

Universality in the Anticoncentration of Noisy Quantum Circuits at Finite Depths

Arman Sauliere^{1,*}, Guglielmo Lami¹, Corentin Boyer¹, Jacopo De Nardis¹,
and Andrea De Luca^{1,2}

¹*Laboratoire de Physique Théorique et Modélisation, CNRS UMR 8089, CY Cergy Paris Université, Cergy-Pontoise Cedex 95302, France*

²*Laboratoire de Physique de l'École Normale Supérieure, ENS, Université PSL, CNRS, Sorbonne Université, Université Paris-Cité, Paris 75005, France*

 (Received 11 October 2025; revised 3 February 2026; accepted 17 March 2026; published 22 June 2026)

We present universal properties of anticoncentration in *weakly* noisy quantum circuits at finite depth. We develop a generic framework for single- and multi-qubit noise channels in the weak-noise limit and introduce an effective description in terms of a *random matrix product operator*. Within this weak-noise regime, we show that distinct noise mechanisms act in a quantitatively similar way, yielding a universal distribution of bit-string probabilities that is largely independent of the microscopic noise channel and of the circuit architecture. We identify three depth-dependent regimes, each characterized by a distinct scaling of cross-entropy benchmarking (XEB) with rescaled depth. In the shallow-depth regime, noise effects are perturbatively small; in the intermediate regime, circuit-induced fluctuations and noise compete on equal footing; and in the deep-depth regime, the output distribution becomes effectively classical, up to corrections that are exponentially small in the noise strength. We provide quantitative predictions for anticoncentration in generic finite-depth circuits and benchmark them against numerical simulations, finding excellent agreement even at shallow depths. Moreover, we show that, contrary to previous expectations, the late-time value of XEB provides direct access to the global circuit fidelity, even at large noise strengths. Our results are directly applicable to current quantum processors and demonstrate universal behavior beyond the pure random-matrix-theory regime, which only emerges at asymptotically large depths.

DOI: [10.1103/xl16-cdy9](https://doi.org/10.1103/xl16-cdy9)

I. INTRODUCTION

Quantum machines have transformative potential in many scientific fields, but current and near-term devices remain limited by external noise [1–12]. Although advances reduce noise levels, only fault-tolerant error correction can effectively suppress errors [4,13–17]. Meanwhile, the development of methods to characterize noisy quantum circuits and to benchmark their output remains crucial [18–26]. A widely used benchmarking technique is *random circuit sampling* (RCS), in which one samples bit strings $\mathbf{x}$ from the output density matrix ρ of a quantum circuit U [1,27–29]. By comparing the observed output statistics with those from an ideal (noise-free) classical simulation of the same circuit, one quantifies the hardware global fidelity. This procedure, known as *cross-entropy*

benchmarking (XEB), is experimentally accessible and has attracted significant interest [1,3,30–33]. Since XEB compares two replicas of the system—one noisy and one ideal—it can be readily analyzed using effective circuit averaging techniques developed for random unitary circuits [34]. However, it captures only partial information about the *anticoncentration* properties of the output distribution in the computational basis [33,35–48]. The latter are instead fully encoded in the so-called *probability-of-probabilities distribution* (PoP) [1,49,50] $P(w)$, i.e., the average distribution of output probabilities $w = D\langle \mathbf{x} | \rho | \mathbf{x} \rangle$ (for convenience rescaled by the Hilbert space dimension D). The moments of the PoP are proportional to the celebrated *inverse participation ratios* (IPRs), often used to signal localization and nonergodicity effects [42,51].

In the absence of noise, the PoP for Haar-random quantum states follows the Porter-Thomas (PT) distribution [52], that for $D \rightarrow \infty$ takes the simple form $P_{\text{PT}}(w) = e^{-w}$. The same distribution is obtained for generic circuits of sufficiently large depth in the absence of any conserved quantities. By contrast, generic classical stochastic processes produce uniform bit-string sampling, implying $P_C(w) = \delta(w - 1)$, a distribution that also characterizes

*Contact author: armansauliere@gmail.com

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.

quantum devices overwhelmed by noise. In general, a noisy quantum circuit is expected to interpolate between these two extremes as a function of the noise strength. Additionally, these forms hold strictly in the infinite-depth limit [1,49,50]; capturing the quantum-classical crossover requires incorporating finite-depth corrections.

II. SUMMARY OF RESULTS

In this work, we analyze the combined impact of finite circuit depth and noise on both the linear cross-entropy benchmarking (XEB) estimator and the full probabilities-of-probabilities (PoP) distribution in a short-range, $1 + 1$ -dimensional quantum circuit with open boundary conditions. We focus on the *weak-noise regime* [30,31,53,54], where η , a measure of the total number of errors in the circuit, can be treated as an *intensive* parameter, i.e., it does not grow linearly with the system size. We obtain explicit and exact results in the thermodynamic limit $N \rightarrow \infty$ at fixed η . It is clear that the most physical assumption in realistic setups is that the number of errors η in a circuit is at least proportional to the number of gates and qubits. However, state-of-the-art platforms operate at a very low error rate [3,55–58], so that for the N currently accessible, one effectively has $\eta \ll N$, a regime in which, as we shall see, our predictions prove useful and accurate.

We show that, in this weak-noise regime, both the XEB and the PoP distributions exhibit a *universal scaling regime* under *generic forms of noise*. This universal scaling is characterized by two microscopic parameters that can be accessed experimentally from just the XEB: η , directly related to the global fidelity $F = e^{-\eta}$, and x , a rescaled inverse depth defined as

$$x = \frac{N}{L(t)}, \quad L(t) = L_0 e^{t/\tau}. \quad (1)$$

The effective length $L(t)$ is the characteristic scale controlling the approach to random-matrix behavior, often referred to in quantum chaos as a *Thouless length*: it marks the scale at which the circuit locally behaves as if it were fully random [44,45,59–61]. The growth of $L(t)$ in time—which can be viewed as a reparameterization of circuit depth—occurs on the same timescale τ that governs the decay of the half-chain purity. Denoting by $\rho_{N/2}$ the reduced density matrix of half the system and by $P_{N/2} = \text{Tr}(\rho_{N/2}^2)$ the corresponding purity, a standard entanglement diagnostic, we have that, starting from a generic product state, $P_{N/2} \sim e^{-t/\tau} \sim L(t)^{-1}$ [43,45,62]. The rescaled depth x in Eq. (1) thus targets a shallow-circuit regime, where time scales logarithmically with system size, namely $t = O(\log N)$.

Within this scaling regime, we identify *three dynamical regimes*, determined by the interplay between inverse depth x and noise η , and we fully characterize the PoP

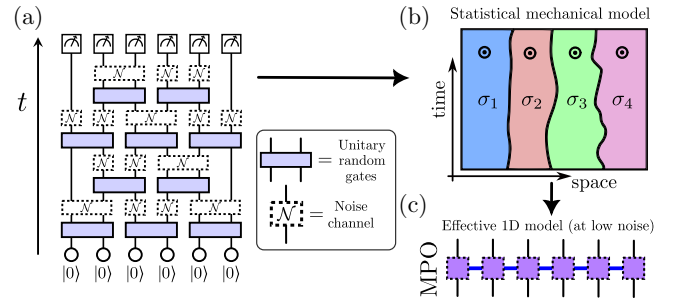


FIG. 1. Sketch of the work. (a) We consider the task of random circuit sampling. The circuit is composed by random unitary gates in the usual brickwall geometry, and it is affected by noise represented by a quantum channel $\mathcal{N}$. We explore noise acting either on single or two-adjacent qubits, all yielding similar conclusions. (b) Theoretical arguments show that the gate average of the replicated circuit gives rise to an effective one-dimensional statistical model of permutations. In this model, noise acts as an external field that biases the system toward permutations close to the identity. (c) As a result, the original spatiotemporal circuit becomes equivalent to a one-dimensional random matrix product operator which incorporate noise.

distribution $P_{x,\eta}(w)$ in each of them. This yields a quantitative description of the quantum-to-classical crossover at finite depth, i.e., the progressive loss of distinctly quantum sampling features as noise and finite depth compete.

Our *universal results* rely on the observation that, in the weak-noise limit and within the scaling regime (1), generic chaotic circuits subject to different noise mechanisms can be reduced to an effective *random matrix product operator* (RMPO), a tractable random-tensor description of the noisy evolution (see also Fig. 1). The RMPO construction extends the notion of *random matrix product states* (RMPS)—random tensor-network models used to capture typical properties of pure states [63–68]—to noisy channels. This is achieved by doubling the Hilbert space (a standard purification technique to represent mixed states and noise processes) and by incorporating the noisy evolution step by step along the circuit.

We establish this mapping explicitly in a controlled setting, namely the random phase model (RPM) [45,59], in the limit of large local Hilbert-space dimension. In this limit we obtain explicit expressions for XEB and for the moments of the outcome probabilities, and we introduce an efficient numerical procedure to reconstruct the full PoP distribution. Through extensive numerical simulations, we show that the scaling regime provides a quantitatively accurate description also for random circuits with brick-wall geometry, a standard architecture widely used in many-body quantum chaos studies [34,69–80], and for several experimentally relevant noise models.

Finally, we consider the more realistic situation in which the number of errors grows with circuit depth, $\eta = \lambda t$,

while remaining within the regime $\lambda t \ll N$, corresponding either to shallow circuits or sufficiently weak noise, where our results remain applicable. In this setting, we use our theory to derive a universal late-intermediate-time form for XEB that captures the crossover in its decay rate as the noise strength is varied [3]. Contrary to what was previously believed, we also show that, even at generic noise strength, XEB can still be used to extract the circuit's global fidelity.

Overall, our results provide a unified framework for understanding how noise and finite depth jointly shape sampling statistics and benchmarking metrics in near-term quantum processors.

III. SETUP AND METHODS

This section introduces the main notation, quantities, and setup used throughout the paper. Table I at the end of this section collects the notation used throughout the rest of the manuscript (with some symbols defined in subsequent sections).

A. Setup

We consider a system of N qudits with local dimension d , and total Hilbert-space dimension $D = d^N$. The computational basis is labeled by bit strings $|\mathbf{x}\rangle = |x_1, \dots, x_N\rangle$, where $x_i \in \{0, \dots, d-1\}$. The task of RCS is to apply a random (but fixed) quantum circuit U to the initial state $|\mathbf{0}\rangle = |0, \dots, 0\rangle$, which yields $\rho(U) = U|\mathbf{0}\rangle\langle\mathbf{0}|U^\dagger$, and then use quantum measurements to sample bit strings $\mathbf{x}$ from the output distribution $p(\mathbf{x}; \rho(U))$, where we define

$$p(\mathbf{x}; \rho) := \langle \mathbf{x} | \rho | \mathbf{x} \rangle. \quad (2)$$

To fix ideas, we consider a brickwall circuit with alternating layers of two-qudit gates, all independently drawn from the Haar measure $\text{Haar}(d^2)$ (see Fig. 1).

TABLE I. Notation used throughout the manuscript.

Symbol	Meaning
N	System size (number of qudits)
d	Local Hilbert-space dimension
D	Total Hilbert-space dimension
q	Generic Hilbert-space dimension
t	Circuit depth
τ	Decay rate of half-system purity (noiseless)
x	Rescaled inverse circuit depth $x = e^{-t/\tau} \cdot N/L_0$
γ	Microscopic noise parameter/strength
λ	Number of errors per circuit layer
η	Total number of errors, fixed by global fidelity $F = e^{-\eta}$

B. Noise models

In realistic settings, gates are subject to noise, which we model by a quantum channel of the form

$$\mathcal{N}(\rho) = \sum_{\alpha} K_{\alpha} \rho K_{\alpha}^{\dagger}, \quad \sum_{\alpha} K_{\alpha}^{\dagger} K_{\alpha} = \mathbb{1}, \quad (3)$$

where K_{α} are Kraus operators. Under this noise, the system evolves into a noisy version $\rho_{\mathcal{N}}(U)$ of the final state, and consequently the output distribution becomes $p(\mathbf{x}; \rho_{\mathcal{N}}(U))$. To discuss the weak-noise limit, we assume that $\mathcal{N}$ belongs to a smooth family of completely positive and trace-preserving (CPTP) map passing through the identity, parameterized via a formal microscopic positive parameter γ . Consequently, both the channel and its Kraus operators depend explicitly on this parameter, i.e., $\mathcal{N} \rightarrow \mathcal{N}_{\gamma}$ and $K_{\alpha} \rightarrow K_{\alpha}(\gamma)$. The parameter γ here represents the dominant scale for noise in the setup quantifying the *strength* of the noise channel $\mathcal{N}$. Consequently, for $\gamma = 0$, the noise channel reduces to the identity, namely $\lim_{\gamma \rightarrow 0+} \mathcal{N}_{\gamma}(\rho) = \rho$, and $\lim_{\gamma \rightarrow 0+} K_{\alpha}(\gamma) = \delta_{\alpha,0} \mathbb{1}$. As we will see explicitly, this parameterization does not restrict the ensemble of possible noise sources we consider, nor the coexistence of different noises with comparable intensities, as long as they all remain weak. Indeed, in Sec. IV we provide an explicit example for a gate depending on two noise parameters. As an explicit example, the depolarizing noise channel is given by

$$\mathcal{N}_{\gamma}^{\text{dep.}}(\rho) = (1 - \gamma)\rho + \gamma \frac{\mathbb{1}}{q}, \quad (4)$$

where $\gamma \in [0, 1]$ and q is the dimension of the Hilbert space on which the noisy channel acts. In general, we consider scenarios where the noise channel acts on single qudits ($q = d$) or on multiple qudits (for example, two qudits with $q = d^2$).

Since we focus on the weak-noise regime, we expand $\mathcal{N}_{\gamma}$ at first-order in γ , so that the action of the noisy channel becomes $\mathcal{N}_{\gamma}(\rho) = \sum_{\alpha} K_{\alpha}(\gamma) \rho K_{\alpha}^{\dagger}(\gamma) = \rho + \gamma \tilde{\mathcal{N}}(\rho) + O(\gamma^2)$, with $\tilde{\mathcal{N}}(\rho)$ a Lindblad operator [81]. Explicitly, from the standard first-order expansion of the Kraus operators

$$\begin{cases} K_0(\gamma) = \mathbb{1} - \gamma \tilde{K}_0 + O(\gamma^2), \\ K_{\alpha}(\gamma) = \sqrt{\gamma} \tilde{K}_{\alpha} + O(\gamma^{3/2}) \quad \text{for } \alpha \geq 1, \end{cases} \quad (5)$$

which implies

$$\tilde{\mathcal{N}}(\rho) = \sum_{\alpha \geq 1} \tilde{K}_{\alpha} \rho \tilde{K}_{\alpha}^{\dagger} - \tilde{K}_0 \rho - \rho \tilde{K}_0^{\dagger}. \quad (6)$$

The trace-preserving condition for the noise channel implies $\tilde{K}_0 + \tilde{K}_0^{\dagger} = \sum_{\alpha \geq 1} \tilde{K}_{\alpha}^{\dagger} \tilde{K}_{\alpha}$. The Kraus operators are

not uniquely defined, and we fix this freedom by choosing the $\tilde{K}_\alpha$ to be traceless. Notice also that one has the freedom to rescale γ and the operators $\tilde{K}$. In particular, the transformation $\gamma \rightarrow \gamma/a$, $\tilde{K}_0 \rightarrow a\tilde{K}_0$, $\tilde{K}_\alpha \rightarrow \sqrt{a}\tilde{K}_\alpha$ (for any value of a) leaves Eq. (5) unchanged. In any case, as we anticipated, the expansion to the first order in γ still leaves great freedom in the choice of an arbitrary Lindblad operator $\tilde{\mathcal{N}}$ and the corresponding jump operators $\tilde{K}_\alpha$.

C. Anticoncentration properties

A crucial quantity to assess the performance of RCS is the *linear cross-entropy benchmark* (XEB),

$$\text{XEB}(U) = D \sum_{\mathbf{x}} p(\mathbf{x}; \rho_{\mathcal{N}}(U)) p(\mathbf{x}; \rho(U)) - 1, \quad (7)$$

which quantifies the correlation between the ideal and the noisy distributions. For intermediate system sizes, XEB can be estimated by combining experimental data with a classical simulation that gives direct access to $p(\mathbf{x}; \rho(U))$ [1] and, for deep circuits ($t \gg \log N$, i.e., $x \rightarrow 0$), it has been used as an experimentally friendly proxy for the fidelity $F(U) = \text{Tr}[\rho_{\mathcal{N}}(U)\rho(U)]$ between the ideal target state and its noisy counterpart; however, recent works have pointed out that this correspondence may only hold when the noise is sufficiently weak [3,30,31,53].

Other key quantities we analyze are the PoP and its moments, represented by the IPRs [50], defined respectively as

$$\begin{aligned} P(w; U) &= D^{-1} \sum_{\mathbf{x}} \delta(w - Dp(\mathbf{x}; \rho_{\mathcal{N}}(U))), \\ I_k(U) &= \sum_{\mathbf{x}} p(\mathbf{x}; \rho_{\mathcal{N}}(U))^k \quad \text{with } k \geq 1. \end{aligned} \quad (8)$$

These quantities probe the anticoncentration of the final state in the computational basis. Up to a constant factor, the IPR equals the k th moment of the PoP: $\mathbb{E}[w^k] = D^{k-1} I_k$.

D. Weingarten calculus

While the quantities defined above depend explicitly on the specific circuit instance U , one can average them over an ensemble of random circuits, thereby defining $\text{XEB} = \mathbb{E}_U[\text{XEB}(U)]$, $F = \mathbb{E}_U[F(U)]$, $I_k = \mathbb{E}_U[I_k(U)]$, and $P(w) = \mathbb{E}_U[P(w; U)]$. To compute such circuit averages $\mathbb{E}_U[\dots]$, we employ Weingarten calculus [82,83], which expresses the Haar average of k copies of a unitary gate as a sum over permutation operators $\sigma \in S_k$, where S_k is the symmetric group of order k . Specifically, using the vectorized representation of operators $O \rightarrow |O\rangle\rangle$ and of their inner product $\text{tr}(O^\dagger Q) = \langle\langle O|Q\rangle\rangle$, the Weingarten

formula reads [82,83]

$$\mathbb{E}_{U \sim \text{Haar}(q)} [(U^* \otimes U)^{\otimes k}] = \sum_{\pi, \sigma \in S_k} \text{Wg}_{\pi, \sigma}(q) |\pi\rangle\rangle \langle\langle \sigma|. \quad (9)$$

Here, $\text{Wg}(q)$ denotes the $k! \times k!$ Weingarten matrix, which is the pseudo-inverse of the overlap matrix $G_{\pi, \sigma}(q) := \langle\langle \pi | \sigma \rangle\rangle_q$ (the subscript indicates the Hilbert-space dimension). For a q -dimensional Hilbert space, vectorized permutations can be formally expanded in the computational basis as

$$|\sigma\rangle\rangle = \sum_{x_1, \dots, x_k=0}^{q-1} |x_1, x_{\sigma^{-1}(1)}\rangle \dots |x_k, x_{\sigma^{-1}(k)}\rangle. \quad (10)$$

We can represent these states graphically using standard tensor-network notation. For example, with $k = 4$ replicas, Eq. (10) for the permutation $\sigma = (1)(2\ 3\ 4)$ reads

$$|\sigma\rangle\rangle = \begin{array}{c} x_1 \quad x_1 \quad x_2 \quad x_4 \quad x_3 \quad x_2 \quad x_4 \quad x_3 \\ \text{---} \cup \text{---} \quad \text{---} \cup \text{---} \quad \text{---} \cup \text{---} \quad \text{---} \cup \text{---} \end{array} \quad (11)$$

where we explicitly label the indices represented by the lines. In the rest of the paper—particularly in Sec. V—we will make use of the large- q expansion of the overlap and Weingarten matrices to derive explicit universal forms for the IPRs and XEB. This expansion reads

$$\begin{aligned} G(q) &= q^k (\mathbb{1} + q^{-1}A + O(q^{-2})), \\ \text{Wg}(q) &= q^{-k} (\mathbb{1} - q^{-1}A + O(q^{-2})), \end{aligned} \quad (12)$$

where A is a matrix whose coefficients $A_{\pi, \sigma}$ are nonvanishing and equal to one only if τ, σ that differ by a single transposition.

E. Porter-Thomas distribution

In the absence of noise and for global random unitaries $U \sim \text{Haar}(D)$ —or, equivalently, for circuits of sufficiently large depth—the PoP follows the well-known *Porter-Thomas* (PT) distribution

$$P_{\text{PT}}(w) = e^{-w}, \quad (13)$$

and correspondingly the IPRs are $I_k^{\text{PT}} = k! D^{1-k}$ (both equalities hold for $D \gg 1$). The absence of noise results in a perfect cross-entropy score $\text{XEB}_{\text{PT}} = DI_2 - 1 \simeq 2 - 1 = 1$.

To illustrate how noise affects these results, we consider the simple scenario of a global depolarizing channel [Eq. (4)] $\rho_{\text{noise}} = \mathcal{N}_{\gamma_{\text{glob}}}^{\text{dep}}(U|\mathbf{0}\rangle\langle\mathbf{0}|U^\dagger)$, with $U \sim \text{Haar}(D)$ as before. In this case, the bit-string distribution is $p(\mathbf{x}; \rho_{\mathcal{N}}(U)) = (1 - \gamma_{\text{glob}})p(\mathbf{x}; \rho(U)) + (\gamma_{\text{glob}}/D)$, and therefore the IPR is $I_k(\gamma_{\text{glob}}) = D \sum_{j=0}^k \binom{k}{j} (1 -$

$\gamma_{\text{glob}})^j (\gamma_{\text{glob}}/D)^{k-j} \mathbb{E}_U[p(\mathbf{0}; \rho(U))^j]$. Using the moments of the PT distribution, we have $\mathbb{E}_U[p(\mathbf{0}; \rho(U))^j] = j! D^{-j}$. Moreover, it is useful to expand the factorial as $j! = \sum_{i=0}^j i! \cdot \binom{j}{i}$, where $i!$ denotes the number of derangements of a set of size i , i.e., the number of permutations with no fixed points [84]. Exchanging the sums over j and i yields

$$I_k(\gamma_{\text{glob}}) = D^{1-k} \text{Tr}[e^{Q \ln(1-\gamma_{\text{glob}})}], \quad (14)$$

with

$$Q_{\sigma'\sigma} = \delta_{\sigma'\sigma} (k - n_F(\sigma)) \quad (15)$$

a diagonal $k! \times k!$ matrix that depends on the number of fixed points $n_F(\sigma)$ of the permutation, and plays a central role in our subsequent analysis. As an explicit example, for $k = 4$ the permutation $\sigma = (1)(2\ 3\ 4)$ shown in Eq. (11) has a single fixed point (the element 1), implying $n_F(\sigma) = 1$ and $Q_{\sigma\sigma} = 3$. More generally, the weight $e^{Q \ln(1-\gamma_{\text{glob}})}$ favors permutations with more fixed points, and in the strong-noise limit it suppresses all but the identity e (for which $Q_{ee} = 0$). For any value of γ_{glob} , the PoP associated with the moments in Eq. (14) is a *shifted Porter-Thomas* (SPT) distribution, given by

$$P_{\gamma_{\text{glob}}}^{\text{SPT}}(w) = (1 - \gamma_{\text{glob}})^{-1} e^{-\frac{w - \gamma_{\text{glob}}}{1 - \gamma_{\text{glob}}}} \theta(w - \gamma_{\text{glob}}), \quad (16)$$

where $\theta(w)$ is the Heaviside function. The noise strength γ_{glob} is directly related to the XEB score, as $\text{XEB}_{\text{SPT}} = 1 - \gamma_{\text{glob}}$. In the strong-noise limit $\gamma_{\text{glob}} \rightarrow 1$, the SPT distribution converges to $P_C(w) = \delta(w - 1)$, with corresponding IPRs $I_k^C = D^{1-k}$, reflecting the dominance of just the identity permutation e over all others. This delta distribution $P_C(w)$ can thus be interpreted as the purely classical (i.e., nonquantum) form of the PoP.

IV. NOISY WEINGARTEN CALCULUS AND WEAK-NOISE EXPANSION

To evaluate the IPRs and XEB, one must average over ensembles of noisy random quantum circuits. This requires a noisy version of the standard Weingarten formula, Eq. (9). Here, we explicitly derive it for the first time for a specific noise model (depolarizing noise). We further argue that the *weak-noise expansion* is universal, i.e., it holds *independently of the particular noise model*.

A. Depolarizing noise

As a first example, we consider a single unitary gate U subjected to depolarizing noise [Eq. (4)]. In vectorized

form, the combined effect of the gate and noise is

$$\begin{aligned} \mathcal{N}_\gamma^{\text{dep.}}((U \otimes U^*)|\rho\rangle\rangle) \\ = \left((1 - \gamma)U \otimes U^* + \frac{\gamma}{q}|e\rangle\rangle\langle\langle e| \right) |\rho\rangle\rangle, \end{aligned} \quad (17)$$

where $|e\rangle\rangle$ is the (folded) identity permutation. To carry out replica calculations, it is necessary to average over k replicas of the noisy gate, i.e., to compute $\mathbb{E}_{U \sim \text{Haar}(q)}[(1 - \gamma)U \otimes U^* + (\gamma/q)|e\rangle\rangle\langle\langle e|]^{\otimes k}$. Since the depolarizing channel commutes with unitary conjugation, the average therefore mirrors that of the noiseless case

$$\begin{aligned} \mathbb{E}_{U \sim \text{Haar}(q)} \left[\left((1 - \gamma)U \otimes U^* + \frac{\gamma}{q}|e\rangle\rangle\langle\langle e| \right)^{\otimes k} \right] \\ = \sum_{\pi, \sigma \in S_k} \tilde{W}_{\pi, \sigma}^{\text{dep.}}(q, \gamma) |\pi\rangle\rangle\langle\langle \sigma|, \end{aligned} \quad (18)$$

where $\tilde{W}_{\pi, \sigma}^{\text{dep.}}(q, \gamma)$ denotes *noisy Weingarten coefficients* that remain to be determined. To find these coefficients, we fix σ and π , and observe that each common fixed point of the two permutations can arise in the expansion of the left-hand side of Eq. (18) from a term $|e\rangle\rangle\langle\langle e|$ inserted at a specific replica, while the remaining part of the permutations must come from the Haar average of $U \otimes U^*$. Taking these observations into account, we arrive at

$$\begin{aligned} \tilde{W}_{\pi, \sigma}^{\text{dep.}}(q, \gamma) \\ = \sum_{i=0}^{n_F(\pi, \sigma)} \binom{n_F(\pi, \sigma)}{i} \frac{\gamma^i}{q^i} (1 - \gamma)^{k-i} W_{\tilde{\pi}(i), \tilde{\sigma}(i)}(q), \end{aligned} \quad (19)$$

where $n_F(\pi, \sigma)$ denotes the number of common fixed points of $\sigma\pi^{-1}$, and $\tilde{\pi}(i)$ and $\tilde{\sigma}(i)$ denote the permutations obtained from π and σ , respectively, by removing i of these common fixed points (which ones is immaterial, because Wg depends only on the cycle structure).

We now consider the weak-noise expansion of the noisy Weingarten coefficients. At first order in γ , Eq. (19) reads

$$\begin{aligned} \tilde{W}_{\pi, \sigma}^{\text{dep.}}(q, \gamma) \\ = (1 - k\gamma) W_{\pi, \sigma}(q) + n_F(\pi, \sigma) \frac{\gamma}{q} W_{\tilde{\pi}(1), \tilde{\sigma}(1)}(q) \\ + O(\gamma^2). \end{aligned} \quad (20)$$

This provides the leading nontrivial correction to the standard Weingarten matrix induced by noise. Beyond the weak-noise assumption, in the remainder of the paper—particularly in Sec. V—we will also consider the regime in which the relevant Hilbert-space dimension q is large; in our effective RMPO model, this corresponds to a

large-bond-dimension limit. Specifically, we take q^{-1} to be of the same order as γ . Using Eq. (12) to expand Eq. (20) for large q yields

$$\tilde{W}g^{\text{dep.}}(q, \gamma) = q^{-k} \left(1 - \frac{A}{q} - \gamma Q + O(q^{-2}, \gamma^2, q^{-1}\gamma) \right), \quad (21)$$

with the matrix Q defined as in Eq. (15). The expansion in Eq. (21) represents the main result of this section.

B. Generic noise

To show explicitly that the weak-noise expansion of the noisy Weingarten coefficients in Eq. (21) is model independent, we repeat the derivation using a *generic noise channel* $\mathcal{N}_\gamma$. As in the previous case, we consider the averaged action of a gate U followed by noise. However, as we are interested in the regime of rare noise events, we assume that the noise is subsequently scrambled by a second gate V , with both U and V Haar random. The averaged combined action of the unitary gates and noise on k copies of the system is described by $\mathbb{E}_{U,V \sim \text{Haar}(q)} \left[((V \otimes V^*) \cdot \mathcal{N}_\gamma \cdot (U \otimes U^*))^{\otimes k} \right]$, which we aim to evaluate. Note that in the noiseless case ($\gamma = 0$), Haar invariance implies that the product UV is still Haar random, so the expression reduces to Eq. (9). For generic values of γ , we can apply Eq. (9) to average over both U and V , obtaining

$$\begin{aligned} & \mathbb{E}_{V \sim \text{Haar}(q)} [(V \otimes V^*)^{\otimes k}] \mathcal{N}_\gamma^{\otimes k} \mathbb{E}_{U \sim \text{Haar}(q)} [(U \otimes U^*)^{\otimes k}] \\ &= \sum_{\pi, \sigma \in S_k} \tilde{W}g_{\pi, \sigma}(q, \gamma) |\pi\rangle\langle\sigma|, \end{aligned} \quad (22)$$

where $\tilde{W}g = Wg\tilde{G}Wg$ are modified Weingarten coefficients and

$$\tilde{G}_{\sigma\pi}(q, \gamma) = \langle\langle \sigma | \mathcal{N}_\gamma^{\otimes k} | \pi \rangle\rangle_q \quad (23)$$

is a modified version of the standard overlap matrix $G(q)$. In particular, $\tilde{G}$ reduces to G in the noiseless limit $\gamma = 0$: $G(q) = \tilde{G}(q, 0)$.

Technically, note that the action of the noise operator $\mathcal{N}_\gamma^{\otimes k}$ on a permutation state $|\sigma\rangle\rangle$ can generate components outside the vector space spanned by permutations. This does not occur for depolarizing noise (Sec. IV A), but it can arise for more general noise models. However, in the present construction, averaging over an additional Haar-random unitary—i.e., applying $\mathbb{E}_{V \sim \text{Haar}(q)} [(V \otimes V^*)^{\otimes k}]$ —projects the state back onto the permutation subspace, thereby simplifying the analysis. We will adopt this construction to define suitable random matrix product operators in Sec. V.

We now expand the modified overlap matrix $\tilde{G}$ as follows:

$$\begin{aligned} \tilde{G}_{\sigma\pi}(q, \gamma) &= G_{\sigma\pi}(q) + \gamma \sum_{a=1}^k \langle\langle \sigma | \mathbb{1}^{\otimes(a-1)} \otimes \tilde{\mathcal{N}}_a \otimes \mathbb{1}^{\otimes(k-a)} | \pi \rangle\rangle_q \\ &\quad + O(\gamma^2), \end{aligned} \quad (24)$$

where $\tilde{\mathcal{N}}$ [defined in Eq. (6)] represents the first-order action of noise in γ , $\mathbb{1}$ is the identity superoperator, and the index $a \in \{1, \dots, k\}$ labels the replicas.

As before, we consider a regime in which both q^{-1} and γ are small (and of the same order). In this regime, the term

$$\langle\langle \sigma | \mathbb{1}^{\otimes(a-1)} \otimes \tilde{\mathcal{N}}_a \otimes \mathbb{1}^{\otimes(k-a)} | \pi \rangle\rangle_q, \quad (25)$$

appearing in the second line of Eq. (24), contributes only to diagonal matrix elements, i.e., only when $\sigma = \pi$. Indeed, off-diagonal terms ($\sigma \neq \pi$) are suppressed by at least a factor q^{-1} relative to the diagonal ones, since the superoperator $\tilde{\mathcal{N}}_a$ in Eq. (25), as it acts only on one replica, cannot restore the reduced overlap between σ and π (recall that for any $\sigma \neq \pi$, $G_{\sigma\pi} q^{-k} = O(q^{-1})$). Therefore, for $\sigma \neq \pi$ the second term in Eq. (24) is subleading, specifically $O(\gamma q^{-1})$. As a result, we only need to evaluate the diagonal matrix elements $\langle\langle \sigma | \mathbb{1}^{\otimes(a-1)} \otimes \tilde{\mathcal{N}}_a \otimes \mathbb{1}^{\otimes(k-a)} | \sigma \rangle\rangle_q$. To visualize these elements, we set $k = 4$, consider the specific permutation $\sigma = (1)(2\ 3\ 4)$, and use a graphical notation. If the noise channel $\tilde{\mathcal{N}}$ is placed on the first replica, i.e., the $a = 1$ term in the sum, then we have

$$\langle\langle \sigma | \tilde{\mathcal{N}} \otimes \mathbb{1}^{\otimes 3} | \sigma \rangle\rangle_q = \tilde{\mathcal{N}} \begin{array}{c} \text{diagram of 4 loops with the first loop shaded} \end{array}. \quad (26)$$

where the permutation σ has been folded as in Eq. (11). We observe that since $a = 1$ corresponds to a fixed point of σ , i.e., $\sigma(a) = a$, Eq. (26) contains the factor

$$\tilde{\mathcal{N}} \begin{array}{c} \text{diagram of 1 loop shaded} \end{array} = \langle\langle e | \tilde{\mathcal{N}} | e \rangle\rangle_q = \text{Tr}[\tilde{\mathcal{N}}(\mathbb{1})] = 0. \quad (27)$$

Therefore, all fixed points of σ give a vanishing contribution to the sum in Eq. (24). Conversely, when the noise channel $\tilde{\mathcal{N}}$ acts on a replica that is not a fixed point of

σ , i.e., $\sigma(a) \neq a$, nonvanishing contributions arise. For example, setting $a = 2$ we obtain the matrix element

$$\langle\langle \sigma | \mathbb{1} \otimes \tilde{\mathcal{N}} \otimes \mathbb{1}^{\otimes 2} | \sigma \rangle\rangle_q = \langle\langle \sigma | \text{diagram} | \sigma \rangle\rangle. \quad (28)$$

This term is proportional to

$$\begin{aligned} \langle\langle \sigma | \text{diagram} | \sigma \rangle\rangle &= \sum_{\alpha \geq 1} \text{Tr}[\tilde{K}_\alpha] \text{Tr}[\tilde{K}_\alpha^\dagger] - q \text{Tr}[\tilde{K}_0] - q \text{Tr}[\tilde{K}_0^\dagger] = \\ &= -q \sum_{\alpha \geq 1} \text{Tr}[\tilde{K}_\alpha^\dagger \tilde{K}_\alpha], \end{aligned} \quad (29)$$

where we used that the $\tilde{K}_\alpha$ are traceless, together with the trace-preserving condition for $\tilde{K}_0$. The proportionality constant can be determined by noting that the noise channel acts as the identity on the remaining $k - 2$ lines, yielding a factor q^{k-2} . For instance, Eq. (28) (with $k = 4$) includes the term in Eq. (29) along with two additional loops, each corresponding to a free index summed over q values, giving a factor $q^2 = q^{k-2}$. Combining these observations and defining

$$\kappa = q^{-1} \sum_{\alpha \geq 1} \text{Tr}[\tilde{K}_\alpha^\dagger \tilde{K}_\alpha], \quad (30)$$

we obtain

$$\begin{aligned} \sum_{a=1}^k \langle\langle \sigma | \mathbb{1}^{\otimes(a-1)} \otimes \tilde{\mathcal{N}}_a \otimes \mathbb{1}^{\otimes(k-a)} | \sigma \rangle\rangle_q &= -q^k \kappa \sum_{a=1}^k (1 - \delta_{a,\sigma(a)}) \\ &= -q^k \kappa (k - n_F(\sigma)). \end{aligned} \quad (31)$$

Notice that κ is necessarily positive, since $\tilde{K}_\alpha^\dagger \tilde{K}_\alpha$ is positive. Moreover, using the freedom to rescale the $\tilde{K}_\alpha$ (see Sec. III B), one can always set $\kappa = 1$, which we will do in the following. Including the expansion of the first term in Eq. (24), namely $G_{\sigma\pi}(q) = q^k (\delta_{\sigma\pi} + q^{-1} A_{\sigma\pi} + O(q^{-2}))$ [see Eq. (12)], which provides corrections to the off-diagonal terms, we obtain

$$\tilde{G}(q, \gamma) = q^k \left(1 + \frac{A}{q} - \gamma Q + O(q^{-2}, \gamma^2, q^{-1} \gamma) \right). \quad (32)$$

Finally, using Eqs. (32) and (12), we find

$$\begin{aligned} \tilde{W}g(q, \gamma) &= Wg(q) \tilde{G}(q, \gamma) Wg(q) \\ &= q^{-k} \left(\mathbb{1} - \frac{A}{q} - \gamma Q + O(q^{-2}, \gamma^2, q^{-1} \gamma) \right). \end{aligned} \quad (33)$$

This form is identical to the one derived for depolarizing noise [Eq. (21)]. Working with generic Kraus channels thus confirms the universal validity of the weak-noise expansion for $\gamma \ll 1$ and large q (with q^{-1} and γ taken to be of the same order).

C. Universality of the weak-noise expansion

While the derivation of the modified Weingarten coefficients assumed a single noise parameter γ , it extends straightforwardly to more general noise channels depending on an arbitrary number of parameters in the weak-noise limit, i.e., when all parameters are small enough and the channel remains close to the identity superoperator. As a concrete example, consider a noise channel characterized by two microscopic parameters γ_1 and γ_2 , obtained as the successive action of two noise maps: $\mathcal{N}_{\gamma_2, \gamma_1}(\rho) = \mathcal{N}_{\gamma_2}^{(2)}(\mathcal{N}_{\gamma_1}^{(1)}(\rho))$. Assuming that γ_1 and γ_2 are both small and of the same order, we can expand the channel to linear order as

$$\begin{aligned} \mathcal{N}_{\gamma_2, \gamma_1}(\rho) &= \rho + \gamma_1 \tilde{\mathcal{N}}^{(1)}(\rho) + \gamma_2 \tilde{\mathcal{N}}^{(2)}(\rho) \\ &\quad + O(\gamma_1^2, \gamma_1 \gamma_2, \gamma_2^2). \end{aligned}$$

Repeating the same steps as before, one arrives at the modified overlap matrix $\tilde{G}_{\sigma\pi}(q, \gamma) = \langle\langle \sigma | \mathcal{N}_{\gamma_1, \gamma_2}^{\otimes k} | \pi \rangle\rangle_q$, whose first-order correction is now

$$\sum_{a=1}^k \langle\langle \sigma | \mathbb{1}^{\otimes(a-1)} \otimes (\gamma_1 \tilde{\mathcal{N}}_a^{(1)} + \gamma_2 \tilde{\mathcal{N}}_a^{(2)}) \otimes \mathbb{1}^{\otimes(k-a)} | \pi \rangle\rangle_q. \quad (34)$$

Since the two noise contributions combine linearly at this order, the previous derivation applies independently to the terms involving $\tilde{\mathcal{N}}_a^{(1)}$ and $\tilde{\mathcal{N}}_a^{(2)}$. This again leads to Eqs. (32) and (33), with an effective noise parameter given by $\gamma = \gamma_1 + \gamma_2$. We therefore conclude that our weak-noise treatment remains valid whenever each source of noise is associated with a small strength γ_i , a physically realistic scenario in several experimental settings.

V. RANDOM MATRIX PRODUCT OPERATOR MODEL AND UNIVERSALITY

We now consider the moments of the PoP distribution, i.e., the IPRs, at finite depth, $t \sim \log N$, in generic chaotic circuits. We begin by summarizing a general argument showing that, in the absence of noise, the calculation

of these moments can be reduced to that of an equivalent one-dimensional problem. To this purpose, let us consider a random brickwall circuit composed of independently Haar-distributed local gates. The computation of the k th IPR involves k replicas of the system, and performing the Haar average over each local gate [via repeated use of Eq. (9)] replaces the original sum over microscopic degrees of freedom (qudits) with a sum over effective degrees of freedom taking values in permutations [34,69,85]. In this way one obtains a classical, Ising-like statistical mechanics model defined on a rectangular space-time domain (of linear sizes N and t), whose local variables are permutations and whose Boltzmann weights are given by Weingarten functions Wg and overlap G matrices. The resulting interaction is generically *ferromagnetic* [see, e.g., Eq. (12)], and therefore favors the formation of spatiotemporal domains characterized by a uniform permutation (see, e.g., Refs. [69,71,85–88]).

For shallow circuits, where $t \ll N$, one can envision a coarse-graining procedure in which all microscopic degrees of freedom below a scale $\sim t$ are decimated. In the spirit of renormalization group theory, this procedure introduces new interactions among the remaining degrees of freedom. Following conventional arguments, we know that such renormalization flows push the system *deeper into the ferromagnetic phase*. Consequently, in the coarse-grained picture, the model reduces to a one-dimensional strip consisting of distinct ferromagnetic domains, each characterized by a single permutation state [see Fig. 1(b)]. The only remaining parameters are the *interface cost* between adjacent domains carrying different permutations σ, σ' , which we denote $C_{\sigma\sigma'}(t)$. That such a description is possible ultimately follows from the *membrane picture* [34,87–89]. As we will see explicitly in the next section, universality arises because, under very general assumptions [89], the minimal-cost interfaces are the *elementary* ones, namely those between permutations σ and σ' that differ by a single transposition (so that $A_{\sigma\sigma'} = 1$). Defining $L(t) \sim C_{\sigma\sigma'}^{-1}(t)$ for such pairs (the precise choice being immaterial) and taking the scaling limit $N, t \rightarrow \infty$ at fixed $x = N/L(t)$, all other interfaces become negligible. One is then left with a dilute gas of elementary interfaces with fugacity $L(t)^{-1}$. This, therefore, also clarifies why the half-chain purity, corresponding to the cost of the interface between the swap $s = (12)$ and the identity e permutations, provides a direct measure of $L(t)^{-1}$ [44,45,60].

Building on this perspective, in the next section we will evaluate these moments in a simpler class of quantum circuits, namely *staircase circuits* [see, e.g., Eq. (37)]. Despite their simplicity, staircase circuits retain the same coarse-grained features and therefore allow for a direct calculation of universal quantities, while also enabling a controlled inclusion of noise. Without noise, these circuits are known to correspond to the sequential generation of one-dimensional tensor networks, namely matrix product

states (MPS) [63,67,90–92]. Their bond dimension (i.e., the matrix size) χ is uniform (except near the boundaries) and is set by the support of the unitary gates (for instance, a staircase circuit featuring gates acting on $r + 1$ qubits generates an MPS of bond dimension $\chi = 2^r$). This construction can be straightforwardly extended to mixed states, where it generates matrix product operators (MPOs), as we will see later. By introducing randomness, for instance in the form of Haar distributed unitary gates, one can use this constructions to naturally define notions of random matrix product states (RMPS) and RMPOs [63,67,91]. While we use the staircase architecture for explicit calculations, our conclusions extend beyond it: we will show that the resulting behavior is in fact *universal*. To corroborate this claim, in Appendix B, we analyze a different solvable random circuit model, specifically the RPM, and demonstrate that it exhibits the same universal properties.

A. Universal distribution in noiseless circuits

We now summarize the derivation of the universal form of the moments of the PoP for generic noiseless chaotic circuits. One way to understand this is to represent the time-evolved state $U|\mathbf{0}\rangle$ as an equivalent random matrix product state (RMPS). Following the discussion of the previous section, the bond dimension χ must be chosen in order to reproduce the exponential decay of the purity. This implies a bond dimension χ growing exponentially with the circuit depth t . More precisely, the Thouless length defined in Eq. (1) can be identified with the MPS bond dimension, $L(t) \sim \chi d/(d - 1)$ leading to the scaling limit [44],

$$x = \frac{d - 1}{d} \frac{N}{\chi}. \quad (35)$$

Then, the k -replica average expresses the IPRs as the partition function of a one-dimensional statistical-mechanical model with local degrees of freedom in permutation space [43–45]. Such a partition function can be expressed as a product of transfer matrices of size $k! \times k!$, ultimately related, in the scaling limit, to the matrix A introduced above, leading to [44]

$$I_k \stackrel{\text{s.l.}}{=} D^{1-k} (\mathbf{1} | e^{xA} | \mathbf{1}) = I_k^{\text{PT}} \exp \left(x \frac{k(k-1)}{2} \right). \quad (36)$$

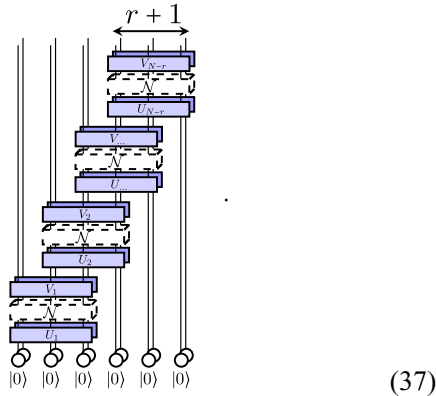
Here, we use a bra $(\dots |$ and ket $|\dots)$ notation to indicate vectors of length $k!$ with components associated to permutations and $|\mathbf{1}\rangle = (1, 1, \dots, 1)$ denotes the vector with all unit components of length $k!$. This is also the eigenvector of the matrix A with maximal eigenvalue, namely $k(k - 1)/2$, which explains the second equality in Eq. (36). This expression confirms in quantitative terms the discussion in the previous section: the exponential factor is the partition

function of a gas of noninteracting domain walls, corresponding to the transpositions. The parameter $x = N/L(t)$ emerges from the balance between fugacity and system size and plays the role of the Poisson intensity controlling the number of domain walls, thereby leading to the characteristic exponential form of the partition function. As a paradigmatic example, the case $k = 2$ has a complete analogy with the classical 1D Ising model, since $S_2 \simeq \mathbb{Z}_2$ as it is evident by mapping the two permutations identity e and swap s to standard Ising variables ± 1 . In this case, domain walls have a Boltzmann cost $e^{-\beta J} = L(t)^{-1} \sim e^{-t/\tau}$, with β the inverse temperature and J the Ising coupling. Therefore, in the scaling limit Eq. (1), where $t \sim \log N \rightarrow \infty$, this corresponds to very low temperature, in which domain walls are *diluted* but their cost is balanced by the entropic factor N .

Finally, one can reconstruct the PoP distribution at any x by recognizing that the moments in Eq. (36) correspond to those of a product of two independent random variables, namely $w = w_{\text{PT}}g$, where w_{PT} is drawn from the Porter-Thomas distribution and g from a log-normal one [such that $\mathbb{E}[g^k] = \exp(xk(k-1)/2)$]. As an explicit confirmation of the emergent universality, these results coincide perfectly with those obtained for the RPM in the limit of large local Hilbert-space dimension d [45].

B. Universal distribution in noisy circuits

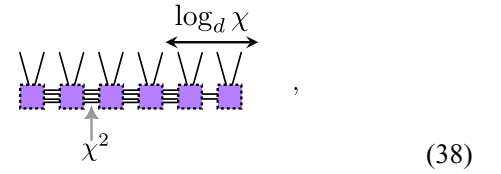
The ideas and results presented above can be successfully extended to circuits in the presence of noise, yielding quantitative predictions for finite-depth circuits. To facilitate the presentation of the argument, and following the previous argument regarding the reduction to effective one-dimensional models, we model a generic noisy circuit as a noisy staircase quantum circuit. Specifically, our construction is as follows:



Each step of the staircase consists of a unitary gate U_i , followed by a noise channel $\mathcal{N}_\gamma$, and then a second unitary gate V_i , similarly to the global gates we discussed in Sec. IV B. Since $\mathcal{N}_\gamma$ acts on a density matrix, both the

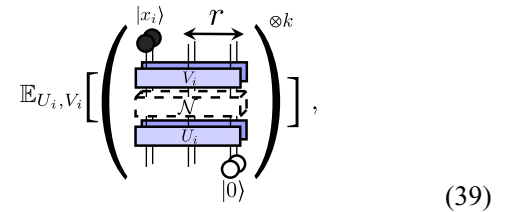
state and its conjugate appear in Eq. (37), with conjugate operators represented by darker colors. For now, we do not need to specify the precise form of $\mathcal{N}_\gamma$ beyond Eq. (3). The gates U_i , V_i , and the noise channel act on $r+1$ qudits, each of local dimension d [for illustrative purposes, $r = 2$ in Eq. (37)]. The full set of gates is denoted by $U := \{U_i, V_i\}_{i=1}^{N-r}$. In the absence of noise, Eq. (37) reduces to the standard sequential generation of an MPS with bond dimension χ , starting from qudits initialized in the state $|0\rangle$ [90]. Specifically, the parameter r is related to the MPS bond dimension χ through, as specified also in the previous section, $r = \log_d \chi$.

The circuit of Eq. (37) generates a mixed density matrix $\rho_{\mathcal{N}}(U)$, which can be represented as a RMPO, see also Fig. 1. This becomes evident by suitably reshaping Eq. (37) and by fusing the gates U_i and V_i (as well as their conjugates) with the corresponding noise channels. This yields



which is an MPO of bond dimension χ^2 , shown in its vectorized form (the bond dimension decreases by a factor d^2 after each of the last r sites near the right boundary).

To compute the overlap, Eq. (2), $w_x = D\langle \mathbf{x} | \rho_{\mathcal{N}}(U) | \mathbf{x} \rangle = D\langle \mathbf{x}, \mathbf{x} | \rho_{\mathcal{N}}(U) \rangle$, the open physical legs in Eq. (37) should be contracted with a bit string $|\mathbf{x}\rangle = |x_1, \dots, x_N\rangle$. Afterward, to evaluate $I_k = D^{-k} \sum_{\mathbf{x}} \mathbb{E}_U[w_x^k]$, one should replicate this network k times. At this stage, the fundamental repeated block that needs to be evaluated can be represented as follows:



where the gray dots represent computational basis states $|x_i\rangle$. Next, Eq. (22) must be applied to compute the average $\mathbb{E}_{U_i, V_i \sim \text{Haar}(d\chi)}[\dots]$. This enables us to express the IPR I_k in terms of powers of an appropriate transfer matrix in replica space, i.e., in the $k!$ -dimensional space spanned by the permutation states $|\sigma\rangle$. Specifically, we arrive at (see details in Appendix A)

$$I_k = D(L|T^{N-r-1}|R), \quad (40)$$

where the transfer matrix

$$T := \tilde{W}g(d\chi, \gamma)G(\chi) \quad (41)$$

acts in the space of permutations and therefore has size $k! \times k!$. The boundary vectors in Eq. (40) read $\langle L| = \langle \mathbf{1}|$ and $|R\rangle = \tilde{W}g(d\chi, \gamma)|\mathbf{1}\rangle$. For small replica numbers k , Eq. (40) enables the exact evaluation of I_k for arbitrary noise channels, via direct multiplication of matrices of size $k! \times k!$.

We now focus on the regime in which the noise strength γ decreases as the system size grows, according to a specific *scaling limit*. The scaling limit (s.l.) is defined as follows: first, in accordance with the noiseless case discussed above and Eq. (35), we keep $x = N/\chi \cdot (d-1)/d$ fixed as $N, \chi \rightarrow \infty$; second, we scale the noise strength as

$$\gamma = \frac{\eta}{N}. \quad (42)$$

In this regime, χ^{-1} and γ are both $O(N^{-1})$, and we can therefore use the expansion of the noisy Weingarten coefficients derived in the previous section. Specifically, using Eqs. (33) and (12), we arrive at

$$\begin{aligned} T &= d^{-k} \left(\mathbb{1} - \frac{x}{N} \frac{1}{d-1} A - \frac{\eta}{N} Q \right) \left(\mathbb{1} + \frac{x}{N} \frac{d}{d-1} A \right) \\ &\quad + O(N^{-2}) \\ &= d^{-k} \left(\mathbb{1} + \frac{x}{N} A - \frac{\eta}{N} Q \right) + O(N^{-2}). \end{aligned} \quad (43)$$

By using this equality and neglecting all terms of order $O(N^{-2})$, which are irrelevant in the scaling limit, we finally obtain the following expression for the transfer matrix:

$$T \stackrel{\text{s.l.}}{=} d^{-k} \exp \left(\frac{x}{N} A - \frac{\eta}{N} Q \right). \quad (44)$$

We stress that our weak-noise expansion leading to Eq. (44) is completely *independent of the specific details of the noise, including whether the channel $\mathcal{N}$ is unital or not*. Using this equation and taking the scaling limit, Eq. (40) reduces to

$$I_k \stackrel{\text{s.l.}}{=} I_k(x, \eta) := D^{1-k} \langle \mathbf{1} | e^{xA - \eta Q} | \mathbf{1} \rangle. \quad (45)$$

The form (45) for the IPRs is one of the main results of this paper. Consistently, it reduces to the noiseless case, Eq. (36), when $\eta = 0$. As discussed above, the resulting expression admits a natural interpretation in terms of the partition function of a one-dimensional gas of domain walls in permutation space. In this picture, the inverse Thouless length $L(t)^{-1}$ plays the role of a fugacity, while the parameter $x = N/L(t)$ corresponds to the associated extensive Poisson intensity controlling the total number of

domain walls. The noise parameter η introduces an additional bias, effectively acting as a magnetic field that favors permutations with a larger number of fixed points, i.e., configurations closer to the identity e . For $k = 2$, the analogy with the classical one-dimensional Ising model is once again useful. In this case, the matrix A reduces to the Pauli matrix X , while $Q = \text{diag}(0, 2)$, making explicit its connection to the local magnetization and the interpretation of η as a magnetic field.

The universality of Eq. (45) rests on two facts: first, as in the noiseless case, in the scaling regime, the domain walls are effectively diluted and only the elementary ones (transpositions) associated with the matrix A are relevant; second, while noise generally favors the identity permutation, the generic measure of distance from the identity emerging in the weak-noise expansion is simply the number of fixed points. The latter is a consequence of the fact that any effect of weak, local noise manifests itself as the insertion of a quantum channel close to identity; in the analysis of replicas, it manifests itself as a term immersed in a ferromagnetic domain of a certain permutation σ ; we can therefore use the expression (31) which, in the scaling limit, leads to (45). From these considerations, it also follows that η is a bulk property that does not change the cost x associated with each interface, which therefore remains the same as in the model without noise. Thus, only two parameters x and η encode all microscopic interactions and the specific form of the noise, where x is still expressed in terms of the Thouless length [Eq. (35)], in turn linked to the half-chain purity of the noiseless model. This hypothesis is verified in the models we solve analytically and in the numerical data we present below. Note that in our RMPO model, Eq. (37), approximately r noise channels are applied per site. As a result, the total number of noisy gates scales as $N_g \propto N$, and $\gamma = \eta/N$. In a realistic quantum circuit extended in time, the number of errors is naturally proportional to the circuit depth. In that case, the time variable t plays the role of r , leading to $N_g \sim Nt$. The relevance of our results for this case will be addressed in Sec. VIII.

The PoP $P_{x,\eta}(w)$ is uniquely determined by the moments in Eq. (45). In the noiseless case ($\eta = 0$), Eq. (45) reduces to Eq. (36) and $P_{x,0}(w)$ is a log-normal convolved with PT. Moreover, for infinite circuit depth ($x = 0$), one recovers the moments of the SPT distribution in Eq. (16) (see the previous section). Thus, Eq. (45) interpolates between these two regimes, capturing the combined effects of noise and finite depth through the scaling parameters x and η . For small k , Eq. (45) provides an efficient way to compute the scaling form of the IPRs for arbitrary x and η . In particular, in the same framework and scaling limit, we can derive a closed form for XEB, as it is closely related to I_2 . Technically, the discrepancy between XEB and I_2 lies in the fact that in Eq. (7) only one of the two replicas contains noise,

thus leading to

$$\begin{aligned} \text{XEB} &\stackrel{\text{s.l.}}{=} (1|e^{xA-\frac{\eta}{2}Q}|1) - 1 \\ &= 2e^{-\eta/2} \left[\cosh(\theta(x, \eta)) + \frac{x \sinh(\theta(x, \eta))}{\theta(x, \eta)} \right] - 1, \end{aligned} \quad (46)$$

where $\theta(x, \eta) = \sqrt{x^2 + (\eta/2)^2}$ and η appears divided by a factor of two compared to Eq. (45), since noise affects only one of the two replicas. The second line of Eq. (46) follows directly from the first by diagonalizing the 2×2 symmetric matrix $xA - (\eta/2)Q$ and computing the overlap of the eigenvectors with the boundary vector $|1\rangle = (1, 1)$. For $x = 0$, XEB reduces to the circuit fidelity F , as expected from the white-noise approximation that becomes valid in the weak-noise limit [54]. In fact, we find (see Appendix A for the explicit calculation of the average fidelity)

$$\lim_{x \rightarrow 0^+} \text{XEB} = e^{-\eta} = F. \quad (47)$$

This relation shows that, in general, the parameter η is fixed by the global fidelity of the circuit. The parameter x is fixed by the system size and by the spreading of purity in the noiseless circuit [see Eq. (1) and Fig. 8]. In the case of a one-dimensional Haar-random brickwall circuit with d -dimensional qudits, the scaling of the half-chain purity

is explicitly known [85] and is given by

$$\tau_{\text{bw}}^{-1}(d) = \log \left(\frac{d^2 + 1}{2d} \right), \quad (48)$$

which for qubits gives $\tau_{\text{bw}}^{-1}(2) \simeq 0.223$.

In contrast to the $\eta = 0$ case, an analytic evaluation of the moments in Eq. (45) for arbitrary k —and thus an explicit reconstruction of the distribution—is technically challenging. Nonetheless, the behavior of the IPRs in different regimes of x can be extracted perturbatively, as we show below. Additionally, we provide an efficient numerical procedure to extract the PoP.

VI. THE THREE DYNAMICAL REGIMES

For fixed η , Eqs. (45) and (46) reveal three distinct dynamical regimes as the circuit depth, and hence x , is varied, and which are manifested from the scaling of the XEB, see Fig. 2. These can be naturally grouped into a short-depth case and an intermediate-to-long-depth one. Notice that, for practical reasons, we define the deviation from the asymptotic value of the XEB in log-scale as $\Delta \log(1 + \text{XEB}) := \log(1 + \text{XEB}) - \lim_{x \rightarrow 0^+} \log(1 + \text{XEB})$.

A. Short-depth $x \gg \eta$

A simple inspection of Eq. (46) shows that $\Delta \log(1 + \text{XEB}) \propto x$ in this regime. Instead, Eq. (45) can be analyzed using perturbation theory in the small parameter η/x ,

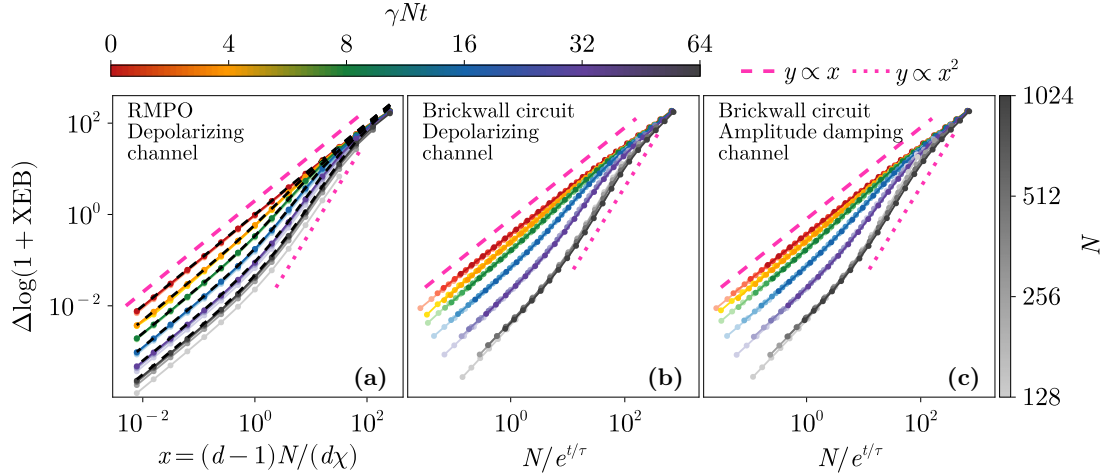


FIG. 2. Plot of the cross-entropy benchmarking for noisy random matrix product operator (a), and random brickwall circuits [(b) and (c)]. Different noise models are considered: random matrix product operator with depolarizing channel (a), see Eq. (37), and random brickwall circuits with two-qubit depolarizing noise (b), or single-qubit amplitude-damping noise (c). The noise strength scales as $\gamma \propto 1/(Nr)$ for random matrix product operator (a), and as $\gamma \propto 1/(Nt)$ for the circuits [(b) and (c)]. The proportionality constant is indicated by the horizontal color bar at the top. Note that in panel (a) the label is γNr rather than γNt . We plot $\Delta \log(1 + \text{XEB}) = \log(1 + \text{XEB}) - \lim_{x \rightarrow 0^+} \log(1 + \text{XEB})$ as a function of the rescaled circuit depth $x \sim Ne^{-t/\tau}$. We identify three different scaling regimes fixed by the fitting parameter $\eta = -\log F$. For $x \gg \eta$, the scaling is with x and is independent of η , showing that the system is effectively noiseless. For $x^2 \sim \eta$, the scaling is with x^2 , characteristic of strong noise. For $x < 1$, it approaches its asymptotic value linearly in x . In panel (a) we also plot the prediction of Eq. (46) (dashed black line).

expanding around the noiseless limit $\eta = 0$ [Eq. (36)]. At first order, we find $I_k \simeq I_k^{\text{PT}} \exp(x(k(k-1)/2) - \eta(k-1) + O(\eta^2/x))$. The fact that the right-hand side of this expression does not yield $\lim_{k \rightarrow 0} I_k = D$, as it trivially should according to the definition in Eq. (8), indicates that the first-order expansion at small η/x cannot be analytically continued to $k < 1$. Nevertheless, it allows us to extract the PoP tails at $w \gg 1$ as a convolution of the PT with a (unnormalized) log-normal.

B. Large/intermediate-depth

To simplify the analysis, we assume that η is large enough that we can focus in practice only on the ground state of Q , which consists of the identity permutation e , perturbed by the matrix A . Without perturbation, and neglecting terms of order $O(e^{-2\eta})$, one simply obtains the classical value of the IPR, I_k^C . The perturbation introduces two relevant types of corrections. First, from second-order perturbation theory, one can obtain the correction: $x^2/(2\eta) \sum_{\sigma} A_{e,\sigma}^2 = x^2 k(k-1)/(4\eta)$. This expression can be interpreted as describing a region in the bulk where the state is locally excited to σ instead of remaining in the ground state e . This excited region is bounded by two domain walls, corresponding to identical transpositions that separate it from the surrounding ground state regions. Pictorially, we can represent this correction as follows:


(49)

Second, the boundary state $|1\rangle$ in Eq. (45) contains contributions from all permutations. As a result, transpositions σ can remain localized near the left or right boundaries before the perturbation given by the A matrix causes them to “jump” to the ground state e , creating a domain wall at cost $O(x)$. This correction can be pictorially represented as


(50)

and similarly on the left boundary. Putting these two contributions together, we arrive at

$$I_k = I_k^C \exp\left(\frac{x^2 k(k-1)}{\eta}\right) \left(1 + \frac{x k(k-1)}{\eta} + O(x^2/\eta^2)\right). \quad (51)$$

The first term of this expression again indicates a log-normal PoP (the convolution with the delta distribution $P_C(w)$ is irrelevant). This expression also determines two distinct regimes by increasing the depth of the circuit (decreasing x). When $1 \ll x = \sqrt{\eta} \ll \eta$, the exponential factor dominates giving rise to $\log \mathbb{E}[w^k] \propto x^2/\eta$ behavior. In contrast, at larger depths for which $x \ll 1$, $\log \mathbb{E}[w^k] \propto x/\eta$. This behavior is explicitly confirmed by the analytical expression of the XEB (46).

The existence of these dynamic regimes is a direct consequence of the universal form (45). Figure 2 presents numerical results for both noisy RMPO and random brick-wall circuits confirming this analysis. For the brickwall circuit, we consider two types of noise: a two-qubit depolarizing channel and a single-qubit amplitude-damping channel $\mathcal{N}_\gamma^{\text{amp}}(\rho) = K_0 \rho K_0^\dagger + K_1 \rho K_1^\dagger$ with

$$K_0 = \begin{pmatrix} 1 & 0 \\ 0 & \sqrt{1-\gamma} \end{pmatrix}$$

and

$$K_1 = \begin{pmatrix} 0 & \sqrt{\gamma} \\ 0 & 0 \end{pmatrix}.$$

Notably, the latter is nonunitary, confirming that the emergence of these regimes does not depend on the unitality of the noise. We compute XEB by contracting the tensor network obtained by replicating the system and averaging all the unitary gates. The resulting replica tensor network lives in a $k!$ -dimensional space, spanned by the permutations; for $k = 2$, which is the case for XEB, a basis is given by the identity e and the transposition $s = (12)$. While for RMPS (and RMPO) the exact form of the rescaled depth $x = (N/\chi)(d-1)/d$ is derived analytically above, for the circuit we take the definition (1) and fit the parameter L_0 [the value of τ is provided in Eq. (48) but it could also be fitted]. We plot $\Delta \log(1 + \text{XEB})$ as a function of x , or $N/e^{l/\tau}$ (circuit). For the circuit, we observe that plotting as a function of this variable yields good collapse of the curves corresponding to different system sizes N . Consistently, in all three cases, we clearly observe the predicted three regimes $O(x)$, $O(x^2)$ and $O(x)$.

VII. UNIVERSAL FORM OF THE PoP DISTRIBUTION

While the derived form of the PoP $P_{x,\eta}(w)$ is fully determined by its moments in Eq. (45), a fully analytical expression in closed form unfortunately remains inaccessible for generic x and η . Nevertheless, using that for sufficiently large $\eta \gg x$, the distribution approaches a log-normal, namely a Gaussian in the variable $\log w$, allows for a systematic expansion. Specifically, we employ the Edgeworth series [93,94], which approximate a probability distribution in terms of its cumulants. The latter can be computed numerically from Eq. (45) for sufficiently small k (e.g., $k \in \{1, 2, 3, 4, 5, 6\}$), since the sizes of the matrices A and Q grow as $k!$. To benchmark our results, in Fig. 3 we compare the distribution predicted by our formula using the Edgeworth series (black lines), with the empirical distribution of overlaps w between random bit strings and the RMPO, constructed as in Eq. (37) (blue histogram). For this construction, we use the multi-qubit depolarizing channel. We consider a specific value of x and various values of η . We observe an excellent agreement between two

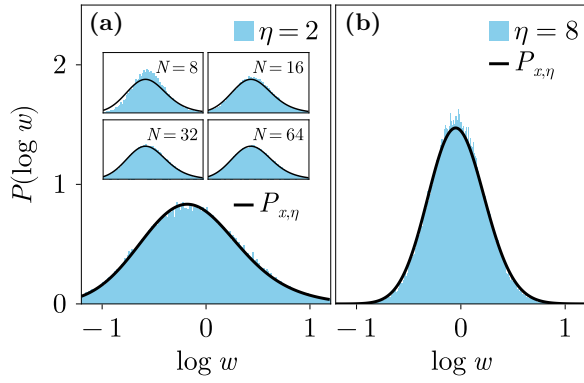


FIG. 3. Probability-of-probabilities of random matrix product operator for construction of Eq. (37) for $N = 128$, $\chi = 128$ ($x = 0.5$), and $2 \leq \eta \leq 8$. For each set of parameters, the histogram is built from 100 independent circuit realizations, yielding a total of 10^5 samples per histogram. The distribution of $\log w$ can be approximated as a Edgeworth series (black line) where only the first two terms of the expansion have been considered and the cumulants of the distribution have been fitted from I_k computed from Eq. (45) for $k \in \{1, \dots, 6\}$. Inset of panel (a) shows the approach to the universal prediction by increasing N .

distributions. Furthermore, in Fig. 3(a), we illustrate how the data approach the Edgeworth expansion as the system size N increases.

Far from being limited to this case, the PoP $P_{x,\eta}(w)$ is *universal*: it applies to generic circuit architectures and noise models and is fully determined by the parameters x and η . In generic cases, these should be regarded as model-dependent fitting parameters [while for RMPO their

explicit form in terms of χ , N , d and γ is known; see Eqs. (35) and (42)]. To show the universality of our form, in Fig. 4 we compare our prediction (45) (black lines) with the empirical PoP distribution (blue histogram) for a *Haar-random brickwall circuit* with two-qubit depolarizing noise on each gate. The empirical overlaps are obtained with tensor network simulations [95–98]. We explore different values of depth t , while we fix N and γNt . Extracting the exact values of x and η is nontrivial, as no analytical expressions are available for the brickwall circuit case. To determine these parameters, we therefore propose an experimentally friendly approach that can be applied to generic circuits at finite depth. The approach proceeds as follows. First, one computes XEB (which can also be measured experimentally) and extracts η by extrapolating its value at large depth [ideally $x = 0$, see Eq. (47)]. Next, x can be determined by equating Eq. (46) to the numerically or experimentally obtained XEB data. These parameters are then used to reconstruct the full distribution $P_{x,\eta}(w)$ with moments given by Eq. (45). The fact that the values of x and η extracted solely from XEB yield good agreement is a nontrivial indication that all moments of the anticoncentration follow the same universal form, depending only on these two parameters. In Appendix C, this is demonstrated by reconstructing the I_3 moment from XEB (see Fig. 7).

In Fig. 4, the empirical distribution and our analytical prediction shows excellent agreement, with small deviations only due the finite size N of the simulation. Dotted red line represents the noiseless distribution $P_{x,0}(w)$ (log-normal convolved with PT), which closely match the empirical distribution for short time t (i.e., $x \gg 1$). At

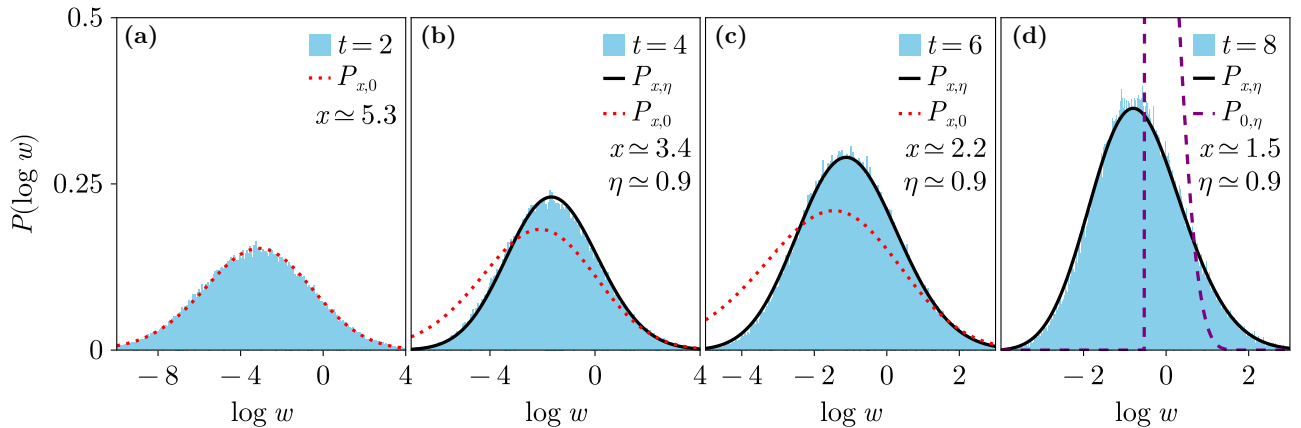


FIG. 4. Plot of the probability-of-probabilities for a Haar-random brickwall circuit subject to two-qubit depolarizing noise with $\gamma Nt = 2$ and system size $N = 32$. Different depths t are explored (see values reported in the plots). For each set of parameters, the histogram is built from 100 independent circuit realizations, yielding a total of 10^5 samples per histogram. At the shortest depth, the distribution is well approximated by the noiseless distribution (dotted red curve) $P_{x,0} \simeq P_{x,\eta}$, whereas the latter fails to accurately describe the data at larger depths. In contrast, the theoretically predicted distribution $P_{x,\eta}$ (black line) accurately reproduces the data. The shifted Porter-Thomas, Eq. (16), corresponding to the distribution at infinite times, $x = 0$, is also plotted for reference in the last plot. The values of x and η are obtained by the fitting procedure involving the cross-entropy benchmarking data explained in the main text (Sec. VII).

late time instead the distribution tends toward a shifted Porter-Thomas, i.e., $P_{0,\eta}(w)$ (purple dashed line), yet, Fig. 4 shows that at intermediate times the two distributions remain clearly distinct, indicating that the shifted Porter-Thomas form should be regarded as a genuinely asymptotic distribution.

VIII. CIRCUITS WITH FIXED NOISE RATE IN TIME AND XEB TRANSITION

We shall now consider a more realistic scenario in which the noise strength γ does not scale inversely to the circuit depth t , but still decreases with the system size N . This corresponds to setting $\gamma = O(N^{-1})$, so that the total number of errors scales linearly with time, i.e., $\eta = O(t)$. Specifically, we define

$$\eta = \lambda t, \quad (52)$$

with fixed λ , related to the global fidelity, as before, via $F = e^{-\lambda t}$. All our results remain applicable in the regime $\lambda t \ll N$, namely whenever the total number of errors remain subleading. Expanding Eq. (46) for the XEB in the small- x limit, retaining only the first-order term in x , substituting $\eta \rightarrow \lambda t$, and using $x = e^{-t/\tau} \cdot N/L_0$, we obtain

$$\text{XEB} = e^{-\lambda t} + 2 \frac{e^{-t/\tau}}{\lambda t} \frac{N}{L_0} + O(x^2). \quad (53)$$

This expression can be easily interpreted viewing XEB as the partition function of a 1D system of size N in permutation space, expressed by [see first line of Eq. (46)]: $\text{XEB} = \langle 1 | (\exp((x/N)A) \exp(-\frac{1}{2}(\lambda t/N)Q))^N | 1 \rangle - 1$. As the XEB is a quantity expressed only in terms of two replicas, the permutation space is limited to identity e and swap s . The term $\exp(-\frac{1}{2}(\lambda t/N)Q)$ weights each site with a factor 1 for identity and $\exp(-\lambda t/N)$ for the swap. Each domain wall between the two permutations carry a factor $x/N = \mathcal{L}^{-1}(t) = \mathcal{L}_0^{-1} e^{-t/\tau}$, therefore at large times (i.e., for small x), domain walls between different permutations are suppressed. In this limit, the leading contributions come from configurations where the permutation is uniform across the system. There are two such trivial contributions. The configuration in which the permutation is e is exactly canceled by the -1 in the definition of XEB. The configuration where the permutation is s , or pictorially

$$\overline{s}, \quad (54)$$

yields $(\exp(-\lambda t/N))^N = e^{-\lambda t}$, i.e., the first term in Eq. (53). The second term in Eq. (53) (of order x) corresponds to the configurations with a single domain wall, pinned either on the left or on the right boundary of the

system,

$$\begin{array}{cc} \text{---} e \text{---} s \text{---} & \text{---} s \text{---} e \text{---} \end{array} \quad (55)$$

In fact, each of these two terms contributes to the XEB as

$$\frac{1}{\mathbb{L}_0} \sum_{\ell=1}^N \left(\exp\left(-\frac{\lambda t}{N}\right) \right)^\ell = \frac{N}{\lambda t \mathbb{L}_0}, \quad (56)$$

where ℓ denotes the length of the region occupied by the s permutation. A direct inspection of Eq. (53) shows that $\text{XEB}(t)$ is the sum of two exponentials, decaying in time with distinct rates. For noise rates $\lambda < \lambda_c = 1/\tau$, the first contribution dominates and coincides with the global fidelity $F(t) = e^{-\lambda t}$. In this regime, the late-time behavior of XEB provides direct access to the fidelity of the *global* circuit. Conversely, for $\lambda > 1/\tau$ the second exponential takes over, yielding a decay governed by the noise-independent rate $1/\tau$. This is precisely the transition reported in Refs. [3,31], where the sharp crossover at $\lambda = \lambda_c$ is interpreted as the breakdown of XEB as a faithful proxy for fidelity. Moreover, it coincides with a transition in the spatial structure of correlations: from an extended, system-spanning regime to a local regime effectively pinned to the boundaries. In the latter phase, the dynamics can be viewed as fragmenting the system into mutually incoherent patches, which can be efficiently reproduced by classical patch-based simulations [3].

We clarify three points. First, we stress that the critical value of λ is determined precisely by the scaling of the half-system purity, $\lambda_c = 1/\tau$. Second, *the XEB at late times does provide information about the fidelity F even in the regime $\lambda > 1/\tau$* . In practice, it suffices to determine λ by fitting λ , τ , and L_0 simultaneously using the second term of Eq. (53), and then to compute the fidelity in the usual way as $F = e^{-\lambda t}$. (If available, τ may alternatively be extracted independently from the scaling of the purity, which can further stabilize the fit). Finally, an analogous transition occurs for all moments k of the distribution Eq. (45). In fact, in complete analogy with the discussion above for the XEB, the small x expansion of I_k yields

$$I_k = D^{1-k} \left[1 + \frac{k(k-1)}{2} \left(e^{-2\lambda t} + \frac{e^{-t/\tau}}{\lambda t} \frac{N}{L_0} \right) + O(x^2, e^{-3\lambda t}) \right], \quad (57)$$

so that the approach to the classical distribution $P_C(w) = \delta(w - 1)$ is characterized by either log-normal corrections with parameter $e^{-2\lambda t}$ or $e^{-t/\tau}$. This transition occurs at a critical noise value $\lambda_c^{(k)} = 1/(2\tau)$, i.e., at half the noise strength associated with the XEB transition.

As done before in Sec. VII, we again use the XEB at different values of depth t to determine x and λ to reconstruct

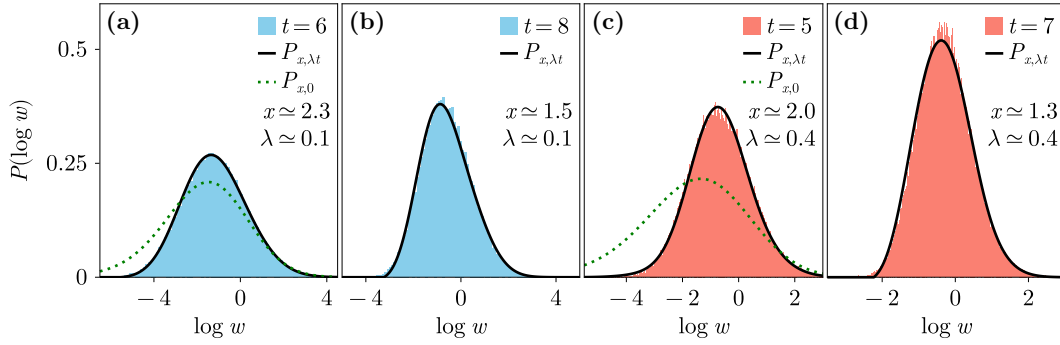


FIG. 5. Plot of the probability-of-probabilities for a Haar-random brickwall circuit with $N = 32$, affected by *both* two-qubit depolarizing noise (strength γ) and single-qubit amplitude-damping noise (strength γ'). For each set of parameters, the histogram is built from ten independent circuit realizations, yielding a total of 10^5 samples per histogram. We show two noise settings: $\gamma = 0.00313$ and $\gamma' = 0.00442$ in panels (a) and (b), which yield a fitted $\lambda \simeq 0.1$, and $\gamma = 0.016$ and $\gamma' = 0.022$ in panels (c) and (d), which yield $\lambda \simeq 0.4$. These correspond to two distinct cross-entropy benchmarking regimes [see Eq. (53)]: $\lambda < 1/\tau$ for (a) and (b) and $\lambda > 1/\tau$ for (c) and (d). For each setting we consider two circuit depths t , and report the associated rescaled inverse depth x in the plots. The parameters x and λ are extracted via the cross-entropy benchmarking-based fitting procedure described in Sec. VIII.

the PoP $P_{x,\lambda t}(w)$, see Fig. 5, and with the XEB data shown in Fig. 6. Specifically, we first fit the numerically (experimentally) obtained values XEB at large times t using Eq. (53), which allows us to extract λ , τ , and L_0 . While one could then directly compute $x(t)$ by assuming the form $x = (N/L_0)e^{-t/\tau}$, using the fitted values of L_0 and τ , this expression may have extra corrections at early times. In practice, we determine λ by fitting the large time behavior of the XEB using Eq. (53). This provides an estimation for $\eta = \lambda t$, which we then use to solve for x from the XEB at time t via Eq. (46). With both x and λ determined in

this way, we reconstruct the analytical distribution $P_{x,\lambda t}(w)$ with moments given by Eq. (45). We stress that the fact that the values of x and η obtained from the XEB are also the ones fixing the distribution $P_{x,\eta}$ is a nontrivial check of the universality and the validity of Eqs. (46) and (45) for all moments k . Moreover, our results extend beyond the specific setting considered here to include general chaotic circuits and various forms of noise. In fact, Appendix C provides an example based on a *deterministic gate set*, thus going beyond the paradigm of random quantum circuits. Specifically, in Fig. 9 we consider the chaotic kicked Ising model with dephasing noise and observe good agreement with our predictions.

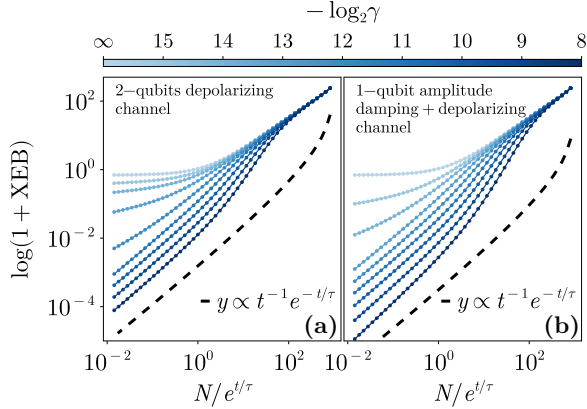


FIG. 6. Plot of the cross-entropy benchmarking for the brickwall circuit with (a) two-qubits depolarizing (b) single-qubit amplitude-damping and depolarizing channels, with different γ , and $N = 1024$. We observe the transition at large times in the scaling of $\log(1 + \text{XEB})$, from $-\lambda t$ to $-t/\tau$ as the noise strength γN is increased. The crossing happens at $\lambda = 1/\tau$ as predicted by Eq. (53) [$\tau = \tau_{\text{bw}}(2)$ in this model]. In the regime $\lambda > 1/\tau$, while the slope of the decay is fixed by the noise-independent τ , a shift proportional to $1/\lambda$ as predicted by Eq. (53) is clearly visible, the latter allowing to determine the global fidelity.

IX. CONCLUSIONS

In this work we have investigated the universal properties of anticoncentration in generic, chaotic 1D quantum many-body systems subject to weak noise. Whereas the influence of noise on infinite-depth circuits, characterized by the Porter-Thomas distribution of overlaps, has been recently studied, see, for example, Ref. [50], here we have focused on *the interplay between finite circuit depth and noise*, a question of central relevance for near-term quantum devices. We derived a universal form of the distribution of bit-string probabilities (PoP distribution), as a function of two dimensionless parameters x and η determined by circuit depth, system size, and the global fidelity. As we have shown, these parameters can be extracted in an experimental setting by evaluating the linear cross-entropy XEB across different circuit depths, for any values of noise strength, at least in the regime where the number of error remains subleading $-\log F \ll N$. Given that our interest is in shallow circuits, this regime is achieved even for very strong noise strengths at small depths, and indeed our

universal PoP distribution provides an excellent fit of our numerical data for quite shallow circuits and strong noise.

Our analysis is based on an exact mapping to a one-dimensional model of random matrix product operator that is generated by the action of a generic noisy staircase circuit. Beyond serving as a tractable theoretical toy model, noisy RMPOs are themselves experimentally accessible on current quantum platforms [99–103], whereas random MPOs can also be used to benchmark of quantum experiments [104]. Thus, our results not only furnish clear predictions for state-of-the-art quantum machines that can be directly used for useful benchmarking but also clarify the persistence of genuine quantum effects in the presence of realistic noise.

While the universality of our results encompasses numerous realistic systems, it may also be interesting to explore PoP shapes beyond this universality class, which requires a certain amount of fine-tuning: dual-unitary circuits [47], the presence of conserved charges [105–107], integrability would have implications on our results (even in the absence of noise) [108]. We leave these settings to future works.

ACKNOWLEDGMENTS

We acknowledge discussions with Xhek Turkeshi and for collaborations on related topics. J.D.N. and G.L. are funded by the ERC Starting Grant No. 101042293 (HEPIQ) and the ANR-22-CPJ1-0021-01. A.D.L. acknowledges support by the ANR JCJC Grant No. ANR-21-CE47-0003 (TamEnt). This work was granted access to the HPC resources of IDRIS under the allocation AD010914149R2 and AD010613967R2.

DATA AVAILABILITY

The data that support the findings of this article are openly available [109].

APPENDIX A: DERIVATION OF THE TRANSFER MATRIX FOR RMPO

Here, we provide a full derivation of the replica tensor network for the noisy RMPO model giving Eq. (40). In Eq. (37), unitaries U_i and V_i , of size $d\chi \times d\chi$, act on $r + 1$ qudits, with $r = \log_d \chi$. To compute the overlap $w_x = D\langle x, x | \rho_N(U) \rangle$ for a fixed bit-string x , one must contract the open physical legs in Eq. (37) with the state $\langle x, x |$; that is,

$$w_x = D\langle x, x | \rho_N(U) \rangle = D\langle x, x | \prod_{i=1}^{N-r} W_i | \mathbf{0}, \mathbf{0} \rangle, \quad (\text{A1})$$

where we defined

$$W_i = ((V_i \otimes V_i^*) \cdot \mathcal{N}_i(U_i \otimes U_i^*)) \quad (\text{A2})$$

and where we omit dependence on the noise strength γ for readability. The IPR $I_k = D^{-k} \sum_x \mathbb{E}[w_x^k]$ is then obtained by replicating the network k times, summing over all x and performing the average over the U_i and V_i . This results in

$$I_k = \sum_x \left(\langle x, x | \prod_{i=1}^{N-r} \mathbb{E}[W_i] | \mathbf{0}, \mathbf{0} \rangle \right)^{\otimes k}. \quad (\text{A3})$$

Now we employ the modified Weingarten formula [Eq. (22)]

$$\begin{aligned} \mathbb{E}[W_i] &= \mathbb{E}_{U, V \sim \text{Haar}(d\chi)} \left[((V_i \otimes V_i^*) \cdot \mathcal{N}_i(U_i \otimes U_i^*))^{\otimes k} \right] \\ &= \sum_{\pi, \sigma \in S_k} \tilde{W}_{g_{\pi, \sigma}}(d\chi, \gamma) |\sigma\rangle_d \langle \sigma|_d \chi \langle \pi|_d \langle \pi|_d, \end{aligned} \quad (\text{A4})$$

where we explicitly used the factorization of permutation states over distinct Hilbert spaces $|\sigma\rangle_d \chi = |\sigma\rangle_d \langle \sigma|_d \chi$. This equation has to be applied for each W_i , i.e., for each pair U_i and V_i of matrices. In the d -dimensional spaces, permutations are either contracted with the initial state $|0, 0\rangle$ or with the final state $|x_i, x_i\rangle$. Since these states are normalized to one, the result is an irrelevant factor: $(\langle 0, 0 | \sigma \rangle_d^{\otimes k} = \langle \pi | (|x_i, x_i\rangle)_d^{\otimes k} = 1$. The sum over x_i simply yields a multiplicative factor of d for each qubit, i.e., an overall contribution $D = d^N$. Because of these free boundary conditions, the left and right boundary vectors can be written simply as $|L\rangle = |\mathbf{1}\rangle$ and $|R\rangle = \tilde{W}g(d\chi, \gamma)|\mathbf{1}\rangle$. The Weingarten matrix in the right boundary vector comes from the average of the last gate of the staircase. In the bulk, permutations resulting from consecutive gates V_i, U_{i+1} are contracted together, giving rise to the overlaps matrix $G_{\sigma, \pi}(\chi) = \langle \sigma | \pi \rangle_\chi$. Putting together all these terms one finally obtain Eq. (40), with the transfer matrix defined as in Eq. (41). Notice that in Eq. (40) there are in total $N - r$ Weingarten matrices (one incorporated in $|R\rangle$ and $N - r - 1$ in the term T^{N-r-1}), as should be as the staircase circuit Eq. (37) features exactly $N - r$ unitaries.

The derivation of the XEB for this model follows similarly, now involving only $k = 2$ replicas: one noisy and one noiseless. The overlaps matrix remains unchanged, while the Weingarten matrix is now defined by

$$\begin{aligned} \mathbb{E}_{U, V} [((V \otimes V^*) \cdot \mathcal{N}_\gamma \cdot (U \otimes U^*)) \otimes ((V \otimes V^*) \cdot (U \otimes U^*))] \\ = \sum_{\pi, \sigma \in S_2} [\tilde{W}_{g_{\text{XEB}}}]_{\pi, \sigma}(q, \gamma) |\sigma\rangle \langle \pi|, \end{aligned} \quad (\text{A5})$$

and can be computed as before using $\tilde{W}_{g_{\text{XEB}}} = Wg\tilde{G}_{\text{XEB}}Wg$, and defining

$$[\tilde{G}_{\text{XEB}}]_{\sigma', \sigma}(q, \gamma) = \langle \sigma' | \mathcal{N}_\gamma \otimes \mathbb{1} | \sigma \rangle_q, \quad (\text{A6})$$

with $\sigma', \sigma \in S_2$. Calculations analogous to those in Sec. IV B then yield

$$\tilde{G}_{\text{XEB}}(q, \gamma) = q^2 \begin{pmatrix} 1 & 1/q \\ 1/q & 1 - \gamma \end{pmatrix} + O(\gamma^2), \quad (\text{A7})$$

where again we assumed $\kappa = 1$. This equation is the analogous of Eq. (32). We can now obtain the expansion at first order in q^{-1} and γ for the noisy Weingarten matrix

$$\tilde{W}_{\text{gXEB}}(q, \gamma) = q^{-2} \left(\begin{pmatrix} 1 & -\frac{1}{q} \\ -\frac{1}{q} & 1 - \gamma \end{pmatrix} + O(\gamma^2, q^{-2}, \gamma q^{-1}) \right). \quad (\text{A8})$$

As for I_2 , we can write the replica contraction $\text{XEB} = D(L|T_{\text{XEB}}^{N-r-1}|R) - 1$ with $T_{\text{XEB}} = \tilde{W}_{\text{gXEB}}(d\chi, \gamma)G(\chi)$, $(L| = (1|$ and $|R) = \tilde{W}_{\text{gXEB}}(d\chi, \gamma)|1)$. Scaling the noise strength as previously done $\gamma = \eta/N$, we get

$$T_{\text{XEB}} = d^{-1} \left(\mathbb{1} + \frac{x}{N}A - \frac{\eta}{2N}Q + O(N^{-2}) \right), \quad (\text{A9})$$

which, in the scaling limit and after exponentiation, leads to Eq. (46).

The calculation of the average state fidelity $F = \mathbb{E}_U[\text{Tr}(\rho_N(U)\rho(U))]$ is very similar to the one of XEB. The trace in the fidelity expression can be evaluated using the “swap trick” by rewriting $\text{Tr}(\rho_N \rho) = \text{Tr}((\rho_N \otimes \rho)s) = \langle \rho_N \otimes \rho | s \rangle$, where s denotes here the swap operator (or equivalently the transposition permutation). This adds boundary terms in the replica contraction, which we encapsulate inside a matrix $[\Lambda_F]_{\sigma'\sigma}(d) = \delta_{\sigma'\sigma} \langle s | \sigma \rangle_d = \text{diag}(d, d^2)$, so that the replica contraction now reads $F = (L|T_F^{N-r-1}|R)$ with $T_F = \tilde{W}_{\text{gXEB}}(d\chi, \gamma)\Lambda_F(d)G(\chi)$, $(L| = (1|$ and $|R) = \tilde{W}_{\text{gXEB}}(d\chi, \gamma)\Lambda_F(d\chi)|1)$. In the scaling limit, we notice that the right boundary vector becomes simply $|R) = (0, 1)$. Finally, after replacing $\gamma = \eta/N$, we get

$$T_F \stackrel{\text{s.l.}}{=} \begin{pmatrix} 1/d & 0 \\ (d+1)x/(dN) & 1 - \eta/N \end{pmatrix}, \quad (\text{A10})$$

which in the scaling limit yields $F \stackrel{\text{s.l.}}{=} e^{-\eta}$.

APPENDIX B: RANDOM PHASE MODEL

To begin, let us briefly revisit the standard definition of the RPM [45,72]. This model consists of a random circuit acting on N qudits and defined the subsequent action of two different layers at time step t : first, on each site i , one acts with single-site Haar-random unitary, $U_i^{(1)}(t)$; then, on each bond between neighboring sites $(j, j+1)$, one acts with a diagonal two-site gate whose entries are phases $[U_{j,j+1}^{(2)}(t')]_{x_j x_{j+1}, x_j x_{j+1}} =$

$\exp[i\phi_{x_j x_{j+1}}^{(j)}(t')]$ ($x_j, x_{j+1} \in \{0, 1, \dots, d-1\}$). Each phase $\phi_{x_j x_{j+1}}^{(j)}(t')$ is an independent Gaussian random real variable with mean zero and variance ϵ , which controls the coupling strength between neighboring spins. A given realization of the circuit is obtained by drawing independently for each j and t' the single-site unitaries and the phases $\phi_{x_j x_{j+1}}^{(j)}(t')$. Consider the k th moment of the output distribution $w^k = D(\mathbf{x}|W_{\text{RPM}}(t)|\mathbf{x})$, where $W_{\text{RPM}}(t)$ denotes the global RPM circuit with t repeated action of single-site and two-site layers. As explained in Ref. [45], the circuit average $\mathbb{E}[w^k]$ in the absence of noise can be computed at large d in two steps: one first averages over the single-site unitaries using once again the Weingarten formula

$$\begin{aligned} \mathbb{E}_{\text{RPM}}[(U_i^{(1)}(t') \otimes U_i^{(1)}(t')^*)^{\otimes k}] \\ = \sum_{\pi, \sigma \in S_k} W_{\pi, \sigma}(d) |\sigma\rangle_d \langle \pi|_d = d^{-k} \sum_{\sigma \in S_k} |\sigma\rangle_d \langle \sigma|_d. \end{aligned} \quad (\text{B1})$$

Note that this equation is formally analogous to Eq. (9), but it is restricted to the action on the d -dimensional local physical Hilbert space. Beyond allowing the expansion of Weingarten functions, the large d limit also implies that the same permutation $\sigma_i \in S_k$ at a given site i propagates at all times, as the overlaps $G_{\sigma', \sigma}(d) = \langle \sigma' | \sigma \rangle_d \sim d^k \delta_{\sigma', \sigma}$. Thus, the effect of the 2-site gates can be accounted for computing

$$\langle \sigma_j, \sigma_{j+1} | U_{j,j+1}^{(2)}(t') | \sigma_j \sigma_{j+1} \rangle = d^{2k} e^{-\epsilon(k - n_F(\sigma_j, \sigma_{j+1}))}, \quad (\text{B2})$$

where $n_F(\sigma_j, \sigma_{j+1})$ denotes the number of common fixed points between σ_j and σ_{j+1} . Collecting the contribution from the t layers of phases, we have that

$$\mathbb{E}[w^k] = \sum_{\sigma, \sigma'} [T_{\text{RPM}}^N]_{\sigma, \sigma'}, \quad [T_{\text{RPM}}]_{\sigma', \sigma} = e^{-\epsilon t(k - n_F(\sigma, \sigma'))}. \quad (\text{B3})$$

In order to include the effect of noise, we assume for simplicity that it affects each single-site Haar unitary, with the presence of a depolarizing channel $\mathcal{N}_{\gamma}^{\text{dep.}}$. Consequently, as in Eq. (18), we replace $Wg(d)$ with $\tilde{W}g(d, \gamma)$ in (B1). For large d , Eq. (21) yields

$$\tilde{W}g(d, \gamma) = d^{-k} (1 - \gamma Q + O(\gamma^2) + O(d^{-1})). \quad (\text{B4})$$

We thus arrive at

$$\mathbb{E}[w^k] = \sum_{\sigma, \sigma'} [\tilde{T}_{\text{RPM}}^N]_{\sigma, \sigma'}, \quad \tilde{T}_{\text{RPM}} = T_{\text{RPM}} e^{-Q\gamma t + O(\gamma^2)}. \quad (\text{B5})$$

Now, we recall that for the RPM, the Thouless length can be identified with $L(t) = e^{2\epsilon t}$ and at large time, one can

rewrite

$$T_{\text{RPM}} = 1 + \frac{A}{L(t)} + O(L(t)^{-2}) \sim e^{A/L(t)}. \quad (\text{B6})$$

Finally, in the scaling limit (1) and $\gamma = \eta/(Nt)$, we can combine the two factors in $\tilde{T}_{\text{RPM}}^N \sim [e^{Ax/N} e^{-Q\eta/N}]^N \rightarrow e^{Ax-Q\eta}$, in agreement with (45) in the main text.

APPENDIX C: ADDITIONAL NUMERICAL DATA

In Fig. 7, we show I_3 for the brickwall circuit. The plot is analogous to the plot of XEB in Fig. 2. We show that the x extracted from the XEB, with the procedure detailed in the main text, gives good predictions beyond 2-replicas observables through Eq. (45). This provides an additional check that all the moments of the anticoncentration, namely the whole PoP distribution, depend only on the two fitting parameters η and x .

In Fig. 8, we show that the average half-chain purity

$$P_{N/2} = \mathbb{E}[\text{Tr}(\rho_{N/2}^2)], \quad (\text{C1})$$

with $\rho_{N/2} = \text{Tr}_{1,\dots,N/2}(\rho_N(U))$, scales as $e^{-t(\lambda+1/\tau)}$. This observation comes from the fact that we can write $P_{N/2} = \text{Tr}(\rho^{\otimes 2}(s^{\otimes N/2} \otimes \mathbb{1}^{\otimes N/2}))$, where s is the swap operator. This forces a unique domain configuration

$$\begin{array}{c} \text{---} s \text{---} e \text{---} \\ \text{---} s \text{---} e \text{---} \end{array} \quad (\text{C2})$$

The contribution of the identity e is 1 while the contribution of the swap s is $\lim_{N \rightarrow \infty} (1 - (\eta t/N) Q_{ss})^{N/2} = e^{-\eta t}$. The unique domain wall gives $L^{-1} = L_0^{-1} e^{-t/\tau}$. Collecting all these contributions gives the correct scaling.

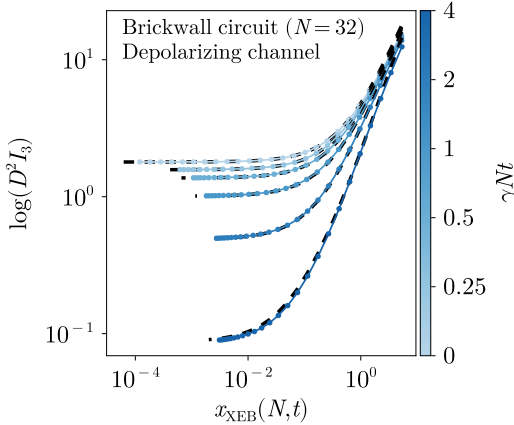


FIG. 7. Plot of I_3 for the brickwall circuit with two-qubits depolarizing channel with noise strength $\gamma \propto 1/(Nt)$, with $N = 32$. We show it as a function of $x_{\text{XEB}}(N, t)$ which is obtained by computing first the XEB at large time to estimate η and then by computing the cross-entropy benchmarking at time t , and inverting Eq. (46) to obtain x . We show also the analytical prediction Eq. (45) (black dashed line).

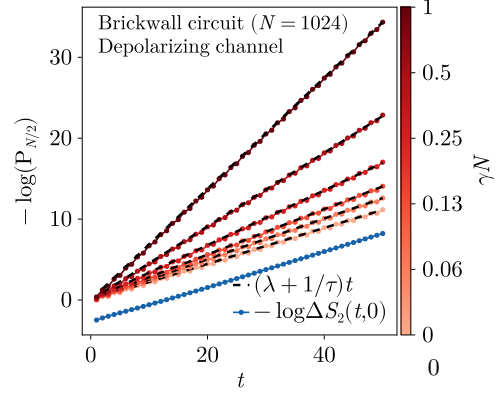


FIG. 8. Scaling of the half-chain purity for the brickwall circuit with two-qubits depolarizing channel and noise strength $\gamma \propto 1/N$. The system size is $N = 1024$. We show that this scaling is $\exp(-t \cdot (\lambda + 1/\tau))$, where the parameter λ is obtained through a fit from the cross-entropy benchmarking at large times. τ , given by Eq. (48), is the scaling of the purity in the noiseless case $\lambda = 0$. It also corresponds to the scaling of $\Delta S_2(t, 0) = S_2(\infty, 0) - S_2(t, 0)$ (see blue line), where $S_2(t, \lambda) = -\log I_2(t, \lambda)$ is the participation entropy and $S_2(\infty, 0) = -\log I_2^{\text{PT}}$. (For visualization purposes, the blue line is shifted by a constant of 3.)

To test the robustness of our universality claims beyond ensembles of random unitaries, we consider the kicked Ising model [41,110–112]. The dynamics is generated by the Floquet operator

$$U_F = e^{-ib \sum_j X_j} e^{-i(h \sum_j Z_j + J \sum_j Z_j Z_{j+1})}, \quad (\text{C3})$$

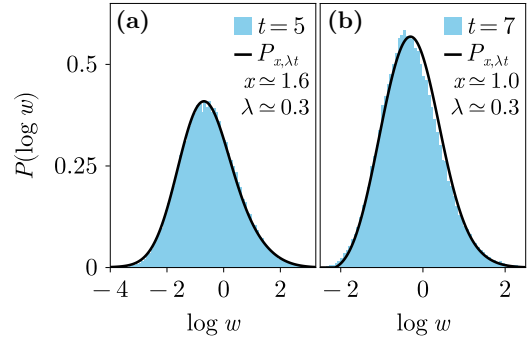


FIG. 9. Plot of the probability-of-probabilities for the kicked Ising model with $N = 32$ subject to single-qubit dephasing noise of strength $\gamma = 0.00781$. For each set of parameters, the histogram is built from ten independent circuit realizations, yielding a total of 10^5 samples per histogram. We present the distribution at two different times and extract the parameters x and λ through a maximum likelihood procedure. The resulting values are shown in the figure. The close agreement between the estimate of λ and of the half-chain purity timescale $\tau_{\text{Floquet}}^{-1}$, reveals that the system operates at the crossover between the two cross-entropy benchmarking regimes.

applied at each time step. We fix the parameters to be $J = 1$, $b = (\sqrt{5} + 5)/8$, and $h = (\sqrt{5} + 1)/4$. Our conclusions are not expected to depend sensitively on these particular values, provided the model remains in a nonintegrable regime. To mitigate basis-dependent effects, the Floquet evolution is initialized from a random product state. We consider as a source of noise a single-qubit quantum dephasing channel

$$\mathcal{N}(\rho) = (1 - \gamma)\rho + \gamma Z\rho Z. \quad (\text{C4})$$

Figure 9 shows the empirical and analytical distribution for this model. Due to the absence of efficient numerical techniques to estimate the XEB at large time, the two parameters of the analytical distribution are fitted through maximum likelihood estimation. The fitted λ can be compared with the half-chain purity timescale $\tau_{\text{Floquet}}^{-1} = 0.28$ obtained for the same model in a recent study [43].

-
- [1] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. Brandao, D. A. Buell *et al.*, Quantum supremacy using a programmable superconducting processor, *Nature* **574**, 505 (2019).
 - [2] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, Characterizing quantum supremacy in near-term devices, *Nat. Phys.* **14**, 595 (2018).
 - [3] A. Morvan *et al.*, Phase transitions in random circuit sampling, *Nature* **634**, 328 (2024).
 - [4] L. Daguerre, R. Blume-Kohout, N. C. Brown, D. Hayes, and I. H. Kim, Experimental demonstration of high-fidelity logical magic states from code switching, *Phys. Rev. X* **15**, 041008 (2025).
 - [5] M. Liu *et al.*, Certified randomness using a trapped-ion quantum processor, *Nature* **640**, 343 (2025).
 - [6] Z. Yin, I. Agresti, G. de Felice, D. Brown, A. Toumi, C. Pentangelo, S. Piacentini, A. Crespi, F. Ceccarelli, R. Osellame, B. Coecke, and P. Walther, Experimental quantum-enhanced kernel-based machine learning on a photonic processor, *Nat. Photonics* **19**, 1020 (2025).
 - [7] R. Haghshenas *et al.*, Digital quantum magnetism at the frontier of classical simulations, [arXiv:2503.20870](#).
 - [8] L. Shirizly, G. Misguich, and H. Landa, Dissipative dynamics of graph-state stabilizers with superconducting qubits, *Phys. Rev. Lett.* **132**, 010601 (2024).
 - [9] M. C. Smith, A. D. Leu, K. Miyanishi, M. F. Gely, and D. M. Lucas, Single-qubit gates with errors at the 10^{-7} level, *Phys. Rev. Lett.* **134**, 230601 (2025).
 - [10] B. Rost, L. Del Re, N. Earnest, A. F. Kemper, B. Jones, and J. K. Freericks, Long-time error-mitigating simulation of open quantum systems on near term quantum computers, *npj Quantum Inf.* **11**, 10 (2025).
 - [11] M. Ringbauer, M. Hinsche, T. Feldker, P. K. Faehrmann, J. Bermejo-Vega, C. L. Edmunds, L. Postler, R. Stricker, C. D. Marciniak, M. Meth, I. Pogorelov, R. Blatt, P. Schindler, J. Eisert, T. Monz, and D. Hangleiter, Verifiable measurement-based quantum random sampling with trapped ions, *Nat. Commun.* **16**, 106 (2025).
 - [12] J. A. Jacoby, D. A. Huse, and S. Gopalakrishnan, Spectral gaps of local quantum channels in the weak-dissipation limit, *Phys. Rev. B* **111**, 104303 (2025).
 - [13] C. Cafaro and S. Mancini, Quantum stabilizer codes for correlated and asymmetric depolarizing errors, *Phys. Rev. A* **82**, 012306 (2010).
 - [14] J. Preskill, Quantum computing in the NISQ era and beyond, *Quantum* **2**, 79 (2018).
 - [15] B. M. Terhal, Quantum error correction for quantum memories, *Rev. Mod. Phys.* **87**, 307 (2015).
 - [16] D. Aasen *et al.*, Roadmap to fault tolerant quantum computation using topological qubit arrays, [arXiv:2502.12252](#).
 - [17] T. Peham, L. Schmid, L. Berent, M. Müller, and R. Wille, Automated synthesis of fault-tolerant state preparation circuits for quantum error-correction codes, *PRX Quantum* **6**, 020330 (2025).
 - [18] J. C. Napp, R. L. La Placa, A. M. Dalzell, F. G. S. L. Brandão, and A. W. Harrow, Efficient classical simulation of random shallow 2D quantum circuits, *Phys. Rev. X* **12**, 021021 (2022).
 - [19] J. Chen, F. Zhang, C. Huang, M. Newman, and Y. Shi, Classical simulation of intermediate-size quantum circuits, [arXiv:1805.01450](#).
 - [20] I. L. Markov, A. Fatima, S. V. Isakov, and S. Boixo, Quantum supremacy is both closer and farther than it appears, [arXiv:1807.10749](#).
 - [21] C. Huang, F. Zhang, M. Newman, J. Cai, X. Gao, Z. Tian, J. Wu, H. Xu, H. Yu, B. Yuan, M. Szegedy, Y. Shi, and J. Chen, Classical simulation of quantum supremacy circuits, [arXiv:2005.06787](#).
 - [22] S. Aaronson and L. Chen, Complexity-theoretic foundations of quantum supremacy experiments, in *32nd Computational Complexity Conference (CCC 2017)*, edited by R. O'Donnell, Vol. 79 (Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl, Germany, 2017), pp. 22:1–22:67.
 - [23] V. Pednault, J. A. Gunnels, G. Nannicini, L. Horesh, and R. Wisnieff, Leveraging secondary storage to simulate deep 54-qubit sycamore circuits, [arXiv:1910.09534](#).
 - [24] J. Gray and S. Kourtis, Hyper-optimized tensor network contraction, *Quantum* **5**, 410 (2021).
 - [25] S. Bravyi, D. E. Browne, P. Calpin, E. Campbell, D. Gosset, and M. Howard, Simulation of quantum circuits by low-rank stabilizer decompositions, *Quantum* **3**, 181 (2019).
 - [26] T. Vincent, L. J. O'Riordan, M. Andrenkov, J. Brown, N. Killoran, H. Qi, and I. Dhand, Fast quantum circuit simulations with parallel task-based tensor-network contraction, *Quantum* **6**, 709 (2022).
 - [27] R. Harper, S. T. Flammia, and J. J. Wallman, Efficient learning of quantum noise, *Nat. Phys.* **16**, 1184 (2020).
 - [28] S. Hohloch and J. Palmer, Extending compact Hamiltonian S^1 -spaces to integrable systems with mild degeneracies in dimension four, [arXiv:2105.00523](#).
 - [29] J. Choi, A. L. Shaw, I. S. Madjarov, X. Xie, R. Finkelstein, J. P. Covey, J. S. Cotler, D. K. Mark, H.-Y. Huang, A. Kale, H. Pichler, F. G. S. L. Brandão, S. Choi, and M.

- Endres, Preparing random states and benchmarking with many-body quantum chaos, *Nature* **613**, 468 (2023).
- [30] 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, *PRX Quantum* **5**, 010334 (2024).
- [31] 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, [arXiv:2305.04954](https://arxiv.org/abs/2305.04954).
- [32] S. Liu, M.-R. Li, S.-X. Zhang, S.-K. Jian, and H. Yao, Noise-induced phase transitions in hybrid quantum circuits, *Phys. Rev. B* **110**, 064323 (2024).
- [33] D. K. Mark, J. Choi, A. L. Shaw, M. Endres, and S. Choi, Benchmarking quantum simulators using ergodic quantum dynamics, *Phys. Rev. Lett.* **131**, 110601 (2023).
- [34] M. P. Fisher, V. Khemani, A. Nahum, and S. Vijay, Random quantum circuits, *Annu. Rev. Condens. Matter Phys.* **14**, 335 (2023).
- [35] A. W. Harrow and R. A. Low, Random quantum circuits are approximate 2-designs, *Commun. Math. Phys.* **291**, 257 (2009).
- [36] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, Local random quantum circuits are approximate polynomial-designs, *Commun. Math. Phys.* **346**, 397 (2016).
- [37] D. Hangleiter, J. Bermejo-Vega, M. Schwarz, and J. Eisert, Anticoncentration theorems for schemes showing a quantum speedup, *Quantum* **2**, 65 (2018).
- [38] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits anticoncentrate in log depth, *PRX Quantum* **3**, 010333 (2022).
- [39] D. J. Luitz, N. Laflorencie, and F. Alet, Participation spectroscopy and entanglement Hamiltonian of quantum spin models, *J. Stat. Mech.: Theor. Exp.* **2014**, P08007 (2014).
- [40] D. J. Luitz, F. Alet, and N. Laflorencie, Universal behavior beyond multifractality in quantum many-body systems, *Phys. Rev. Lett.* **112**, 057203 (2014).
- [41] E. Tirrito, X. Turkeshi, and P. Sierant, Anticoncentration and nonstabilizerness spreading under ergodic quantum dynamics, *Phys. Rev. Lett.* **135**, 220401 (2025).
- [42] N. Macé, F. Alet, and N. Laflorencie, Multifractal scalings across the many-body localization transition, *Phys. Rev. Lett.* **123**, 180601 (2019).
- [43] A. Sauliere, B. Magni, G. Lami, X. Turkeshi, and J. De Nardis, Universality in the anticoncentration of chaotic quantum circuits, *Phys. Rev. B* **112**, 134312 (2025).
- [44] G. Lami, J. De Nardis, and X. Turkeshi, Anticoncentration and state design of random tensor networks, *Phys. Rev. Lett.* **134**, 010401 (2025).
- [45] A. Christopoulos, A. Chan, and A. De Luca, Universal distributions of overlaps from generic dynamics in quantum many-body systems, *Phys. Rev. Res.* **7**, 043035 (2025).
- [46] B. Magni, A. Christopoulos, A. De Luca, and X. Turkeshi, Anticoncentration in Clifford circuits and beyond: From random tensor networks to pseudo-magic states, *Phys. Rev. X* **15**, 031071 (2025).
- [47] P. W. Claeys and G. De Tomasi, Fock-space delocalization and the emergence of the Porter-Thomas distribution from dual-unitary dynamics, *Phys. Rev. Lett.* **134**, 050405 (2025).
- [48] D. K. Mark, F. Surace, A. Elben, A. L. Shaw, J. Choi, G. Refael, M. Endres, and S. Choi, Maximum entropy principle in deep thermalization and in Hilbert-space ergodicity, *Phys. Rev. X* **14**, 041051 (2024).
- [49] A. Kaufman, J. Corona, Z. Ozzello, B. Senseman, M. Asaduzzaman, and Y. Meurice, Improved entanglement entropy estimates from filtered bitstring probabilities, *Phys. Rev. A* **112**, 032430 (2025).
- [50] A. L. Shaw, D. K. Mark, J. Choi, R. Finkelstein, P. Scholl, S. Choi, and M. Endres, Universal fluctuations and noise learning from Hilbert-space ergodicity, *Phys. Rev. X* **15**, 031001 (2025).
- [51] A. De Luca and A. Scardicchio, Ergodicity breaking in a model showing many-body localization, *Europhys. Lett.* **101**, 37003 (2013).
- [52] C. E. Porter and R. G. Thomas, Fluctuations of nuclear reaction widths, *Phys. Rev.* **104**, 483 (1956).
- [53] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits transform local noise into global white noise, *Commun. Math. Phys.* **405**, 78 (2024).
- [54] A. M. Dalzell, N. Hunter-Jones, and F. G. S. L. Brandão, Random quantum circuits transform local noise into global white noise, *Commun. Math. Phys.* **405**, 78 (2024).
- [55] P. V. Klimov *et al.*, Optimizing quantum gates towards the scale of logical qubits, *Nat. Commun.* **15**, 2442 (2024).
- [56] D. A. Abanin *et al.*, Observation of constructive interference at the edge of quantum ergodicity, *Nature* **646**, 825 (2025).
- [57] H. J. Manetsch, G. Nomura, E. Bataille, X. Lv, K. H. Leung, and M. Endres, A tweezer array with 6100 highly coherent atomic qubits, *Nature* **647**, 60 (2025).
- [58] N.-C. Chiu, E. C. Trapp, J. Guo, M. H. Abobeih, L. M. Stewart, S. Hollerith, P. L. Stroganov, M. Kalinowski, A. A. Geim, S. J. Evered, S. H. Li, X. Lyu, L. M. Peters, D. Bluvstein, T. T. Wang, M. Greiner, V. Vuletić, and M. D. Lukin, Continuous operation of a coherent 3000-qubit system, *Nature* **646**, 1075 (2025).
- [59] A. Chan, A. De Luca, and J. T. Chalker, Spectral statistics in spatially extended chaotic quantum many-body systems, *Phys. Rev. Lett.* **121**, 060601 (2018).
- [60] S. Shivam, A. De Luca, D. A. Huse, and A. Chan, Many-body quantum chaos and emergence of Ginibre ensemble, *Phys. Rev. Lett.* **130**, 140403 (2023).
- [61] A. Chan, S. Shivam, D. A. Huse, and A. De Luca, Many-body quantum chaos and space-time translational invariance, *Nat. Commun.* **13**, 7484 (2022).
- [62] X. Turkeshi and P. Sierant, Hilbert space delocalization under random unitary circuits, *Entropy* **26**, 471 (2024).
- [63] S. Garnerone, T. R. de Oliveira, and P. Zanardi, Typicality in random matrix product states, *Phys. Rev. A* **81**, 032336 (2010).
- [64] S. Garnerone, T. R. de Oliveira, S. Haas, and P. Zanardi, Statistical properties of random matrix product states, *Phys. Rev. A* **82**, 052312 (2010).
- [65] D. Haag, F. Baccari, and G. Styliaris, Typical correlation length of sequentially generated tensor network states, *PRX Quantum* **4**, 030330 (2023).
- [66] C. Lancien and D. Pérez-García, Correlation length in random MPS and PEPS, *Ann. Henri Poincaré* **23**, 141 (2022).

- [67] G. Lami, T. Haug, and J. De Nardis, Quantum state designs with Clifford-enhanced matrix product states, *PRX Quantum* **6**, 010345 (2025).
- [68] Z. Cheng, X. Feng, and M. Ippoliti, Pseudoentanglement from tensor networks, *Phys. Rev. Lett.* **135**, 020403 (2025).
- [69] A. Nahum, S. Vijay, and J. Haah, Operator spreading in random unitary circuits, *Phys. Rev. X* **8**, 021014 (2018).
- [70] L. Logarić, S. Dooley, S. Pappalardi, and J. Goold, Quantum many-body scars in dual-unitary circuits, *Phys. Rev. Lett.* **132**, 010401 (2024).
- [71] A. Chan, A. De Luca, and J. T. Chalker, Solution of a minimal model for many-body quantum chaos, *Phys. Rev. X* **8**, 041019 (2018).
- [72] A. Chan, A. De Luca, and J. T. Chalker, Spectral statistics in spatially extended chaotic quantum many-body systems, *Phys. Rev. Lett.* **121**, 060601 (2018).
- [73] A. Chan, A. De Luca, and J. T. Chalker, Eigenstate correlations, thermalization, and the butterfly effect, *Phys. Rev. Lett.* **122**, 220601 (2019).
- [74] B. Bertini, P. Kos, and T. Prosen, Exact correlation functions for dual-unitary lattice models in $1 + 1$ dimensions, *Phys. Rev. Lett.* **123**, 210601 (2019).
- [75] L. Piroli, B. Bertini, J. I. Cirac, and T. Prosen, Exact dynamics in dual-unitary quantum circuits, *Phys. Rev. B* **101**, 094304 (2020).
- [76] D. E. Parker, X. Cao, A. Avdoshkin, T. Scaffidi, and E. Altman, A universal operator growth hypothesis, *Phys. Rev. X* **9**, 041017 (2019).
- [77] J. S. Cotler, N. Hunter-Jones, J. Liu, B. Yoshida, and X.-L. Qi, Chaos, complexity, and random matrices, *J. High Energy Phys.* **2017**, 48 (2017).
- [78] K. Khanna, A. Kumar, R. Vasseur, and A. W. W. Ludwig, Random quantum circuits with time-reversal symmetry, *Phys. Rev. Res.* **8**, 013165 (2026).
- [79] F. Barratt, U. Agrawal, S. Gopalakrishnan, D. A. Huse, R. Vasseur, and A. C. Potter, Field theory of charge sharpening in symmetric monitored quantum circuits, *Phys. Rev. Lett.* **129**, 120604 (2022).
- [80] N. Hunter-Jones, J. Cotler, J. Liu, B. Yoshida, and X.-L. Qi, Chaos, complexity, and random matrices, *J. High Energy Phys.* **2018**, 121 (2018).
- [81] H.-P. Breuer and F. Petruccione, *The Theory of Open Quantum Systems* (OUP, Oxford, 2002).
- [82] G. Köstenberger, Weingarten calculus, [arXiv:2101.00921](https://arxiv.org/abs/2101.00921).
- [83] B. Collins, S. Matsumoto, and J. Novak, The Weingarten calculus, *Not. Am. Math. Soc.* **69**, 1 (2022).
- [84] R. Brualdi, *Introductory Combinatorics*, Pearson Education (Pearson Prentice Hall, London, United Kingdom, 2012).
- [85] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, Quantum entanglement growth under random unitary dynamics, *Phys. Rev. X* **7**, 031016 (2017).
- [86] 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).
- [87] T. Zhou and A. Nahum, Emergent statistical mechanics of entanglement in random unitary circuits, *Phys. Rev. B* **99**, 174205 (2019).
- [88] T. Zhou and A. Nahum, Entanglement membrane in chaotic many-body systems, *Phys. Rev. X* **10**, 031066 (2020).
- [89] A. De Luca, C. Liu, A. Nahum, and T. Zhou, Universality classes for purification in nonunitary quantum processes, *Phys. Rev. X* **15**, 041024 (2025).
- [90] C. Schon, K. Hammerer, M. M. Wolf, J. I. Cirac, and E. Solano, Sequential generation of matrix-product states in cavity QED, *Phys. Rev. A* **75**, 032311 (2007).
- [91] J. Haferkamp, C. Bertoni, I. Roth, and J. Eisert, Emergent statistical mechanics from properties of disordered random matrix product states, *PRX Quantum* **2**, 040308 (2021).
- [92] M. Collura, G. Lami, N. Ranabhat, and A. Santini, *Tensor Network Techniques for Quantum Computation* (SISSA Medialab s.r.l., Trieste, Italy, 2024).
- [93] P. Hall, *The Bootstrap and Edgeworth Expansion*, 1st ed., Springer Series in Statistics (Springer, New York, 1992).
- [94] F. Y. Edgeworth, The law of error, *Nature* **36**, 482 (1887).
- [95] G. Vidal, Efficient simulation of one-dimensional quantum many-body systems, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [96] M. Zwolak and G. Vidal, Mixed-state dynamics in one-dimensional quantum lattice systems: A time-dependent superoperator renormalization algorithm, *Phys. Rev. Lett.* **93**, 207205 (2004).
- [97] 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).
- [98] M. Fishman, S. R. White, and E. M. Stoudenmire, The Itensor software library for tensor network calculations, *SciPost Phys. Codebases*, 004 (2022).
- [99] L. Piroli, G. Styliaris, and J. I. Cirac, Quantum circuits assisted by local operations and classical communication: Transformations and phases of matter, *Phys. Rev. Lett.* **127**, 220503 (2021).
- [100] D. Malz, G. Styliaris, Z.-Y. Wei, and J. I. Cirac, Preparation of matrix product states with log-depth quantum circuits, *Phys. Rev. Lett.* **132**, 040404 (2024).
- [101] K. C. Smith, A. Khan, B. K. Clark, S. M. Girvin, and T.-C. Wei, Constant-depth preparation of matrix product states with adaptive quantum circuits, *PRX Quantum* **5**, 030344 (2024).
- [102] D. T. Stephen and O. Hart, Preparing matrix product states via fusion: Constraints and extensions, *Phys. Rev. B* **112**, 235127 (2025).
- [103] Y. Zhang, S. Gopalakrishnan, and G. Styliaris, Characterizing MPS and PEPS preparable via measurement and feedback, *PRX Quantum* **5**, 040304 (2024).
- [104] M. Votto, M. Ljubotina, C. Lancien, J. I. Cirac, P. Zoller, M. Serbyn, L. Piroli, and B. Vermersch, Learning mixed quantum states in large-scale experiments, *Phys. Rev. Lett.* **136**, 090801 (2026).
- [105] A. J. Friedman, A. Chan, A. De Luca, and J. T. Chalker, Spectral statistics and many-body quantum chaos with conserved charge, *Phys. Rev. Lett.* **123**, 210603 (2019).
- [106] E. McCulloch, J. De Nardis, S. Gopalakrishnan, and R. Vasseur, Full counting statistics of charge in chaotic

- many-body quantum systems, *Phys. Rev. Lett.* **131**, 210402 (2023).
- [107] L. Hruza and D. Bernard, Coherent fluctuations in noisy mesoscopic systems, the open quantum SSEP, and free probability, *Phys. Rev. X* **13**, 011045 (2023).
- [108] F. H. L. Essler and A. J. J. M. de Klerk, Statistics of matrix elements of local operators in integrable models, *Phys. Rev. X* **14**, 031048 (2024).
- [109] A. Saulière, Data for “Universality in the anticoncentration of noisy quantum circuits at finite depths,” <https://doi.org/10.5281/zenodo.19094338> (2026).
- [110] T. Prosen, Time evolution of a quantum many-body system: Transition from integrability to ergodicity in the thermodynamic limit, *Phys. Rev. Lett.* **80**, 1808 (1998).
- [111] T. Prosen, Ergodic properties of a generic nonintegrable quantum many-body system in the thermodynamic limit, *Phys. Rev. E* **60**, 3949 (1999).
- [112] T. Prosen, General relation between quantum ergodicity and fidelity of quantum dynamics, *Phys. Rev. E* **65**, 036208 (2002).