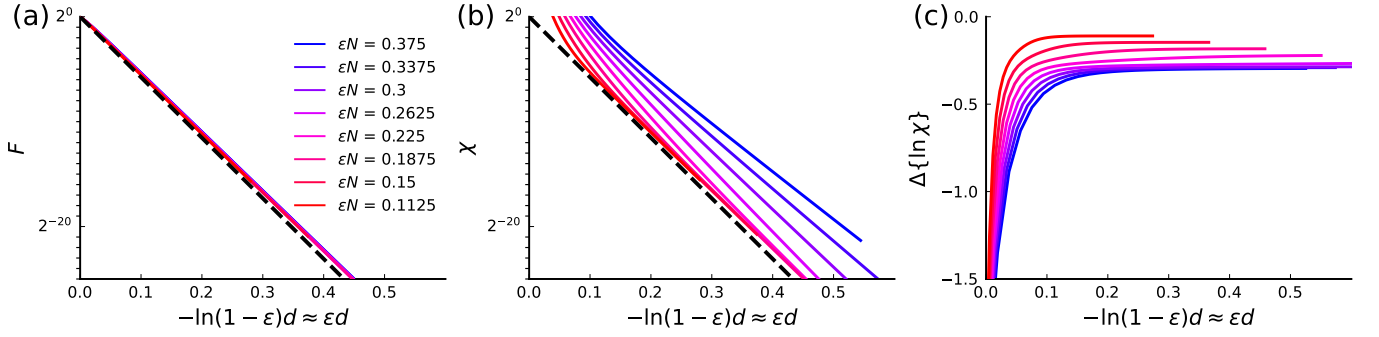


FIG. 7: Fidelity and XEB in 1D circuits with  $N = 40$  qubits and Haar-random gates, for various values of the noise rate  $\epsilon$  (see color code in (a)), plotted against rescaled circuit depth. (a) Fidelity  $F$  curves collapse onto the global-white noise prediction  $(1 - \epsilon)^{Nd}$  until saturating at  $q^{-N}$ . (b) Unnormalized XEB  $\chi$ . For low noise  $\epsilon N \lesssim 0.22$ , the XEB decay rate matches that of fidelity. (c) Decay rate of the XEB. For  $\epsilon N \gtrsim 0.22$ , the decay rate saturates to a  $\epsilon N$ -independent constant.

where

$$\delta_2 = \frac{\mu}{q(q^2 - 1)}, \quad (\text{A50})$$

$$\gamma_2 = 1 - u + \frac{\mu}{q^2 - 1}. \quad (\text{A51})$$

### Appendix B: Purity and collision probabilities for noisy circuits

In this Appendix, we consider second-moment properties of the density matrix in which both copies experience noise. As such, we no longer have the liberty to only consider the infidelity  $\gamma$  of the noise channel  $\mathcal{N}$ .

The averaged purity of the noisy evolution has the same form as fidelity except both copies of the circuit are evolved with the noisy evolution:

$$\begin{aligned} P(\mathcal{N}) &= \mathbb{E}_{\mathbf{U}} \text{Tr}(\rho_{\mathbf{U}}(\mathcal{N})\rho_{\mathbf{U}}(\mathcal{N})) \\ &= \mathbb{E}_{\mathbf{U}} \text{Tr}(\mathcal{S}^{\otimes N} \rho_{\mathbf{U}}(\mathcal{N}) \otimes \rho_{\mathbf{U}}(\mathcal{N})). \end{aligned} \quad (\text{B1})$$

Similarly, the collision probability of the noisy circuit has the same form as the linear cross entropy benchmark except both copies are evolved with the noisy evolution:

$$\begin{aligned} Z(\mathcal{N}) &= \mathbb{E}_{\mathbf{U}} \sum_{\mathbf{x}} \langle \mathbf{x} | \rho_{\mathbf{U}}(\mathcal{N}) | \mathbf{x} \rangle \langle \mathbf{x} | \rho_{\mathbf{U}}(\mathcal{N}) | \mathbf{x} \rangle \\ &= \mathbb{E}_{\mathbf{U}} \text{Tr}(\mathcal{P}^{\otimes N} \rho_{\mathbf{U}}(\mathcal{N}) \otimes \rho_{\mathbf{U}}(\mathcal{N})). \end{aligned} \quad (\text{B2})$$

When both copies of the density matrix experience noise, the update rules for noise take the form of a Markov matrix  $N^{(2)}$ :

$$N^{(2)} = \begin{bmatrix} 1 - \delta_2 & \gamma_2 \\ \delta_2 & 1 - \gamma_2 \end{bmatrix}. \quad (\text{B3})$$

Derivations of formulas for  $\gamma_2$  and  $\delta_2$  are shown in Appendix A.

For unital noise,  $\delta_2 = 0$ , and thus  $N^{(2)}$  has the same form as  $N(\gamma)$  considered throughout the paper. In this

case, the update rules for computing purity and the collision probability are identical to those of computing fidelity and the linear XEB, but with  $\gamma_2$  in place of  $\gamma$ :

$$P(\gamma_2, \delta_2 = 0) = F(\gamma = \gamma_2), \quad (\text{B4})$$

$$Z(\gamma_2, \delta_2 = 0) = X(\gamma = \gamma_2). \quad (\text{B5})$$

Thus, the phase transition described above for XEB can be seen additionally in the collision probability of noisy circuits with unital noise.

However, realistic digital quantum devices experience non-unital noise. Non-zero  $\delta_2$  dramatically changes the physics of the statistical model, as the maximally mixed density matrix is no longer the equilibrium state reached at large depths. We leave further considerations of non-unital noise to future work.

### Appendix C: Additional numerical data

In Fig. 7, we illustrate that, like the all-to-all circuits, one dimensional circuits show fidelity that decays like  $(1 - \epsilon)^{Nd}$  for nearly all depths before saturation, while the XEB decay rate goes through a transition. The corresponding decay rate of XEB is shown in Fig. 3(c) to match the leading eigenvalue  $\Lambda_1$  of the transfer matrix and to experience a kink at the eigenvalue crossing, just as in the all-to-all circuit. Thus, we see that the same phenomenology that describes the all-to-all circuit appears in the 1D circuit.

While the fidelity decays like  $(1 - \epsilon)^N$  until saturation at  $q^{-N}$ , we can additionally study the shifted-and-rescaled fidelity  $f = q^N F - 1$ . By subtracting off the saturation value, we see in Fig. 8, at sufficiently large depths, the  $(1 - \epsilon)^N$  decay is dominated by an  $N$ -independent decay. This confirms that the coupling constants between fidelity and locally-gapped eigenvectors is non-zero.

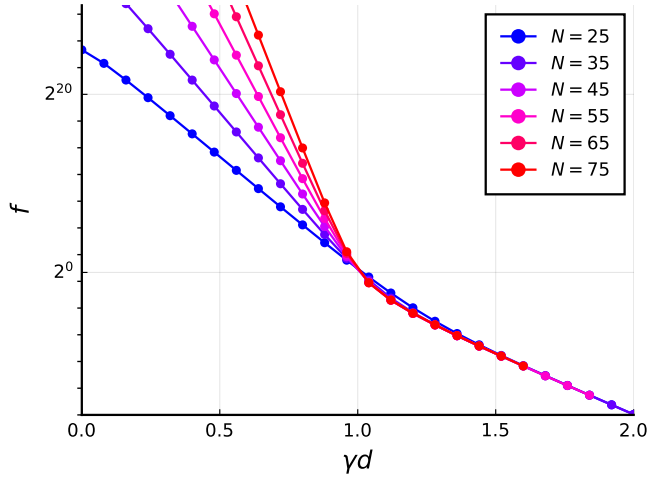


FIG. 8: Shifted fidelity in one-dimensional circuits with Haar random gates and  $\varepsilon = 0.03$ , for various system sizes  $N$ , plotted versus rescaled depth  $\gamma d = \frac{4}{3}\varepsilon d$ .