


Efficient Sampling of Noisy Shallow Circuits Via Monitored Unraveling

Zihan Cheng^{1,*} and Matteo Ippoliti^{1,2}

¹*Department of Physics, University of Texas at Austin, Austin, Texas 78712, USA*

²*Department of Physics, Stanford University, Stanford, California 94305, USA*

 (Received 18 July 2023; accepted 19 October 2023; published 15 November 2023)

We introduce a classical algorithm for sampling the output of shallow, noisy random circuits on two-dimensional qubit arrays. The algorithm builds on the recently proposed “space-evolving block decimation” (SEBD) [Napp *et al.*, *Phys. Rev. X* **12**, 021021 (2022)] and extends it to the case of noisy circuits. SEBD is based on a mapping of two-dimensional unitary circuits to one-dimensional *monitored* ones, which feature measurements alongside unitary gates; it exploits the presence of a measurement-induced entanglement phase transition to achieve efficient (approximate) sampling below a finite critical depth T_c . Our noisy-SEBD algorithm unravels the action of noise into measurements, further lowering entanglement and enabling efficient classical sampling up to larger circuit depths. We analyze a class of physically relevant noise models (unital qubit channels) within a two-replica statistical mechanics treatment, finding weak measurements to be the optimal (i.e., most disentangling) unraveling. We then locate the noisy-SEBD complexity transition as a function of circuit depth and noise strength in realistic circuit models. As an illustrative example, we show that circuits on heavy-hexagon qubit arrays with noise rates of approximately equal to 2% per CNOT, based on IBM Quantum processors, can be efficiently sampled up to a depth of five *i*SWAP (or ten CNOT) gate layers. Our results help sharpen the requirements for practical hardness of simulation of noisy hardware.

DOI: [10.1103/PRXQuantum.4.040326](https://doi.org/10.1103/PRXQuantum.4.040326)

I. INTRODUCTION

Before quantum computers can do anything useful, they must be able to reliably beat classical computers at some task, ideally one that is quantifiable, well-understood theoretically, and has reasonable experimental requirements. *Random circuit sampling* (RCS) has emerged as one of the leading candidates for this role: it is well suited to the architectures of present-day gate-based quantum processors, and—at least in the ideal case of noiseless computation—its hardness is firmly established in complexity theory [1–5]. As a result, it has become the focus of pioneering experimental efforts in the last few years [6–9]. All such experiments, however, are by necessity carried out on present-day noisy intermediate-scale quantum (NISQ) processors, where the question of classical simulation hardness is much more nuanced. This has spurred much interest in the exploration of how noise affects the boundary between “easy” and “hard” simulation problems.

It is now established that RCS in the presence of a finite noise rate can be simulated in polynomial time [10] in the regime of anticoncentration (i.e., informally, at sufficiently large depth [11]). However the algorithm of Ref. [10] is not practical, leaving open the question of simulability of finite-sized, noisy RCS experiments with reasonable classical resources. This issue is subtle and depends on many variables, such as circuit architecture, size and depth, details of the noise models, choice of target metrics, etc. A powerful classical approach is based on *tensor networks*, which leverage limited entanglement and work best in one dimension (1D); for this reason, experiments have focused on two-dimensional (2D) qubit arrays and picked highly entangling gate sets. Large circuit depth also generates more entanglement and thus makes simulation harder. However, in practice, depth is limited by the presence of noise, as more gates also cause the accumulation of more errors. Additionally, the presence of noise in the quantum experiment lowers the bar for classical simulation: an apples-to-apples comparison requires that we tolerate similar levels of error from the classical algorithm as well. This opens the door to various strategies based on tensor networks that can outperform sufficiently noisy quantum computers [12–14].

Characterizing the practical hardness of noisy RCS problems is thus a pressing question in quantum

*zihan_cheng@utexas.edu

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

information science. It is also closely related to recent developments in nonequilibrium many-body physics regarding dynamical phases of quantum information in open systems. In particular, the fate of unitary circuits that are sampled *during* the dynamics (a special case of open-system evolution) has attracted much interest due to the discovery of entanglement phases that occur as a function of the measurement rate [15–44]. The emergence of a phase with limited entanglement (area law) at high measurement rate in these circuits has been used to develop an efficient algorithm for sampling the output of shallow unitary circuits in 2D [45]. The algorithm, dubbed “space-evolving block decimation” (SEBD) and sketched in Fig. 1(a), works at depths T below a finite critical depth T_c (model dependent, but approximately equal to 4–5 in typical models). Circuits with $T = 3$ in 2D are capable of universal quantum computation and thus hard to sample in the worst case [46], making this result surprising. Another very recent development in dynamical phases of quantum information is the discovery of a sharp noise-induced phase transition in RCS [9,47], which was identified at a fixed number of errors per circuit layer (i.e., noise strength $\varepsilon \sim 1/N$, N being the number of qubits). In the strong-noise phase, it was argued [9] that the processor’s output

can be classically “spoofed,” while in the weak-noise phase simulation is conjectured to be practically hard.

In this work, we consider the problem of noisy RCS on shallow 2D circuits, Figs. 1(a)–1(b), where the results of Ref. [10] are not applicable. Unlike other works, our goal is not to sample from the ideal output within some tolerance determined by the noise; instead, we aim to accurately sample the noisy output itself—a closely related but different problem. To this end, we develop a classical algorithm, dubbed noisy SEBD, and show that it undergoes a complexity phase transition (from quadratic to exponential in the linear size of the system) as a function of circuit depth T and noise strength ε . The physical principle behind the algorithm is that noise can be viewed as a sequence of (fictitious) measurements done by the environment on the system, Fig. 1(c). Measurements can lower the amount of entanglement in the system and drive a phase transition to an area-law entangled phase, where tensor-network simulation is efficient; as a consequence, noise can drive a phase transition in the complexity of noisy SEBD. Since the unraveling of noise into measurements is not unique, we are free to optimize it in order to lower this threshold as much as possible. In this work we use a two-replica statistical mechanics model to optimize the unraveling, finding

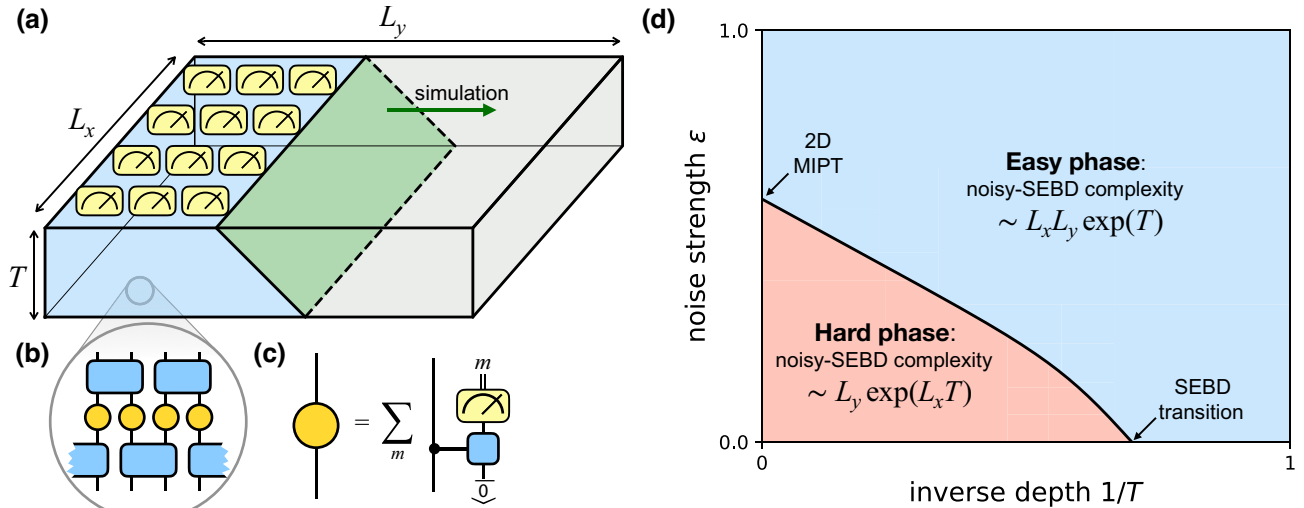


FIG. 1. Main ideas of this work. (a) Sampling of 2D shallow unitary circuits. The qubit array has linear dimensions L_x , L_y and the circuit has depth T . The SEBD algorithm is based on an MPS simulation carried out along a spatial direction (e.g., L_y), approximating the wave function of a 1D subsystem of qubits on a light cone surface (green). It exploits the measurements on the top boundary of the circuit to disentangle the wave function. (b) We consider noisy circuits with uncorrelated local noise: unitary gates (blue rectangles) are interspersed with single-qubit noise channels (orange circles). (c) The noise channels are unraveled as measurements (generically weak), represented here as interactions with a fresh ancilla, which is then measured. These fictitious measurements can further disentangle the wave function. (d) Qualitative sketch of the complexity phase diagram of the noisy-SEBD algorithm. An entanglement phase transition separates an “easy” phase (low depth and/or strong noise), where the computational cost of sampling the noisy circuit output via noisy SEBD scales as $L_x L_y \exp(T)$, from a “hard” phase where the same cost scales as $L_y \exp(L_x T)$. The location of the phase boundary is model dependent. The line $\varepsilon = 0$ yields the finite-depth complexity transition of Ref. [45] (noiseless SEBD) while the $1/T = 0$ line [which stands for $T = O(L)$] yields the standard measurement-induced phase transition in two spatial dimensions [24]. We conjecture the scaling $\varepsilon_c(T) \sim \varepsilon_{c,2D} + O(1/T)$ at large T . We emphasize that this is a transition in the complexity of the noisy-SEBD algorithm, *not* of the sampling task itself [10].

weak measurements to be more disentangling than stochastic projective measurements; the effect of this optimized choice lowers the threshold noise rate by as much as a factor of 2, thus significantly expanding the “easy” phase.

Combined with the depth-induced complexity transition in the original (noiseless) SEBD algorithm [45], this defines a phase boundary in the space of circuit depth T and noise strength ε , sketched in Fig. 1(d). This adds to our growing understanding of the boundaries of “practical” simulability of noisy quantum systems. Moreover, it places sharp constraints on the possibility of achieving beyond-classical computation by scaling RCS experiments in space only—i.e., by growing quantum processor size at fixed circuit depth. This highlights the importance of further improvements in error rates of NISQ hardware.

The paper is structured as follows. Section II reviews background material on random circuit sampling, matrix-product state simulation methods, the SEBD algorithm and monitored dynamics. In Sec. III we discuss the unraveling of noise into monitored trajectories and the choice of entanglement-optimal unravelings, including explicit solutions for unital qubit channels. Section IV presents the noisy-SEBD algorithm, numerical simulations of its complexity phase transition, and an illustrative application to circuits based on IBM Quantum’s heavy-hexagon qubit arrays. We conclude in Sec. V by summarizing our results, their implications and connections with other works, and directions for future research.

II. BACKGROUND

A. Random circuit sampling

Random circuit sampling (RCS) has emerged as a leading candidate for early demonstrations of quantum computational supremacy, combining good fit with existing hardware capabilities and robust complexity-theoretic arguments for classical hardness [1–5]. The idea is to draw a random instance U from an ensemble of local unitary circuits of depth T , run it on a quantum computer prepared in the initial state $|\mathbf{0}\rangle \equiv |0\rangle^{\otimes N}$, and measure the state of each qubit in the computational basis, obtaining a bitstring $\mathbf{z} \in \{0, 1\}^N$. Ideally, this process samples bitstrings from the probability distribution $P_U(\mathbf{z}) = |\langle \mathbf{z} | U | \mathbf{0} \rangle|^2$, which is computationally hard to do for classical computers. Intuitively, this is due to the production of extensive entanglement in the system over the course of typical instances of the unitary evolution U [48], which causes tensor-network classical algorithms to fail. At the same time, sufficiently generic ensembles of unitary gates in U ensure that various other strategies for efficient classical simulation (such as stabilizers [49] or matchgates [50–52]) are not viable.

These theoretical insights have motivated pioneering experimental efforts in the past few years to demonstrate RCS-based quantum computational supremacy on

present-day NISQ hardware [6–9]. Verification of successful RCS is nontrivial. The experiments employ a linear cross-entropy diagnostic

$$\text{XEB} = 2^N \sum_{\mathbf{z}} p_{\text{expt.}}(\mathbf{z}) p_U(\mathbf{z}) - 1, \quad (1)$$

where $p_U(\mathbf{z})$ is the previously defined ideal distribution (to be computed classically), and $p_{\text{expt.}}(\mathbf{z})$ is the distribution of bitstrings obtained from the experiment, which generally differs from the ideal one due to imperfect implementation and uncontrolled noise. This quantity is convenient as it can be estimated by sampling from the experiment:

$$\text{XEB} = 2^N \langle p_U(\mathbf{z}) \rangle_{\mathbf{z} \sim p_{\text{expt.}}} - 1. \quad (2)$$

Furthermore, it is designed in such a way that $\text{XEB} = 1$ if the experiment is perfect ($p_{\text{expt.}} = p_U$, assuming the Porter-Thomas distribution), while $\text{XEB} = 0$ if it is completely noisy ($p_{\text{expt.}}(\mathbf{z}) = 2^{-N}$ for all \mathbf{z}). In real-world conditions, with finite noise per gate, it has been argued that $\text{XEB} \sim f^{NT}$ in the regime of interest to the experiment [6], where f is the average fidelity per gate (more detailed results on the conversion of local noise into global depolarizing noise have been obtained subsequently [53,54]). This lowers the bar for classical algorithms: they do not need to achieve a perfect score $\text{XEB} = 1$, but only to exceed the fidelity of the NISQ experiment.

This has led to a flurry of activity in the past few years to develop approximate classical algorithms that can efficiently simulate RCS with XEB scores and runtimes comparable to those of the NISQ experiments [12–14,55,56]. Furthermore, it was recently shown that there exists a polynomial-time algorithm for RCS with constant noise per gate [10], which uses a Feynman path-integral representation for the amplitudes $\langle \mathbf{z} | U | \mathbf{0} \rangle$ wherein noise damps the contribution of most paths. While this settles the complexity of noisy RCS formally, at least in the regime of anticoncentration [11], the algorithm comes with a very large exponent and is expected to be impractical at the relevant noise strengths. Thus the *practical* issue of efficient classical simulation of finite-sized noisy circuits remains open.

B. MPS simulation and the entanglement barrier

The interplay of entanglement and noise and its effects on the complexity of classical simulation become especially transparent in the case of matrix-product state (MPS) simulations of random circuits, limited to 1D (or quasi-1D) geometries [57–59]. The idea is to represent a wave function $|\psi\rangle = \sum_{\mathbf{z}} c_{\mathbf{z}} |\mathbf{z}\rangle$ as a product of three-index tensors via $c_{\mathbf{z}} = \text{Tr}(A_1^{z_1} A_2^{z_2} \cdots A_N^{z_N})$, where each A_i^z is a $\chi \times \chi$ matrix (i labels position in the chain, z is the physical state $|z\rangle$, and two virtual indices are implicit). The *bond dimension* χ is a cutoff on the Schmidt rank of the state $|\psi\rangle$, which

makes the state classically representable. The MPS ansatz allows approximate simulations with high accuracy whenever the entanglement entropy [60] of $|\psi\rangle$ about each cut obeys $S \ll \ln(\chi)$. Given the linear growth of entanglement in random circuits [48], the MPS method enables accurate simulation for short circuit depths $T \lesssim \ln(\chi)$, after which the truncation of bond dimension incurs a large error. One can nonetheless carry out finite- χ simulations for larger depths; the effect of MPS truncation error on the fidelity with the true state is found to be qualitatively similar to the effect of noise in the quantum experiments [12,14]. Thus if noise strength is large enough, classical MPS simulations may beat NISQ experiments at the task of approximating a given ideal random circuit.

A different task is to classically simulate (or sample from) noisy circuits themselves. Here, classical simulation needs to incorporate noise and thus mixed states. This can be accomplished by using matrix-product operators (MPOs): a mixed state $\rho = \sum_{\mathbf{z}, \mathbf{z}'} c_{\mathbf{z}, \mathbf{z}'} |\mathbf{z}\rangle\langle \mathbf{z}'|$ is represented as a product of four-index tensors via $c_{\mathbf{z}, \mathbf{z}'} = \text{Tr}(A_1^{z_1, z'_1} A_2^{z_2, z'_2} \cdots A_N^{z_N, z'_N})$, where each $A_i^{z, z'}$ is a $\chi \times \chi$ matrix. [A representation that guarantees positivity is given by *matrix-product density operators* (MPDOs) [61].] In the presence of noise and unstructured random gates, the unique steady state is expected to be the maximally mixed state $\rho = \mathbb{I}/2^N = \bigotimes_i (\mathbb{I}/2)_i$, which is manifestly disentangled, and can be written as an MPO of bond dimension $\chi = 1$ (i.e., the tensors $A_i^{z, z'} \equiv \delta_{z, z'}/2$ carry no virtual indices). Thus entanglement grows initially due to random unitary interactions, $S \sim T$; this persists until the effects of noise are felt, at depth $T \sim 1/p$, p being noise strength; after that, S decreases to zero. Thus to accurately simulate the noisy dynamics one has to overcome an “entanglement barrier” of height $S \sim 1/p$, with computational cost $\text{poly}(N, \exp(1/p))$ [62,63]. While polynomial in the system size N , for realistic noise strength $p \sim 10^{-2}$ the cost can still be prohibitively large. Moreover, the efficient MPO simulation only applies to 1D; in higher dimension, small entanglement generally does not guarantee efficient simulation [59].

C. 2D shallow circuits and the SEBD algorithm

Due to MPS simulations, 1D circuit architectures require large depths [diverging faster than $\log(N)$] for hardness. In contrast, 2D circuits could be hard already for finite depth T , as tensor-network methods in two or more dimensions are generally not efficient. As hardness of simulation generally scales exponentially in the treewidth of the tensor network, experimental RCS works in two-dimensional circuits [6–9] set $T \sim \sqrt{N}$ to maximize classical hardness. Still, it is natural to ask at what depth T the classical simulation would become hard (asymptotically in large N).

It is straightforward to note that $T \leq 2$ is easy, as the output state $U|\mathbf{0}\rangle$ is given by a tensor product of decoupled

dimers ($T = 1$) or one-dimensional subsystems each hosting an MPS of finite bond dimension ($T = 2$). However, starting at depth $T = 3$, it is possible to prepare states whose exact simulation is provably hard [46].

Surprisingly, Ref. [45] has shown that *approximate sampling* from 2D circuits up to a finite depth $T_c \geq 3$ is in fact possible with polynomial time classical algorithms. One of the methods they propose, dubbed *space-evolving block decimation*, is based on reducing the sampled $(2 + 1)$ D circuit to a $(1 + 1)$ D circuit featuring measurements alongside unitary gates. Below a critical depth T_c , the state that needs to be simulated classically obeys an area law for entanglement [64] and can thus be accurately simulated via MPS methods. The emergence of this low-entanglement phase is an instance of a general phenomenon taking place in *monitored dynamics*, i.e., time evolution that combines unitary interactions and measurements, which we review next.

D. Monitored dynamics

Measurements can disentangle a quantum state. Given a many-body wave function $|\psi\rangle$, measuring a qubit i in the computational basis leaves behind a product state between i and the rest of the system: $\propto |z\rangle_i \otimes \langle z|\psi\rangle_{\neg i}$ ($z \in \{0, 1\}$ is the measurement outcome, obtained randomly from the Born rule, and $\neg i$ denotes the rest of the system). This not only disentangles i , but may also destroy entanglement globally—as an extreme example, measuring any one qubit of a GHZ state $|\psi\rangle = \frac{1}{\sqrt{2}}(|0\rangle^{\otimes N} + |1\rangle^{\otimes N})$ in the computational basis leaves behind a global product state. At the same time, in systems with local interactions, entanglement is generated locally. This asymmetry between entanglement creation and destruction suggests that monitored dynamics, featuring a finite rate p of measurements alongside local unitary interactions, should generically lead to states with low entanglement, for any rate $p > 0$ [18].

Surprisingly, it was found that monitored dynamics can instead successfully stabilize a highly entangled phase as long as the rate of measurement p is below a critical threshold $p_c > 0$ [15–17]. The phases are characterized by the structure of entanglement in pure output states of the dynamics, $|\psi_{\mathbf{m}}\rangle$, which are indexed by the measurement record \mathbf{m} (the sequence of classical outcomes collected during the dynamics). In particular, the scaling of entanglement entropy S_A for a subsystem A in these states undergoes a transition from an area law $S_A \sim |\partial A|$ (∂A is the boundary of A) to a volume law $S_A \sim |A|$. These scalings are generally washed out in the mixed state $\rho = \sum_{\mathbf{m}} p_{\mathbf{m}} |\psi_{\mathbf{m}}\rangle\langle \psi_{\mathbf{m}}|$ obtained by discarding the measurement record.

The stability of entanglement in the volume-law phase ($p < p_c$) is explained qualitatively by the emergence of a quantum error-correcting code that manages to hide information from future measurements for a long time

[21,25,65]. This coding perspective is illustrated concretely by the behavior of a reference qubit R initially entangled with a monitored many-body system [26]. Let $S_R(t)$ be the entanglement entropy of the reference qubit as a function of time t in the monitored dynamics; at late times one generally has $S_R(t) \sim e^{-t/\tau}$, with a time constant τ dubbed the *purification time*. The behavior of τ changes sharply at the transition. In the volume-law phase it obeys $\tau \sim \exp(N)$, signifying a successful encoding of at least some information about the state of R , which is protected from measurements for a long time. In the area-law phase τ becomes $O(1)$, showing that the encoding fails. Finally, at the critical point ($p = p_c$) τ diverges algebraically, $\tau \sim N^{z/d}$, with z a dynamical critical exponent (typically $z = 1$ [40]) and d the spatial dimension. In the following we will make use of the purification time as a practical diagnostic for the underlying entanglement phase, as it is numerically more efficient to compute than the entanglement entropy of large subsystems.

In one dimension, the entanglement phase transition also corresponds to a transition in the classical simulation complexity of the dynamics: area-law states in 1D have constant entanglement, and can be efficiently simulated with MPS methods. This is the principle behind the SEBD algorithm [45], in which a 2D sampling task is reduced to a $(1+1)$ D monitored dynamics and simulated by MPS methods in the area-law phase. Reference [66] has extended this approach to continuous-time Markovian open-system dynamics: the system-environment coupling, in the form of a Lindbladian master equation, can be “unraveled” into trajectories [67] (stochastic pure-state evolutions) that contain quantum measurements, which in turn can lower entanglement and help the accuracy of MPS simulation. As the unraveling is a simulation artifact and is not physical, it can be optimized to minimize the amount of entanglement. Reference [66] proposes an adaptive scheme that chooses the unraveling with the lowest expected value of the entropy at each position and time during the evolution; above a threshold noise strength, the trajectories enter an area-law phase and efficient MPS simulation becomes possible.

In this work we build on the approaches of Refs. [45,66] to address the problem of sampling from *noisy, shallow* circuits in 2D.

III. UNRAVELING NOISE INTO MONITORED TRAJECTORIES

In this section we review the basics noise models and their unraveling and discuss the entanglement-optimal unraveling to use in the noisy-SEBD algorithm.

A. Noise models and unraveling

In quantum computers, it is often a good approximation to treat the inevitable interactions with the environment

as Markovian noise, modeled by quantum channels (completely positive trace-preserving maps on the space of density matrices [68]). Mathematically a quantum channel Φ can be represented by as a sum of Kraus operators $\{M_i\}$,

$$\Phi(\rho) = \sum_i M_i \rho M_i^\dagger, \quad (3)$$

which must obey the trace-preservation condition:

$$\sum_i M_i^\dagger M_i = \mathbb{I}. \quad (4)$$

As examples, the dephasing channel can be represented by

$$\Phi(\rho) = (1 - \varepsilon)\rho + \varepsilon Z \rho Z, \quad (5)$$

i.e., with Kraus operators $\{\sqrt{1 - \varepsilon}\mathbb{I}, \sqrt{\varepsilon}Z\}$, and the depolarizing channel by

$$\Phi(\rho) = (1 - \varepsilon)\rho + \frac{\varepsilon}{3}(X\rho X + Y\rho Y + Z\rho Z), \quad (6)$$

with Kraus operators $\{\sqrt{1 - \varepsilon}\mathbb{I}, \sqrt{\varepsilon/3}X, \sqrt{\varepsilon/3}Y, \sqrt{\varepsilon/3}Z\}$. In both cases ε is the noise strength.

One important property of the Kraus operators’ representation is that it is not unique. For a given quantum channel, different sets of Kraus operators are equivalent under unitary transformations U :

$$M'_j = \sum_i U_{ji} M_i. \quad (7)$$

This equivalence also holds for nonsquare, semiunitary transformations U that satisfy only $U^\dagger U = \mathbb{I}$. As an important consequence, even when a channel Φ can be unraveled into unitary processes (such as the dephasing and depolarizing channels above), this equivalence allows the freedom to choose an unraveling into nonunitary operators, which correspond to measurements. For example, the dephasing channel can be unraveled into a stochastic projective measurement,

$$M_0 = \sqrt{1 - 2\varepsilon}\mathbb{I}, \quad M_1 = \sqrt{2\varepsilon}|0\rangle\langle 0|, \quad M_2 = \sqrt{2\varepsilon}|1\rangle\langle 1| \quad (8)$$

(i.e., a projective measurement of Z taking place with probability 2ε). It can also be unraveled into a weak measurement of Z ,

$$\begin{aligned} M_0 &= \sqrt{1 - \varepsilon} \cos \theta \mathbb{I} - \sqrt{\varepsilon} \sin \theta Z, \\ M_1 &= \sqrt{1 - \varepsilon} \sin \theta \mathbb{I} + \sqrt{\varepsilon} \cos \theta Z, \end{aligned} \quad (9)$$

where θ is a free parameter tuning the strength of the measurement ($\theta = 0$ returns a unitary unraveling).

B. Sampling noisy circuits

Returning to the noisy circuit problem, we wish to sample from the distribution $P_{\mathcal{N}}(\mathbf{z}) = \langle \mathbf{z} | \mathcal{N}(|\mathbf{0}\rangle\langle\mathbf{0}|) | \mathbf{z} \rangle$, with \mathcal{N} the channel describing the noisy circuit evolution:

$$\mathcal{N} = \Phi^{\otimes N} \circ \mathcal{U}_T \circ \Phi^{\otimes N} \circ \mathcal{U}_{T-1} \circ \dots \circ \Phi^{\otimes N} \circ \mathcal{U}_1. \quad (10)$$

Here Φ is a single-qubit noise of strength ε as before, while $\mathcal{U}_i(\rho) = u_i \rho u_i^\dagger$ describes the i th layer of the (ideal) shallow unitary circuit, $U = u_T u_{T-1} \dots u_1$.

By expanding each Φ into its Kraus operators, $\Phi(\rho) = \sum_i M_i \rho M_i^\dagger$, we can rewrite the probability distribution $P_{\mathcal{N}}(\mathbf{z})$ as

$$P_{\mathcal{N}}(\mathbf{z}) = \langle \mathbf{z} | \mathcal{N}(|\mathbf{0}\rangle\langle\mathbf{0}|) | \mathbf{z} \rangle = \sum_{\mathbf{m}} |\langle \mathbf{z} | \mathbb{M}_{\mathbf{m}} | \mathbf{0} \rangle|^2, \quad (11)$$

$$\mathbb{M}_{\mathbf{m}} = \left(\bigotimes_{x=1}^N M_{m_{T,x}} \cdot u_T \right) \dots \left(\bigotimes_{x=1}^N M_{m_{1,x}} \cdot u_1 \right), \quad (12)$$

where each index $m_{t,x}$ labels the Kraus operator for the noise channel Φ acting on qubit x at step t , and \mathbf{m} is a shorthand for the whole collection of indices $\{m_{t,x}\}$. Thus the operators $\{\mathbb{M}_{\mathbf{m}}\}$ are a set of Kraus operators for the channel \mathcal{N} . We can view each \mathbf{m} as a *quantum trajectory* of the evolution: the initial state $|\mathbf{0}\rangle$ evolves into the pure state $|\psi_{\mathbf{m}}\rangle \equiv \mathbb{M}_{\mathbf{m}} |\mathbf{0}\rangle / \langle \mathbf{0} | \mathbb{M}_{\mathbf{m}}^\dagger \mathbb{M}_{\mathbf{m}} | \mathbf{0} \rangle^{1/2}$ with probability $\langle \mathbf{0} | \mathbb{M}_{\mathbf{m}}^\dagger \mathbb{M}_{\mathbf{m}} | \mathbf{0} \rangle$; the true (mixed) state $\mathcal{N}(|\mathbf{0}\rangle\langle\mathbf{0}|)$ is recovered as a stochastic mixture of the trajectories. It is straightforward to see that $|\langle \mathbf{z} | \mathbb{M}_{\mathbf{m}} | \mathbf{0} \rangle|^2 \equiv P_{\mathcal{N}}(\mathbf{z}, \mathbf{m})$ is the *joint* probability of drawing trajectory \mathbf{m} and sampling bitstring \mathbf{z} at the end. Then, Eq. (11) can be written as

$$P_{\mathcal{N}}(\mathbf{z}) = \sum_{\mathbf{m}} P_{\mathcal{N}}(\mathbf{z}, \mathbf{m}), \quad (13)$$

i.e., the marginal distribution obtained by summing over trajectories.

This insight is widely used in simulations of open-system dynamics [67,69–71]. The trajectory method allows one to simulate pure states rather than density matrices, which is often much more memory efficient. This comes at the expense of having to average over many trajectories, e.g., in order to Monte Carlo sample the expectation value of a target operator. In our case, however, we only aim to *sample* from the distribution $P_{\mathcal{N}}(\mathbf{z})$, so there is no need for trajectory averaging: any joint sample (\mathbf{z}, \mathbf{m}) drawn from a simulation of the trajectory dynamics yields a valid sample \mathbf{z} from the desired distribution $P_{\mathcal{N}}(\mathbf{z})$. We emphasize that the \mathbf{m} samples and their distribution are purely mathematical artifacts of the method: the decomposition of Φ into Kraus operators includes a gauge degree of freedom that can be fixed arbitrarily. However, the

marginal distribution $P_{\mathcal{N}}(\mathbf{z})$ is gauge invariant and physical, corresponding to the true experimental distribution of bitstrings.

Beyond the practical advantage of simulating pure rather than mixed state evolution, the gauge degree of freedom in the choice of unraveling can be exploited to minimize entanglement within the trajectory [66], thus extending the applicability of tensor-network methods. Below we address the question of which unraveling yields the lowest entanglement for a given channel.

C. Entanglement-optimal unravelings

For a given model of dynamics (e.g., a Hamiltonian or an individual instance of a RUC), one can locally optimize the unraveling of each noise channel Φ separately [66]. Here we take a simpler approach, and look for a general prescription that works well *on average* over RUCs. Specifically, we aim to exploit the area-law phase in monitored dynamics (Sec. IID), which is driven by the density of measurements in the dynamics. Thus we look for an unraveling of Φ into measurements, so as to increase the effective density of measurements and facilitate a transition to the area-law phase [72]. As we have seen in Sec. III A, there are multiple inequivalent ways of unraveling noise into measurements (e.g., weak vs stochastic projective measurements); therefore, we look for the unraveling $\mathbf{M} = \{M_i\}$ of the noise channel Φ that has the strongest disentangling effect on the dynamics.

Working at a “mean-field” level in Haar-random circuits, we consider the scaling of average purity [73]. In trajectories of the dynamics within a two-replica setting. This maps onto the partition function of a \mathbb{Z}_2 Ising magnet, whose ordered and disordered phase corresponds to volume-law and area-law entanglement, respectively. Our goal is to facilitate simulation by minimizing entanglement, therefore we aim to *minimize* the couplings in the magnet. The problem is analyzed in Appendix A, where we show that the minimization of the coupling amounts to maximizing the following objective function:

$$x(\mathbf{M}) = \sum_i \frac{\text{tr}(M_i^\dagger M_i M_i^\dagger M_i)}{2 \text{tr}(M_i^\dagger M_i)}. \quad (14)$$

This has a natural physical interpretation: if we view the Kraus operators $\{M_i\}$ as *instruments* of a generalized measurement (positive operator-valued measure) $\{M_i^\dagger M_i\}$, and apply such a measurement to the fully mixed state $\mathbb{I}/2$, we obtain postmeasurement states $\rho_i \equiv M_i M_i^\dagger / \text{Tr}(M_i M_i^\dagger)$ with probabilities $\pi_i \equiv \text{Tr}(M_i^\dagger M_i) / 2$; the objective function is given by the average purity of the postmeasurement states: $x(\mathbf{M}) = \sum_i \pi_i \text{Tr}(\rho_i^2)$. Thus within this approach, the unraveling that minimizes many-body entanglement in the RUC dynamics is also the one that best purifies

a single mixed qubit. Also this is a particular feature of Haar-random circuits and likely not true of more structured models, since under Haar-random unitaries $\mathbb{I}/2$ is the natural averaged reduced density matrix for a single qubit.

We also note that this cost function (postmeasurement purity of a mixed qubit) is in fact identical to the entanglement of a Bell pair state after measuring one qubit, which was proposed in Ref. [72] as a heuristic measure of the disentangling power of different unravelings.

The optimal unraveling \mathbf{M}_{opt} is given by

$$\mathbf{M}_{\text{opt}} = \operatorname{argmax}_{\mathbf{M}} [x(\mathbf{M})], \quad (15)$$

where $\mathbf{M} = \{M_i\}$ ranges over Kraus decompositions of Φ , and thus is subject to (semi)unitary gauge freedom per Eq. (7). In general, the semiunitary gauge freedom makes the optimization nontrivial as it allows for infinitely many parameters. However, below we show that for a broad and physically relevant class of quantum channels there exists an optimal unraveling of minimal rank, which can be obtained analytically.

D. Unital qubit channels

Let us consider *unital* qubit channels, which are defined as those that leave the fully mixed state $\mathbb{I}/2$ invariant. Up to unitary transformations (which we can ignore in the RUC setting), unital qubit channels have the canonical form [74]

$$\Phi(\rho) = p_0\rho + p_x X\rho X + p_y Y\rho Y + p_z Z\rho Z, \quad (16)$$

with $p_\alpha \geq 0$ and $\sum_\alpha p_\alpha = 1$. We start with a set of n Kraus operators $\mathbf{M} = \{M_i\}$, where each element can be represented as

$$M_i = a_i\mathbb{I} + b_i\sigma \cdot \tilde{\mathbf{u}}_i, \quad (17)$$

where a_i and b_i are real non-negative numbers [75], $\tilde{\mathbf{u}}_i = (e^{i\phi_{i,x}}u_{i,x}, e^{i\phi_{i,y}}u_{i,y}, e^{i\phi_{i,z}}u_{i,z})$ are complex unit vectors (i.e., $\tilde{\mathbf{u}}_i^* \cdot \tilde{\mathbf{u}}_i = 1$), and $\sigma = (X, Y, Z)$. The Kraus operators of Eq. (17) must yield the Choi matrix of the unital channel Φ in Eq. (16), i.e., they must satisfy

$$\sum_i M_i \otimes M_i^* = \sum_{\alpha=0,x,y,z} p_\alpha \sigma^\alpha \otimes (\sigma^\alpha)^* \quad (18)$$

(we set $\sigma^0 = \mathbb{I}$). This gives the following four relations:

$$\sum_i a_i^2 = p_0, \quad \sum_i b_i^2 = 1 - p_0, \quad (19)$$

$$\sum_i a_i b_i \tilde{\mathbf{u}}_i = \mathbf{0}, \quad \sum_i \tilde{u}_{i,\alpha}^* \tilde{u}_{i,\beta} b_i^2 = p_\alpha \delta_{\alpha\beta}. \quad (20)$$

The optimization of the target function $x(\mathbf{M})$ subject to these constraints is carried out in Appendix B. The

objective function is maximized when $a_i = \sqrt{p_0/n}$ and $b_i = \sqrt{(1-p_0)/n}$ for all i , and the unit vectors $\tilde{\mathbf{u}}_i$ are real: $\tilde{\mathbf{u}}_i = \mathbf{u}_i = (u_{i,x}, u_{i,y}, u_{i,z})$. The remaining constraints on the vectors $\{\mathbf{u}_i\}$, Eq. (20), read

$$\sum_i \mathbf{u}_i = \mathbf{0}, \quad \sum_i u_{i,\alpha} u_{i,\beta} = \frac{np_\alpha}{1-p_0} \delta_{\alpha\beta}. \quad (21)$$

As an example, a solution with $n = 4$ is

$$\begin{cases} \mathbf{u}_0 = (\sqrt{p_x}, \sqrt{p_y}, \sqrt{p_z}) / \sqrt{1-p_0}, \\ \mathbf{u}_1 = (\sqrt{p_x}, -\sqrt{p_y}, -\sqrt{p_z}) / \sqrt{1-p_0}, \\ \mathbf{u}_2 = (-\sqrt{p_x}, \sqrt{p_y}, -\sqrt{p_z}) / \sqrt{1-p_0}, \\ \mathbf{u}_3 = (-\sqrt{p_x}, -\sqrt{p_y}, \sqrt{p_z}) / \sqrt{1-p_0}, \end{cases} \quad (22)$$

i.e., the vertices of a regular tetrahedron inscribed in the Bloch sphere, up to a rescaling $\sqrt{3p_\alpha/(1-p_0)}$ of each axis.

In the case of $p_x = p_y = p_z$ [depolarizing noise, Eq. (6)], the conditions in Eq. (21) become

$$\mathbb{E}_i[\mathbf{u}_i] = \mathbf{0}, \quad \mathbb{E}_i[u_{i,\alpha} u_{i,\beta}] = \frac{1}{3} \delta_{\alpha\beta}, \quad (23)$$

where \mathbb{E}_i denotes averaging over i with respect to the uniform probability distribution $\Pr(i) = 1/n$. The conditions in Eq. (23) define a *spherical 2-design*, i.e., a probability distribution on the sphere whose first two moments coincide with those of the uniform distribution. Such distributions exist if $n = 4$ or $n \geq 6$. In particular, the minimal ($n = 4$) optimal unraveling is

$$\begin{cases} M_0 = \sqrt{\frac{1-\varepsilon}{4}}\mathbb{I} + \sqrt{\frac{\varepsilon}{12}}(X + Y + Z), \\ M_1 = \sqrt{\frac{1-\varepsilon}{4}}\mathbb{I} + \sqrt{\frac{\varepsilon}{12}}(X - Y - Z), \\ M_2 = \sqrt{\frac{1-\varepsilon}{4}}\mathbb{I} + \sqrt{\frac{\varepsilon}{12}}(-X + Y - Z), \\ M_3 = \sqrt{\frac{1-\varepsilon}{4}}\mathbb{I} + \sqrt{\frac{\varepsilon}{12}}(-X - Y + Z), \end{cases} \quad (24)$$

i.e., a weak measurement along four directions corresponding to the vertices of a regular tetrahedron. Similarly, a solution to Eq. (23) with $n = 6$ is given by the vertices of a regular octahedron, corresponding to weak measurements of the Pauli X , Y , and Z operators.

For the dephasing channel Eq. (5) ($n = 2$), an optimal unraveling is

$$\begin{cases} M_0 = \sqrt{\frac{1-\varepsilon}{2}}\mathbb{I} + \sqrt{\frac{\varepsilon}{2}}Z, \\ M_1 = \sqrt{\frac{1-\varepsilon}{2}}\mathbb{I} - \sqrt{\frac{\varepsilon}{2}}Z. \end{cases} \quad (25)$$

In all these cases, the most disentangling unraveling takes the form of *weak measurements* rather than stochastic

projective measurements [e.g., Eq. (8) for the case of dephasing].

In Appendix C we verify that, for the optimal unraveling Eq. (25), the measurement-induced phase transition occurs at a lower value of ε ($\varepsilon_c \simeq 0.044$) compared to the unraveling into stochastic projective measurements Eq. (8) ($\varepsilon_c \simeq 0.084$ —note the probability of doing a measurement is $p = 2\varepsilon$ and the well-known MIPT critical point is at $p_c \simeq 0.168$ [27]). We see that our simple mean-field approach is already sufficient to lower the critical noise strength by almost a factor of 2. It would be interesting to test whether locally and adaptively optimized unravelings as in Ref. [66] could further improve this threshold. For the rest of this work, we will use the optimal unraveling in Eqs. (24) and (25) for depolarizing and dephasing noise channels unless otherwise specified.

Before moving on, we note that, although the above solution works only for unital channels, other physically relevant models may also be analytically tractable. In particular, it is straightforward to find an optimized unraveling for the amplitude damping channel, a paradigmatic nonunital channel defined by Kraus operators

$$\begin{cases} M_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\varepsilon} \end{pmatrix}, \\ M_1 = \begin{pmatrix} 0 & \sqrt{\varepsilon} \\ 0 & 0 \end{pmatrix}. \end{cases} \quad (26)$$

In Appendix D we consider the optimization over 2×2 unitary rotations of these Kraus operators and derive the optimized unraveling

$$\begin{cases} M_0 = \begin{pmatrix} 1 & \sqrt{\varepsilon} \\ 0 & \sqrt{1-\varepsilon} \end{pmatrix} / \sqrt{2}, \\ M_1 = \begin{pmatrix} -1 & \sqrt{\varepsilon} \\ 0 & -\sqrt{1-\varepsilon} \end{pmatrix} / \sqrt{2}. \end{cases} \quad (27)$$

Furthermore, we numerically verify the disentangling effect of this unraveling by studying the critical noise strength of the measurement-induced phase transition: $\varepsilon_c \approx 0.17$ for the optimized unraveling, compared with $\varepsilon_c \approx 0.29$ for the original unraveling, again almost a factor of 2 improvement. Finally, beyond analytically tractable cases, one can always perform the optimization numerically for arbitrary single-qubit noise models, by restricting to a finite number of Kraus operators.

IV. NOISY-SEBD ALGORITHM

A. Description of the algorithm

We consider noisy random circuits in 2D with finite depth T acting on a grid with $N = L_x \times L_y$ qubits. The goal of our algorithm is to sample bitstrings \mathbf{z} from the distribution $P_{\mathcal{N}}(\mathbf{z})$, Eq. (11), determined by the circuit instance

and noise channel Φ (whose parameters we assume are known), with a small error.

Let us first review the noiseless case, studied in Ref. [45] and illustrated in Fig. 1(a). Due to locality, the outcome z_i on any given qubit depends only on the evolution within its past lightcone. Thus to sample all the outcomes z_i on the first row of qubits, $y = 1$, we need only to apply gates and channels on qubits within the past lightcone of the line $\{(x, y = 1, t = T)\}$, which includes qubits with $y \leq T$. This corresponds to a circuit of depth T on $\leq L_x T$ qubits, which can be simulated efficiently via MPS methods. At this point one can successfully sample the outcomes z_i for the first row of qubits, and move on to the second row ($y = 2$) iterating the same approach, performing only the gates and channels within the past lightcone of $\{(x, y = 2, t = T)\}$, etc. This effectively maps the 2D shallow circuit to an equivalent 1D circuit, where the readout measurements are converted to midcircuit measurements and resets [see Fig. 2(c) as an example]. As a result of this mapping, the spatial direction along L_y becomes the time direction for the 1D circuit (an idea also known as *space-time duality* in quantum circuits [37–39]). The SEBD algorithm is based on MPS simulation of this effective 1D dynamics for the purpose of sampling the \mathbf{z} outcomes.

The issue with iterating this approach indefinitely is that the entanglement in the quasi-1D state of $L_x \times T$ qubits, in principle, grows with each step, up to the point where MPS simulation fails. However, Ref. [45] observed that, due to the midcircuit measurements, the effective dynamics may enter an area-law phase, wherein the entanglement remains finite and approximate sampling can be carried out efficiently with high accuracy. In general, the effective 1D circuit consists of $N_{\text{1D}} = cTL_x$ qubits, where c is a constant depending on the lattice geometry (the slope of the lightcone, in the models considered here $c \approx 1/2$), and the spatial range of the two-qubit gates is proportional to T . Equivalently, one may view the effective system as a quasi-1D strip of size $L_x \times cT$ with nearest-neighbor gates in both directions. Either way, each circuit layer on cTL_x qubits is followed by the measurement of L_x qubits (a full row), giving a ratio of measurements to unitary operation of approximately $1/T$. When this ratio is sufficiently high (i.e., T sufficiently low), the system enters an area-law phase and SEBD is efficient.

Let us now add noise to the picture. As discussed in Sec. III B, we can unravel the noise channels Φ into an arbitrary set of Kraus operators, simulate the pure-state trajectories, sample from the joint distribution $P_{\mathcal{N}}(\mathbf{z}, \mathbf{m})$ and keep only the \mathbf{z} samples. We adopt the entanglement-optimal unraveling of unital noise channels discussed in Sec. III C to suppress entanglement in the effective 1D state at the level of quantum trajectories. The simulated dynamics now features midcircuit measurements with two distinct origins: a density $\propto 1/T$ coming from the final sampling step in the 2D circuit, and a density $\propto \varepsilon$ coming from the unraveling

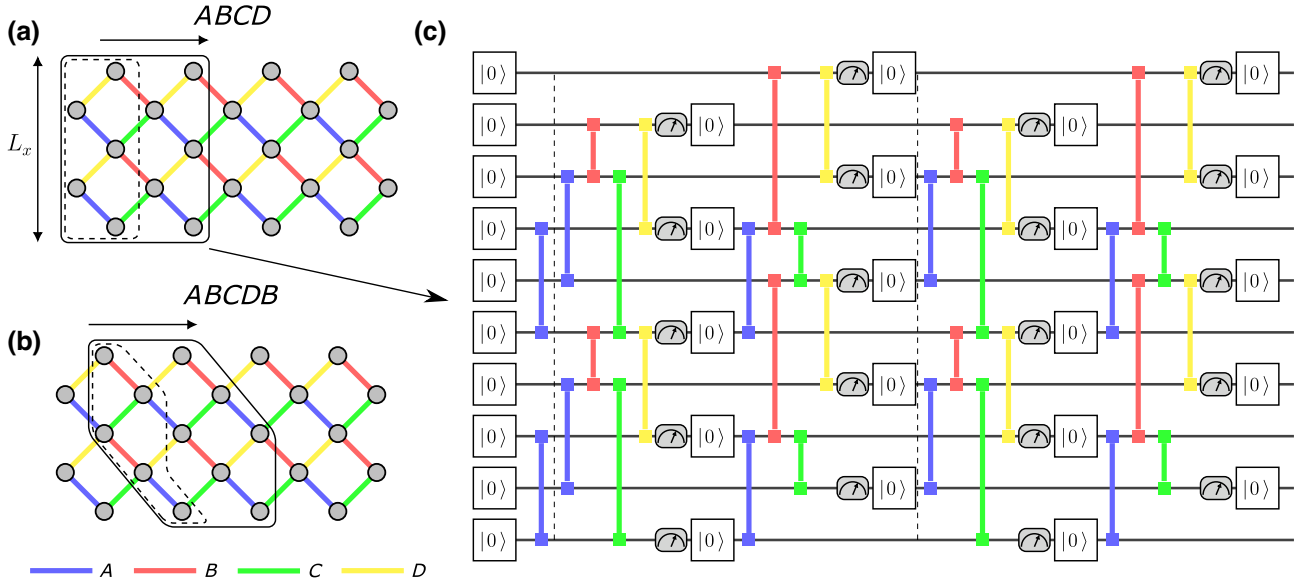


FIG. 2. Schematic depiction of the 2D qubit array and effective 1D subsystem used in the SEBD algorithm. (a),(b) Example of 2D qubit array with $L_x = 5$, showing the gate sequences used: $ABCD$ for depth $T = 4$ (a), $ABCDB$ for depth $T = 5$ (b). The region enclosed by the solid loop corresponds to the effective 1D subsystem, and the region enclosed by the dashed loop corresponds to sites, which are measured after applying gates in the past lightcone. (c) Equivalent 1D circuit for $T = 4$ and $L_x = 5$ with gate sequence $ABCD$. The effective 1D system has size $2L_x = 10$, with gates up to the third nearest neighbor, and each measurement is followed by a reset to the $|0\rangle$ state. The dashed lines enclose a unit cell whose architecture repeats periodically in time.

of noise. Below a critical noise rate ε_c , dependent on the model and the circuit depth T , the entanglement of effective 1D state satisfies area law, and thus allows efficient classical simulation. These ideas are schematically summarized in Fig. 1.

B. Numerical results: entanglement phase transition

In this section, we numerically study the entanglement phase transition in the effective 1D subsystem that is used in the noisy-SEBD algorithm [Fig. 2(c)]. As an example, we choose shallow circuits that act on a 2D square lattice [Figs. 2(a) and 2(b)] with unitary gates similar to those employed in Google's RCS experiment [6]. The two-qubit gates are i SWAP-like fermionic simulation gates

$$\text{fSim}(\pi/2, \pi/6) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 \\ 0 & -i & 0 & 0 \\ 0 & 0 & 0 & e^{-i\pi/6} \end{pmatrix}, \quad (28)$$

sandwiched between single-qubit rotations randomly chosen from the set $\{X^{\pm 1/2}, Y^{\pm 1/2}, W^{\pm 1/2}, V^{\pm 1/2}\}$, where $W = (X + Y)/\sqrt{2}$ and $V = (X - Y)/\sqrt{2}$ are Hadamard-like gates (note that $W^{\pm 1/2}$ and $V^{\pm 1/2}$ are non-Clifford). The two-qubit gates are applied to bonds of the square lattice according to the sequence $ABCD$ for $T = 4$ [Fig. 2(a)] and $ABCDB$ for $T = 5$ [Fig. 2(b)]. These specific sequences are chosen to be the most entangling for the effective

1D dynamics, defined as giving the longest single-qubit purification time as discussed in the following. The noise channel Φ is taken to be the depolarizing channel, Eq. (6), with strength ε .

As a diagnostic for the entanglement phase transition, which underpins the efficiency of noisy SEBD, we use the single-qubit purification time τ discussed in Sec. II D. This is advantageous from the numerical point of view, relative to a direct calculation of the bipartite entanglement entropy, since it does not suffer from the large finite-size drifts arising from the logarithmic divergence of entropy at the critical point [27]. The phase transition between area-law and volume-law entanglement scaling is probed by the mutual information between the single reference qubit and the rest of the system. In practice, we introduce a reference qubit, which initially forms a Bell pair with a system qubit. At later time the average entropy of the reference qubit is captured by the exponential decaying relation $S_R(t) \sim e^{-t/\tau}$. In the volume-law phase (low measurement and noise rate), S_R can remain nonzero for a long time $\tau \sim \exp(L_x)$. Physically this implies that there is finite measurement-induced entanglement between opposite ends of the system as one takes L_x, L_y to infinity jointly with $L_y = \text{poly}(L_x)$; this can be interpreted as emergent quantum teleportation [76] and has recently been explored experimentally [44]. On the other hand, in the area-law phase (high measurement and noise rate), S_R decays rapidly to zero with $\tau = O(1)$. At the critical

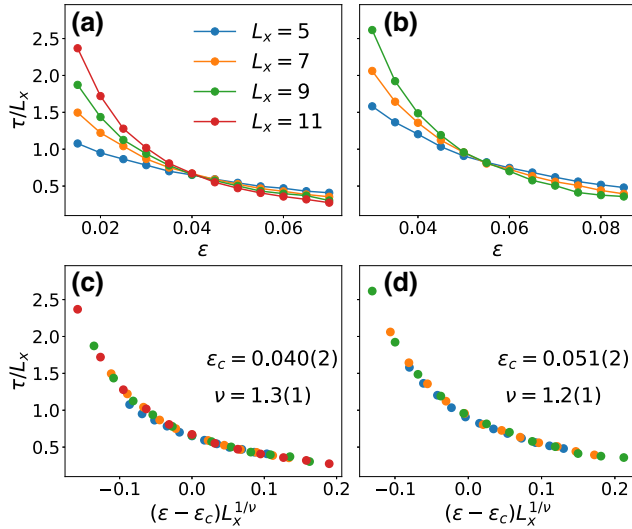


FIG. 3. Purification time τ of a single probe qubit as a function of the noise rate ε for (a) depth $T = 4$ with gate sequence $ABCD$ and (b) depth $T = 5$ with gate sequence $ABCDB$. Scaling collapse of the data for (c) $T = 4$ and (d) $T = 5$. The data are averaged over 5×10^3 – 3×10^4 realizations.

point, we expect $\tau \sim L_x^z$, where z is the dynamical critical exponent (typically $z = 1$ for measurement-induced transitions in short-range interacting 1D systems [26,27]).

In Fig. 3, we show a finite-size scaling analysis of τ/L_x for both $T = 4$ (gate sequence $ABCD$) and $T = 5$ (gate sequence $ABCDB$), obtained from MPS simulation of the spacewise dynamics up to $L_y = 2L_x$, where only $L_x \leq L_y \leq 2L_x$ are used for fitting to avoid the early-time transient effect. The existence of a finite-size crossing point of the ratio τ/L_x for all system sizes [Figs. 3(a) and 3(b)] indicates $z = 1$, which is consistent with the emergence of $1 + 1$ D conformal symmetry at the transition. Therefore, we use the scaling ansatz

$$\tau(\varepsilon, L_x) = L_x F[(\varepsilon - \varepsilon_c)L_x^{1/\nu}], \quad (29)$$

to determine the location of the critical point ε_c and the correlation length critical exponent ν . From the data collapse, Figs. 3(b) and 3(d) (see Appendix E for method details), we locate $\varepsilon_c = 0.040(2)$ with $\nu = 1.3(1)$ for $T = 4$ ($ABCD$) and $\varepsilon_c = 0.053(2)$ with $\nu = 1.2(1)$ for $T = 5$ ($ABCDB$). For both values of T we find correlation length exponent $\nu \approx 1.3$, which is consistent with the known value for the measurement-induced phase transition in 1D, as expected. Moreover we see that increasing circuit depth T leads to a larger critical noise rate ε_c , which is the qualitative behavior sketched in Fig. 1(d).

These numerical results support our qualitative expectations for the complexity of noisy SEBD. Since the hardness of the method is exponential in T , obtaining accurate predictions for the phase boundary at larger T becomes

increasingly challenging. However, it is reasonable to conjecture that, for large T ,

- (i) as the quasi-1D system of $cT \times L_x$ qubits approaches a 2D limit, one should recover the 2D measurement-induced phase transition, which occurs at a finite noise rate $\varepsilon_{c,2D}$ [24,77];
- (ii) the transition should occur at a finite *total* noise rate, comprising the unraveled measurements (rate ε) and the sampling of the final state (rate $\sim 1/T$), thus $\varepsilon_c(T) = \varepsilon_{c,2D} + O(1/T)$.

We find evidence in support of these conjectures in Clifford circuits, which can be simulated in polynomial time by the stabilizer method [49] (though note that this limits us to the projective unraveling of the noise channel), see Appendix F.

Finally, in Appendix G we verify the accuracy of the noisy-SEBD algorithm by benchmarking its output against direct MPO simulations of the noisy dynamics and against stabilizer simulations of Clifford circuits.

C. Application: IBM quantum processors

Quantitatively, the results in the previous section show that circuits on square-lattice architectures, on NISQ platforms such as Google’s Sycamore processor with native i SWAP-like gates and $\gtrsim 98\%$ two-qubit gate fidelities (translating to $\varepsilon \lesssim 0.01$ in our parametrization, see Appendix H), are already in the “hard phase” at depth $T = 4$. Since depth $T = 3$ is in the easy phase already for noiseless SEBD (i.e., for $\varepsilon = 0$), the inclusion of noise does not allow efficient simulation of an additional gate layer in this setting. This class of circuits is however a worst-case scenario, representing highly scrambling dynamics optimized for hardness of simulation [6,78].

Here we consider the application of noisy SEBD to circuits on quantum processors with heavy-hexagon geometry and native CNOT gates, such as IBM Quantum’s family of processors, Figs. 4(a) and 4(b). We consider again highly scrambling random circuits with i SWAP two-qubit gates and single-qubit gates randomly chosen from $\{X^{\pm 1/2}, Y^{\pm 1/2}, W^{\pm 1/2}, V^{\pm 1/2}\}$. An i SWAP gate can be compiled into two native CNOT gates, plus single-qubit gates. Thus the effective noise rate for i SWAP gates is twice the CNOT error, giving, e.g., approximately equal to 96% median gate fidelity on the Osprey processor (corresponding to $\varepsilon \approx 0.025$ in our parametrization, see Appendix H).

We consider subsystems of a qubit array based on the Condor processor, Fig. 4(a). The full system has $N = 1,121$ qubits and linear size $L_x = 43$. The gate sequence and unit cell for the SEBD algorithm are shown in Fig. 4(b), for the $N = 65$, $L_x = 11$ Hummingbird processor. We characterize the efficiency of the (noisy-)SEBD algorithm by computing the half-system bipartite entanglement entropy S , Fig. 4(c), and single-qubit purification

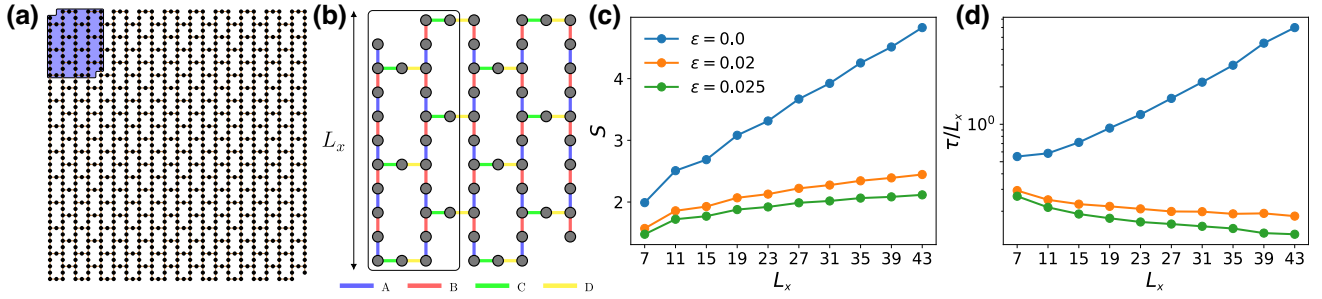


FIG. 4. (a) Layout of the 1121-qubit IBM quantum processor Condor; 65-qubit Hummingbird is shown as the blue region. (b) Gate sequence for circuits with depth $T = 5$ ($ABCD$) on 65-qubit quantum processor Hummingbird, of linear size $L_x = 11$. Other IBM processors have similar layout: Eagle with $L_x = 15$, Osprey with $L_x = 27$, and Condor with $L_x = 43$. The region surrounded by a solid line depicts the 1D effective subsystem for noisy-SEBD simulation of random circuits with $T = 5$ (gate sequence $ABCD$). (c) Bipartite entanglement entropy S and (d) purification time τ of a reference qubit, for noiseless ($\epsilon = 0$) and noisy ($\epsilon = 0.02$ and 0.025) circuits of depth $T = 5$ as a function of linear system size L_x , which refers to subsets of the heavy-hexagon lattice in panel (a). The data are averaged over 10^3 – 10^4 realizations of the random circuits.

time τ , Fig. 4(d), for both noiseless and noisy random circuits with depth $T = 5$ (measured in units of i SWAP gates, thus corresponding to ten CNOT gates on each qubit). As the linear system size is varied between $L_x = 7$ and $L_x = 43$, we observe a clear volume-law phase in the noiseless case, with volume-law entropy $S \propto L_x$ and exponential purification time $\tau \sim \exp(L_x)$. In contrast, for noise rate $\epsilon = 0.02$, we see a weak growth of S with L_x , consistent with critical scaling $S \sim \ln(L_x)$ or eventual saturation to an area law. The ratio τ/L_x also decreases, either to a finite constant (critical behavior) or to zero (area-law behavior). Finally, for $\epsilon = 0.025$, both diagnostics are indicative of an area-law phase.

We conclude that in this case the inclusion of a realistic noise rate in the simulation algorithm is sufficient to drive a transition in complexity of the SEBD algorithm. This makes the difference between an asymptotically efficient and asymptotically hard MPS simulation of the sampling problem. We note that, while circuits with low depth such as $T = 5$ can likely be simulated by brute-force tensor-network methods up to hundreds or thousands of qubits, the cost of such methods remains generically exponential in the linear size of the system L_x . On next-generation processors with $N \sim 10^5$ qubits those methods would become intractable, while noisy SEBD would remain practical in the easy phase.

V. DISCUSSION

We have introduced a classical algorithm, noisy SEBD, to sample from the output distribution of noisy, shallow circuits in two dimensions. The algorithm uses the insight of mapping a 2D RCS problem to a 1D monitored dynamics problem (space-evolving block decimation [45]), while also unraveling the action of noise into additional measurements on the system. At sufficiently low depth and sufficiently strong noise, this enables efficient MPS

simulation of the monitored quantum trajectories, and thus efficient sampling from the appropriate noisy output distribution.

Given that the unraveling of noise into measurements is arbitrary, it can be optimized so as to reduce the amount of entanglement [66]. Here we have focused on a “mean-field” approach where single-qubit noise channels at all positions and times are unraveled in the same way, chosen based on the two-replica statistical mechanics description of the circuit upon averaging over random gates. We have found that, for unital channels (such as dephasing and depolarizing), the optimal unraveling is based on uniform weak measurements, rather than stochastic projective measurements. The difference between unravelings is substantial—in the standard model of brickwork circuits in 1D, the noise threshold corresponding to the measurement-induced entanglement transition is reduced by a factor of about 2 for the optimal weak-measurement unraveling ($\epsilon_c \approx 0.04$) compared to the usual projective measurement unraveling ($\epsilon_c \approx 0.08$, i.e., measurement rate $p_c = 2\epsilon_c \approx 0.16$). This is consistent with prior observations in Ref. [72], and the optimization technique could be of independent interest for the study of measurement-induced entanglement transitions.

While noisy RCS in the anticoncentration regime was shown to be classically simulable in polynomial time based on a sampling of “Feynman paths” in Pauli operator space [10], the polynomial scaling of the algorithm features a large exponent (proportional to $1/\epsilon$, i.e., of order 100 in present-day experiments) that makes the algorithm impractical. This leaves open the question of “practical hardness” for finite-sized RCS experiments. Furthermore, the requirement of anticoncentration in Ref. [10] is not met at constant depth T [depth diverging as at least $\log(N)$ is needed [11]], so that even the in-principle hardness of sampling noisy shallow circuits is an open problem. Our results contribute to sharpen the requirements for such

hardness by identifying a phase in the parameter space of depth T and noise strength ε [Fig. 1(d)] where noisy RCS can be classically simulated via a straightforward MPS algorithm in time [79] scaling as $N \exp(T)$. We have located the entanglement phase transitions in circuit architectures based on real-world quantum processors. In square lattices with native i SWAP-like gates, we found that noisy SEBD allows the efficient sampling of circuits of depth $T = 3$, like SEBD in the noiseless case; but for realistic noise rates of $\varepsilon \lesssim 1\%$ it does not increase the depth threshold (i.e., $T = 4$ remains in the hard phase). On heavy-hexagon lattices with native CNOT gates, we have instead found that the inclusion of realistic noise rates can increase the depth threshold, as shown in Figs. 4(c) and 4(d).

Our results add to the growing body of work on noise-induced phase transitions in computational complexity. Recent works have studied the simulation of noisy RCS via MPS simulations truncated to constant bond dimension χ [12,14], finding that the accumulated truncation error behaves similarly to noise in the quantum experiment—i.e., causes an exponential decay of the linear cross-entropy, Eq. (1); to beat this classical simulation method, the quantum processor must be below a finite noise threshold. We remark that the task considered in those works is different from the one considered here. Namely the goal in Refs. [12,14] is to simulate the *ideal* (noiseless) bitstring distribution better than the noisy quantum experiment, as quantified, e.g., by fidelity or linear cross-entropy. Notably this allows for an exponentially small (in N and T) fidelity between simulation and experiment, as long as the former is closer to the ideal result. In contrast, our goal is to simulate with high accuracy the *noisy* bitstring distribution itself. This is a significant distinction physically as the effect of uncontrolled MPS truncation is *a priori* very different from that of local noise, even at the same level of fidelity (e.g., MPS truncation does not obey locality).

Even more recently, a noise-induced phase transition in RCS has been reported [9,47] in deep quantum circuits with noise rates scaled as $\varepsilon \sim 1/N$, i.e., a constant number of errors per layer. The scaling of linear cross-entropy [Eq. (1)] in these models was predicted to sharply change as a function of εN , from an “easy phase” where the system appears to break into finite-sized clusters from the point of view of linear cross-entropy, to a phase where it behaves as a single large cluster. The former phase is easy to spoof classically, while the latter is conjectured to be practically hard. This transition is also different from the one studied here. For one, it applies to a vanishing noise rate $\varepsilon = O(1/N)$ rather than a finite $\varepsilon = O(1)$. Additionally, and more importantly, it is a phase transition in an observable (albeit a complex one like linear cross-entropy) that reflects an intrinsic property of the system, whereas the transition studied here is a property of a

simulation algorithm (noisy SEBD) that is not intrinsic to the system.

A distinctive aspect of the problem we study here is that the simulation task requires precise knowledge of the noise model on the quantum processor, which is not fully controlled or programmable, and whose characterization is a separate challenge [80]. For this reason, the sampling task simulated by noisy SEBD is quite different from the standard (noiseless) one. In other words, *given a noise model*, there is a fully well-defined simulation task; however, comparison with real-world noisy devices is more subtle, as the “true” noise model for such devices is never completely known. At the same time, this feature of the problem (which is generic of simulation algorithms for noisy systems) opens up an interesting direction for future research, namely using noisy SEBD as a noise benchmarking tool. One could run noisy SEBD (or other algorithms for the same task) with a parametrized noise model, compare its output with that of real quantum hardware, and use the comparison to optimize the noise parameters. To this end, it would be useful to generalize our discussion to more complex noise models beyond uncorrelated single-qubit errors. We leave these ideas as directions for future research.

Finally, we emphasize again that noisy SEBD is only one possible strategy for classical sampling. It remains an interesting open question to identify other simulation algorithms with polynomial cost $O(N^c)$, with c an ε -independent constant (unlike in the Feynman path sampling approach of Ref. [10] or the direct MPO approach for 1D circuits of Ref. [62]) below a finite noise threshold ε_c . The possibility of an intrinsically hard phase at finite depth (before anticoncentration and the results of Ref. [10] apply) is also an interesting open question, with noisy SEBD providing a constraint on this hypothetical phase boundary. Another interesting direction for future work is the possibility of extending these results to higher dimension or all-to-all connected systems, e.g., trapped ion quantum computers. While MIPTs are known to arise in all these cases, the ensuing area-law for entanglement can only be exploited for efficient MPS simulation in 1D (including in the effective 1D subsystems of shallow 2D circuits used in SEBD). Whether other formulations of the MIPT, e.g., the dynamical purification approach [25,26], can be exploited for efficient simulation in more general systems is an interesting open question.

ACKNOWLEDGMENTS

M.I. thanks Aram Harrow for helpful discussions. We thank Soonwon Choi for bringing Ref. [72] to our attention. Numerical simulations were carried out on computational resources provided by Texas Advanced Computing Center (TACC) and on Stanford Research Computing Center’s Sherlock cluster. M.I. was partly supported by the

Gordon and Betty Moore Foundation’s EPIQS Initiative through Grant No. GBMF8686.

Note added.—Recently, we became aware of an independent work on related topics appearing in the same arXiv posting [81]. Our works are complementary and our results agree where they overlap.

APPENDIX A: DERIVATION OF THE UNRAVELING COST FUNCTION FROM STATISTICAL MECHANICAL MODEL

Here we study the most disentangling unraveling of a single-qubit noise channel Φ in the context of 1D brickwork random circuits, by mapping the entropy calculation of random circuits to a classical statistical mechanical model [22,45].

1. Quasientropy

Consider a subsystem A of a one-dimensional qudit chain. The k -Renyi entropy for the reduced density matrix ρ_A is defined as

$$S_k(A) = \frac{1}{1-k} \log \left(\frac{Z_{k,A}}{Z_{k,\emptyset}} \right), \quad (\text{A1})$$

where

$$Z_{k,\emptyset} = \text{Tr}(\rho)^k, \quad Z_{k,A} = \text{Tr}(\rho_A^k). \quad (\text{A2})$$

The von Neumann entropy is given by the $k \rightarrow 1$ limit

$$S_1(A) = -\text{Tr} \left(\frac{\rho_A}{\text{tr} \rho} \log \frac{\rho_A}{\text{tr} \rho} \right). \quad (\text{A3})$$

For a hybrid random circuit, we are interested in the trajectory-averaged behavior of k -Renyi entropy

$$\langle S_k(A) \rangle = \frac{\mathbb{E}_C [\text{Tr} \rho S_k(A)]}{\mathbb{E}_C [\text{Tr} \rho]} = \frac{1}{1-k} \frac{\mathbb{E}_C \left[\text{Tr} \rho \log \frac{Z_{k,A}}{Z_{k,\emptyset}} \right]}{\mathbb{E}_C [\text{Tr} \rho]}, \quad (\text{A4})$$

where \mathbb{E}_C represents the combined average of Haar random circuits \mathbb{E}_U and the average over Kraus operators \mathbb{E}_M (i.e., quantum trajectories).

Replica tricks can be applied to cure the average of the logarithm. However, to get direct mapping to the stat-mech model, one can alternatively consider the k -quasientropy $\tilde{S}_k(A)$ [45], i.e., the k th moment of the entanglement spectrum, weighted by the k th power of the measurement

outcome probability,

$$\begin{aligned} \tilde{S}_k(A) &= \frac{1}{1-k} \log \left(\frac{\mathbb{E}_C \left[\text{tr}(\rho)^k \frac{Z_{k,A}}{Z_{k,\emptyset}} \right]}{\mathbb{E}_C \left[\text{tr}(\rho)^k \right]} \right) \\ &= \frac{1}{1-k} \log \left(\frac{\mathbb{E}_C [Z_{k,A}]}{\mathbb{E}_C [Z_{k,\emptyset}]} \right). \end{aligned} \quad (\text{A5})$$

Importantly, in the limit $k \rightarrow 1$, $\tilde{S}_k(A) \rightarrow \langle S_1(A) \rangle$, similar to the k -Renyi entropy. The k quasientropy has a natural mapping with a classical stat-mech model: the quantities $\mathbb{E}_C [Z_{k,A/\emptyset}]$ can be viewed as partition functions for a $(2+0)$ -dimensional spin model with different boundary conditions.

2. Generalized measurement

A general measurement procedure can be represented by a set of Kraus operators $\{M_i\}$ satisfying $\sum_i M_i^\dagger M_i = \mathbb{I}$ [68]. For each quantum trajectory, a specific Kraus operator is chosen with probability $p_i = \text{tr}(M_i \rho M_i^\dagger)$ and gives the updated state $\rho' = M_i \rho M_i^\dagger / p_i$. However, in the calculation of the k quasientropy, it is more convenient to reparametrize the generalized measurement in terms of operators \tilde{M}_i and classical probabilities μ_i satisfying

$$\text{tr}(\tilde{M}_i^\dagger \tilde{M}_i) = q, \quad \mathbb{E}_M M^\dagger M = \sum_i \mu_i \tilde{M}_i^\dagger \tilde{M}_i = \mathbb{I}, \quad (\text{A6})$$

where q is the local Hilbert-space dimension. It is easy to show μ_i is given by

$$\mu_i = \frac{1}{q} \text{tr}(M_i^\dagger M_i), \quad (\text{A7})$$

and that $\mu_i \geq 0$ and $\sum_i \mu_i = 1$ (i.e., μ_i is a probability distribution).

The advantage of this reparametrization is that it renders the k -quasientropy invariant under trivial decomposition of Kraus operators, which is important for optimizing the unraveling. As an example, consider the sets $\mathbf{M}_0 = \{\sigma_z\}$ and $\mathbf{M}_1 = \{\sigma_z/\sqrt{2}, \sigma_z/\sqrt{2}\}$, which differ by a trivial decomposition of σ_z and are hence physically equivalent (both describe the unitary transformation $\rho \mapsto \sigma^z \rho \sigma^z$). However, if we define the average over a Kraus set as $\mathbb{E}_M[f] = \sum_{M \in \mathbf{M}} f(M \rho M^\dagger)$, then one can verify $\mathbb{E}_{\mathbf{M}_0}[Z_{k,\emptyset}] = 1$ but $\mathbb{E}_{\mathbf{M}_1}[Z_{k,\emptyset}] = 2^{1-k}$. Using the reparametrized Kraus operators and defining the average as $\mathbb{E}_M[f] = \sum_i \mu_i f(\tilde{M}_i \rho \tilde{M}_i^\dagger)$ instead yields the consistent results $\mathbb{E}_{\mathbf{M}_0}[Z_{k,\emptyset}] = \mathbb{E}_{\mathbf{M}_1}[Z_{k,\emptyset}] = 1$. It is also worth mentioning that, in the limit of interest ($k \rightarrow 1$), these two formalisms are equivalent.

3. Mapping to a classical statistical mechanical model

The goal of this mapping is to calculate the averaged partition functions $\mathbb{E}_C [Z_{k,X}]$ where $X = A$ or \emptyset :

$$\mathbb{E}_C [Z_{k,X}] = \mathbb{E}_U \mathbb{E}_M \left[\text{tr} \left((\cdots M U \rho_0 U^\dagger M^\dagger \cdots)^{\otimes k} \mathcal{S}_X^{\otimes k} \right) \right], \quad (\text{A8})$$

where \mathcal{S}_X is the operator that implements a cyclical permutation of the k replicas only in the X subsystem. By using the Haar-measure calculus [82], the average over replicated gates $(U \otimes U^*)^{\otimes k}$ can be expanded onto permutations of the replicas, giving a partition function of ‘‘spins’’ valued in the permutation group of k elements S_k . Hence the average over single-site measurement operators can be evaluated by appropriately contracting with connecting permutations. In this sense, each unitary gate can be viewed as two permutation nodes $\{\sigma, \tau\} \in S_k$, which form a honeycomb lattice as shown in Fig. 5(b). The total partition function can be evaluated by contracting the nodes with proper weights $w^{(k)}(\sigma_r, \tau_{r'})$ on the links:

$$\mathbb{E}_C [Z_{k,X}] = \sum_{\{\sigma_r, \tau_r\}} \prod_{\langle r, r' \rangle} w^{(k)}(\sigma_r, \tau_{r'}), \quad (\text{A9})$$

with distinct couplings on dashed and solid links, as shown in Fig. 5(b). The average over Haar random unitary gates gives the coupling for dashed links:

$$w_u^{(k)}(\sigma, \tau) = \text{Wg}_{q^2}(\tau \sigma^{-1}), \quad (\text{A10})$$

where Wg_{q^2} is the Weingarten function [82]. These couplings are independent of measurements. The solid links in Fig. 5(b), on the other hand, are given by

$$w_m^{(k)}(\sigma, \tau) = \mathbb{E}_M \prod_c \text{tr} \left((M^\dagger M)^{\lambda_c} \right), \quad (\text{A11})$$

where c denotes the number of cycles in permutation $\tau \sigma^{-1} \in S_k$ and λ_c is the length of cycle c . Using the convention above for averaging over Kraus operators, $\mathbb{E}_M[f] = \sum_i \mu_i f(\tilde{M}_i \rho \tilde{M}_i)$, we obtain

$$\begin{aligned} w_m^{(k)}(\sigma, \tau) &= \sum_i \mu_i \prod_c \text{tr}[(\tilde{M}_i^\dagger \tilde{M}_i)^{\lambda_c}] \\ &= \sum_i \mu_i^{1-k} \prod_c \text{tr}[(M_i^\dagger M_i)^{\lambda_c}]. \end{aligned} \quad (\text{A12})$$

4. Two replicas: classical Ising model

Although the combination of Eqs. (A10) and (A12) gives the exact expression for the partition function, the possible negative weights of $w_u(\sigma, \tau)$ impede the direct mapping to a physical system with real interactions at a real temperature. For the case of $k = 2$, this sign problem can

be circumvented by integrating out all τ variables, which gives rise to a classical Ising model defined on a triangular lattice as shown in Fig. 5(c):

$$\begin{aligned} \mathbb{E}_C [Z_{k,X}] &= \sum_{\{\sigma\}} \prod_{\langle \sigma_1, \sigma_2, \sigma_3 \rangle} \\ &\times \left[\sum_{\tau=\pm 1} w_m^{(2)}(\sigma_1, \tau) w_m^{(2)}(\sigma_2, \tau) w_u^{(2)}(\sigma_3, \tau) \right] \\ &\equiv \sum_{\{\sigma\}} \prod_{\langle \sigma_1, \sigma_2, \sigma_3 \rangle} w^{(2)}(\sigma_1, \sigma_2, \sigma_3), \end{aligned} \quad (\text{A13})$$

where $\langle \sigma_1, \sigma_2, \sigma_3 \rangle$ denotes a lower-facing triangle with three neighboring vertices $\sigma_1, \sigma_2, \sigma_3$. For $k = 2$, we have

$$w_u^{(2)}(\sigma, \tau) = \begin{cases} \frac{1}{q^4 - 1} & \text{if } \sigma = \tau, \\ -\frac{1}{q^2(q^4 - 1)} & \text{if } \sigma \neq \tau, \end{cases} \quad (\text{A14})$$

and denote

$$w_m^{(2)}(\sigma, \tau) = \begin{cases} u & \text{if } \sigma = \tau, \\ v & \text{if } \sigma \neq \tau, \end{cases} \quad (\text{A15})$$

where u and v are determined by specific Kraus operators via Eq. (A12). Using the definition of μ_i , one can see that

$$u = \sum_i \mu_i^{-1} \text{tr} \left(M_i^\dagger M_i \right)^2 = q^2, \quad (\text{A16})$$

$$v = \sum_i \mu_i^{-1} \text{tr} \left((M_i^\dagger M_i)^2 \right) = q \sum_i \frac{\text{tr} \left(M_i^\dagger M_i M_i^\dagger M_i \right)}{\text{tr} \left(M_i^\dagger M_i \right)}. \quad (\text{A17})$$

With these, one can express the three-body interaction $w^{(2)}(\sigma_1, \sigma_2, \sigma_3)$ explicitly. Due to the permutation symmetry $\sigma \rightarrow \bar{\sigma}$ (where $\bar{\sigma}$ denotes the other permutation with $S_2 \equiv \mathbb{Z}_2$) and spatial reflection symmetry $\sigma_1 \leftrightarrow \sigma_2$, one only needs to specify

$$w^{(2)}(\sigma, \sigma, \sigma) = \frac{u^2}{q^4 - 1} - \frac{v^2}{q^2(q^4 - 1)}, \quad (\text{A18})$$

$$w^{(2)}(\sigma, \sigma, \bar{\sigma}) = \frac{v^2}{q^4 - 1} - \frac{u^2}{q^2(q^4 - 1)}, \quad (\text{A19})$$

$$w^{(2)}(\sigma, \bar{\sigma}, \sigma) = \frac{uv}{q^2(q^2 + 1)}. \quad (\text{A20})$$

Furthermore one may re-express the three-body interaction as the product of three two-body interaction $w^{(2)}(\sigma_1, \sigma_2, \sigma_3) = C e^{-J_h \sigma_1 \sigma_2 - J_d \sigma_1 \sigma_3 - J_d \sigma_2 \sigma_3}$, where J_d and J_h

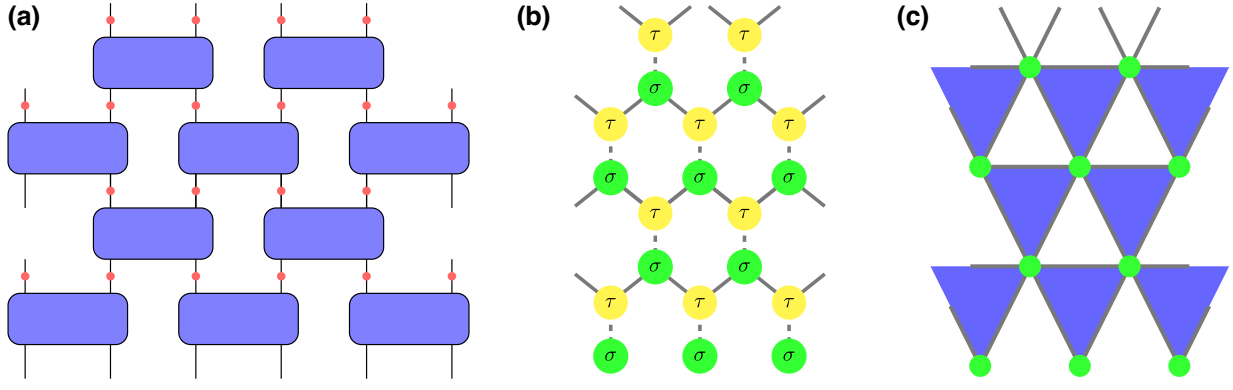


FIG. 5. (a) The quantum circuit consists of brick-wall unitaries (blue rectangles) and generalized measurements (red dots). (b) The mapped classical statistical mechanical model on the honeycomb lattice, where dashed links are weighted by the Weingarten function and the solid links are weighted by the generalized measurements. (c) Classical Ising model for $k = 2$, after integrating out τ nodes.

are the two-body interaction strength for diagonally and horizontally neighboring sites, respectively, whose expressions are

$$J_d = \frac{1}{4} \log \left(\frac{q^2 x^2 - 1}{q^2 - x^2} \right), \quad (\text{A21})$$

$$J_h = \frac{1}{4} \log \left(\frac{x^2 (q^2 - 1)^2}{(q^2 x^2 - 1)(q^2 - x^2)} \right), \quad (\text{A22})$$

where $x \equiv v/u$ is a dimensionless parameter given by

$$x = \frac{1}{q} \sum_i \frac{\text{tr}(M_i^\dagger M_i M_i^\dagger M_i)}{\text{tr}(M_i^\dagger M_i)}. \quad (\text{A23})$$

Since $\text{tr}(M^\dagger M)^2 \geq \text{tr}((M^\dagger M)^2)$ for all Kraus operators, one has $u \geq v$, hence $J_d \leq 0$ and $J_h \geq 0$ for $q \geq 2$.

The statistical mechanics of this classical Ising model can be solved exactly [83,84] and the critical point x_c , separating paramagnet and ferromagnet phases, is determined by the relation $2e^{2J_h} = e^{-2J_d} - e^{2J_d}$. Solving for x_c gives

$$\frac{1}{x_c} = \frac{q^2 - 1}{q^2 + 1} + \frac{\sqrt{2q^4 + 2}}{q^2 + 1}. \quad (\text{A24})$$

This phase transition between an ordered and disordered phase in the statistical mechanical model then maps onto the area-law and volume-law phase transition of the 2 quasientropy in the hybrid random circuit. As we see, within this two-replica analysis the transition is controlled exclusively by the parameter x , Eq. (A23), with an ordered phase (volume law) for $x < x_c$ and disordered phase for $x > x_c$ (area law). Therefore, to obtain the optimal unraveling for a given quantum channel, we aim to maximize the function $x(\mathbf{M})$. This justifies the use of $x(\mathbf{M})$ as a cost function for the unraveling of Φ in the main text.

APPENDIX B: MAXIMIZATION OF THE TARGET FUNCTION FOR UNITAL CHANNELS

To compute Eq. (14) for the parametrization of Kraus operators in Eq. (17), we first evaluate $M_i^\dagger M_i$. Omitting the subscript i for ease of notation, we have

$$M^\dagger M = (a^2 + b^2)\mathbb{I} + 2b(a\mathbf{u}_R + b\mathbf{u}_I \wedge \mathbf{u}_R) \cdot \sigma, \quad (\text{B1})$$

with $\tilde{\mathbf{u}} = \mathbf{u}_R + i\mathbf{u}_I$. (We used the identity $\sigma^\alpha \sigma^\beta = \sigma^0 \delta_{\alpha,\beta} + i\varepsilon_{\alpha\beta\gamma} \sigma^\gamma$.)

Letting $\mathbf{v} \equiv a\mathbf{u}_R + b\mathbf{u}_I \wedge \mathbf{u}_R$, the operator $M^\dagger M$ has eigenvalues $\lambda_\pm = (a^2 + b^2) \pm 2b\|\mathbf{v}\|$. It follows that the ratio in the definition of $x(\mathbf{M})$, Eq. (14), reads

$$\frac{\text{Tr}[(M_i^\dagger M_i)^2]}{2\text{Tr}(M_i^\dagger M_i)} = \frac{a_i^2 + b_i^2}{2} + \frac{2b_i^2 \|\mathbf{v}_i\|^2}{a_i^2 + b_i^2}. \quad (\text{B2})$$

Focusing on the second term, we have $\|\mathbf{v}\|^2 = a^2 \cos^2(\theta) + b^2 \cos^2(\theta) \sin^2(\theta) \sin^2(\chi)$, where $\|\mathbf{u}_R\| \equiv \cos(\theta)$ (note $\tilde{\mathbf{u}}$ is a unit vector, so $\|\mathbf{u}_R\|^2 + \|\mathbf{u}_I\|^2 = 1$) and χ is the angle between \mathbf{u}_R and \mathbf{u}_I . Since we aim to maximize $x(\mathbf{M})$, we can take $\chi = \pi/2$ and then maximize over θ the function $f(\theta) = a^2 \cos^2(\theta) + b^2 \cos^2(\theta) \sin^2(\theta)$. The maximum depends on the ratio of a and b :

$$\max_\theta f(\theta) = \begin{cases} a^2 & \text{if } a > b, \\ \frac{(a^2 + b^2)^2}{4b^2} & \text{if } a \leq b. \end{cases} \quad (\text{B3})$$

It follows that

$$x(\mathbf{M}) = 1 - \frac{1}{2} \sum_{i: a_i > b_i} \frac{(a_i^2 - b_i^2)^2}{a_i^2 + b_i^2}, \quad (\text{B4})$$

where a_i and b_i are subject to the constraints $\sum_i a_i^2 = p_0$ and $\sum_i b_i^2 = 1 - p_0$, Eq. (19). For strong noise, $p_0 \leq 1/2$, it is possible to choose $a_i \leq b_i$ for all i . This gets rid of the

negative term in the sum and gives $x(\mathbf{M}) = 1$, i.e., complete purification, which is optimal. However, for weak noise ($p_0 > 1/2$) we have $\sum_i a_i^2 > \sum_i b_i^2$, so it is not possible to avoid the sum over i in the above expression.

We first handle the case in which $a_i > b_i$ for all i . Introducing $c_i = a_i^2 + b_i^2$ and $d_i = a_i^2 - b_i^2$, we have

$$x(\mathbf{M}) = 1 - \sum_i \frac{d_i^2}{2c_i} \leq 1 - \frac{(\sum_i d_i)^2}{2\sum_i c_i}, \quad (\text{B5})$$

where we used Sedrakyan's inequality (i.e., Cauchy-Schwartz applied to the vectors $\{\sqrt{c_i}\}$ and $\{d_i/\sqrt{c_i}\}$). The constraints in Eq. (19) impose $\sum_i c_i = 1$ and $\sum_i d_i = 2p_0 - 1$, therefore

$$x(\mathbf{M}) \leq 1 - \frac{(2p_0 - 1)^2}{2}. \quad (\text{B6})$$

This is saturated by setting $a_i = \sqrt{p_0/n}$ and $b_i = \sqrt{(1-p_0)/n}$ for all i .

Next, we consider the case in which $a_i \leq b_i$ for some i . We use primed sums to denote sums over i that are restricted only to those indices: $\sum_{i: a_i \leq b_i} = \sum'_i$. Let us define $\gamma \equiv \sum'_i c_i$ (we have $0 \leq \gamma \leq 1$, where $\gamma = 0$ recovers the previous case); then, by the same reasoning as above, we have

$$x(\mathbf{M}) \leq 1 - \frac{(2p_0 - 1 - \sum'_i d_i)^2}{2(1 - \gamma)}. \quad (\text{B7})$$

Now $-\sum'_i d_i = \sum'_i b_i^2 - a_i^2$ is non-negative by definition, so

$$x(\mathbf{M}) \leq 1 - \frac{(2p_0 - 1)^2}{2(1 - \gamma)}, \quad (\text{B8})$$

where the right-hand side is maximized for $\gamma = 0$. Thus the symmetric solution $a_i = \sqrt{p_0/n}$ and $b_i = \sqrt{(1-p_0)/n}$ for all i is optimal in general.

Finally, note that when $a_i > b_i$, the maximum in Eq. (B3) is achieved at $\theta = 0$, i.e., $\|\mathbf{u}_I\| = 0$; thus in the solution above, the unit vectors are real: $\tilde{\mathbf{u}} = \mathbf{u}_R$.

APPENDIX C: MEASUREMENT-INDUCED PHASE TRANSITION FOR THE OPTIMAL WEAK MEASUREMENT IN 1D RANDOM CIRCUITS

Here we analyze measurement-induced phase transition corresponding to weak measurements from optimal unraveling in Eq. (25) against stochastic projective measurements in Eq. (8).

We consider 1D Haar random brick-wall circuits with periodic boundary condition and optimal weak measurements applied between layers of unitaries. We use the

tripartite mutual information [85]

$$I_{3,n} = S_n(A) + S_n(B) + S_n(C) - S_n(A \cup B) - S_n(A \cup C) - S_n(B \cup C) + S_n(A \cup B \cup C), \quad (\text{C1})$$

where $S_n(X)$ is the Renyi entropy and A, B, C are contiguous subsystems of size $L/4$ (i.e., the system is divided into quarters and three such subsystems are used). $I_{3,n}$ was argued to be system-size-independent constant at criticality and thus can be used to accurately locate the critical point [27].

In Fig. 6 we show the $I_{3,n}$ for $n = 0.8, 1.0, 2.0$ at late time $t = 4L$, varying the system size L from 8 to 24. Since the crossing points drift to lower ε as system size increases, we only consider the data collapse for $L = 16, 20, 24$. The obtained critical point $\varepsilon_{c,\text{weak}} \approx 0.044$ is smaller than $\varepsilon_{c,\text{projective}} \approx 0.084$ [27] (note that in our convention ε is interpreted as noise rate, which is half of the measurement rate in the context of the measurement-induced phase transition: $p = 2\varepsilon$). This is consistent with our argument that for a given noise rate, optimal weak measurements are more disentangling than projective measurements.

APPENDIX D: OPTIMIZATION FOR THE AMPLITUDE DAMPING CHANNEL

In Sec. III D, we give an analytical solution to the optimization for unital qubit channels. Nonunital channels, on the other hand, are generally hard to solve exactly due to the absence of spherical symmetry, thus one usually seeks for a numerical solution using optimization algorithms. However, for a common nonunital channel, the amplitude damping channel with Kraus operators

$$\begin{cases} M_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\varepsilon} \end{pmatrix}, \\ M_1 = \begin{pmatrix} 0 & \sqrt{\varepsilon} \\ 0 & 0 \end{pmatrix}, \end{cases} \quad (\text{D1})$$

the optimization problem can be highly simplified since its Kraus rank is only 2. Here we consider the optimization over generic 2×2 unitary matrices taking the form

$$U = e^{i\varphi/2} \begin{pmatrix} e^{i\psi} & 0 \\ 0 & e^{-i\psi} \end{pmatrix} \begin{pmatrix} \cos \theta & \sin \theta \\ -\sin \theta & \cos \theta \end{pmatrix} \begin{pmatrix} e^{i\Delta} & 0 \\ 0 & e^{-i\Delta} \end{pmatrix}. \quad (\text{D2})$$

Substituting Eqs. (D1) and (D2) into Eq. (14), one can obtain

$$x = \frac{1}{2 + \frac{4\varepsilon^2}{1+3\varepsilon-(1-\varepsilon)\cos 4\theta}}, \quad (\text{D3})$$

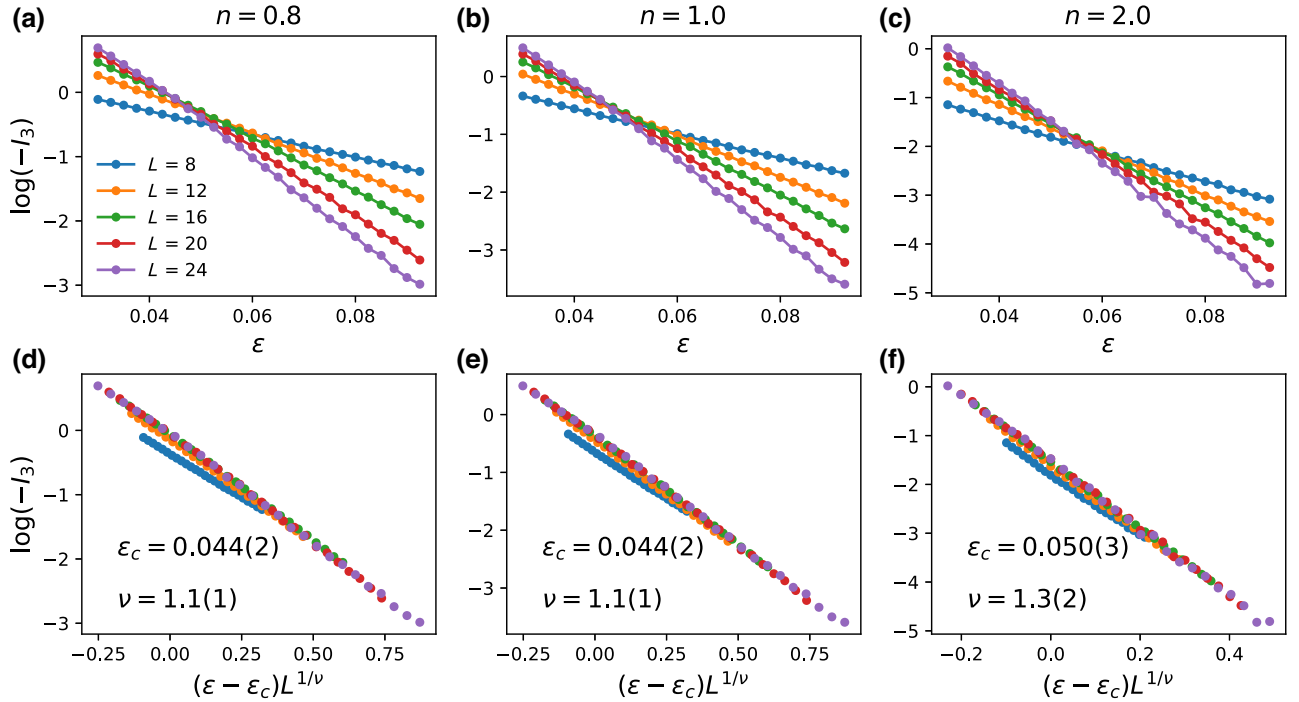


FIG. 6. (a)–(c) Tripartite mutual information I_3 [shown as $\log(-I_3)$] in 1D random circuits with dephasing noise of strength ε unraveled into optimal weak measurements, for Renyi indices (a) $n = 0.8$, (b) $n = 1.0$, and (c) $n = 2.0$. (d)–(f) Data collapse of $\log(-I_3)$ versus $(\varepsilon - \varepsilon_c)L^{1/\nu}$, with fit parameters indicated. The data are averaged over 5×10^2 – 5×10^3 realizations.

reaching maximum when $\theta = \pi/4$, which gives the optimized unraveling

$$\begin{cases} M_0 = \begin{pmatrix} 1 & \sqrt{\varepsilon} \\ 0 & \sqrt{1-\varepsilon} \end{pmatrix} / \sqrt{2}, \\ M_1 = \begin{pmatrix} -1 & \sqrt{\varepsilon} \\ 0 & -\sqrt{1-\varepsilon} \end{pmatrix} / \sqrt{2}. \end{cases} \quad (\text{D4})$$

We also numerically verify the effect of disentangling from using the optimized unraveling in the context of measurement-induced phase transition in 1D random circuits. In Fig. 7 we show the tripartite mutual information I_3 at late time $t = 4L$ in the 1D Haar random circuits subject to the amplitude damping channel. The critical point for optimized unraveling Eq. (D4) $\varepsilon_c \approx 0.17$ is much smaller than that for original unraveling Eq. (D1) $\varepsilon_c \approx 0.29$, indicating the disentangling effect from unraveling optimization also works for nonunitary channels.

APPENDIX E: DATA COLLAPSE

In Sec. IV B we determine the critical point ε_c and correlation length exponent ν by finding a data collapse satisfying the scaling ansatz

$$\tau(\varepsilon, L_x) = L_x F[(\varepsilon - \varepsilon_c)L_x^{1/\nu}]. \quad (\text{E1})$$

To quantify the collapse we consider the similar objective function as in Ref. [27]. We first sort the data by $x_i =$

$(\varepsilon_i - \varepsilon_c)L_{x,i}^{1/\nu}$ where $i = \{1, \dots, n\}$ labels sorted data points and define $y_i = \tau(x_i)/L_{x,i}$. Then the objective function $R(\varepsilon_c, \nu)$ is defined as

$$R(\varepsilon_c, \nu) = \frac{1}{n-2} \sum_{i=2}^{n-1} (y_i - \bar{y}_i)^2, \quad (\text{E2})$$

where

$$\bar{y}_i = \frac{(x_{i+1} - x_i)y_{i-1} - (x_{i-1} - x_i)y_{i+1}}{x_{i+1} - x_{i-1}}, \quad (\text{E3})$$

is the estimation of y_i given by linear interpolating (x_{i-1}, y_{i-1}) and (x_{i+1}, y_{i+1}) . Then the aim is to minimize Eq. (E2) over (ε_c, ν) . As an example, a color plot of R is shown in Fig. 8 for data set of $T = 5$ ($ABCDB$) shown in Fig. 3(c). We estimate the error by consider a region enclosed by $R = 1.3 \times R_{\min}$, which is shown as a black contour in Fig. 8.

APPENDIX F: PHASE BOUNDARY AT LARGE T

Here we present the results of stabilizer simulations of Clifford circuits to study the entanglement phase transition in circuits of large depth T . The goal is to qualitatively study the phase boundary in Fig. 1(d) at values of T that are beyond what can easily be studied via exact simulation of generic circuits (e.g., $T = 4, 5$ in Fig. 3). Due to

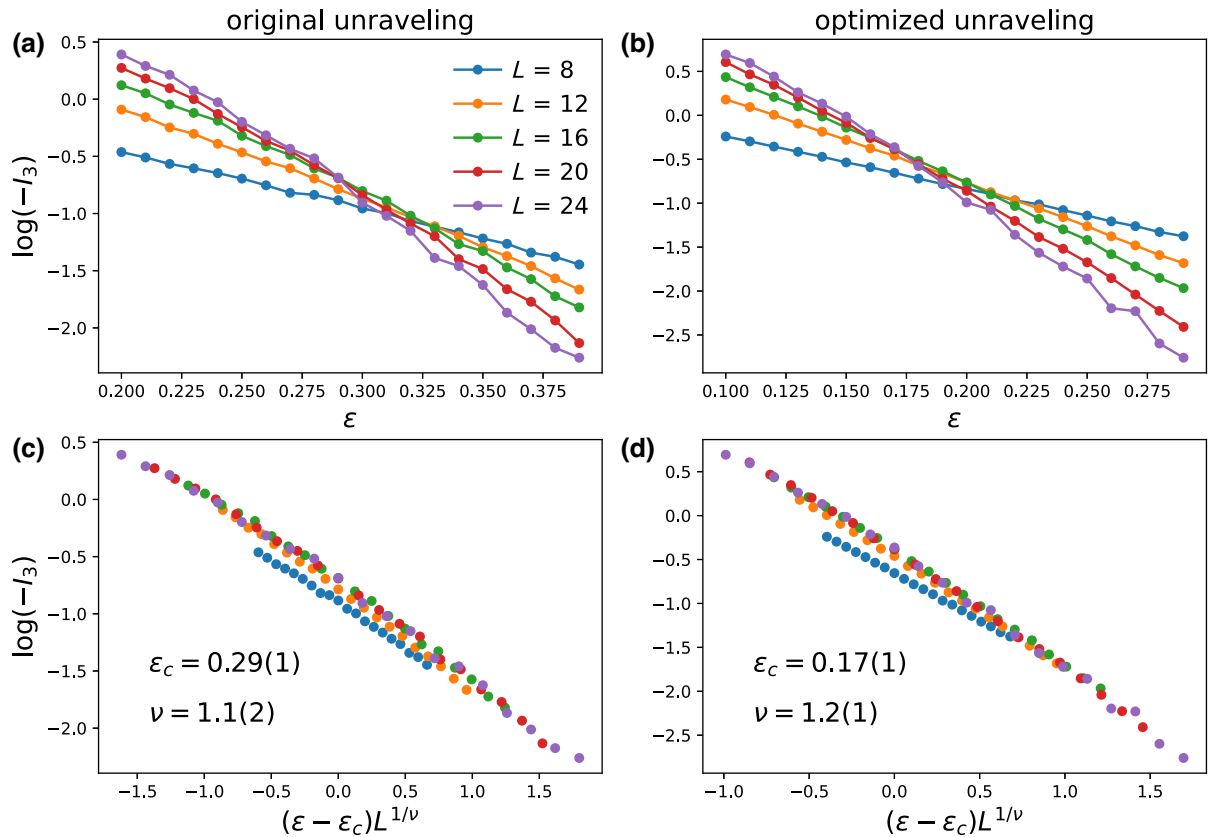


FIG. 7. Tripartite mutual information I_3 [shown as $\log(-I_3)$] for $n = 1.0$ in 1D random circuits with damping of strength ϵ for (a) original unraveling and (b) optimized unraveling. (c), (d) Corresponding data collapse of $\log(-I_3)$ versus $(\epsilon - \epsilon_c)L^{1/\nu}$, with fit parameters indicated. The data are averaged over 2×10^2 – 5×10^3 realizations.

the restriction of stabilizer simulation, we cannot unravel depolarizing noise into weak measurements. For this reason we use stochastic projective measurements, bearing in mind that this will give a *larger* numerical value of the

threshold noise strength ϵ_c (cf. Appendix C). We expect the qualitative behavior of $\epsilon_c(T)$ to be similar across weak and projective unravelings.

We consider Clifford circuits that mimic the models analyzed in Sec. IV B. To approximate the *i*SWAP-like gates, we sample two-qubit Clifford gates that are *i*SWAP with 90% probability and SWAP otherwise (corresponding to a uniform sampling of the *dual-unitary* half of the two-qubit Clifford group [40]). These gates are sandwiched between random single-qubit Clifford rotations. A projective measurement of Z is applied with probability $p = 2\epsilon$ on each qubit after each gate. Specifically, we simulate the space-wise evolution of the circuit as in, e.g., Fig. 2(c). This involves a quasi-1D subsystem, which is a strip of length L_x and width $1 + T/4$. As a diagnostic of the phase transition we use the tripartite mutual information $I_3(A : B : C)$ between contiguous regions that make up strips of length $L_x/4$. Results in Fig. 9 show a transition at critical noise rate $\epsilon_c(T)$ that increases with T , as expected. We have, e.g., $\epsilon_c(T = 8) \simeq 0.070$, which increases to $\epsilon_c(T = 24) \simeq 0.133$.

We additionally simulate the conventional MIPT in square 2D circuits, whose dimensions $L_x = L_y = L$ are jointly increased. The tripartite mutual information I_3 , for a

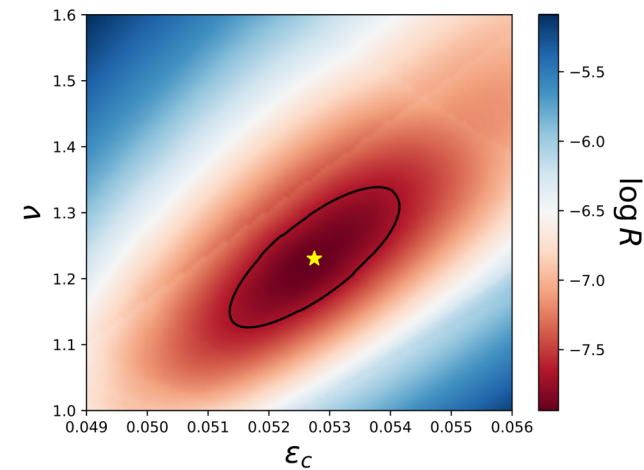


FIG. 8. Color map of data collapse object function R , Eq. (E2). The black contour encloses a region satisfying $R \leq 1.3 \times R_{\min}$, which gives an estimation of the error.

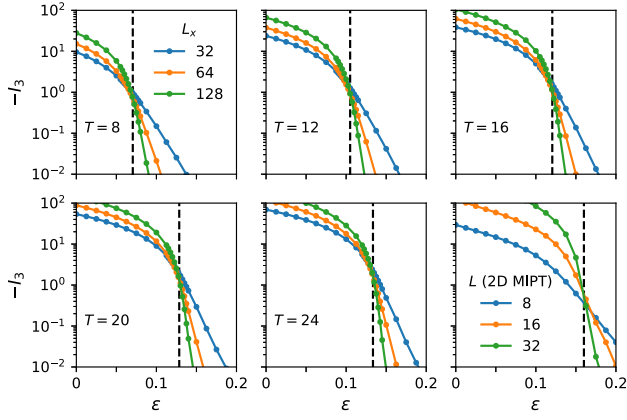


FIG. 9. Entanglement phase transition in Clifford circuits with dephasing noise of strength ε . First five panels refer to quasi-1D subsystems that mimic the noisy-SEBD simulation of 2D circuits of depth $T = 8, 12, 16, 20, 24$ ($T = 4$, not shown, is found to be in the area-law phase for all ε). The last panel shows data for the conventional MIPT in 2D square lattices of size $L \times L$, $L = 8, 16, 32$, with circuits of depth $O(L)$. Vertical dashed lines indicate estimates of the critical point. Data obtained by averaging between 4×10^2 and 2×10^4 realizations of the random circuits depending on system size.

partition of the square into four rectangles of size $L/4 \times L$, also shows a transition. We observe $\varepsilon_{c,2D} \simeq 0.16$, consistent with the value $p_{c,2D} = 0.3116(1)$ reported in Ref. [77] for this circuit architecture [86] (recall $p = 2\varepsilon$).

Finally, in Fig. 10 we compare the observed critical points for shallow circuits of depth T , $\varepsilon_c(T)$, with the 2D MIPT $\varepsilon_{c,2D}$. We see good agreement with the conjectured form $\varepsilon_c(T) \simeq \varepsilon_{c,2D} + O(1/T)$, sketched in Fig. 1(d).

APPENDIX G: BENCHMARKS

Here we benchmark our noisy-SEBD algorithm against stabilizer simulations of Clifford random circuits and MPO simulation with controlled errors. Since the depolarizing channel with general ε is not a Clifford operation, we consider the probabilistic trace setup [87], i.e., replacing the depolarizing channel by a probabilistic mixture of

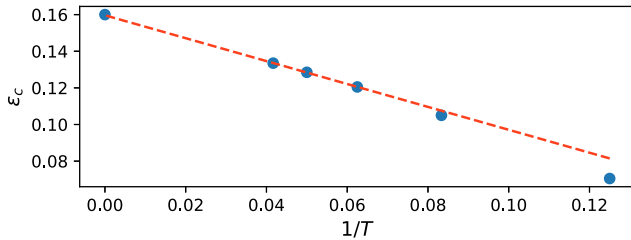


FIG. 10. Noise threshold for the entanglement phase transition as a function of circuit depth T (data from Fig. 9). Dashed line indicates a linear fit of datapoints $T = 16, 20, 24$ to $a + b/T$; the extrapolation to $T = \infty$ is in good agreement with the 2D MIPT datapoint.

an identity operation with probability $1 - \frac{4}{3}\varepsilon$ and a trace channel (or erasure) $\rho \mapsto (\mathbb{I}/2)_i \otimes \text{Tr}_i \rho$ with probability $\frac{4}{3}\varepsilon$. With this replacement, each instance or trajectory in this random ensemble is classically simulatable [49], and on average reproduces the effect of the depolarizing noise with strength ε .

We first consider the noisy sampling problem in the architecture given in Sec. IV B with $\varepsilon = 0.02$, $L_x = 5$, $L_y = 5$, and $T = 4$. In this case the system size is small enough that direct MPO simulation of the density matrix can be implemented with small error and can thus serve as a benchmark for noisy SEBD. In the upper row of Fig. 11, we show the averaged final probability $P_{\mathcal{N}}(\mathbf{z})$ of ten randomly chosen output bitstrings \mathbf{z} from noisy Clifford circuits, varying the number of sampled trajectories K from 10^3 to 10^5 (note we unravel depolarizing noise into probabilistic erasure for the stabilizer simulation, and into weak measurements for noisy SEBD; we use the same K for both methods). The probabilities $P_{\mathcal{N}}(\mathbf{z})$ are scaled by the exact reference value obtained from MPO simulation, whose truncation error is kept below 10^{-10} . As the trajectory number increases, one can see both SEBD and Clifford simulation show a good convergence to the exact result.

We then consider the same circuit architecture but with $L_x = 9$, $L_y = 9$, where exact MPO computation would require large computational effort. Therefore, in this case we directly compare SEBD results against the Clifford results (with 10^6 trajectories for the latter). Results are shown in the bottom row of Fig. 11, again displaying good agreement.

APPENDIX H: CONVERTING GATE FIDELITY TO NOISE RATE

The convention for noise strength ε used in this work is in terms of single-qubit channels as in Eq. (6). The usual figure of merit for two-qubit gates is the average fidelity f . To convert between the two, we note that

$$\begin{aligned} f &= \int d\psi \langle \psi | \Phi^{\otimes 2} [|\psi\rangle\langle\psi|] |\psi\rangle \\ &= (1 - \varepsilon)^2 + [1 - (1 - \varepsilon)^2] \int d\psi \langle \psi | P | \psi \rangle^2, \end{aligned} \quad (\text{H1})$$

where the integral is over the Haar measure on the two-qubit Hilbert space, P is any traceless Pauli operator, and the formula applies equally to dephasing and depolarizing noise [in fact to any unital noise channel upon setting $1 - \varepsilon \mapsto p_0$, cf. Eq. (16)]. The Haar integral yields $\text{Tr}(P^2)/20 = 1/5$, so

$$f = (1 - \varepsilon)^2 + \frac{\varepsilon(2 - \varepsilon)}{5} \simeq 1 - \frac{8}{5}\varepsilon, \quad (\text{H2})$$

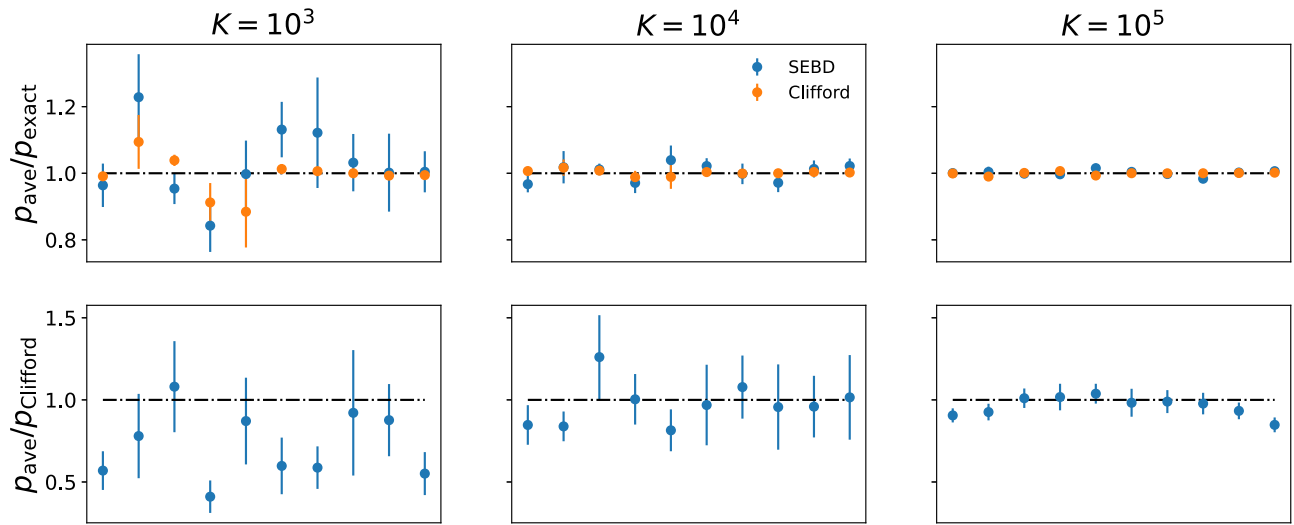


FIG. 11. Estimated probability of ten randomly generated output bitstrings of noisy Clifford circuits of depth $T = 4$, with $L_x = 5, L_y = 5$ (top row) and $L_x = 9, L_y = 9$ (bottom row). Probabilities are estimated by averaging over K trajectories, with $K = 10^3$ – 10^5 indicated on top. In the top row, the averaged probabilities p_{ave} (from noisy-SEBD and stabilizer simulations) are normalized by the exact value obtained from MPO simulation. In the bottom row, the averaged probabilities p_{ave} from noisy SEBD are normalized by the average of 10^6 trajectories of stabilizer simulations. The ratios converge towards 1 (horizontal lines) with increasing K .

at small ε , thus $\varepsilon \simeq (5/8)(1 - f)$. In particular, this means that a gate fidelity of $f = 96\%$ translates to $\varepsilon \simeq 0.025$, which is used in Sec. IV C.

-
- [1] S. Aaronson and L. Chenin *Proceedings of the 32nd Computational Complexity Conference, CCC '17* (Dagstuhl, DEU, 2017), p. 1.
 - [2] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, and J. M. Martinis, *et al.*, Characterizing quantum supremacy in near-term devices, *Nat. Phys.* **14**, 595 (2018).
 - [3] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, and B. Chiaro, *et al.*, A blueprint for demonstrating quantum supremacy with superconducting qubits, *Science* **360**, 195 (2018).
 - [4] A. Bouland, B. Fefferman, C. Nirkhe, and U. Vazirani, On the complexity and verification of quantum random circuit sampling, *Nat. Phys.* **15**, 159 (2019).
 - [5] R. Movassagh, Quantum supremacy and random circuits, Eprints [ArXiv:1909.06210](https://arxiv.org/abs/1909.06210) (2020).
 - [6] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, and S. Boixo, *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
 - [7] Y. Wu, W.-S. Bao, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, and H. Deng, *et al.*, Strong quantum computational advantage using a superconducting quantum processor, *Phys. Rev. Lett.* **127**, 180501 (2021).
 - [8] Q. Zhu, S. Cao, F. Chen, M.-C. Chen, X. Chen, T.-H. Chung, H. Deng, and Y. Du, *et al.*, Quantum computational advantage via 60-qubit 24-cycle random circuit sampling, *Sci. Bull.* **67**, 240 (2022).
 - [9] A. Morvan, B. Villalonga, X. Mi, S. Mandra, A. Bengtsson, P. V. Klimov, Z. Chen, and S. Hong, *et al.*, Phase transition in Random Circuit Sampling, Eprints [ArXiv:2304.11119](https://arxiv.org/abs/2304.11119) (2023).
 - [10] D. Aharonov, X. Gao, Z. Landau, Y. Liu, and U. Vazirani, in *Proceedings of the 55th Annual ACM Symposium on Theory of Computing, STOC 2023* (Association for Computing Machinery, New York, NY, USA, 2023), p. 945.
 - [11] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits anticentralize in log depth, *PRX Quantum* **3**, 010333 (2022).
 - [12] Y. Zhou, E. M. Stoudenmire, and X. Waintal, What limits the simulation of quantum computers?, *Phys. Rev. X* **10**, 041038 (2020).
 - [13] G. Kalachev, P. Pantelev, P. Zhou, and M.-H. Yung, Classical sampling of random quantum circuits with bounded fidelity, Eprints [ArXiv:2112.15083](https://arxiv.org/abs/2112.15083) (2021).
 - [14] T. Ayril, T. Louvet, Y. Zhou, C. Lambert, E. M. Stoudenmire, and X. Waintal, A density-matrix renormalization group algorithm for simulating quantum circuits with a finite fidelity, *PRX Quantum* **4**, 020304 (2023).
 - [15] Y. Li, X. Chen, and M. P. A. Fisher, Quantum Zeno effect and the many-body entanglement transition, *Phys. Rev. B* **98**, 205136 (2018).
 - [16] Y. Li, X. Chen, and M. P. A. Fisher, Measurement-driven entanglement transition in hybrid quantum circuits, *Phys. Rev. B* **100**, 134306 (2019).
 - [17] B. Skinner, J. Ruhman, and A. Nahum, Measurement-induced phase transitions in the dynamics of entanglement, *Phys. Rev. X* **9**, 031009 (2019).
 - [18] A. Chan, R. M. Nandkishore, M. Pretko, and G. Smith, Unitary-projective entanglement dynamics, *Phys. Rev. B* **99**, 224307 (2019).
 - [19] M. Szytniszewski, A. Romito, and H. Schomerus, Entanglement transition from variable-strength weak measurements, *Phys. Rev. B* **100**, 064204 (2019).

- [20] X. Cao, A. Tilloy, and A. D. Luca, Entanglement in a fermion chain under continuous monitoring, *SciPost Phys.* **7**, 024 (2019).
- [21] S. Choi, Y. Bao, X.-L. Qi, and E. Altman, Quantum error correction in scrambling dynamics and measurement-induced phase transition, *Phys. Rev. Lett.* **125**, 030505 (2020).
- [22] Y. Bao, S. Choi, and E. Altman, Theory of the phase transition in random unitary circuits with measurements, *Phys. Rev. B* **101**, 104301 (2020).
- [23] C.-M. Jian, Y.-Z. You, R. Vasseur, and A. W. W. Ludwig, Measurement-induced criticality in random quantum circuits, *Phys. Rev. B* **101**, 104302 (2020).
- [24] X. Turkeshi, R. Fazio, and M. Dalmonte, Measurement-induced criticality in $(2 + 1)$ -dimensional hybrid quantum circuits, *Phys. Rev. B* **102**, 014315 (2020).
- [25] M. J. Gullans and D. A. Huse, Dynamical purification phase transition induced by quantum measurements, *Phys. Rev. X* **10**, 041020 (2020).
- [26] M. J. Gullans and D. A. Huse, Scalable probes of measurement-induced criticality, *Phys. Rev. Lett.* **125**, 070606 (2020).
- [27] A. Zabalo, M. J. Gullans, J. H. Wilson, S. Gopalakrishnan, D. A. Huse, and J. H. Pixley, Critical properties of the measurement-induced transition in random quantum circuits, *Phys. Rev. B* **101**, 060301 (2020).
- [28] Y. Bao, S. Choi, and E. Altman, Symmetry enriched phases of quantum circuits, *Ann. Phys. (N. Y.)* **435**, 168618 (2021).
- [29] Y. Li, X. Chen, A. W. W. Ludwig, and M. P. A. Fisher, Conformal invariance and quantum nonlocality in critical hybrid circuits, *Phys. Rev. B* **104**, 104305 (2021).
- [30] Y. Li, S. Vijay, and M. P. A. Fisher, Entanglement domain walls in monitored quantum circuits and the directed polymer in a random environment, *PRX Quantum* **4**, 010331 (2023).
- [31] O. Alberton, M. Buchhold, and S. Diehl, Entanglement transition in a monitored free-fermion chain: From extended criticality to area law, *Phys. Rev. Lett.* **126**, 170602 (2021).
- [32] A. Lavasani, Y. Alavirad, and M. Barkeshli, Measurement-induced topological entanglement transitions in symmetric random quantum circuits, *Nat. Phys.* **17**, 342 (2021).
- [33] M. Ippoliti, M. J. Gullans, S. Gopalakrishnan, D. A. Huse, and V. Khemani, Entanglement phase transitions in measurement-only dynamics, *Phys. Rev. X* **11**, 011030 (2021).
- [34] A. Nahum, S. Roy, B. Skinner, and J. Ruhman, Measurement and entanglement phase transitions in all-to-all quantum circuits, on quantum trees, and in Landau-Ginsburg theory, *PRX Quantum* **2**, 010352 (2021).
- [35] S. Sang and T. H. Hsieh, Measurement-protected quantum phases, *Phys. Rev. Res.* **3**, 023200 (2021).
- [36] A. Lavasani, Y. Alavirad, and M. Barkeshli, Topological order and criticality in $(2 + 1)$ D monitored random quantum circuits, *Phys. Rev. Lett.* **127**, 235701 (2021).
- [37] M. Ippoliti and V. Khemani, Postselection-free entanglement dynamics via spacetime duality, *Phys. Rev. Lett.* **126**, 060501 (2021).
- [38] T.-C. Lu and T. Grover, Spacetime duality between localization transitions and measurement-induced transitions, *PRX Quantum* **2**, 040319 (2021).
- [39] M. Ippoliti, T. Rakovszky, and V. Khemani, Fractal, logarithmic, and volume-law entangled nonthermal steady states via spacetime duality, *Phys. Rev. X* **12**, 011045 (2022).
- [40] A. Zabalo, M. J. Gullans, J. H. Wilson, R. Vasseur, A. W. W. Ludwig, S. Gopalakrishnan, D. A. Huse, and J. H. Pixley, Operator scaling dimensions and multifractality at measurement-induced transitions, *Phys. Rev. Lett.* **128**, 050602 (2022).
- [41] Z. Weinstein, Y. Bao, and E. Altman, Measurement-induced power-law negativity in an open monitored quantum circuit, *Phys. Rev. Lett.* **129**, 080501 (2022).
- [42] C. Noel, P. Niroula, D. Zhu, A. Risinger, L. Egan, D. Biswas, M. Cetina, and A. V. Gorshkov, *et al.*, Measurement-induced quantum phases realized in a trapped-ion quantum computer, *Nat. Phys.* **18**, 760 (2022).
- [43] J. M. Koh, S.-N. Sun, M. Motta, and A. J. Minnich, Measurement-induced entanglement phase transition on a superconducting quantum processor with mid-circuit readout, *Nat. Phys.* **19**, 1314 (2023).
- [44] Google Quantum AI and Collaborators, Measurement-induced entanglement and teleportation on a noisy quantum processor, *Nature* **622**, 481 (2023).
- [45] J. C. Napp, R. L. La Placa, A. M. Dalzell, F. G. S. L. Brandao, and A. W. Harrow, Efficient classical simulation of random shallow 2D quantum circuits, *Phys. Rev. X* **12**, 021021 (2022).
- [46] B. M. Terhal and D. P. DiVincenzo, Adaptive quantum computation, constant depth quantum circuits and Arthur-Merlin games, Eprints [ArXiv:quant-ph/0205133](https://arxiv.org/abs/quant-ph/0205133) (2004).
- [47] B. Ware, A. Deshpande, D. Hangleiter, P. Niroula, B. Fefferman, A. V. Gorshkov, and M. J. Gullans, A sharp phase transition in linear cross-entropy benchmarking, Eprints [ArXiv:2305.04954](https://arxiv.org/abs/2305.04954) (2023).
- [48] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, Quantum entanglement growth under random unitary dynamics, *Phys. Rev. X* **7**, 031016 (2017).
- [49] S. Aaronson and D. Gottesman, Improved simulation of stabilizer circuits, *Phys. Rev. A* **70**, 052328 (2004).
- [50] L. G. Valiant, Quantum circuits that can be simulated classically in polynomial time, *SIAM J. Comput.* **31**, 1229 (2002).
- [51] B. M. Terhal and D. P. DiVincenzo, Classical simulation of noninteracting-fermion quantum circuits, *Phys. Rev. A* **65**, 032325 (2002).
- [52] R. Jozsa and A. Miyake, Matchgates and classical simulation of quantum circuits, *Proc. R. Soc. A: Math. Phys. Eng. Sci.* **464**, 3089 (2008).
- [53] 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, *PRX Quantum* **3**, 040329 (2022).
- [54] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandao, Random quantum circuits transform local noise into global white noise, Eprints [ArXiv:2111.14907](https://arxiv.org/abs/2111.14907) (2021).
- [55] 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, Eprints [ArXiv:2112.01657](https://arxiv.org/abs/2112.01657) (2021).

- [56] F. Pan, K. Chen, and P. Zhang, Solving the sampling problem of the sycamore quantum circuits, *Phys. Rev. Lett.* **129**, 090502 (2022).
- [57] D. Perez-Garcia, F. Verstraete, M. M. Wolf, and J. I. Cirac, Matrix product state representations, *Quantum Inf. Comput.* **7**, 401 (2007).
- [58] F. Verstraete, V. Murg, and J. Cirac, Matrix product states, projected entangled pair states, and variational renormalization group methods for quantum spin systems, *Adv. Phys.* **57**, 143 (2008).
- [59] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, Matrix product states and projected entangled pair states: Concepts, symmetries, theorems, *Rev. Mod. Phys.* **93**, 045003 (2021).
- [60] In fact, compression into MPS form depends on the behavior of small Schmidt eigenvalues, which is captured by Renyi entropies S_n with $n < 1$. In typical random circuits the von Neumann and Renyi entropies have similar scaling for all values of n away from 0.
- [61] F. Verstraete, J. J. García-Ripoll, and J. I. Cirac, Matrix product density operators: Simulation of finite-temperature and dissipative systems, *Phys. Rev. Lett.* **93**, 207204 (2004).
- [62] K. Noh, L. Jiang, and B. Fefferman, Efficient classical simulation of noisy random quantum circuits in one dimension, *Quantum* **4**, 318 (2020).
- [63] S. Cheng, C. Cao, C. Zhang, Y. Liu, S.-Y. Hou, P. Xu, and B. Zeng, Simulating noisy quantum circuits with matrix product density operators, *Phys. Rev. Res.* **3**, 023005 (2021).
- [64] J. Eisert, M. Cramer, and M. B. Plenio, Colloquium: Area laws for the entanglement entropy, *Rev. Mod. Phys.* **82**, 277 (2010).
- [65] R. Fan, S. Vijay, A. Vishwanath, and Y.-Z. You, Self-organized error correction in random unitary circuits with measurement, *Phys. Rev. B* **103**, 174309 (2021).
- [66] T. Vovk and H. Pichler, Entanglement-optimal trajectories of many-body quantum Markov processes, *Phys. Rev. Lett.* **128**, 243601 (2022).
- [67] J. Dalibard, Y. Castin, and K. Mølmer, Wave-function approach to dissipative processes in quantum optics, *Phys. Rev. Lett.* **68**, 580 (1992).
- [68] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
- [69] R. Dum, P. Zoller, and H. Ritsch, Monte Carlo simulation of the atomic master equation for spontaneous emission, *Phys. Rev. A* **45**, 4879 (1992).
- [70] C. W. Gardiner, A. S. Parkins, and P. Zoller, Wave-function quantum stochastic differential equations and quantum-jump simulation methods, *Phys. Rev. A* **46**, 4363 (1992).
- [71] L. Tian and H. J. Carmichael, Quantum trajectory simulations of two-state behavior in an optical cavity containing one atom, *Phys. Rev. A* **46**, R6801 (1992).
- [72] M. Kolodrubetz, Optimality of Lindblad unfolding in measurement phase transitions, *Phys. Rev. B* **107**, L140301 (2023).
- [73] Note that the purity $\mathcal{P} = \text{Tr}(\rho^2)$ is related to the second Renyi entropy S_2 via $\mathcal{P} = e^{-S_2}$. Using convexity and the fact that $S_2 \leq S_\alpha$ for all $\alpha < 2$, we have a lower bound $\overline{S_\alpha} \geq -\ln \overline{\mathcal{P}}$ between the average purity and the average Renyi entropies $S_{\alpha < 1}$. The latter determine compressibility of the state into MPS form [88].
- [74] C.-K. Li and M.-D. Choi, On unital qubit channels, Eprints [ArXiv:2301.01358](https://arxiv.org/abs/2301.01358) (2023).
- [75] Note we can assume a_i real and non-negative by redefining M_i by an overall phase; the phase of b_i can then be absorbed into \tilde{u}_i .
- [76] Y. Bao, M. Block, and E. Altman, Finite time teleportation phase transition in random quantum circuits, [ArXiv:2110.06963](https://arxiv.org/abs/2110.06963) (2022).
- [77] P. Sierant, M. Schirò, M. Lewenstein, and X. Turkeshi, Measurement-induced phase transitions in $(d + 1)$ -dimensional stabilizer circuits, *Phys. Rev. B* **106**, 214316 (2022).
- [78] X. Mi, *et al.*, Information scrambling in quantum circuits, *Science* **374**, 1479 (2021).
- [79] This scaling corresponds to an MPS with a number $\propto L_x$ of tensors, of constant bond dimension and physical dimension $\propto 2^{cT}$, evolved for $\propto L_y$ steps.
- [80] R. Harper, S. T. Flammia, and J. J. Wallman, Efficient learning of quantum noise, *Nat. Phys.* **16**, 1184 (2020).
- [81] Z. Chen, Y. Bao, and S. Choi, Optimized trajectory unraveling for classical simulation of noisy quantum dynamics, [ArXiv:2306.17161](https://arxiv.org/abs/2306.17161) (2023).
- [82] G. Kostenberger, Weingarten calculus, Eprints [ArXiv:2101.00921](https://arxiv.org/abs/2101.00921).
- [83] T. P. Eggarter, Triangular antiferromagnetic Ising model, *Phys. Rev. B* **12**, 1933 (1975).
- [84] Y. Tanaka and N. Uryû, Triangular Ising lattice with anisotropic interactions, *J. Phys. Soc. Jpn.* **44**, 1091 (1978).
- [85] For $n \neq 1$ this quantity is not a proper mutual information.
- [86] See Appendix A 1 therein. While the circuit architecture is the same, the gate set is slightly different—random Clifford gates, not restricted to SWAP and *i*SWAP.
- [87] Z. Li, S. Sang, and T. H. Hsieh, Entanglement dynamics of noisy random circuits, *Phys. Rev. B* **107**, 014307 (2023).
- [88] N. Schuch, M. M. Wolf, F. Verstraete, and J. I. Cirac, Entropy scaling and simulability by matrix product states, *Phys. Rev. Lett.* **100**, 030504 (2008).