Minimal Model for Fast Scrambling

Ron Belyansky[*], Przemyslaw Bienias, Yaroslav A. Kharkov, Alexey V. Gorshkov, and Brian Swingle

Joint Center for Quantum Information and Computer Science, NIST/University of Maryland,
College Park, Maryland 20742, USA and
Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742, USA

We study quantum information scrambling in spin models with both long-range all-to-all and short-range interactions. We argue that a simple global, spatially homogeneous interaction together with local chaotic dynamics is sufficient to give rise to fast scrambling, which describes the spread of quantum information over the entire system in a time that is logarithmic in the system size. This is illustrated in two tractable models: (1) a random circuit with Haar random local unitaries and a global interaction and (2) a classical model of globally coupled nonlinear oscillators. We use exact numerics to provide further evidence by studying the time evolution of an out-of-time-order correlator and entanglement entropy in spin chains of intermediate sizes. Our results pave the way towards experimental investigations of fast scrambling and aspects of quantum gravity with quantum simulators.

Introduction.—The study of quantum information scrambling has recently attracted significant attention due to its relation to quantum chaos and thermalization of isolated many-body systems [1–3] as well as the dynamics of black holes [4–7]. Scrambling refers to the spread of initially local quantum information over the many-body degrees of freedom of the entire system, rendering it inaccessible to local measurements. Scrambling is also related to the Heisenberg dynamics of local operators, and can be probed via the squared commutator of two local and Hermitian operators W_1, V_r, at positions 1 and r, respectively,

$$\mathcal{C}(r,t) = -\frac{1}{2}\langle [W_1(t), V_r]^2 \rangle, \qquad (1)$$

where $W_1(t)$ is the Heisenberg evolved operator. The growth of the squared commutator corresponds to $W_1(t)$ increasing in size and complexity, leading it to fail to commute with V_r. In a local quantum chaotic system, $\mathcal{C}(r,t)$ typically spreads ballistically, exhibiting rapid growth ahead of the wavefront and saturation behind, at late times [8–10].

Of particular interest are the so-called fast scramblers, systems where $\mathcal{C}(r,t)$ reaches $O(1)$ for all r in a time $t_s \propto \log(N)$, with N being the number of degrees of freedom. Among the best known examples are black holes, which are conjectured to be the fastest scramblers in nature [5–7,11], as well as the Sachdev-Ye-Kitaev (SYK) [12,13] model and other related holographic models [14–17].

Recent advances in the development of coherent quantum simulators have enabled the study of out-of-equilibrium dynamics of spin models with controllable interactions [18], making them ideal platforms to experimentally study information scrambling. Several experiments have already been performed [19–24], probing scrambling in either local or nonchaotic systems. The experimental observation of fast scrambling remains challenging however, particularly because few systems are known to be fast scramblers, and those that are, like the SYK model, are highly nontrivial, involving random couplings and many-body interactions. Some recent proposals suggested that spin models with nonlocal interactions can exhibit fast scrambling [25–27], albeit with complicated and inhomogeneous interactions.

In this Letter, we argue that the simplest possible global interaction, together with chaotic dynamics, are sufficient to make a spin model fast scrambling. We consider spin-1/2 chains with Hamiltonians of the form

$$\mathcal{H} = \mathcal{H}_{\text{local}} - \frac{g}{\sqrt{N}} \sum_{i<j} Z_i Z_j, \qquad (2)$$

where Z_i is the Pauli z operator acting on site i and $\mathcal{H}_{\text{local}}$ is a Hamiltonian with only local interactions that ensures that the full $\mathcal{H}$ is chaotic. We note that such global interactions are ubiquitous in ultracold atoms in optical cavities [28–32], and also in ion traps [33–36].

We first show that this effect is generic, by studying two models, a random quantum circuit and a classical model, both designed to mimic the universal dynamics of Eq. (2). We then provide numerical evidence for fast scrambling for a particular time-independent quantum Hamiltonian. Finally, we discuss possible experimental realizations.

Random circuit model.—As a proof of principle, we consider a system of N spin-1/2 sites, with dynamics generated by a random quantum circuit (see Fig. 1) inspired
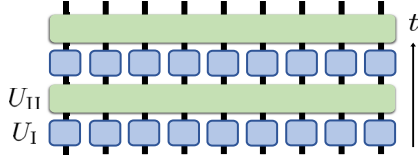
FIG. 1. Diagram of the random circuit. As given in Eq. (3), each blue square is an independent Haar-random unitary $U_{H,i}$ acting on site i, and the green rectangle is the global interaction U_{II}.

by the Hamiltonian in Eq. (2). While less physical than the Hamiltonian model, it has the advantage of being exactly solvable while providing intuition about generic many-body chaotic systems with similar features.

The time-evolution operator is $U(t) = (U_{II}U_I)^t$ where a single-time-step update consists of the two layers

$$U_I = \prod_{i=1}^{N} U_{H,i}, \qquad U_{II} = e^{-i\frac{g}{2\sqrt{N}}\sum_{i<j} Z_i Z_j}, \qquad (3)$$

where each $U_{H,i}$ is an independent Haar-random single-site unitary. The two layers in Eq. (3) are motivated by the two terms in Eq. (2), with the Haar-random unitaries replacing $\mathcal{H}_{\text{local}}$.

We are interested in the operator growth of an initially simple operator $\mathcal{O}$. At any point in time, the Heisenberg operator $\mathcal{O}(t) = U^\dagger(t)\mathcal{O}U(t)$ can be decomposed as $\mathcal{O}(t) = \sum_{\mathcal{S}} a_{\mathcal{S}}(t)\mathcal{S}$, where $\mathcal{S}$ is a string composed of the Pauli matrices and the identity, forming a basis for $\text{SU}(2^N)$. As in random brickwork models [37,38] and random Brownian models [9], the Haar-averaged probabilities $\langle a_{\mathcal{S}}^2(t)\rangle$, encoding the time evolution of $\mathcal{O}(t)$, themselves obey a linear equation

$$\langle a_{\mathcal{S}}^2(t+1)\rangle = \sum_{\mathcal{S}'} W_{\mathcal{S},\mathcal{S}'}\langle a_{\mathcal{S}'}^2(t)\rangle. \qquad (4)$$

Here, $W_{\mathcal{S},\mathcal{S}'}$ is a $4^N \times 4^N$ stochastic matrix describing a fictitious Markov process [39,40]. The average probabilities $\langle a_{\mathcal{S}}^2(t)\rangle$ fully determine the growth of the average of $\mathcal{C}(t)$ in Eq. (1) [see Supplemental Material (SM) [41]]. Because of the Haar unitaries and the simple uniform interaction in Eq. (3), $W_{\mathcal{S},\mathcal{S}'}$ is highly degenerate and only depends on the total weights of the strings $\mathcal{S}, \mathcal{S}'$, counting the number of nonidentity operators, i.e., $w(\mathcal{S}) = \sum_i(1 - \delta_{\mathcal{S}_i,1})$, and on the number of sites where both $\mathcal{S}$ and $\mathcal{S}'$ are nonidentity, i.e., $v(\mathcal{S},\mathcal{S}') = \sum_i(1 - \delta_{\mathcal{S}_i,1})(1 - \delta_{\mathcal{S}'_i,1})$, and is given by (see SM for derivation [41]) [43]

$$W(w,w',v) = \left(\frac{1}{3}\right)^{w+w'} \sum_{k=0}^{v}\binom{v}{k}\sum_{l=0}^{k}\binom{k}{l}$$
$$\times \left[\cos^2\left(\frac{2l-k}{\sqrt{N}}g\right)\right]^{N-k-(w+w'-2v)}\left[\sin^2\left(\frac{2l-k}{\sqrt{N}}g\right)\right]^{w+w'-2v}. \qquad (5)$$

If we further assume that $\mathcal{O}$ starts out as a single site operator on site 1, then throughout the evolution, $\langle a_{\mathcal{S}}^2(t)\rangle$ only depend on the total operator weight w, and the weight on site 1, which we denote by $w_1 \in \{0, 1\}$. We thus introduce the operator weight probability h_t at time t,

$$h_t(w, w_1) = \langle a_{\mathcal{S}}^2(t)\rangle 3^w \binom{N-1}{w-w_1}, \qquad (6)$$

which gives the probability of $\mathcal{O}(t)$ having total weight w and weight w_1 on site 1.

The time evolution of $h_t(w, w_1)$ is given by the master equation

$$h_{t+1}(w, w_1) = \sum_{w_1'=0,1}\sum_{w'=w_1'}^{N-1+w_1'} \mathcal{R}(w,w_1,w',w_1')h_t(w',w_1'), \qquad (7)$$

where the $2N \times 2N$ matrix $\mathcal{R}$ is

$$\mathcal{R}(w,w_1,w',w_1') = 3^w \sum_{m=0}^{\min\{w-w_1,w'-w_1'\}}\binom{w'-w_1'}{m}$$
$$\times \binom{N-1-w'+w_1'}{w-w_1-m}W(w,w',m+w_1 w_1'). \qquad (8)$$

The transition matrix $\mathcal{R}$, scaling only linearly with N, allows us to efficiently simulate the dynamics for large system sizes (see Fig. 2).

To proceed analytically, we Taylor expand Eq. (5) to leading order in g, which gives rise to a closed master equation for the total operator weight probability $h_t(w) \equiv h_t(w, 0) + h_t(w, 1)$,
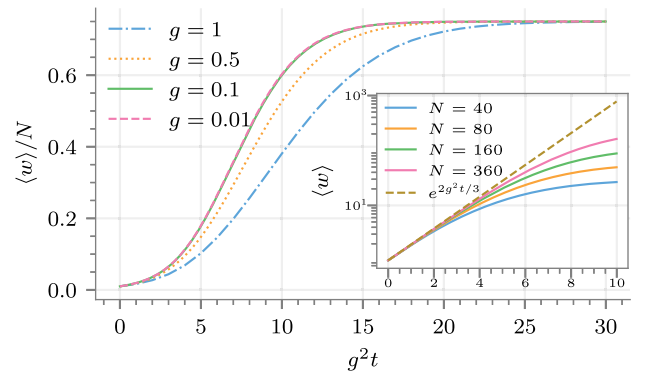

FIG. 2. Normalized mean operator weight $\langle w(t)\rangle/N = (1/N)\sum_w wh_t(w)$ as a function of time for different g and $N = 100$, computed using Eq. (7). For small enough g, all the curves collapse to a single curve as a function of g^2t, as implied by Eq. (10). The inset shows the initial exponential increase of $\langle w(t)\rangle$ for different system sizes N and $g = 0.1$.

$$\frac{h_{t+1}(w) - h_t(w)}{g^2} = \frac{2w}{9N}(1 - 3N + 2w)h_t(w)$$
$$+ \frac{2w(w+1)}{9N}h_t(w+1)$$
$$+ \frac{N-w+1}{3N}2(w-1)h_t(w-1), \quad (9)$$

which is similar to random Brownian models [9,44] and shows that, at $O(g^2)$, w can change by at most ± 1 in a single step. Assuming that $h(w,t)$ varies slowly with respect to $g^2 t$ and w, we can approximate the above equation by a Fokker-Planck equation (rescaling time $\tau = g^2 t$)

$$\partial_\tau h(w,\tau) = -\partial_w[D_1(w)h(w,\tau)] + \partial_w^2[D_2(w)h(w,\tau)], \quad (10)$$

where the drift and diffusion coefficients are [dropping higher order terms $O(1/N, w/N)$]

$$D_1(w) = \frac{2}{3}\left(w - \frac{4w^2}{3N}\right), \qquad D_2(w) = \frac{w}{3} - \frac{2w^2}{9N}. \quad (11)$$

This equation describes the rapid growth of an initially localized distribution, followed by a broadening and finally saturation (see Fig. 2 and SM [41] for more details). At early time, the $\frac{2}{3}w$ term in the drift coefficient dominates, giving rise to exponential growth of the mean operator weight $\langle w(t) \rangle \sim e^{2g^2 t/3}$, which agrees with the full numerical solution of the master equation, as can be seen in Fig. 2. The mean weight is related to the infinite-temperature squared-commutator in Eq. (1) (averaged over different circuits) via $\langle \mathcal{C}(t) \rangle = \frac{4}{3}\langle w(t) \rangle / N$ [41]. Since $\langle w(t) \rangle$ grows exponentially with time, $\langle w(t) \rangle$ reaches $O(N)$ and $\langle \mathcal{C}(t) \rangle$ reaches $O(1)$ when $t = (3/2g^2)\log(N)$, thus establishing that this model is fast scrambling. Note that the $1/\sqrt{N}$ normalization in Eqs. (2) and (3) is crucial. Had we chosen instead $1/N (g \to g/\sqrt{N})$, the Lyapunov exponent would have been $2g^2/3N$ and the scrambling time would have been $t \sim N \log(N)$.

Classical model.—Let us now consider a different setting that also allows us to probe the basic timescales involved, and shows that randomness is not required. A convenient tractable choice is a classical model consisting of globally coupled nonlinear oscillators. Note that the analogs of out-of-time-order correlators (OTOCs) have been studied in a variety of classical models [26,45–49] and have been shown to capture the scrambling dynamics of quantum models like the SYK model [50–52].

Consider a 2N-dimensional phase space with coordinates q_r (positions) and p_r (momenta) for $r = 1, \dots, N$ with canonical structure specified by the Poisson brackets $\{q_r, p_s\}_{\mathrm{PB}} = \delta_{rs}$. The Hamiltonian is $\mathcal{H}_c = K + V_2 + V_4$ where

$$K = \sum_{r=1}^{N} \frac{p_r^2}{2}, \qquad V_4 = \frac{\Omega_3^2}{4}\sum_{r=1}^{N} q_r^4, \quad (12)$$

$$V_2 = \frac{\Omega_1^2}{2}\sum_{r=1}^{N-1}(q_{r+1} - q_r)^2 + \frac{\Omega_2^2}{2\sqrt{N}}\left(\sum_{r=1}^{N} q_r\right)^2. \quad (13)$$

The timescales for the growth of perturbations under $\mathcal{H}_c$ dynamics may be understood in two stages. First, $K + V_2$ can be solved exactly; this combination of terms provides the nonlocality. The remaining V_4 term renders the dynamics chaotic, provided Ω_3 is large enough. The dynamics of $K + V_2$ causes a localized perturbation to spread to every oscillator with nonlocal amplitude $1/N$ in a time of order $1/N^{1/4}\Omega_2$. Then conventional local chaos can amplify this $1/N$-sized perturbation to order-one size in a time of order $\lambda^{-1} \ln N$, where λ is some typical Lyapunov exponent.

At the quadratic level, the uniform mode, $Q = (1/N)\sum_r q_r$, is decoupled from the remaining modes of the chain. Hence, the propagation of any perturbation is a superposition of the motion due to the local Ω_1 terms and the special dynamics of the uniform mode. Since the local terms cannot induce nonlocal perturbations, we may focus on the dynamics of the uniform mode. The uniform mode's equation of motion is $(d^2Q/dt^2) = -\sqrt{N}\Omega_2^2 Q$ with solution

$$Q(t) = Q(0)\cos(N^{\frac{1}{4}}\Omega_2 t) + \frac{\frac{dQ}{dt}(0)}{N^{\frac{1}{4}}\Omega_2}\sin(N^{\frac{1}{4}}\Omega_2 t). \quad (14)$$

A localized perturbation on site 1 with zero initial time derivative can be written as $\delta\vec{q}(0) = \epsilon([\hat{e}_1 - \hat{u}_0] + \hat{u}_0)$, where $\hat{u}_0 = [1, \dots, 1]^T/N$ represents the uniform mode, $\hat{e}_1 = [1, 0, \dots, 0]^T$, and $\hat{e}_1 - \hat{u}_0$ is orthogonal to the uniform mode. The orthogonal mode evolves in a local fashion, hence $\delta\vec{q}(t) = \epsilon[\text{local piece} + \hat{u}_0\cos(N^{\frac{1}{4}}\Omega_2 t)]$. For oscillators far from the initial local perturbation, the dynamics is given by

$$\delta q_{r\gg 1}(t) = \frac{\epsilon}{N}\left[\cos\left(N^{\frac{1}{4}}\Omega_2 t\right) - 1\right]. \quad (15)$$

Thus, after a time $\pi/N^{\frac{1}{4}}\Omega_2$, any localized perturbation has spread to distant sites with amplitude ϵ/N.

The inclusion of V_4 renders the equations of motion nonlinear and the system chaotic in at least part of the phase space. We leave a detailed study of the classical chaotic dynamics of this model to the future, but as can be seen in Fig. 3, a numerical solution of the equations of motion displays sensitivity to initial conditions.

The precise protocol is as follows. We compare the dynamics of two configurations, $\vec{q}^{(1)}$ and $\vec{q}^{(2)}$, averaged over many initial conditions. The initial condition of configuration one has each oscillator start at rest from a
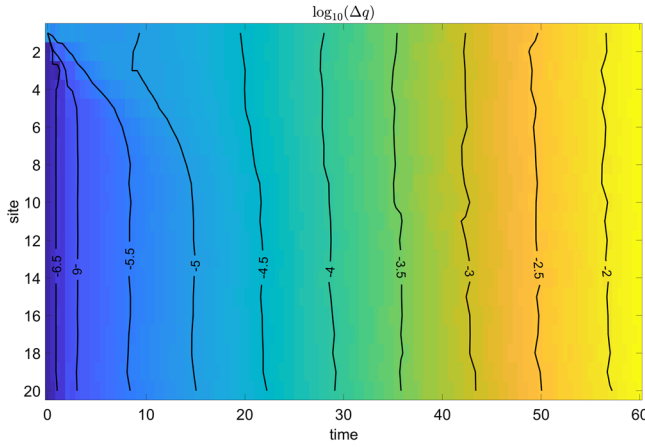
FIG. 3. $\log_{10} \Delta q_r(t)$ for $N = 20$, $\epsilon = 10^{-5}$, $\Omega_1 = \Omega_2 = 1$, and $\Omega_3 = 2$. The labeled black lines are contours of constant $\log_{10} \Delta q$. Early time ballistic growth is visible in the upper left corner while at later times the system exhibits spatially uniform exponential growth in time.

random amplitude drawn uniformly and independently from $[-1, 1]$. Configuration two is identical to configuration one except that $q_1^{(2)}(0) = q_1^{(1)}(0) + \epsilon$ for $\epsilon = 10^{-5}$. Both configurations are evolved in time and the difference $\Delta q_r(t) = |q_r^{(2)}(t) - q_r^{(1)}(t)|$ is computed and averaged over 4000 different initial conditions. Figure 3 shows this average of Δq_r for $N = 20$ with $\Omega_1 = 1$, $\Omega_2 = 1$, and $\Omega_3 = 2$. Because the system can generate an ϵ/N-sized perturbation on all sites in a short time, the subsequent uniform exponential growth implies that any local perturbation will become order one on all sites after a time $\sim \lambda^{-1} \log(N/\epsilon)$.

The above analysis corresponds to the classical limit of coupled quantum oscillators where some effective dimensionless Planck's constant vanishes, $\hbar_{\text{eff}} \to 0$. In the opposite limit of large N at fixed $\hbar_{\text{eff}}$, the dynamics of quantum OTOCs can be obtained from the corresponding classical Lyapunov growth up to a timescale of order $\log(1/\hbar_{\text{eff}}) \ll \log N$. At later times, one needs to consider fully quantum local dynamics. If one imagines breaking the system up into local clusters and if each cluster can be viewed as a quantum chaotic system with random-matrix-like energy levels, a dynamical system not unlike the random circuit model above is obtained.

Chaos and level statistics.—Having established fast scrambling in both the random circuit and the classical model, we now return to the quantum spin model of Eq. (2). We first examine whether such a model is chaotic, which is a necessary condition for it being fast scrambling. For the local Hamiltonian part, we consider the mixed-field Ising chain

$$\mathcal{H}_{\text{local}} = -J\sum_i Z_i Z_{i+1} - h_x\sum_i X_i - h_z\sum_i Z_i. \quad (16)$$

A standard approach to identify a transition from integrability to quantum chaos is based on a comparison of energy-level-spacing statistics with Poisson and Wigner-Dyson distributions. Another convenient metric is the average ratio of consecutive level spacings [53] $\langle r \rangle$, where $r = \min(r_n, 1/r_n)$, $r_n = \delta_n/\delta_{n-1}$, $\delta_n = E_n - E_{n-1}$, and E_n are the eigenvalues ordered such that $E_n \geq E_{n-1}$.

As was already suggested in Ref. [54] for a similar model, we find that the longitudinal field is unnecessary, and the full system can have Wigner-Dyson statistics even for $h_z = 0$, in which case $\mathcal{H}_{\text{local}}$ is integrable. The resulting Hamiltonian reads

$$\mathcal{H} = -J\sum_i Z_i Z_{i+1} - h_x\sum_i X_i - \frac{g}{\sqrt{N}}\sum_{i<j} Z_i Z_j. \quad (17)$$

The average adjacent-level-spacing ratio changes from $\langle r \rangle_{\text{Pois}} \approx 0.38$ for Poisson level statistics to $\langle r \rangle_{\text{GOE}} \approx 0.53$ for Wigner-Dyson level statistics in the Gaussian orthogonal ensemble (GOE) [53]. In the vicinity of $g \to 0$, $\langle r \rangle$ (see Fig. 4) shows proximity to Poisson statistics, while, for $|g| \gtrsim 0.25$, the level statistics agree with those of the GOE.

Out-of-time-order correlator and entanglement growth.— We now study the dynamics of an OTOC and entanglement entropy in the spin chain. We consider the following OTOC:

$$F(r, t) = \Re[\langle Z_1(t)Z_r Z_1(t)Z_r \rangle], \quad (18)$$

which is related to Eq. (1) by $\mathcal{C}(r, t) = 1 - F(r, t)$. The expectation value is evaluated in a Haar-random pure state, which approximates the infinite-temperature OTOC, but enables us to reach larger system sizes [55].

In Fig. 5(a), we show the OTOC for an open chain of $N = 20$ spins for both the local model, governed by $\mathcal{H}_{\text{local}}$ only, and the nonlocal model in Eq. (17), which includes the global interaction. In the local case, the OTOC spreads ballistically, forming a linear light cone. In contrast, in the nonlocal case, the spreading is superballistic and $F(r \gg 1, t)$ is approximately independent of r, as expected
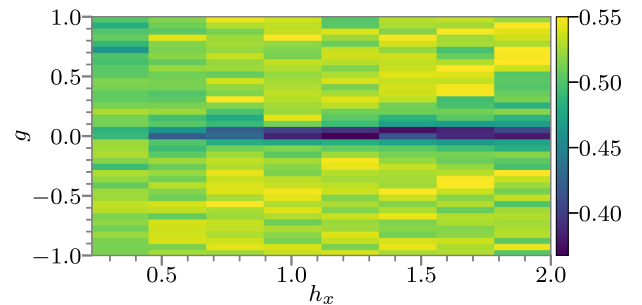

FIG. 4. Average adjacent-level-spacing ratio $\langle r \rangle$ for the model in Eq. (17) with $J = 1$. Data corresponds to a system of $N = 15$ spins with periodic boundary conditions for fixed momentum and Z-reflection symmetry blocks of the Hamiltonian.
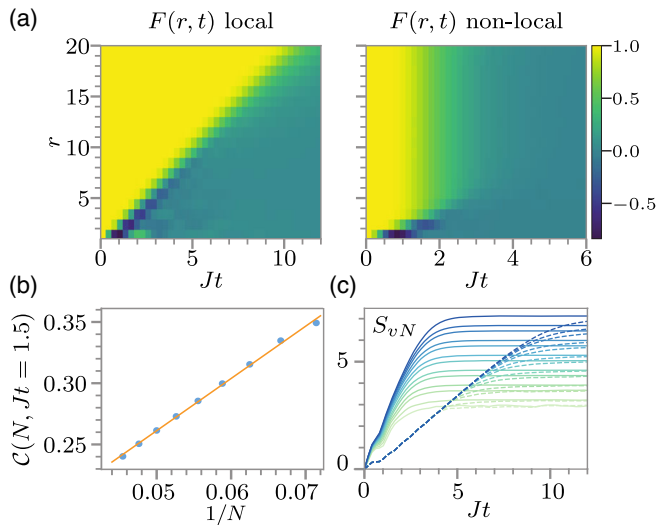
FIG. 5. (a) Time evolution of the OTOC for the (left) local and (right) nonlocal models. (b) $1 - F(N, t)$ after a fixed evolution time for the nonlocal model (for different system sizes N), showing a linear dependence on $1/N$. The orange line is a linear fit. (c) Half-cut entanglement-entropy growth starting from the $+\hat{y}$ state for local (dashed lines) and nonlocal (solid lines) models. The color indicates the system size, starting from $N = 10$ (light green) until $N = 22$ (dark blue). For all plots, $J = 1, h_x = 1.05$, and $h_z = 0.5, g = 0$ ($h_z = 0, g = -1$) for the local (nonlocal) models.

for a fast scrambler. As we discussed in the context of the classical model, a necessary condition for fast scrambling is that, before the onset of exponential growth, the decay of correlations with N should be at most algebraic ($\mathcal{C} \propto N^{-\alpha}$) and not exponential. In Fig. 5(b), we verify that this is the case for the nonlocal model, showing that $\mathcal{C} \propto N^{-1}$ between the two ends of the chain after a fixed time.

Figure 5(c) shows the half-cut entanglement entropy following a quench starting from the $+\hat{y}$ state for both models. For the local model, the entanglement grows linearly in time before saturating, whereas the nonlocal model shows a significant speed up. Moreover, in the nonlocal model, the growth rate clearly increases with the system size, further supporting our claim.

Experimental realization.—The Hamiltonian in Eq. (17), and many variations of it, can be experimentally realized in a variety of platforms. A natural realization is with Rydberg dressing of neutral atoms [56–59]. The spin can be encoded in two ground states with one of them dressed to two Rydberg states such that one of the Rydberg states leads to all-to-all interactions and the second to nearest-neighbor interactions. Other similar spin models can be realized with cavity-QED setups, using photon-mediated all-to-all interactions [28,31,60,61] of the *XX* or *XXZ*-Heisenberg form [25,27] together with nearest-neighbor interactions achieved by Rydberg dressing one of the grounds states [62,63]. Other possibilities include a chain of coupled superconducting qubits, with all-to-all flip-flop interactions

mediated via a common bus [64–66] or trapped ions [33–36,67].

Conclusion and outlook.—In this Letter, we argued that a single global interaction together with local chaotic dynamics is sufficient to give rise to fast scrambling. While fast scrambling is intrinsically difficult to study numerically, our numerical evidence, together with the semiclassical analysis and the exactly solvable random circuit, provide a compelling argument in favor of our claim. Our models do not require disordered or inhomogeneous couplings and are within reach of current state-of-the-art quantum simulators. Thus, an experimental implementation of the spin model could test our claims on much larger system sizes, something that may very well be impossible to do on a classical computer. This can pave the way towards experimental investigations of aspects of quantum gravity.

Future theoretical work may include a more systematic analysis of the N dependence of various timescales, e.g., for entanglement growth, and of the behavior of the OTOC at low temperatures. It is also interesting to investigate whether similar conclusions can be reached without perfectly uniform global interactions, for example with power-law decaying interactions.

Note added.—We would like to draw the reader's attention to two related parallel works which appeared recently: by Li, Choudhury, and Liu [68], on fast scrambling with similar spin models; and by Yin and Lucas [69], on lower bounds of the scrambling time in similar spin models.

---

[1] J. M. Deutsch, Quantum statistical mechanics in a closed system, Phys. Rev. A **43**, 2046 (1991).

[2] M. Srednicki, Chaos and quantum thermalization, Phys. Rev. E **50**, 888 (1994).

[3] M. Rigol, V. Dunjko, and M. Olshanii, Thermalization and its mechanism for generic isolated quantum systems, Nature (London) 452, 854 (2008).

[4] P. Hayden and J. Preskill, Black holes as mirrors: Quantum information in random subsystems, J. High Energy Phys. 09 (2007) 120.

[5] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, Towards the fast scrambling conjecture, J. High Energy Phys. 4 (2013) 022.

[6] S. H. Shenker and D. Stanford, Black holes and the butterfly effect, J. High Energy Phys. 3 (2014) 067.

[7] J. Maldacena, S. H. Shenker, and D. Stanford, A bound on chaos, J. High Energy Phys. 8 (2016) 106.

[8] A. A. Patel, D. Chowdhury, S. Sachdev, and B. Swingle, Quantum Butterfly Effect in Weakly Interacting Diffusive Metals, Phys. Rev. X 7, 031047 (2017).

[9] S. Xu and B. Swingle, Locality, Quantum Fluctuations, and Scrambling, Phys. Rev. X 9, 031048 (2019).

[10] S. Xu and B. Swingle, Accessing scrambling using matrix product operators, Nat. Phys. 16, 199 (2020).

[11] Y. Sekino and L. Susskind, Fast scramblers, J. High Energy Phys. 10 (2008) 065.

[12] S. Sachdev and J. Ye, Gapless Spin-Fluid Ground State in a Random Quantum Heisenberg Magnet, Phys. Rev. Lett. 70, 3339 (1993).

[13] A. Kitaev, A simple model of quantum holography, in KITP Program: Entanglement in Strongly-Correlated Quantum Matter (2015), http://online.kitp.ucsb.edu/online/entangled15/kitaev/.

[14] I. Danshita, M. Hanada, and M. Tezuka, Creating and probing the Sachdev-Ye-Kitaev model with ultracold gases: Towards experimental studies of quantum gravity, Prog. Theor. Exp. Phys. 2017, 83 (2017).

[15] A. Chen, R. Ilan, F. de Juan, D. I. Pikulin, and M. Franz, Quantum Holography in a Graphene Flake with an Irregular Boundary, Phys. Rev. Lett. 121, 036403 (2018).

[16] A. Chew, A. Essin, and J. Alicea, Approximating the Sachdev-Ye-Kitaev model with Majorana wires, Phys. Rev. B 96, 121119(R) (2017).

[17] Y. Gu, X.-L. Qi, and D. Stanford, Local criticality, diffusion and chaos in generalized Sachdev-Ye-Kitaev models, J. High Energy Phys. 05 (2017) 125.

[18] J. I. Cirac and P. Zoller, Goals and opportunities in quantum simulation, Nat. Phys. 8, 264 (2012).

[19] J. Li, R. Fan, H. Wang, B. Ye, B. Zeng, H. Zhai, X. Peng, and J. Du, Measuring Out-of-Time-Order Correlators on a Nuclear Magnetic Resonance Quantum Simulator, Phys. Rev. X 7, 031011 (2017).

[20] M. Gärttner, J. G. Bohnet, A. Safavi-Naini, M. L. Wall, J. J. Bollinger, and A. M. Rey, Measuring out-of-time-order correlations and multiple quantum spectra in a trapped-ion quantum magnet, Nat. Phys. 13, 781 (2017).

[21] K. X. Wei, C. Ramanathan, and P. Cappellaro, Exploring Localization in Nuclear Spin Chains, Phys. Rev. Lett. 120, 070501 (2018).

[22] K. A. Landsman, C. Figgatt, T. Schuster, N. M. Linke, B. Yoshida, N. Y. Yao, and C. Monroe, Verified quantum information scrambling, Nature (London) 567, 61 (2019).

[23] M. K. Joshi, A. Elben, B. Vermersch, T. Brydges, C. Maier, P. Zoller, R. Blatt, and C. F. Roos, Quantum information scrambling in a trapped-ion quantum simulator with tunable range interactions, Phys. Rev. Lett. 124, 240505 (2020).

[24] M. S. Blok, V. V. Ramasesh, T. Schuster, K. O'Brien, J. M. Kreikebaum, D. Dahlen, A. Morvan, B. Yoshida, N. Y. Yao, and I. Siddiqi, Quantum Information Scrambling in a Superconducting Qutrit Processor, arXiv:2003.03307.

[25] B. Swingle, G. Bentsen, M. Schleier-Smith, and P. Hayden, Measuring the scrambling of quantum information, Phys. Rev. A 94, 040302(R) (2016).

[26] J. Marino and A. M. Rey, Cavity-QED simulator of slow and fast scrambling, Phys. Rev. A 99, 051803(R) (2019).

[27] G. Bentsen, T. Hashizume, A. S. Buyskikh, E. J. Davis, A. J. Daley, S. S. Gubser, and M. Schleier-Smith, Treelike Interactions and Fast Scrambling with Cold Atoms, Phys. Rev. Lett. 123, 130601 (2019).

[28] A. S. Sørensen and K. Mølmer, Entangling atoms in bad cavities, Phys. Rev. A 66, 022314 (2002).

[29] S. Chaudhury, S. Merkel, T. Herr, A. Silberfarb, I. H. Deutsch, and P. S. Jessen, Quantum Control of the Hyperfine Spin of a Cs Atom Ensemble, Phys. Rev. Lett. 99, 163002 (2007).

[30] T. Fernholz, H. Krauter, K. Jensen, J. F. Sherson, A. S. Sørensen, and E. S. Polzik, Spin Squeezing of Atomic Ensembles Via Nuclear-Electronic Spin Entanglement, Phys. Rev. Lett. 101, 073601 (2008).

[31] I. D. Leroux, M. H. Schleier-Smith, and V. Vuletić, Implementation of Cavity Squeezing of a Collective Atomic Spin, Phys. Rev. Lett. 104, 073602 (2010).

[32] M. H. Schleier-Smith, I. D. Leroux, and V. Vuletić, Squeezing the collective spin of a dilute atomic ensemble by cavity feedback, Phys. Rev. A 81, 021804(R) (2010).

[33] B. C. Sawyer, J. W. Britton, A. C. Keith, C.-C. J. Wang, J. K. Freericks, H. Uys, M. J. Biercuk, and J. J. Bollinger, Spectroscopy and Thermometry of Drumhead Modes in a Mesoscopic Trapped-Ion Crystal Using Entanglement, Phys. Rev. Lett. 108, 213003 (2012).

[34] J. W. Britton, B. C. Sawyer, A. C. Keith, C.-C. J. Wang, J. K. Freericks, H. Uys, M. J. Biercuk, and J. J. Bollinger, Engineered two-dimensional Ising interactions in a trapped-ion quantum simulator with hundreds of spins, Nature (London) 484, 489 (2012).

[35] C.-C. J. Wang, A. C. Keith, and J. K. Freericks, Phonon-mediated quantum spin simulator employing a planar ionic crystal in a Penning trap, Phys. Rev. A 87, 013422 (2013).

[36] J. G. Bohnet, B. C. Sawyer, J. W. Britton, M. L. Wall, A. M. Rey, M. Foss-Feig, and J. J. Bollinger, Quantum spin dynamics and entanglement generation with hundreds of trapped ions, Science 352, 1297 (2016).

[37] A. Nahum, S. Vijay, and J. Haah, Operator Spreading in Random Unitary Circuits, Phys. Rev. X 8, 021014 (2018).

[38] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems Without Conservation Laws, Phys. Rev. X 8, 021013 (2018).

[39] O. C. O. Dahlsten, R. Oliveira, and M. B. Plenio, The emergence of typical entanglement in two-party random processes, J. Phys. A 40, 8081 (2007).

[40] M. Žnidarič, Exact convergence times for generation of random bipartite entanglement, Phys. Rev. A **78**, 032324 (2008).

[41] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.125.130601 for additional details concerning the random circuit including the derivation, which includes Ref. [42].

[42] L. Zhang, Matrix integrals over unitary groups: An application of Schur-Weyl duality, arXiv:1408.3782.

[43] The term with $2l = k, w + w' - 2v = 0$ is assumed to be $0^0 = 1$. See Ref. [41].

[44] T. Zhou and X. Chen, Operator dynamics in a Brownian quantum circuit, Phys. Rev. E **99**, 052212 (2019).

[45] E. B. Rozenbaum, S. Ganeshan, and V. Galitski, Lyapunov Exponent and Out-of-Time-Ordered Correlator's Growth Rate in a Chaotic System, Phys. Rev. Lett. **118**, 086801 (2017).

[46] T. Bilitewski, S. Bhattacharjee, and R. Moessner, Temperature Dependence of the Butterfly Effect in a Classical Many-Body System, Phys. Rev. Lett. **121**, 250602 (2018).

[47] J. Chávez-Carlos, B. López-del Carpio, M. A. Bastarrachea-Magnani, P. Stránský, S. Lerma-Hernández, L. F. Santos, and J. G. Hirsch, Quantum and Classical Lyapunov Exponents in Atom-Field Interaction Systems, Phys. Rev. Lett. **122**, 024101 (2019).

[48] R. A. Jalabert, I. García-Mata, and D. A. Wisniacki, Semiclassical theory of out-of-time-order correlators for low-dimensional classically chaotic systems, Phys. Rev. E **98**, 062218 (2018).

[49] B. Yan, L. Cincio, and W. H. Zurek, Information Scrambling and Loschmidt Echo, Phys. Rev. Lett. **124**, 160603 (2020).

[50] J. Kurchan, Quantum bound to chaos and the semiclassical limit, J. Stat. Phys. **171**, 965 (2018).

[51] M. Schmitt, D. Sels, S. Kehrein, and A. Polkovnikov, Semiclassical echo dynamics in the Sachdev-Ye-Kitaev model, Phys. Rev. B **99**, 134301 (2019).

[52] T. Scaffidi and E. Altman, Chaos in a classical limit of the Sachdev-Ye-Kitaev model, Phys. Rev. B **100**, 155128 (2019).

[53] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, Distribution of the Ratio of Consecutive Level Spacings in Random Matrix Ensembles, Phys. Rev. Lett. **110**, 084101 (2013).

[54] A. Lerose, J. Marino, B. Žunkovič, A. Gambassi, and A. Silva, Chaotic Dynamical Ferromagnetic Phase Induced by Nonequilibrium Quantum Fluctuations, Phys. Rev. Lett. **120**, 130603 (2018).

[55] D. J. Luitz and Y. B. Lev, Information propagation in isolated quantum systems, Phys. Rev. B **96**, 020406(R) (2017).

[56] J. Honer, H. Weimer, T. Pfau, and H. P. Büchler, Collective Many-Body Interaction in Rydberg Dressed Atoms, Phys. Rev. Lett. **105**, 160404 (2010).

[57] T. Graß, P. Bienias, M. J. Gullans, R. Lundgren, J. Maciejko, and A. V. Gorshkov, Fractional Quantum Hall Phases of Bosons with Tunable Interactions: From the Laughlin Liquid to a Fractional Wigner Crystal, Phys. Rev. Lett. **121**, 253403 (2018).

[58] G. Pupillo, A. Micheli, M. Boninsegni, I. Lesanovsky, and P. Zoller, Strongly Correlated Gases of Rydberg-Dressed Atoms: Quantum and Classical Dynamics, Phys. Rev. Lett. **104**, 223002 (2010).

[59] A. W. Glaetzle, R. Nath, B. Zhao, G. Pupillo, and P. Zoller, Driven-dissipative dynamics of a strongly interacting Rydberg gas, Phys. Rev. A **86**, 043403 (2012).

[60] S. Gopalakrishnan, B. L. Lev, and P. M. Goldbart, Frustration and Glassiness in Spin Models with Cavity-Mediated Interactions, Phys. Rev. Lett. **107**, 277201 (2011).

[61] P. Strack and S. Sachdev, Dicke Quantum Spin Glass of Atoms and Photons, Phys. Rev. Lett. **107**, 277202 (2011).

[62] J. Gelhausen, M. Buchhold, A. Rosch, and P. Strack, Quantum-optical magnets with competing short- and long-range interactions: Rydberg-dressed spin lattice in an optical cavity, SciPost Phys. **1**, 004 (2016).

[63] B. Zhu, J. Marino, N. Y. Yao, M. D. Lukin, and E. A. Demler, Dicke time crystals in driven-dissipative quantum many-body systems, New J. Phys. **21**, 073028 (2019).

[64] J. Majer, J. M. Chow, J. M. Gambetta, J. Koch, B. R. Johnson, J. A. Schreier, L. Frunzio, D. I. Schuster, A. A. Houck, A. Wallraff, A. Blais, M. H. Devoret, S. M. Girvin, and R. J. Schoelkopf, Coupling superconducting qubits via a cavity bus, Nature (London) **449**, 443 (2007).

[65] G. Zhu, M. Hafezi, and T. Grover, Measurement of many-body chaos using a quantum clock, Phys. Rev. A **94**, 062329 (2016).

[66] T. Onodera, E. Ng, and P. L. McMahon, A quantum annealer with fully programmable all-to-all coupling via Floquet engineering, npj Quantum Inf. **6**, 48 (2020).

[67] S. Genway, W. Li, C. Ates, B. P. Lanyon, and I. Lesanovsky, Generalized Dicke Nonequilibrium Dynamics in Trapped Ions, Phys. Rev. Lett. **112**, 023603 (2014).

[68] Z. Li, S. Choudhury, and W. V. Liu, Fast scrambling without appealing to holographic duality, arXiv:2004.11269.

[69] C. Yin and A. Lucas, Bound on quantum scrambling with all-to-all interactions, Phys. Rev. A **102**, 022402 (2020).

Supplemental Material: A minimal model for fast scrambling

Ron Belyansky,[1,2] Przemyslaw Bienias,[1,2] Yaroslav A. Kharkov,[1,2] Alexey V. Gorshkov,[1,2] and Brian Swingle[1,2]

[1]*Joint Center for Quantum Information and Computer Science,*
NIST/University of Maryland, College Park, Maryland 20742 USA
[2]*Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742 USA*
(Dated: August 24, 2020)

In this Supplemental Material we present additional details concerning the random circuit. In Sec. I, we derive the general transition rate matrix W, given in Eq. (5) of the main text. In Sec. II we specialize it to the case of an initial single-site operator, deriving Eqs. (7) and (8) of the main text. In Sec. III, we present the continuum approximation for small g, deriving the Fokker-Planck equation, Eqs. (9-11) of the main text. In Sec. IV, we derive the relation between the average squared commutator and the mean operator weight. In Sec. V, we provide additional details on the dynamics and steady-state of the probability weight distribution. In Sec. VI, we derive an analytical expression for the probability weight distribution after one step of the random circuit and show that if the interactions are strong enough, the scrambling time is $O(1)$.

I. DERIVATION OF THE STOCHASTIC MATRIX W

To be slightly more general, we consider a system of N sites, each of local dimension q. As discussed in the main text, we are interested in the time evolution of a simple initial operator $\mathcal{O}(t) = U^\dagger(t)\mathcal{O}U(t)$

$$\mathcal{O}(t) = \sum_{\mathcal{S}} a_{\mathcal{S}}(t)\mathcal{S}, \tag{S1}$$

where the strings $\mathcal{S}$ form a basis for $SU(q^N)$, normalized as $\operatorname{tr}(\mathcal{S}) = q^N \delta_{\mathcal{S},1}, \operatorname{tr}(\mathcal{S}\mathcal{S}') = q^N \delta_{\mathcal{S}\mathcal{S}'}$. We take $U(t) = \prod_{i=1}^t U_i$ where $U_i = U_{\mathrm{I}}U_{\mathrm{II}}U_{\mathrm{I}}$ and U_{I} is a product of single site Haar random unitaries while U_{II} is the global interaction. Note that the two U_{I} appearing on either side of the U_{II} are different, i.e the random unitaries are random in both space in time. Here we inserted an additional layer of the Haar unitaries, as compared to the main text. This is completely equivalent, as this extra layer can always be absorbed into the Haar layer of either the step before or the step after, but it simplifies calculations.

Using $a_{\mathcal{S}}(t) = q^{-N}\operatorname{tr}(\mathcal{O}(t)\mathcal{S})$, we can write $a_{\mathcal{S}}^2(t)$ in terms of the coefficients at the previous time step

$$a_{\mathcal{S}}^2(t) = q^{-2N} \sum_{\mathcal{S}',\mathcal{S}''} a_{\mathcal{S}'}(t-1)a_{\mathcal{S}''}(t-1)\operatorname{tr}(U^\dagger\mathcal{S}'U\mathcal{S})\operatorname{tr}(U^\dagger\mathcal{S}''U\mathcal{S}). \tag{S2}$$

Thus, we want to evaluate the quantity

$$\left\langle \operatorname{tr}(U^\dagger\mathcal{S}'U\mathcal{S})\operatorname{tr}(U^\dagger\mathcal{S}''U\mathcal{S}) \right\rangle, \tag{S3}$$

where $\langle ... \rangle$ denotes Haar average over the random unitaries.

Using properties of trace, we can write

$$\left\langle \operatorname{tr}(U^\dagger\mathcal{S}'U\mathcal{S})\operatorname{tr}(U^\dagger\mathcal{S}''U\mathcal{S}) \right\rangle = \left\langle \operatorname{tr}(U^\dagger\mathcal{S}'U\mathcal{S} \otimes U^\dagger\mathcal{S}''U\mathcal{S}) \right\rangle. \tag{S4}$$

In doing so, we now have a trace over two copies of the system, which could still be thought as a N-site system, where every site is now of dimension q^2 instead of q. In the following, we will denote operators acting on the right system by an overbar. For example $Z_i\bar{Z}_i$ corresponds to the Pauli Z operator acting on site i of both copies, i.e $Z_i \otimes Z_i$.

For our choice of U, Eq. (S4) becomes

$$\left\langle \operatorname{tr}(U^\dagger\mathcal{S}'U\mathcal{S})\operatorname{tr}(U^\dagger\mathcal{S}''U\mathcal{S}) \right\rangle = \operatorname{tr}\left(\left\langle (U_{\mathrm{I}} \otimes U_{\mathrm{I}})(U_{\mathrm{II}} \otimes U_{\mathrm{II}}) \left\langle U_{\mathrm{I}}\mathcal{S}'U_{\mathrm{I}}^\dagger \otimes U_{\mathrm{I}}\mathcal{S}''U_{\mathrm{I}}^\dagger \right\rangle (U_{\mathrm{II}} \otimes U_{\mathrm{II}})^\dagger (U_{\mathrm{I}} \otimes U_{\mathrm{I}})^\dagger \right\rangle (\mathcal{S} \otimes \mathcal{S}) \right). \tag{S5}$$

We will calculate the above in several steps, working from inside out

$$\mathcal{I}_1 = \left\langle U_{\mathrm{I}}\mathcal{S}'U_{\mathrm{I}}^\dagger \otimes U_{\mathrm{I}}\mathcal{S}''U_{\mathrm{I}}^\dagger \right\rangle, \tag{S6}$$

$$\mathcal{I}_2 = (U_{\mathrm{II}} \otimes U_{\mathrm{II}})\mathcal{I}_1(U_{\mathrm{II}} \otimes U_{\mathrm{II}})^\dagger, \tag{S7}$$

$$\mathcal{I}_3 = \left\langle (U_{\mathrm{I}} \otimes U_{\mathrm{I}})\mathcal{I}_2(U_{\mathrm{I}} \otimes U_{\mathrm{I}})^\dagger \right\rangle, \tag{S8}$$

with $\mathrm{tr}(\mathcal{I}_3(\mathcal{S} \otimes \mathcal{S}))$ being our quantity of interest.

Before proceeding, let us introduce an important formula for calculating the Haar averages. Consider a $d^2 \times d^2$ matrix A, and a $d \times d$ Haar random unitary matrix U. Then, we have the following formula [S1, S2]

$$
\begin{aligned}
\langle (U \otimes U)A(U \otimes U)^\dagger \rangle &\equiv \int_{U(d)} (U \otimes U)A(U \otimes U)^\dagger d\mu(U) \\
&= \left(\frac{\mathrm{tr}(A)}{d^2-1} - \frac{\mathrm{tr}(AF)}{d(d^2-1)} \right) \mathbb{1}_{d^2} - \left(\frac{\mathrm{tr}(A)}{d(d^2-1)} - \frac{\mathrm{tr}(AF)}{d^2-1} \right) F,
\end{aligned}
\tag{S9}
$$

where $F = \sum_{ij} |ij\rangle \langle ji|$ is the swap operator.

From this, it follows that

$$
\mathcal{I}_1 = \prod_r \langle U_r^\dagger \mathcal{S}_r' U_r \otimes U_r^\dagger \mathcal{S}_r'' U_r \rangle = \delta_{\mathcal{S}',\mathcal{S}''} \prod_r \left(\frac{q^2 \delta_{\mathcal{S}_r',1} - 1}{q^2-1} \mathbb{1}_{q^2} + \frac{q - q\delta_{\mathcal{S}_r',1}}{q^2-1} F_r \right)
\tag{S10}
$$

where we used $\mathrm{tr}(\mathcal{S}_r) = q\delta_{\mathcal{S}_r,1}$ and $\mathrm{tr}(\mathcal{S}_r\mathcal{S}_r') = q\delta_{\mathcal{S}_r,\mathcal{S}_r'}$. Here F_r swaps site r of the left system with the corresponding site r of the right system.

The overall delta function $\delta_{\mathcal{S}',\mathcal{S}''}$ immediately implies that the Haar average of Eq. (S2) may be written as

$$
\langle a_\mathcal{S}^2(t+1) \rangle = \sum_{\mathcal{S}'} W_{\mathcal{S},\mathcal{S}'} \langle a_{\mathcal{S}'}^2(t) \rangle,
\tag{S11}
$$

where $W_{\mathcal{S},\mathcal{S}'} = q^{-2N} \mathrm{tr}(\mathcal{I}_3(\mathcal{S} \otimes \mathcal{S}))$.

To proceed, we specialize to qubits, i.e. $q = 2$, in which case the swap operator can be written as $F_r = \frac{1}{2}(\mathbb{1}_r \otimes \bar{\mathbb{1}}_r + \boldsymbol{\sigma}_r \cdot \bar{\boldsymbol{\sigma}}_r) = \frac{1}{2}(\mathbb{1}_r\bar{\mathbb{1}}_r + X_r\bar{X}_r + Y_r\bar{Y}_r + Z_r\bar{Z}_r)$ where bar denotes operators acting on the second system. We can combine all the $\mathbb{1}$s together, giving

$$
\begin{aligned}
\mathcal{I}_1 &= \delta_{\mathcal{S}',\mathcal{S}''} \prod_i \left(\delta_{\mathcal{S}_i',1} \mathbb{1}_{2^2} + \frac{1 - \delta_{\mathcal{S}_i',1}}{3} \boldsymbol{\sigma}_i \cdot \bar{\boldsymbol{\sigma}}_i \right) \\
&= \delta_{\mathcal{S}',\mathcal{S}''} \sum_{\Omega_{\mathcal{S}'} \subset \{1,2,\cdots,N\}} \prod_{i \in \{1,2,\cdots,N\}/\Omega_{\mathcal{S}'}} \delta_{\mathcal{S}_i',1} \mathbb{1}_4 \prod_{j \in \Omega_{\mathcal{S}'}} \frac{1 - \delta_{\mathcal{S}_j',1}}{3} \boldsymbol{\sigma}_j \cdot \bar{\boldsymbol{\sigma}}_j,
\end{aligned}
\tag{S12}
$$

where in the second equality the sum is over the powerset of $\{1,2,\cdots,N\}$, i.e, all the (2^N) subsets of $\{1,2,\cdots,N\}$. The sum above essentially contains every possible string of the form $\mathcal{S} \otimes \mathcal{S}$, i.e the same operator appears on both copies of the system. Note that for a given string $\mathcal{S}'$, there is only one nonzero term in the sum. For each site i, we either put an $\mathbb{1}_4$ if $\mathcal{S}_i' = \mathbb{1}$ or we place $\frac{1}{3}\boldsymbol{\sigma}_i \cdot \bar{\boldsymbol{\sigma}}_i$, if $\mathcal{S}_i'$ is any other generator. The set $\Omega_{\mathcal{S}'}$ therefore represents the support of the string $\mathcal{S}'$.

Before proceeding, let us summarize the high-level idea behind the derivation that follows. Our tasks consist of the following:

1. First, we need to apply the global interaction $U_{\mathrm{II}} \otimes U_{\mathrm{II}}$ on Eq. (S12), giving us $\mathcal{I}_2$.

2. Then, we need to apply the layer of single-site Haar unitaries $U_{\mathrm{I}} \otimes U_{\mathrm{I}}$, and average over the Haar distribution on each site, giving us $\mathcal{I}_3$.

3. Finally, we need to multiply the result by $\mathcal{S} \otimes \mathcal{S}$ and take the trace, giving us $W_{\mathcal{S},\mathcal{S}'}$.

Recall that

$$
U_{\mathrm{II}} = e^{-i\frac{g'}{2} \sum_{i<j} Z_i Z_j},
\tag{S13}
$$

where in the main text we have assumed $g' = \frac{g}{\sqrt{N}}$. To perform the first step, we will make use of the formulas

$$
U_{\mathrm{II}} X_r U_{\mathrm{II}}^\dagger = X_r \cos\left(g' \sum_{i \neq r} Z_i \right) + Y_r \sin\left(g' \sum_{i \neq r} Z_i \right),
\tag{S14}
$$

$$
U_{\mathrm{II}} Y_r U_{\mathrm{II}}^\dagger = Y_r \cos\left(g' \sum_{i \neq r} Z_i \right) - X_r \sin\left(g' \sum_{i \neq r} Z_i \right).
\tag{S15}
$$

Now, note that each term in the sum in Eq. (S12) is a product of single-site operators. By performing our first task, using Eqs. (S14) and (S15), we will obtain complicated operators, like those appearing on the right-hand-side of Eqs. (S14) and (S15), that are supported on a large number of sites. In order to perform the second step, we can make use of Eq. (S9). However, to use Eq. (S9), we need A to be a single-site operator. Thus, we will have to break down complicated operators, like those on the right-hand-side of Eqs. (S14) and (S15), into sums of simple terms consisting of products of single-site operators. This will allow us to use Eq. (S9), after which we can easily perform the last step, 3, since this will only require taking traces of single-site operators.

The result of step 1 and 2 can be simplified by noting that Eq. (S12) contains all possible strings of the form $\mathcal{S} \otimes \mathcal{S}$. Hence, it is instructive to first consider the result of applying $U_{\mathrm{II}} \otimes U_{\mathrm{II}}$ and $U_{\mathrm{I}} \otimes U_{\mathrm{I}}$ to a single string of this form. Note that the result of applying $U_{\mathrm{II}} \otimes U_{\mathrm{II}}$, $U_{\mathrm{I}} \otimes U_{\mathrm{I}}$, and averaging over the Haar unitaries is invariant if we replace any number of Xs in the string by Ys or vice-versa. To see this, we use the fact that we can change a X into a Y (or vice-versa) by applying a rotation about the Z axis, i.e $e^{-i\frac{\pi}{4}Z}Xe^{i\frac{\pi}{4}Z} = Y$. This rotation clearly commutes with U_{II} and can be absorbed into U_{I}, since by definition, the Haar measure is invariant under multiplication by any unitary.

This means that we may calculate the result for a single representative string from each group and multiply by the degeneracy. Let us denote $\Omega_{\mathcal{S}}$ the support of some string $\mathcal{S}$. We can further divide $\Omega_{\mathcal{S}}$ based on the number and location of Zs in the string. Define the subset $\Sigma \subseteq \Omega$ as the set of all sites with Z in them, and the remaining sites (with either Xs or Ys) by $\Lambda = \Omega \setminus \Sigma$. For strings that are supported on k sites (i.e $|\Omega_{\mathcal{S}}| = k$), with fixed number and position of Zs, the degeneracy is $2^{|\Lambda|}$.

Without loss of generality, we can therefore consider strings composed of either Xs or Zs. Consider the string $\prod_{i \in \Lambda} X_i \bar{X}_i \prod_{j \in \Sigma} Z_j \bar{Z}_j$. To apply U_{II}, we can use the fact that $[X_i X_j, Z_i Z_j] = 0$. We get

$$(U_{\mathrm{II}} \otimes U_{\mathrm{II}})(\prod_{i \in \Lambda} X_i \bar{X}_i \prod_{j \in \Sigma} Z_j \bar{Z}_j)(U_{\mathrm{II}} \otimes U_{\mathrm{II}})^{\dagger} = \prod_{i \in \Lambda} \left[(X_i \cos(Q_\Lambda) + Y_i \sin(Q_\Lambda))(\bar{X}_i \cos(\bar{Q}_\Lambda) + \bar{Y}_i \sin(\bar{Q}_\Lambda)) \right] \prod_{j \in \Sigma} Z_j \bar{Z}_j$$

(S16)

where we used Eq. (S14). Here, Q_Λ acts on all sites except those in Λ, i.e $Q_\Lambda \equiv g' \sum_{l \notin \Lambda} Z_l$.

We see that we can safely apply the Haar unitaries and perform the Haar average on sites inside of Λ, since all the cosines and sines and the $Z\bar{Z}$ act on sites outside of Λ. With slight abuse of notation, let us denote $\langle (U_{\mathrm{I}} \otimes U_{\mathrm{I}})A(U_{\mathrm{I}} \otimes U_{\mathrm{I}})^{\dagger} \rangle$ by simply $\langle A \rangle$ where it is understood that the Haar unitaries act only on the support of A.

From Eq. (S9), one can easily check that $\langle X_i \bar{Y}_i \rangle = 0$, so the cross terms in the above expression will vanish. Only $\langle X_i \bar{X}_i \rangle = \langle Y_i \bar{Y}_i \rangle \equiv V_i$ will remain. Here the single site operator V_i is defined as $V_j = -\frac{1}{3}\mathbb{1}_4 + \frac{2}{3}F$. Explicitly, we find

$$\left\langle (U_{\mathrm{II}} \otimes U_{\mathrm{II}})(\prod_{i \in \Lambda} X_i \bar{X}_i \prod_{j \in \Sigma} Z_j \bar{Z}_j)(U_{\mathrm{II}} \otimes U_{\mathrm{II}})^{\dagger} \right\rangle = \prod_{i \in \Lambda} V_i \left\langle \cos^{|\Lambda|}(R_\Lambda) \prod_{j \in \Sigma} Z_j \bar{Z}_j \right\rangle$$

(S17)

where $R_\Lambda = \bar{Q}_\Lambda - Q_\Lambda$.

Combining this with the discussion above, we find that $\mathcal{I}_3$ may be written as

$$\mathcal{I}_3 = \delta_{\mathcal{S}',\mathcal{S}''} \sum_{\Omega_{\mathcal{S}'} \subset \{1,2,\cdots,N\}} \left(\prod_{j \notin \Omega_{\mathcal{S}'}} \delta_{\mathcal{S}'_j,1} \right) \left(\prod_{i \in \Omega_{\mathcal{S}'}} \frac{1 - \delta_{\mathcal{S}'_i,1}}{3} \right) \sum_{\Lambda \subset \Omega_{\mathcal{S}'}} 2^{|\Lambda|} \left(\prod_{m \in \Lambda} V_m \right) \left\langle \cos^{|\Lambda|}(R_\Lambda) \prod_{n \in \Omega_{\mathcal{S}'} \setminus \Lambda} Z_n \bar{Z}_n \right\rangle.$$

(S18)

It remains to compute $\left\langle \cos^{|\Lambda|}(R_\Lambda) \prod_{n \in \Omega_{\mathcal{S}'} \setminus \Lambda} Z_n \bar{Z}_n \right\rangle$. To do so we expand the cosine as follows $\cos^k(x) = \frac{1}{2^k} \sum_{n=0}^{k} \binom{k}{n} \cos[(2n-k)x]$,

$$\left\langle \cos^{|\Lambda|}(R_\Lambda) \prod_{n \in \Omega_{\mathcal{S}'} \setminus \Lambda} Z_n \bar{Z}_n \right\rangle = \frac{1}{2^{|\Lambda|}} \sum_{l=0}^{|\Lambda|} \binom{|\Lambda|}{l} \left\langle \cos((2l - |\Lambda|)R_\Lambda) \prod_{n \in \Omega_{\mathcal{S}'} \setminus \Lambda} Z_n \bar{Z}_n \right\rangle.$$

(S19)

To proceed we can pull a single-site operator out of R_Λ. Since $R_\Lambda = \sum_{k \notin \Lambda} D_k$ where $D_k = g'(\bar{Z}_k - Z_k)$, we can pull out a D_j, $j \in \Omega_{\mathcal{S}'} \setminus \Lambda$ so that $R_\Lambda = R_{\Lambda \cup \{j\}} + D_j$. We then use the trig identity

$$\cos((2l - |\Lambda|)R_\Lambda) = \cos((2l - |\Lambda|)R_{\Lambda \cup \{j\}})\cos((2l - |\Lambda|)D_j) - \sin((2l - |\Lambda|)R_{\Lambda \cup \{j\}})\sin((2l - |\Lambda|)D_j).$$

(S20)

This allows us to perform the Haar average over site j. The sine term will not contribute, since $\left\langle \sin((2l-|\Lambda|)D_j)Z_j\bar{Z}_j \right\rangle = 0$. Repeating this procedure recursively for all sites in $\Omega_{\mathcal{S}'} \setminus \Lambda$, we get

$$\left\langle \cos((2l-|\Lambda|)R_\Lambda) \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} Z_n\bar{Z}_n \right\rangle = \left\langle \cos((2l-|\Lambda|)R_{\Omega_{\mathcal{S}'}}) \right\rangle \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} \left\langle \cos((2l-|\Lambda|)D_n)Z_n\bar{Z}_n \right\rangle. \tag{S21}$$

Continuing the procedure for the $\left\langle \cos((2l-|\Lambda|)R_{\Omega_{\mathcal{S}'}}) \right\rangle$ term, we have

$$\left\langle \cos((2n-|\Lambda|)R_\Lambda) \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} Z_n\bar{Z}_n \right\rangle = \prod_{t\notin\Omega_{\mathcal{S}'}} \left\langle \cos((2l-|\Lambda|)D_t) \right\rangle \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} \left\langle \cos((2l-|\Lambda|)D_n)Z_n\bar{Z}_n \right\rangle. \tag{S22}$$

Using $\cos((2l-|\Lambda|)D) = \cos^2((2l-|\Lambda|)g') + Z\bar{Z}\sin^2((2l-|\Lambda|)g')$ gives

$$\left\langle \cos((2n-|\Lambda|)R_\Lambda) \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} Z_n\bar{Z}_n \right\rangle$$
$$= \prod_{t\notin\Omega_{\mathcal{S}'}} \left(\cos^2((2l-|\Lambda|)g') + V_t\sin^2((2l-|\Lambda|)g')\right) \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} \left(\cos^2((2l-|\Lambda|)g')V_n + \sin^2((2l-|\Lambda|)g')\right). \tag{S23}$$

Putting things together, we find that Eq. (S19) is

$$\left\langle \cos^{|\Lambda|}(R_\Lambda) \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} Z_n\bar{Z}_n \right\rangle$$
$$= \frac{1}{2^{|\Lambda|}} \sum_{l=0}^{|\Lambda|} \binom{|\Lambda|}{l} \prod_{t\notin\Omega_{\mathcal{S}'}} \left(\cos^2((2l-|\Lambda|)g') + V_t\sin^2((2l-|\Lambda|)g')\right) \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} \left(\cos^2((2l-|\Lambda|)g')V_n + \sin^2((2l-|\Lambda|)g')\right), \tag{S24}$$

and finally, $\mathcal{I}_3$ is given by

$$\mathcal{I}_3 = \delta_{\mathcal{S}',\mathcal{S}''} \sum_{\Omega_{\mathcal{S}'}\subset\{1,2,\cdots,N\}} \left(\prod_{j\notin\Omega_{\mathcal{S}'}} \delta_{\mathcal{S}'_j,1}\right) \left(\prod_{i\in\Omega_{\mathcal{S}'}} \frac{1-\delta_{\mathcal{S}'_i,1}}{3}\right) \sum_{\Lambda\subset\Omega_{\mathcal{S}'}} \left(\prod_{m\in\Lambda} V_m\right) \sum_{l=0}^{|\Lambda|} \binom{|\Lambda|}{l}$$
$$\times \prod_{t\notin\Omega_{\mathcal{S}'}} \left(\cos^2((2l-|\Lambda|)g') + V_t\sin^2((2l-|\Lambda|)g')\right) \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} \left(\cos^2((2l-|\Lambda|)g')V_n + \sin^2((2l-|\Lambda|)g')\right). \tag{S25}$$

To compute the $W_{\mathcal{S},\mathcal{S}'}$-matrix from Eq. (S11), it remains to take the trace of Eq. (S25) with $\mathcal{S}\otimes\mathcal{S}$ and divide by 2^{2N}, i.e

$$W_{\mathcal{S},\mathcal{S}'} = \frac{1}{2^{2N}} \operatorname{tr}(\mathcal{I}_3(\mathcal{S}\otimes\mathcal{S})) \tag{S26}$$

Using Eq. (S25) together with $\operatorname{tr}(V_i(\mathcal{S}_i\otimes\bar{\mathcal{S}}_i)) = \frac{4}{3}(1-\delta_{\mathcal{S}_i,1})$, gives

$$W_{\mathcal{S},\mathcal{S}'} = \frac{1}{2^{2N}} \sum_{\Omega_{\mathcal{S}'}\subset\{1,2,\cdots,N\}} \left(\prod_{j\notin\Omega_{\mathcal{S}'}} \delta_{\mathcal{S}'_j,1}\right) \left(\prod_{i\in\Omega_{\mathcal{S}'}} \frac{1-\delta_{\mathcal{S}'_i,1}}{3}\right) \sum_{\Lambda\subset\Omega_{\mathcal{S}'}} \left(\prod_{m\in\Lambda} \frac{4}{3}(1-\delta_{\mathcal{S}_m,1})\right) \sum_{l=0}^{|\Lambda|} \binom{|\Lambda|}{l}$$
$$\times \prod_{t\notin\Omega_{\mathcal{S}'}} \left(\cos^2((2l-|\Lambda|)g')4\delta_{\mathcal{S}_t,1} + \frac{4}{3}(1-\delta_{\mathcal{S}_t,1})\sin^2((2l-|\Lambda|)g')\right)$$
$$\times \prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda} \left(\cos^2((2l-|\Lambda|)g')\frac{4}{3}(1-\delta_{\mathcal{S}_n,1}) + 4\delta_{\mathcal{S}_n,1}\sin^2((2l-|\Lambda|)g')\right). \tag{S27}$$

Note that because of $\prod_{m\in\Lambda}\frac{4}{3}(1-\delta_{\mathcal{S}_m,1})$ in Eq. (S27), Λ is constrained to be in $\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}$. The matrix elements of W are

$$
W = \frac{1}{2^{2N}}\left(\frac{1}{3}\right)^{|\Omega_{\mathcal{S}'}|}\sum_{\Lambda\subset\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}}\left(\frac{4}{3}\right)^{|\Lambda|}\left[\sum_{l=0,2l\neq|\Lambda|}^{|\Lambda|}\binom{|\Lambda|}{l}\left(4\cos^2((2l-|\Lambda|)g')\right)^{N-|\Omega_{\mathcal{S}}\cup\Omega_{\mathcal{S}'}|}\right.
$$

$$
\times\left(\frac{4}{3}\sin^2((2l-|\Lambda|)g')\right)^{|\Omega_{\mathcal{S}}\backslash\Omega_{\mathcal{S}'}|}\times\left(\frac{4}{3}\cos^2((2l-|\Lambda|)g')\right)^{|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|-|\Lambda|}\times\left(4\sin^2((2l-|\Lambda|)g')\right)^{|\Omega_{\mathcal{S}'}\backslash\Omega_{\mathcal{S}}|}
$$

$$
\left.+\delta_{2l,|\Lambda|}\binom{|\Lambda|}{|\Lambda|/2}\prod_{t\notin\Omega_{\mathcal{S}'}}(4\delta_{\mathcal{S}_t,1})\prod_{n\in\Omega_{\mathcal{S}'}\backslash\Lambda}\left(\frac{4}{3}(1-\delta_{\mathcal{S}_n,1})\right)\right]. \quad\text{(S28)}
$$

Note that the last term is only nonzero when both $2l = |\Lambda|$ and $\Omega_{\mathcal{S}} = \Omega_{\mathcal{S}'}$. The last condition is equivalent to $|\Omega_{\mathcal{S}}| + |\Omega_{\mathcal{S}'}| - 2|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}| = 0$.

We can combine all constant factors (with the same result holding for the $2l = |\Lambda|$ term)

$$
\frac{1}{2^{2N}}\left(\frac{1}{3}\right)^{|\Omega_{\mathcal{S}'}|}\left(\frac{4}{3}\right)^{|\Lambda|}4^{N-|\Omega_{\mathcal{S}}\cup\Omega_{\mathcal{S}'}|}\left(\frac{4}{3}\right)^{|\Omega_{\mathcal{S}}\backslash\Omega_{\mathcal{S}'}|}\left(\frac{4}{3}\right)^{|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|-|\Lambda|}4^{|\Omega_{\mathcal{S}'}\backslash\Omega_{\mathcal{S}}|}=\left(\frac{1}{3}\right)^{|\Omega_{\mathcal{S}'}|+|\Omega_{\mathcal{S}}|}. \quad\text{(S29)}
$$

Now, note that Λ only appears in Eq. (S28) as $|\Lambda|$. Thus, we can replace the sum over subsets of $\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}$ as $\sum_{\Lambda\subset\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}} = \sum_{k=0}^{|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|}\binom{|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|}{k}$. Thus, the W matrix can be written as

$$
W_{\mathcal{S},\mathcal{S}'} = W(|\Omega_{\mathcal{S}}|,|\Omega_{\mathcal{S}'}|,|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|) \quad\text{(S30)}
$$

$$
=\left(\frac{1}{3}\right)^{|\Omega_{\mathcal{S}'}|+|\Omega_{\mathcal{S}}|}\sum_{k=0}^{|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|}\binom{|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|}{k}\left[\sum_{l=0,2l\neq k}^{k}\binom{k}{l}\left[\cos^2((2l-k)g')\right]^{N-k-(|\Omega_{\mathcal{S}}|+|\Omega_{\mathcal{S}'}|-2|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|)}\right.
$$

$$
\times\left.\left[\sin^2((2l-k)g')\right]^{|\Omega_{\mathcal{S}}|+|\Omega_{\mathcal{S}'}|-2|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|}+\delta_{2l,k}\delta_{|\Omega_{\mathcal{S}}|+|\Omega_{\mathcal{S}'}|-2|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|,0}\binom{k}{k/2}\right],
$$

which is what appears in Eq. (5) of the main text, with the identification $w = |\Omega_{\mathcal{S}}|, w' = |\Omega_{\mathcal{S}'}|, v = |\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|$. In the main text, we also dropped the $\delta_{2l,k}\delta_{|\Omega_{\mathcal{S}}|+|\Omega_{\mathcal{S}'}|-2|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|,0}$ term and the $2l\neq k$ restriction in the sum which requires one to be careful to identify 0^0 as 1. From this expression it is clear that W is a real symmetric ($W_{\mathcal{S},\mathcal{S}'} = W_{\mathcal{S}',\mathcal{S}}$) matrix with all positive matrix elements.

II. MASTER EQUATION FOR SIMPLE INITIAL OPERATOR

Let us now assume that the initial operator $\mathcal{O}$ starts as a single-site operator on site 1 without loss of generality. We may further assume that we start with X_1, i.e $a_{\mathcal{S}} = \delta_{\mathcal{S},X_1}$. Since the circuit will involve random Haar unitaries, let us consider the result of applying a Haar random unitary on X_1, which, after averaging over the Haar unitary, will be $\frac{1}{3}(X_1 + Y_1 + Z_1)$, which already does not contain any information about the specific generator we picked. Let us therefore pick this as the initial conditions at $t = 0$ for the master equation, Eq. (S11),

$$
\langle a_{\mathcal{S}}^2(t=0)\rangle = \begin{cases} \frac{1}{3} & \text{if } \mathcal{S} = X_1, Y_1, Z_1, \\ 0 & \text{otherwise} \end{cases}. \quad\text{(S31)}
$$

We now claim that for these initial conditions, the probabilities $\langle a_{\mathcal{S}}^2(t)\rangle$ only depend on the string weight $w \equiv |\Omega_{\mathcal{S}}|$ and the weight on site 1, $w_1 \equiv |\Omega_{\mathcal{S}}\cap\{1\}|$. Note that w_1 takes values either 0 or 1. In light of this, it is convenient to account for the number of string configurations with constant w and w_1 by defining the operator weight probability h_t,

$$
h_t(w,w_1) = \langle a_{\mathcal{S}}^2(t)\rangle D(w,w_1), \quad\text{(S32)}
$$

where $D(w,w_1)$ is the number of string configurations for a given w and w_1. Since $\sum_{\mathcal{S}'} = \sum_{w_1=0,1}\sum_{w=w_1}^{N-1+w_1}3^k\binom{N-1}{w-w_1}$, we have

$$
D(w,w_1) = 3^w\binom{N-1}{w-w_1}. \quad\text{(S33)}
$$

Note that $h_t(w, w_1)$ is a valid (normalized) probability distribution since $\sum_{w_1=0,1} \sum_{w=w_1}^{N-1+w_1} h_t(w, w_1) = \sum_{w_1=0,1} \sum_{w=w_1}^{N-1+w_1} \langle a_{\mathcal{S}}^2(t) \rangle D(w, w_1) = \sum_{\mathcal{S}} \langle a_{\mathcal{S}}^2(t) \rangle = 1$, using the fact that $a_{\mathcal{S}}^2$ are probabilities that sum to 1. Thus, $h_t(w, w_1)$ gives the probability of $\mathcal{O}(t)$ being a string of total weight w with a weight of w_1 on the initial site 1.

The claim above can be proved by induction. The base case is trivial to see, by multiplying the initial conditions Eq. (S31) by the transition matrix W from Eq. (S30) (see also Sec. VI). The inductive step proceeds as follows. First, we decompose the sum over strings $\mathcal{S}'$ as $\sum_{\mathcal{S}'} = \sum_{\Omega_{\mathcal{S}'} \subset \{1, \cdots, N\}} 3^{|\Omega_{\mathcal{S}'}|}$, which yields

$$\langle a_{\mathcal{S}}^2(t+1) \rangle = \sum_{\Omega_{\mathcal{S}'} \subset \{1, \cdots, N\}} \frac{1}{D(|\Omega_{\mathcal{S}'}|, |\Omega_{\mathcal{S}'} \cap 1|)} 3^{|\Omega_{\mathcal{S}'}|} W(|\Omega_{\mathcal{S}}|, |\Omega_{\mathcal{S}'}|, |\Omega_{\mathcal{S}} \cap \Omega_{\mathcal{S}'}|) h_t(|\Omega_{\mathcal{S}'}|, |\Omega_{\mathcal{S}'} \cap 1|). \qquad (S34)$$

We then split the sum over terms where $|\Omega_{\mathcal{S}'} \cap \{1\}| = 0$ or $|\Omega_{\mathcal{S}'} \cap \{1\}| = 1$. For each of these terms, we further decompose the sum over terms with equal $|\Omega_{\mathcal{S}'}|$. The remaining sum can be written as a sum over different values of the overlap $|\Omega_{\mathcal{S}} \cap \Omega_{\mathcal{S}'}|$. The final result is

$$\langle a_{\mathcal{S}}^2(t+1) \rangle$$
$$= \sum_{k=0}^{N-1} 3^k \left[\sum_{m=0}^{\min\{|\Omega_{\mathcal{S}}| - |\Omega_{\mathcal{S}} \cap \{1\}|, k\}} \binom{|\Omega_{\mathcal{S}}| - |\Omega_{\mathcal{S}} \cap \{1\}|}{m} \binom{N-1 - |\Omega_{\mathcal{S}}| + |\Omega_{\mathcal{S}} \cap \{1\}|}{k-m} W(|\Omega_{\mathcal{S}}|, k, m) \right] \frac{h_t(k, 0)}{D(k, 0)}$$
$$+ \sum_{k=1}^{N} 3^k \left[\sum_{m=|\Omega_{\mathcal{S}} \cap \{1\}|}^{\min\{|\Omega_{\mathcal{S}}|, k-1+|\Omega_{\mathcal{S}} \cap \{1\}|\}} \binom{|\Omega_{\mathcal{S}}| - |\Omega_{\mathcal{S}} \cap \{1\}|}{m - |\Omega_{\mathcal{S}} \cap \{1\}|} \binom{N-1 + |\Omega_{\mathcal{S}} \cap \{1\}| - |\Omega_{\mathcal{S}}|}{k-m-1+|\Omega_{\mathcal{S}} \cap \{1\}|} W(|\Omega_{\mathcal{S}}|, k, m) \right] \frac{h_t(k, 1)}{D(k, 1)}.$$
$$(S35)$$

Here, the first binomial in each bracket counts the number of ways one can choose the part of $\Omega_{\mathcal{S}'}$ that is overlapping with $\Omega_{\mathcal{S}}$ and the second binomial counts the number of ways to choose the non-overlapping part of $\Omega_{\mathcal{S}'}$. It is clear at this point that the right-hand-side is a function of $w = |\Omega_{\mathcal{S}}|$ and $w_1 = |\Omega_{\mathcal{S}} \cap 1|$. Thus, replacing $\langle a_{\mathcal{S}}^2(t+1) \rangle$ by Eq. (S32) and simplifying gives

$$h_{t+1}(w, w_1) = \sum_{w_1'=0,1} \sum_{w'=w_1'}^{N-1+w_1'} \mathcal{R}(w, w_1, w', w_1') h_t(w', w_1') \qquad (S36)$$

where the $2N \times 2N$ matrix $\mathcal{R}$ is

$$\mathcal{R}(w, w_1, w', w_1') = 3^w \sum_{m=\max\{0, w+w'-N+1-w_1-w_1'\}}^{\min\{w-w_1, w'-w_1'\}} \binom{w' - w_1'}{m} \binom{N-1-w'+w_1'}{w-w_1-m} W(w, w', m + w_1 w_1'), \qquad (S37)$$

where $w_1, w_1' \in \{0, 1\}, w \in [w_1, N-1+w_1], w' \in [w_1', N-1+w_1']$, and for completeness

$$W(w, w', v) = \left(\frac{1}{3}\right)^{w+w'} \sum_{k=0}^{v} \binom{v}{k} \left[\sum_{l=0, 2l \neq k}^{k} \binom{k}{l} \left[\cos^2((2l-k)g') \right]^{N-k-(w+w'-2v)} \right.$$
$$\left. \times \left[\sin^2((2l-k)g') \right]^{w+w'-2v} + \delta_{2l,k} \delta_{w+w'-2v, 0} \binom{k}{k/2} \right]. \qquad (S38)$$

One may verify that $\sum_i \mathcal{R}_{i,j} = 1$ where $i = (w, w_1)$ and $j = (w', w_1')$. This means that if we start with normalized h_0, we will have a valid (normalized) probability distribution at later times.

The initial conditions become

$$h_0(w, w_1) = \begin{cases} 1 & \text{if } w = w_1 = 1, \\ 0 & \text{otherwise} . \end{cases} \qquad (S39)$$

To get the probability of having a specific weight, we can sum over w_1,

$$h(w) = \begin{cases} h(0,0) & \text{if } w = 0, \\ h(N, 1) & \text{if } w = N, \\ h(w, 0) + h(w, 1) & \text{otherwise} . \end{cases} \qquad (S40)$$

Note that $h(0,0)$ does not actually participate in the dynamics since $\mathcal{R}(0, 0, w', w_1') = W(0, w', 0) = \delta_{w', 0}$.

III.   CONTINUUM APPROXIMATION

We assume here the normalization $g' = \frac{g}{\sqrt{N}}$. The first step is to approximate $W(w, w', v)$ for small g. We consider the two cases $w + w' - 2v = 0, 1$ which amount to a change of the string weight by $0, \pm 1$ and give rise to terms up to g^2.

Taylor expanding the factors of cosine and sine appearing in Eq. (S38), up to g^2, gives

$$\left[\cos^2\left((2l - k)\frac{g}{\sqrt{N}}\right)\right]^{N-k-(w+w'-2v)} \left[\sin^2\left((2l - k)\frac{g}{\sqrt{N}}\right)\right]^{w+w'-2v} \approx \begin{cases} \frac{g^2(k-2l)^2(k-N)}{N} + 1 & \text{if} \quad w + w' - 2v = 0, \\ \frac{g^2(k-2l)^2}{N} & \text{if} \quad w + w' - 2v = 1. \end{cases} \tag{S41}$$

In general, the $w + w' - 2v = n$, $n \in \mathbb{N}_{>0}$ case will scale as $O(g^{2n})$. We can now perform the sums over l and k appearing in Eq. (S38). We find

$$W(w, w', v) \approx \left(\frac{1}{3}\right)^{w+w'-v} \begin{cases} 1 + g^2 \frac{2v}{3^2 N}(1 - 3N + 2v) & \text{if} \quad w + w' - 2v = 0, \\ g^2 \frac{2v}{3N} & \text{if} \quad w + w' - 2v = 1. \end{cases} \tag{S42}$$

The higher order terms will scale at most like $O(g^4 N^2/N^2) = O(g^4)$ and so for small g, the above expression for $W(w, w', v)$ is an excellent approximation. In the general case of $g' = \frac{g}{N^a}$, $a \geq 0$, the above Taylor expansion yields a series expression for $W(w, w', v)$ where the nth term scales at most as $O(g^{2n} N^n/N^{2na})$. Thus, for $a < \frac{1}{2}$, the series is not convergent, and Eq. (S42) does not constitute a good approximation. Below, we assume $a = \frac{1}{2}$, but all results and expressions in this section are applicable for $a \geq \frac{1}{2}$ as well, with the appropriate replacement of g. For some discussion of the $a = 0$ case, see Sec. VI.

Let us now consider the $\mathcal{R}$ matrix. The $w + w' - 2v = 0, 1$ cases contribute to the diagonal as well as super- and sub-diagonals of each block of $\mathcal{R}$. These matrix elements are

$$\mathcal{R}(w, 0, w', 0) = \delta_{w,w'} 3^w W(w, w', w') + \delta_{w,w'+1} 3^w (N - w' - 1) W(w, w', w') \tag{S43}$$
$$+ \delta_{w,w'-1} 3^w w' W(w, w', w' - 1) + O(g^4),$$
$$\mathcal{R}(w, 1, w', 0) = \delta_{w,w'+1} 3^w W(w, w', w') + O(g^4), \tag{S44}$$
$$\mathcal{R}(w, 0, w', 1) = \delta_{w,w'-1} 3^w W(w, w', w' - 1) + O(g^4), \tag{S45}$$
$$\mathcal{R}(w, 1, w', 1) = \delta_{w,w'} 3^w W(w, w', w') + \delta_{w,w'+1} 3^w (N - w') W(w, w', w') \tag{S46}$$
$$+ \delta_{w,w'-1} 3^w (w' - 1) W(w, w', w' - 1) + O(g^4).$$

Writing out the master equation, Eq. (S36), within the g^2 approximation, we have two coupled equations for the two $(w_1 = 0, 1)$ blocks:

$$\frac{h_{t+1}(w, 0) - h_t(w, 0)}{g^2} = \frac{2w}{9N} h_t(w + 1, 1) + \frac{2w}{9N}(1 - 3N + 2w)h_t(w, 0) \tag{S47}$$
$$+ \frac{2(N - w)}{3N}(w - 1)h_t(w - 1, 0) + \frac{2w(w + 1)}{9N}h_t(w + 1, 0),$$
$$\frac{h_{t+1}(w, 1) - h_t(w, 1)}{g^2} = \frac{2(w - 1)}{3N} h_t(w - 1, 0) + \frac{2w}{9N}(1 - 3N + 2w)h_t(w, 1) \tag{S48}$$
$$+ \frac{2(w - 1)}{3N}(N - w + 1)h_t(w - 1, 1) + \frac{2w^2}{9N}h_t(w + 1, 1).$$

Note that the coupling between the two w_1 sectors scales as w/N. Since the initial conditions are constrained to the $w_1 = 1$ sector [see Eq. (S39)], the early time dynamics will remain approximately in $h_t(w, 1)$ (i.e $h_t(w, 0) \approx 0$ at early times) until w reaches $O(N)$.

By adding Eqs. (S47) and (S48), we get a closed equation for the total operator weight probability $h_t(w) \equiv h_t(w, 0) + h_t(w, 1)$

$$\frac{h_{t+1}(w) - h_t(w)}{g^2} = \frac{2w(w + 1)}{9N} h_t(w + 1) + \frac{2w}{9N}(1 - 3N + 2w)h_t(w) + \frac{N - w + 1}{3N} 2(w - 1)h_t(w - 1). \tag{S49}$$

Up to now, the only approximation we made was the expansion up to g^2. We now assume that $h(w,t)$ varies slowly with respect to $g^2 t$ and w, and replace finite differences by derivatives which yields a Fokker-Planck equation

$$\partial_\tau h(w,\tau) = -\partial_w(D_1(w)h(w,\tau)) + \partial_w^2(D_2(w)h(w,\tau)), \tag{S50}$$

where we introduced a rescaled time $\tau = g^2 t$. Note that Eqs. (S47) and (S48) individually are not in the form of a Fokker-Planck equation, but their sum is. The drift and diffusion coefficients are

$$D_1(w) = \frac{2(4 + w + 3Nw - 4w^2)}{9N}, \tag{S51}$$

$$D_2(w) = \frac{-3 + 3N(w-1) + 7w - 2w^2}{9N}. \tag{S52}$$

In terms of the scaled weight $\phi \equiv w/N$, the Fokker-Planck equation takes the form

$$\partial_\tau h(\phi,\tau) = -\partial_\phi\left(\frac{2}{3}\left(\phi - \frac{4}{3}\phi^2\right)h(\phi,\tau)\right) + \partial_\phi^2\left(\left(\frac{\phi}{3N} - \frac{2}{9}\frac{\phi^2}{N}\right)h(\phi,\tau)\right), \tag{S53}$$

where we dropped all the $O(1/N)$ terms from the drift coefficient and all the $O(1/N^2)$ terms from the diffusion.

IV. RELATION BETWEEN THE AVERAGE OF THE SQUARED COMMUTATOR AND THE MEAN OPERATOR WEIGHT

In this section, we derive the relation between the average of the squared commutator, defined in Eq. (1) of the main text, and the operator weight probability $h_t(w, w_1)$.

Let us start with Eq. (1) of the main text, and, without loss of generality, pick the two operators to be X_1 at position 1 and Y_r at position $r > 1$

$$\mathcal{C}(r,t) = -\frac{1}{2}\operatorname{tr}\left(\rho_\infty[X_1(t), Y_r]^2\right), \tag{S54}$$

where ρ_∞ is the infinite-temperature Gibbs state, and $X_1(t)$ is the Heisenberg evolved operator. Using Eq. (S1), the commutator in Eq. (S54) can be written as

$$[X_1(t), Y_r]^2 = \left(\sum_{\mathcal{S}} a_{\mathcal{S}}(t)[\mathcal{S}, Y_r]\right)^2 = \left(2\sum_{\mathcal{S}:\mathcal{S}_r=X,Z} a_{\mathcal{S}}(t)\mathcal{S}Y_r\right)^2, \tag{S55}$$

which gives

$$\mathcal{C}(r,t) = -2\sum_{\mathcal{S}:\mathcal{S}_r=X,Z}\sum_{\mathcal{S}':\mathcal{S}'_r=X,Z} a_{\mathcal{S}}(t)a_{\mathcal{S}'}(t)\operatorname{tr}(\rho_\infty \mathcal{S}Y_r\mathcal{S}'Y_r) \tag{S56}$$

$$= 2\sum_{\mathcal{S}:\mathcal{S}_r=X,Z} a_{\mathcal{S}}(t)^2, \tag{S57}$$

where we used $\operatorname{tr}(\rho_\infty \mathcal{S}\mathcal{S}') = \delta_{\mathcal{S}\mathcal{S}'}$ and the fact that different Pauli matrices anti-commute. Here the sum is constrained to be over all strings that have an X or a Z on site r.

The average of Eq. (S56) over many realizations of the random circuit is therefore given by

$$\langle \mathcal{C}(r,t)\rangle = 2\sum_{\mathcal{S}:\mathcal{S}_r=X,Z}\left\langle a_{\mathcal{S}}(t)^2\right\rangle, \tag{S58}$$

where the evolution of $\left\langle a_{\mathcal{S}}^2(t)\right\rangle$ is what we calculated in the previous sections.

Since we have assumed in this section that we start from a single site operator, as we did in Sec. II, we have that the average probabilities $\left\langle a_{\mathcal{S}}^2(t)\right\rangle$ only depend on the total weight w and weight w_1 on site 1, as explained in Sec. II.

Thus, we may rewrite Eq. (S58) in terms of $h_t(w, w_1)$, using Eq. (S32). A similar calculation to the one leading to Eq. (S36) yields

$$\langle \mathcal{C}(r, t) \rangle = 4 \sum_{\substack{\Omega_\mathcal{S} \subset \{1, \cdots, N\} \\ r \in \Omega_\mathcal{S}}} 3^{|\Omega_\mathcal{S}| - 1} \frac{h_t(|\Omega_\mathcal{S}|, |\Omega_\mathcal{S} \cap \{1\}|)}{3^{|\Omega_\mathcal{S}|} \binom{N-1}{|\Omega_\mathcal{S}| - |\Omega_\mathcal{S} \cap \{1\}|}} \tag{S59}$$

$$= \frac{4}{3} \left[\sum_{w=1}^{N-1} \binom{N-2}{w-1} \frac{h_t(w, 0)}{\binom{N-1}{w}} + \sum_{w=2}^{N} \binom{N-2}{w-2} \frac{h_t(w, 1)}{\binom{N-1}{w-1}} \right] \tag{S60}$$

$$= \frac{4}{3(N-1)} \sum_{w=1}^{N} [(w-1) h_t(w) + h_t(w, 0)], \tag{S61}$$

where $h_t(w) \equiv h_t(w, 0) + h_t(w, 1)$, as defined in the main text and in Sec. III [Eq. (S40)].

Using the fact that $h_t(w)$ is normalized (i.e. $\sum_w h_t(w) = 1$) and defining the mean weight $\langle w(t) \rangle = \sum_w w h_t(w)$, we get

$$\langle \mathcal{C}(r, t) \rangle = \frac{4}{3} \frac{\langle w(t) \rangle - 1}{N - 1} + \frac{4}{3(N-1)} \sum_{w=1}^{N} h_t(w, 0). \tag{S62}$$

By the normalization of the probability distribution, we further know that $\sum_{w=1}^{N} h_t(w, 0) < 1$. Hence, the second term in the equation above scales as $O(1/N)$ and is therefore negligible for large N. Thus, in the limit of large N we have

$$\langle \mathcal{C}(r, t) \rangle = \frac{4}{3} \frac{\langle w(t) \rangle}{N} + O(1/N). \tag{S63}$$

V.   ADDITIONAL DETAILS ON THE TIME-EVOLUTION OF $h(w, w_1)$

In this section, we provide additional numerical and analytical details regarding the probability weight distribution.

In Fig. S1, we plot snapshots of $h(w)$ and $h(w, w_1 = 0, 1)$, at different times, computed numerically using the exact master equation. The initial distribution starts in the $w_1 = 1$ sector and quickly (exponentially) expands. At early times, during the exponential growth, the distribution is supported almost exclusively on the $w_1 = 1$ sector. At later times, when $h(w)$ is very broad in weight space and has large support on weights $w \sim O(N)$, the coupling between the two $w_1 = 0, 1$ sectors turns on and $h(w, 0)$ starts to get populated. Finally, $h(w)$ reaches the steady-state, which, as we show below, is, to a good approximation, a Gaussian centered at $w = 3N/4$ with a width $\sim \Delta w / N \propto 1/\sqrt{N}$. The steady-state corresponds to all strings being equally likely, and hence the Gaussian peak in $h(w, 1)$ is three times as large as the one in $h(w, 0)$.

S1.   Stationary solution for $h(w)$

At large t the distribution $h(t, \phi = w/N)$ approaches a stationary solution that obeys following equation

$$- \partial_\phi [D_1(\phi) h(\phi)] + \partial_\phi^2 [D_2(\phi) h(\phi)] = 0, \tag{S64}$$

where

$$D_1(\phi) = \frac{2}{3} \phi \left(1 - \frac{4\phi}{3} \right), \qquad D_2(\phi) = \frac{\phi}{3N} \left(1 - \frac{2\phi}{3} \right). \tag{S65}$$

Integrating out Eq. S64 we obtain

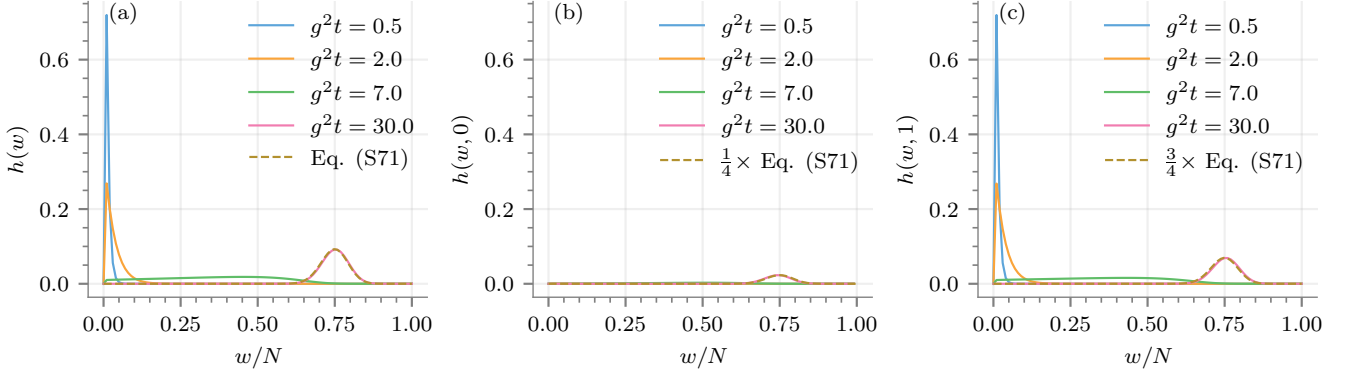$$- D_1(\phi) h(\phi) + \partial_\phi [D_2(\phi) h(\phi)] = C. \tag{S66}$$

FIG. S1: (a) Snapshots of the numerically computed total probability weight distribution $h(w) = h(w,0) + h(w,1)$ for $g = 0.1$ and $N = 100$, together with the analytical expression of the steady-state from Eq. (S71), which essentially agrees with the $g^2t = 30$ numerics. (b) The same plot for $h(w, w_1 = 0)$. While $h(w,0) \approx 0$ for early and intermediate times, the numerics for $g^2t = 30$ essentially agree with the analytical prediction for the steady state. (c) The same plot for $h(w, w_1 = 1)$.

Equation (S66) can be rewritten as

$$\partial_\phi h(\phi) = \left(\frac{D_1(\phi) - \partial_\phi D_2(\phi)}{D_2(\phi)} \right) h(\phi) + \frac{C}{D_2(\phi)}. \tag{S67}$$

Solution of (S67) is straightforward:

$$h(\phi) = const \times e^{J(\phi)} \int_0^\phi \frac{d\phi' e^{-J(\phi')}}{D_2(\phi')},$$
$$J(\phi) = \int d\phi \frac{D_1 - \partial_\phi D_2}{D_2} = 4N\phi - \log\phi + (3N - 1)\log(3 - 2\phi). \tag{S68}$$

As a result we obtain solution for $h(\phi)$ in the form:

$$h(\phi) = const \times \frac{e^{NS(\phi)}}{(3 - 2\phi)\phi} \int_0^\phi d\phi' \ e^{-NS(\phi')}, \tag{S69}$$

where

$$S(\phi) = 4\phi + 3\log(3 - 2\phi). \tag{S70}$$

In the limit $N \to \infty$ the main contribution in the integral (S69) comes from the vicinity of the boundary point $\phi = 0$. Expanding $S(\phi)$ in Taylor series in powers ϕ: $S(\phi) \approx S(0) + 2\phi$ and substituting it inside of the integrand in Eq. (S69) results in

$$h(\phi) \sim \frac{e^{NS(\phi)}}{(3 - 2\phi)\phi} \left[1 - e^{-2N\phi} \right]. \tag{S71}$$

Expression Eq. (S71) can be further simplified since $e^{NS(\phi)}$ is strongly peaked in the vicinity of $\phi_0 = 3/4$ which is the extremum of $S(\phi)$: $S(\phi) \approx S(\phi_0) + \frac{S''(\phi_0)}{2}(\phi - \phi_0)^2 + ...$, that gives

$$h(\phi) \sim \frac{e^{-\frac{8N}{3}(\phi - 3/4)^2}}{\phi(3 - 2\phi)} \left[1 - e^{-2N\phi} \right]. \tag{S72}$$

VI. MEAN-WEIGHT AFTER ONE STEP AND SCRAMBLING IN O(1)

In this section, we derive a simple expression for the mean-weight after a single step of the random circuit. Here, a single step is defined as in Sec. I, i.e $U = U_\mathrm{I} U_\mathrm{II} U_\mathrm{I}$. In doing so, we show that if the global interactions are sufficiently strong (i.e if g' is independent of N) then a single step of the circuit is sufficient to achieve scrambling.

Starting from the initial conditions defined in Eq. (S39), and using the master equation in Eq. (S36), we find after a single step

$$h_{t=1}(w, w_1) = \mathcal{R}(w, w_1, 1, 1). \tag{S73}$$

Using Eqs. (S37) and (S38), we can further simplify

$$h_{t=1}(w, w_1) = 3^w \binom{N-1}{w-w_1} W(w, 1, w_1) = \begin{cases} 0 & \text{if} \quad w_1 = 0, \\ \frac{1}{3}\binom{N-1}{w-1}\left(\delta_{w,1} + 2\left[\cos^2(g')\right]^{2(N-w)}\left[\sin^2(g')\right]^{2(w-1)}\right) & \text{if} \quad w_1 = 1. \end{cases} \tag{S74}$$

The above describes the probability weight distribution after a single step, valid for arbitrary g'.

The mean of the above distribution can be computed exactly,

$$\langle w \rangle = \sum_{w=1}^{N} w h_{t=1}(w, w_1 = 1) = \frac{1}{3} + \frac{2}{3}\cos^2(g') + \frac{2}{3}N\sin^2(g'). \tag{S75}$$

Thus, if g' is independent of N, then $\langle w \rangle$ is $O(N)$ and $\langle \mathcal{C} \rangle$ (see Sec. IV) is $O(1)$ after just a single step.

[S1] L. Zhang, Matrix integrals over unitary groups: An application of Schur-Weyl duality, (2014), arXiv:1408.3782.
[S2] A. Nahum, S. Vijay, and J. Haah, Operator Spreading in Random Unitary Circuits, Phys. Rev. X **8**, 021014 (2018).