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 Swingle1, 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: May 13, 2020)

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 exactly solvable models: (1) a random circuit with Haar random local unitaries and a global interaction and (2) a classical model of globally coupled non-linear 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 can be verified with state-of-the-art 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 an 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, commuting, unitary and Hermitian operators \( W \) can be probed via the squared commutator of two local, commuting, unitary and Hermitian operators

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

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, \( C(r, t) \) 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 \( C(r, t) \) begins to deviate substantially from 0 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 non-chaotic 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 non-trivial, involving random couplings and many-body interactions. Some recent proposals suggested that spin models with non-local interactions can exhibit fast scrambling [25–27], albeit with complicated and inhomogeneous interactions.

In this paper, 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

\[
\mathcal{H} = \mathcal{H}_{\text{local}} - \frac{K}{N} \sum_{i<j} Z_i Z_j, \tag{2}
\]

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 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 \( U(t) = (U_1 U_2)^t \) where a single-time-step update consists of the two layers

\[
U_1 = \prod_{i=1}^{N} U_{H,i}, \quad U_2 = e^{-i \frac{g}{2 \sqrt{N}} \sum_{i<j} Z_i Z_j}, \tag{3}
\]
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 $\mathcal{H}_{\text{loc}}$. 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 $\mathcal{O}(t) = \sum S a_S(t) S$, where $S$ is a string composed of the Pauli matrices and the identity, forming a basis for $SU(2^N)$. As in random brickwork models [37, 38] and random Brownian models [9], the Haar-averaged probabilities $\langle a_S^2(t) \rangle$, encoding the time evolution of $\mathcal{O}(t)$, themselves obey a linear equation 

$$\frac{\partial}{\partial t} h(w,\tau) = -\frac{\partial}{\partial w} (D_1(w) h(w,\tau)) + \frac{\partial^2}{\partial w^2} (D_2(w) h(w,\tau)), \quad (10)$$

where $D_1(w) = \sum_{v=0}^{g-1} (1 - \delta_S) v_{\tau}^m$, counting the number of non-identity operators, i.e $v(S) = \sum_v (1 - \delta_{S,v})$, and on the number of sites where both $\mathcal{S}$ and $\mathcal{S}'$ are non-identity, i.e $v(S, S') = \sum_v (1 - \delta_{S,v}) (1 - \delta_{S',v})$, and is given by (see Supplemental Material (SM) for derivation [39]) [40]

$$W(w, w', v) = \left(\frac{1}{3}\right)^w \sum_{k=0}^w \left(\begin{array}{c} v \\ k \end{array}\right) \sum_{l=0}^k \left(\begin{array}{c} k \\ l \end{array}\right) \left[\cos^2\left(\frac{2l - k}{\sqrt{N}} g\right)\right]^{N - k - (w + w' - 2v)} \left[\sin^2\left(\frac{2l - k}{\sqrt{N}} g\right)\right]^{w + w' - 2v} \quad (5)$$

If we further assume that $\mathcal{O}$ starts out as a single site operator on site 1, then throughout the evolution, the probabilities 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 probability distribution $h_t$ at time $t$,

$$h_t(w, w_1) = \langle a_S^2(t) \rangle 3^w \left(\frac{N - 1}{w - w_1}\right), \quad (6)$$

which is normalized by the number of string configurations for a given $w$ and $w_1$.

![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 $U_{II}$.](image1)

The time evolution of this probability distribution is given by the master equation

$$h_{t+1}(w, w_1) = \sum_{w'=0,1} \sum_{w'=w'_1} R(w, w_1, w', w'_1) h_t(w', w'_1), \quad (7)$$

where the $2N \times 2N$ matrix $R$ is

$$R(w, w_1, w', w'_1) = 3^w \sum_{m=0}^{\min\{w-w_1, w'-w'_1\}} \left(\begin{array}{c} w' - w'_1 \\ m \end{array}\right) \left(\begin{array}{c} N - 1 - w' + w'_1 \\ w - w_1 - m \end{array}\right) W(w, w', m + w_1 w'_1). \quad (8)$$

The transition matrix $R$, 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 $h_t(w) \equiv h_t(w, 0) + h_t(w, 1)$,

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

which is similar to random Brownian models [9, 41] and shows that, at $\mathcal{O}(g^2)$, $w$ can change by at most $\pm 1$ 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 $O(1/N, w/N)$)

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

This equation describes the rapid growth of an initially localized distribution, followed by a broadening and finally saturation (see Fig. 2 and SM [39] 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 $\langle w(t) \rangle \sim e^{3d/4}t^{3/4}$, which agrees with the full numerical solution of the master equation, as can be seen in Fig. 2. For small $g$ and for an infinite-temperature Gibbs state, the mean weight is related to the squared-commutator in Eq. (1) by $C(t) = \langle w(t) \rangle / N$. Together, these two facts establish that this model is fast scrambling, with a Lyapunov exponent of $\lambda = \frac{2}{3}g^2$.

Finally, let us comment on our choice of the normalization of the global interactions. In Eq. (3), we had to suppress the interactions by $1/\sqrt{N}$ and not by $1/N$ that features in Eq. (2). This is an artifact of the random nature of the model, which leads to the cancellation of the linear term in $g$. As we show next, this does not preclude a non-random extensive Hamiltonian from exhibiting fast scrambling.

**Classical Model.** Let us now consider a different setting, without randomness, that also allows to probe the basic timescales involved, particularly with regard to the normalization of the global interaction. A convenient tractable choice is a classical model consisting of globally coupled non-linear oscillators. Note that the analogs of out-of-time-order correlators (OTOCs) have been studied in a variety of classical models [42–45].

Consider a $2N$-dimensional phase space with coordinates $q_r$ (positions) and $p_r$ (momenta) for $r = 1, \cdots, N$ with canonical structure specified by the Poisson brackets $(q_r, p_s)_{PB} = \delta_{rs}$. The Hamiltonian is

$$\mathcal{H}_c = K + V_2 + V_4,$$  \quad (12)

where

$$K = \sum_{r=1}^{N} \frac{p_r^2}{2}, \quad V_4 = \frac{\Omega_2^2}{4} \sum_{r=1}^{N} q_r^4, \quad (13)$$

$$V_2 = \frac{\Omega_2^2}{2} \sum_{r=1}^{N-1} (q_{r+1} - q_r)^2 + \frac{\Omega_2^2}{2N} \left( \sum_{r=1}^{N} q_r \right)^2. \quad (14)$$

The equation of motion for $q_r$ is

$$\frac{d^2q_r}{dt^2} = -\Omega_4^2 (2q_r - q_{r+1} - q_{r-1}) - \Omega_2^2 \sum_{r=1}^{N} \frac{q_r}{N} - \Omega_2^2 q_r^3. \quad (15)$$

To implement the open boundary conditions, we may substitute $q_0 = q_{N+1} = 0$ into the equation of motion. The timescales for the growth of perturbations under $H_c$ dynamics may be understood in two stages. First, $K + V_2$ can be solved exactly; this combination of terms provides the non-locality. The remaining $V_4$ term renders the dynamics chaotic, provided $\Omega_3$ is large enough. The dynamics of $K + V_2$ causes a localized perturbation to spread to every oscillator with non-local amplitude $1/N$ in a time of order $1/\Omega_2$. 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 $\lambda$ is some typical Lyapunov exponent.

At the quadratic level, the uniform mode, $Q = \frac{1}{N} \sum q_r$ is decoupled from the remaining modes of the chain. In particular, when $\Omega_2 = 0$, the uniform mode is a zero mode of the conventional harmonic chain. Hence, at the quadratic level, the propagation of any perturbation is a superposition of the motion due to the local $\Omega_1$ terms and the special dynamics of the uniform mode. Since the local terms cannot induce non-local perturbations, we may focus on the dynamics of the uniform mode.

The uniform mode’s equation of motion is $\frac{d^2Q}{dt^2} = -\Omega_2^2 Q$ with solution

$$Q(t) = Q(0) \cos \Omega_2 t + \frac{dQ(0)}{\Omega_2} \sin \Omega_2 t. \quad (16)$$

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

$$\delta q_{r \gg 1}(t) = \frac{\epsilon}{N} \cos \Omega_2 t - 1. \quad (17)$$

Thus, after a time $\pi/\Omega_2$, any localized perturbation has spread to distant sites with amplitude $\epsilon/N$.

The inclusion of $V_4$ renders the equations of motion non-linear 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 we consider is as follows. We compute the dynamics of two configurations, $\hat{q}^{(1)}$ and $\hat{q}^{(2)}$, averaged over many initial conditions. The initial condition of configuration one has each oscillator start at rest from a 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. This difference is then averaged over 1000 different initial conditions. Fig. 3 shows this average of $\Delta q_N$ for $N = 20$ with $\Omega_1 = 1$, $\Omega_2 = 1$, and $\Omega_3 = 2$; the inset shows $\Delta q_N$ for the most distant site for sizes
FIG. 3. Main: ln $\Delta q_r(t)$ for $N = 20, \epsilon = 10^{-5}, \Omega_1 = \Omega_2 = 1,$ and $\Omega_3 = 2.$ Each trace is a different site as a function of time. The highest trace is site $r = 1$ while the lowest traces are the most distant sites. One sees that the perturbation grows exponentially with time, at a rate that is approximately uniform across the chain after an early time period where the sites near site $r = 1$ are special. Inset: ln $\Delta q_{r=N}(t)$ for $N = 40, 80, 160.$

$N = 40, 2 \times 40, 2^2 \times 40.$ Each curve is shifted horizontally by a constant amount when the system size is doubled, thus confirming that the time for $\Delta q_N(t)$ to plateau depends logarithmically on $N$ as expected if $\Delta q_N(t) \sim \frac{1}{N} e^{\lambda t} + \cdots$ for $\lambda^{-1} \ln N \gg t \gg 1/\Omega_2.$

Chaos and level statistics.—Having established fast scrambling in both the random circuit and the classical model, we now return to the quantum case with the time-independent Hamiltonian 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

$$
\mathcal{H}_{\text{local}} = -J \sum_i Z_i Z_{i+1} - h_x \sum_i X_i - h_z \sum_i Z_i. \quad (18)
$$

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 [46] $\langle r \rangle,$ 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. [47] 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

$$
\mathcal{H} = -J \sum_i Z_i Z_{i+1} - h_x \sum_i X_i - \frac{K}{N} \sum_{i<j} Z_i Z_j. \quad (19)
$$

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) [46]. In the vicinity of $K \rightarrow 0,$ $\langle r \rangle$ (see Fig. 4) shows proximity to Poisson statistics, while, for $|K| \gtrsim 1,$ 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, in the infinite-temperature state $\rho_\infty$:

$$
F(r, t) = \text{Re}[\text{tr}(\rho_\infty Z_1(t) Z_r(t) Z_1(t) Z_r(t))], \quad (20)
$$

which is related to the squared commutator from Eq. (1) by $C(r, t) = 1 - F(r, t).$ We approximate the trace over $\rho_\infty$ by an expectation value in a Haar-random pure state, which enables us to reach larger systems sizes [48]. 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 non-local model in Eq. (19), which includes the global interaction. In the local case, the OTOC spreads ballistically, forming a linear light cone. It is clear that, in the non-local case, the growth is much faster and is super-ballistic, as expected for a fast scrambler.

In Fig. 5(b), we study the $N$-dependence of the OTOC, by looking at $C(r, t)$ between the two ends of the chain for different chain lengths $N,$ after a fixed evolution time. The result is linear with respect to $1/N,$ in agreement with the result of the random circuit and other fast scrambling models.

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 non-local model shows a significant speed up. Moreover, in the non-local model, the growth rate clearly increases with the system size, further supporting our claim.

Experimental realization.—The Hamiltonian in Eq. (19) can be naturally experimentally realized with Rydberg dressing of neutral atoms [49–52]. The spin
can be represented by two ground states where one of them is dressed to two Rydberg states with one of the Rydberg states giving rise to all-to-all interactions and the second to nearest-neighbor interactions.

We also imagine a plethora of other possible experimental realizations of similar spin models. For example, one can use cavity-photon-mediated all-to-all interactions \([28, 31, 53, 54]\) of the \(XX\) or \(XXZ\)-Heisenberg form \([25, 27]\) together with nearest-neighbor interactions achieved by Rydberg-dressing one of the grounds states \([55, 56]\).

The resulting Hamiltonian only differs in the basis of the local interactions (\(XX\) instead of \(ZZ\)), but we have verified that it leads to qualitatively similar scrambling physics. Other possibilities include a chain of coupled superconducting qubits, with all-to-all flip-flop interactions mediated via a common bus \([57–59]\) or trapped ions \([33–36, 60]\).

**Conclusion and outlook.**—In this paper, we demonstrated that a single global interaction together with local chaotic dynamics is sufficient to give rise to fast scrambling. This does not require disordered or inhomogeneous couplings and is within reach of current state-of-the-art quantum simulators. Future theoretical work may include a more systematic analysis of the time scales for large \(N\) as well the behaviour 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.** During the completion of this work, another study of fast scrambling in similar spin chains appeared [61].

We thank Shenglong Xu for discussions and Paraj Titum for pointing out the Hamiltonian (Eq. (19)) to us. R.B., P.B., Y.A.K., and A.V.G. acknowledge funding by the NSF PFCQC program, DoE ASCR Quantum Testbed Pathfinder program (award No. DE-SC0019040), DoE BES Materials and Chemical Sciences Research for Quantum Information Science program (award No. DE-SC0019449), DoE ASCR FAR-QC (award No. DE-SC0020312), AFOSR MURI, AFOSR, ARO MURI, NSF PFC at JQI, and ARL CDQI. R.B. acknowledges support of NSERC and FRQNT of Canada. B.G.S. is supported in part by the U.S. Department of Energy, Office of Science, Office of High Energy Physics QuantISED Award de-sc0019380.

---


[39] See Supplemental Material for additional details concerning the random circuit including the derivation.

[40] The term with $\omega = \omega' = 0$ is assumed to be $0^\omega = 1$. See [39].


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 Swingle1, 2

1Joint Center for Quantum Information and Computer Science, NIST/University of Maryland, College Park, Maryland 20742 USA
2Joint Quantum Institute, NIST/University of Maryland, College Park, Maryland 20742 USA

(Dated: May 13, 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, 10, 11 of the main text. Finally, in Sec. IV, we provide additional details on the dynamics and steady-state of the probability distribution.

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 $O(t) = U^\dagger(t)OU(t)$

$$O(t) = \sum_S a_S(t)S,$$  

(S1)

where the strings $S$ form a basis for $SU(q^N)$, normalized as $tr(S) = q^N\delta_{S,1}$, $tr(SS') = q^N\delta_{SS'}$. We take $U(t) = \prod_{i=1}^t U_i$ where $U_i = U_1U_1U_1$ and $U_1$ is a product of single site Haar random unitaries while $U_1$ is the global interaction. Note that the two $U_1$ appearing on either side of the $U_1$ are different, i.e. the random unitaries are random in both space and 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 $a_S(t) = q^{-N} tr(O(t)S)$, we can write $a_S^2(t)$ in terms of the coefficients at the previous time step

$$a_S^2(t) = q^{-2N} \sum_{S',S''} a_{S'}(t-1)a_{S''}(t-1) tr(U^\dagger S'US) tr(U^\dagger S''US).$$  

(S2)

Thus, we want to evaluate the quantity

$$\langle tr(U^\dagger S'US) tr(U^\dagger S''US) \rangle,$$  

(S3)

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

Using properties of trace, we can write

$$\langle tr(U^\dagger S'US) tr(U^\dagger S''US) \rangle = \langle tr(U^\dagger S'US \otimes U^\dagger S''US) \rangle.$$

(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

$$\langle tr(U^\dagger S'US) tr(U^\dagger S''US) \rangle = \langle \left( (U^\dagger \otimes U)(U^\dagger \otimes U) \right) \left( U^\dagger S' U^\dagger \otimes U^\dagger S'' U^\dagger \right) (U^\dagger \otimes U)(U^\dagger \otimes U) \rangle (S \otimes S).$$

(S5)

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

$$I_1 = \left( U^\dagger S' U^\dagger \otimes U^\dagger S'' U^\dagger \right),$$

(S6)

$$I_2 = (U^\dagger \otimes U)(U^\dagger \otimes U)I_1,$$  

(S7)

$$I_3 = \left( (U^\dagger \otimes U)I_2(U^\dagger \otimes U)^\dagger \right),$$  

(S8)
with \( \text{tr}(I_3(S \otimes S)) \) 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]

\[
\langle (U \otimes U)A(U \otimes U)^\dagger \rangle \equiv \int_{U(d)} (U \otimes U)A(U \otimes U)^\dagger d\mu(U) = \left( \frac{\text{tr}(A)}{d^2 - 1} - \frac{\text{tr}(AF)}{d(d^2 - 1)} \right) \mathbb{I}_{d^2} - \left( \frac{\text{tr}(A)}{d^2 - 1} - \frac{\text{tr}(AF)}{d^2 - 1} \right) F,
\]

where \( F = \sum_{ij} |ij\rangle \langle ji| \) is the swap operator. From this, it follows that

\[
I_1 = \prod_r \langle U_r^1 S_r U_r \otimes U_r^2 S_r U_r \rangle = \delta_{S_r, S_r'} \prod_r \left( q^{2m} \delta_{S_r, 1} - \frac{1}{q^{2m} - 1} \mathbb{I}_{q^{2m}} + \frac{q^m - q^{m} \delta_{S_r, 1}}{q^{2m} - 1} F_r \right)
\]

where we used \( \text{tr}(S_r) = q^m \delta_{S_r, 1} \) and \( \text{tr}(S_r S_r') = q^m \delta_{S_r, S_r'} \). Here \( F_r \) swaps \( r \) of the left system with the corresponding site \( r \) of the right system.

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

\[
\langle a^2_S(t + 1) \rangle = \sum_{S'} W_{S, S'} \langle a^2_{S'}(t) \rangle,
\]

where \( W_{S, S'} = q^{-2N} \text{tr}(I_3(S \otimes S)) \).

To proceed, we specialize to qubits, i.e. \( q = 2 \), in which case the swap operator can be written as \( F_r = \frac{1}{2}(I_r \otimes \bar{I}_r + \bar{\sigma}_r \cdot \bar{\sigma}_r) = \frac{1}{2}(I_r I_r + X_r \bar{X}_r + Y_r \bar{Y}_r + Z_r \bar{Z}_r) \) where bar denotes operators acting on the second system. We can combine all the \( I_1 \)s together, giving

\[
I_1 = \delta_{S_r, S_r'} \prod_i \left( \delta_{S_r, 1} \mathbb{1}_{2^2} + \frac{1 - \delta_{S_r, 1}}{3} \sigma_i \cdot \bar{\sigma}_i \right)
\]

where in the second equality the sum is over the powerset of \( \{1, 2, \ldots, N\} \), i.e. all the \( (2^N) \) subsets of \( \{1, 2, \ldots, N\} \). The sum above essentially contains every possible string of the form \( S \otimes S \), i.e. the same operator appears on both copies of the system. Note that for a given string \( S' \), there is only one nonzero term in the sum. For each site \( i \), we either put an \( I_1 \) if \( S'_i = \mathbb{1} \) or we place \( \frac{1}{2} \bar{\sigma}_i \cdot \bar{\sigma}_i \), if \( S'_i \) is any other generator. The set \( \Omega_S \) therefore represents the support of the string \( S' \).

We now apply \( U_1 \otimes U_{11} \) to the above. Recall that

\[
U_{11} = e^{-i \frac{\pi}{4} Z_1} e^{i \frac{\pi}{4} Z_{11}}.
\]

Since Eq. (S12) contains all possible strings, it is instructive to first consider the result of applying \( U_{11} \otimes U_{11} \) and \( U_1 \otimes U_1 \) to a single string.

Note that the result of applying \( U_1 \otimes U_{11} \) and \( U_{11} \otimes U_1 \) is invariant if we replace any number of \( X \)s in the string by \( Y \)s 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. \( e^{-i \frac{\pi}{4} Z} X e^{i \frac{\pi}{4} Z} = Y \). This rotation clearly commutes with \( U_{11} \) and can be absorbed into \( U_1 \), 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 \( \Omega_S \) the support of some string \( S \). We can further divide \( \Omega_S \) based on the number and location of \( Z \)s 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 \( X \)s or \( Y \)s) by \( \Lambda = \Omega \setminus \Sigma \). For strings that are supported on \( k \) sites (i.e. \( |\Omega_S| = k \)), with fixed number and position of \( Z \)s, the degeneracy is \( 2^{|\Lambda|} \).

Without loss of generality, we can therefore consider strings composed of either \( X \)s or \( Z \)s. Consider the string \( \prod_{i \in \Lambda} X_i \prod_{j \in \Sigma} Z_j \). To apply the \( U_{11} \), we can use the fact that \( [X_i X_j, Z_i Z_j] = 0 \). We get

\[
(U_{11} \otimes U_{11})(\prod_{i \in \Lambda} X_i \bar{X}_i \prod_{j \in \Sigma} Z_j \bar{Z}_j)(U_{11} \otimes U_{11})^\dagger = \prod_{i \in \Lambda} \left( (X_i \cos(QA) + Y_i \sin(QA)) (\bar{X}_i \cos(\bar{Q}A) + \bar{Y}_i \sin(\bar{Q}A)) \right) \prod_{j \in \Sigma} Z_j \bar{Z}_j
\]

(S14)
where \( Q_\Lambda \) acts on all sites except those in \( \Lambda \), i.e. \( Q_\Lambda \equiv g' \sum_{l \notin \Lambda} Z_l \). Here we used the formula 
\[
U_{II} X_r U_{II}^\dagger = X_r \cos(Q_r) + Y_r \sin(Q_r).
\]
We see that we can safely apply and perform the Haar average on sites inside of \( \Lambda \), since all the cosines and sines and the \( ZZ \) act on sites outside of \( \Lambda \). With slight abuse of notation, let us denote \( \langle (U_1 \otimes U_1)A(U_1 \otimes U_1)^\dagger \rangle \) 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 \( \langle X_j Z_j \rangle = 0 \), so the cross terms in the above expression will vanish. Only \( \langle X_j Z_j \rangle \equiv V_j \) will remain. Here the single site operator \( V_i \) is defined as \( V_j = -\frac{1}{2} \mathbb{1}_4 + \frac{2}{3} F \). Explicitly, we find
\[
\left( U_{II} \otimes U_{II} \right) \left( \prod_{i \in \Lambda} X_i \bar{X}_i \prod_{j \in \Sigma} Z_j \bar{Z}_j \right) \left( U_{II} \otimes U_{II} \right)^\dagger = \prod_{i \in \Lambda} V_i \left( \cos^{\left| \Lambda \right|}(R_\Lambda) \prod_{j \in \Sigma} Z_j \bar{Z}_j \right) \tag{S15}
\]
where \( R_\Lambda = \bar{Q}_\Lambda - Q_\Lambda \).

Combining this with the discussion above, we find that \( \mathcal{I}_3 \) may be written as
\[
\mathcal{I}_3 = \delta_{\mathcal{S'}, \mathcal{S''}} \sum_{\Omega_{\mathcal{S'}, \mathcal{S''}} \subset \{1, 2, \ldots, N\}} \left( \prod_{j \notin \Omega_{\mathcal{S'}}} \delta_{\mathcal{S'}, 1} \right) \left( \prod_{i \in \Omega_{\mathcal{S'}}} \frac{1 - \delta_{\mathcal{S'}, i}}{3} \right) \prod_{\Lambda \subset \Omega_{\mathcal{S'}}} 2^{\left| \Lambda \right|} \left( \prod_{m \in \Lambda} V_m \right) \left( \cos^{\left| \Lambda \right|}(R_\Lambda) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} Z_n \bar{Z}_n \right). \tag{S16}
\]

It remains to compute \( \langle \cos^{\left| \Lambda \right|}(R_\Lambda) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} Z_n \bar{Z}_n \rangle \). To do so we expand the cosine as follows \( \cos^k(x) = \frac{1}{2^k} \sum_{n=0}^k \binom{k}{n} \cos((2n - k)x) \),
\[
\left( \cos^{\left| \Lambda \right|}(R_\Lambda) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} Z_n \bar{Z}_n \right) = \frac{1}{2^{\left| \Lambda \right|}} \sum_{l=0}^{\left| \Lambda \right|} \binom{\left| \Lambda \right|}{l} \cos((2l - \left| \Lambda \right|)R_\Lambda) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} Z_n \bar{Z}_n. \tag{S17}
\]

To proceed we can pull a single-site operator out of \( R_\Lambda \). 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
\[
\cos((2l - \left| \Lambda \right|)R_\Lambda) = \cos((2l - \left| \Lambda \right|)R_{\Lambda \cup \{j\}}) \cos((2l - \left| \Lambda \right|)D_j) - \sin((2l - \left| \Lambda \right|)R_{\Lambda \cup \{j\}}) \sin((2l - \left| \Lambda \right|)D_j). \tag{S18}
\]

This allows us to perform the Haar average over site \( j \). The sine term will not contribute, since \( \sin((2l - \left| \Lambda \right|)D_j)Z_j \bar{Z}_j = 0 \). Repeating this procedure recursively for all sites in \( \Omega_{\mathcal{S'}} \setminus \Lambda \), we get
\[
\left( \cos((2l - \left| \Lambda \right|)R_{\Lambda \cup \{j\}}) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} Z_n \bar{Z}_n \right) = \left( \cos((2l - \left| \Lambda \right|)R_{\Omega_{\mathcal{S'}}}) \right) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} \left( \cos((2l - \left| \Lambda \right|)D_n)Z_n \bar{Z}_n \right). \tag{S19}
\]

Continuing the procedure for the \( \langle \cos((2l - \left| \Lambda \right|)R_{\Omega_{\mathcal{S'}}}) \rangle \) term, we have
\[
\left( \cos((2l - \left| \Lambda \right|)R_{\Omega_{\mathcal{S'}}}) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} Z_n \bar{Z}_n \right) = \left( \prod_{t \notin \Omega_{\mathcal{S'}}} \langle \cos((2l - \left| \Lambda \right|)D_t) \rangle \right) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} \left( \cos((2l - \left| \Lambda \right|)D_n)Z_n \bar{Z}_n \right). \tag{S20}
\]

Using \( \cos((2l - \left| \Lambda \right|)D) = \cos^2((2l - \left| \Lambda \right|)g') + 2 \bar{Z} \sin^2((2l - \left| \Lambda \right|)g') \) gives
\[
\left( \cos((2l - \left| \Lambda \right|)R_{\Omega_{\mathcal{S'}}}) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} Z_n \bar{Z}_n \right)
= \left( \prod_{t \notin \Omega_{\mathcal{S'}}} \langle \cos^2((2l - \left| \Lambda \right|)g') + V_t \sin^2((2l - \left| \Lambda \right|)g') \rangle \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} \langle \cos^2((2l - \left| \Lambda \right|)g') V_n + \sin^2((2l - \left| \Lambda \right|)g') \rangle \right). \tag{S21}
\]

Putting things together, we find that Eq. (S17) is
\[
\left( \cos^{\left| \Lambda \right|}(R_\Lambda) \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} Z_n \bar{Z}_n \right)
= \frac{1}{2^{\left| \Lambda \right|}} \sum_{l=0}^{\left| \Lambda \right|} \binom{\left| \Lambda \right|}{l} \left( \prod_{t \notin \Omega_{\mathcal{S'}}} \langle \cos^2((2l - \left| \Lambda \right|)g') + V_t \sin^2((2l - \left| \Lambda \right|)g') \rangle \prod_{n \in \Omega_{\mathcal{S'}} \setminus \Lambda} \langle \cos^2((2l - \left| \Lambda \right|)g') V_n + \sin^2((2l - \left| \Lambda \right|)g') \rangle \right). \tag{S22}
\]
and finally, $\mathcal{I}_3$ is given by

$$
\mathcal{I}_3 = \delta_{S',S''} \sum_{\Omega_{S'} \subseteq \{1,2,\ldots,N\}} \left( \prod_{j \notin \Omega_{S'}} \delta_{S',1} \right) \left( \prod_{l \in \Omega_{S'}} \left( 1 - \frac{\delta_{S',1}}{3} \right) \right) \sum_{A \subseteq \Omega_{S'}} \left( \prod_{m \in A} V_m \right) \sum_{l=0}^{\left| A \right|} \left( \frac{\left| A \right|}{l} \right).
$$

\begin{align*}
&\times \prod_{t \notin \Omega_{S'}} \left( \cos^2((2l - |A|)g') + V_t \sin^2((2l - |A|)g') \right) \prod_{n \in \Omega_{S'} \setminus A} \left( \cos^2((2l - |A|)g') V_n + \sin^2((2l - |A|)g') \right).
\end{align*}

(S23)

To compute the $W_{S,S'}$-matrix from Eq. (S11), it remains to take the trace of Eq. (S23) with $S \otimes S$ and divide by $2^{2N}$, i.e

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

(S24)

Using Eq. (S23) together with $\text{tr}(V_i(S_i \otimes S_i)) = \frac{4}{3}(1 - \delta_{S_i,1})$, gives

$$
W_{S,S'} = \frac{1}{2^{2N}} \sum_{\Omega_{S'} \subseteq \{1,2,\ldots,N\}} \left( \prod_{j \notin \Omega_{S'}} \delta_{S',1} \right) \left( \prod_{l \in \Omega_{S'}} \left( 1 - \frac{\delta_{S',1}}{3} \right) \right) \sum_{A \subseteq \Omega_{S'}} \left( \prod_{m \in A} \frac{4}{3}(1 - \delta_{S_m,1}) \right) \sum_{l=0}^{\left| A \right|} \left( \frac{\left| A \right|}{l} \right)
$$

\begin{align*}
&\times \prod_{t \notin \Omega_{S'}} \left( \cos^2((2l - |A|)g') + \frac{4}{3}(1 - \delta_{S_{l,1}}) \sin^2((2l - |A|)g') \right) \\
&\times \prod_{n \in \Omega_{S'} \setminus A} \left( \cos^2((2l - |A|)g') + \frac{4}{3}(1 - \delta_{S_{n,1}}) + 4\delta_{S_{n,1}} \sin^2((2l - |A|)g') \right).
\end{align*}

(S25)

Note that because of $\prod_{m \in A} \frac{4}{3}(1 - \delta_{S_{m,1}})$ in Eq. (S25), $\Lambda$ is constrained to be in $\Omega_{S} \cap \Omega_{S'}$. The matrix elements of $W$ are

$$
W = \frac{1}{2^{2N}} \left( \frac{4}{3} \right)^{|\Omega_{S'}|} \sum_{A \subseteq \Omega_{S'} \cap \Omega_{S'}'} \left( \frac{4}{3} \right)^{|A|} \sum_{l=0,2l \neq |A|} \left( \frac{\left| A \right|}{l} \right) \left( 4 \cos^2((2l - |A|)g') \right)^{N - |\Omega_{S'} \cap \Omega_{S'}'|}
$$

\begin{align*}
&\times \left( \frac{4}{3} \sin^2((2l - |A|)g') \right)^{|\Omega_{S} \cap \Omega_{S'}|} \left( \frac{4}{3} \cos^2((2l - |A|)g') \right)^{|\Omega_{S} \cap \Omega_{S'}'| - |A|} \left( 4 \sin^2((2l - |A|)g') \right)^{|\Omega_{S'} \setminus \Omega_{S}|} \\
&\quad + \delta_{2l,|A|} \left( \frac{|A|}{2} \right) \prod_{t \notin \Omega_{S'}} \left( 4 \delta_{S_{l,1}} \right) \prod_{n \in \Omega_{S'} \setminus A} \left( \frac{4}{3}(1 - \delta_{S_{n,1}}) \right).
\end{align*}

(S26)

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

We can combine all constant factors (same results holds for the $2l = |A|$ term)

$$
\frac{1}{2^{2N}} \left( \frac{4}{3} \right)^{|\Omega_{S'}|} \left( \frac{4}{3} \right)^{|A|} \left( \frac{4}{3} \right)^{|\Omega_{S} \cap \Omega_{S'}|} \left( \frac{4}{3} \right)^{|\Omega_{S'} \setminus \Omega_{S}|} = \left( \frac{1}{3} \right)^{|\Omega_{S'}| + |\Omega_{S}|}.
$$

(S27)

Now, note that $\Lambda$ only appears in Eq. (S26) as $|A|$. Thus, we can replace the sum over subsets of $\Omega_{S} \cap \Omega_{S'}$ as $\sum_{A \subseteq \Omega_{S} \cap \Omega_{S'}'} = \sum_{k=0}^{\left| \Omega_{S} \cap \Omega_{S'} \right|} \binom{\left| \Omega_{S} \cap \Omega_{S'} \right|}{k}$. Thus, the $W$ matrix can be written as

$$
W_{S,S'} = W(|\Omega_{S}|,|\Omega_{S'}|,|\Omega_{S} \cap \Omega_{S'}|)
$$

(S28)

\begin{align*}
&= \left( \frac{1}{3} \right)^{|\Omega_{S'}| + |\Omega_{S}|} \sum_{k=0}^{\left| \Omega_{S} \cap \Omega_{S'} \right|} \left( \frac{\left| \Omega_{S} \cap \Omega_{S'} \right|}{k} \right) \left[ \sum_{l=0,2l \neq k} \binom{k}{l} \left( 4 \cos^2((2l - k)g') \right)^{N - k - (|\Omega_{S}| + |\Omega_{S'}| - 2|\Omega_{S} \cap \Omega_{S'}|)} \\
&\quad \times \left( \frac{4}{3} \sin^2((2l - k)g') \right)^{|\Omega_{S}| + |\Omega_{S'}| - 2|\Omega_{S} \cap \Omega_{S'}|} + \delta_{2l,k} \delta_{|\Omega_{S}| + |\Omega_{S'}| - 2|\Omega_{S} \cap \Omega_{S'}|,0} \binom{k}{|\Omega_{S}| + |\Omega_{S'}| - 2|\Omega_{S} \cap \Omega_{S'}|} \right].
\end{align*}

which is what appears in Eq. 5 of the main text, with the identification $w = |\Omega_{S}|$, $w' = |\Omega_{S'}|$, $v = |\Omega_{S} \cap \Omega_{S'}|$. In the main text, we also dropped the $\delta_{2l,k} \delta_{|\Omega_{S}| + |\Omega_{S'}| - 2|\Omega_{S} \cap \Omega_{S'}|,0}$ 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 ($W_{S,S'} = W_{S',S}$) matrix with all positive matrix elements.
II. MASTER EQUATION FOR SIMPLE INITIAL OPERATOR

Let us now assume that the initial operator $O$ starts as a single-site operator on site 1 without loss of generality. Then, the initial condition, averaged over the Haar unitaries is

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

(S29)

We now claim that for these initial conditions, the probabilities $\langle a_S^2(t) \rangle$ only depend on the string weight $w \equiv |\Omega_S|$ and the weight on site 1, $w_1 \equiv |\Omega_S \cap \{1\}|$. Note that $w_1$ takes values either 0 or 1. In light of this, it is convenient to define the normalized probability distribution. Let us normalize $h$ so that it can be thought as a probability distribution,

$$
h(w, w_1) = \langle a_S^2(t) \rangle D(w, w_1),
$$

(S30)

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

$$
D(k, q) = 3^k \binom{N-1}{k-q}.
$$

(S31)

The claim above can be proved by induction. The base case is trivial to see, by multiplying the initial conditions Eq. (S29) by the transition matrix $W$ from Eq. (S28). The inductive step proceeds as follows. First, we decompose the sum over strings $S'$ as $\sum_{S'} = \sum_{\Omega_{S'} \subset \{1, \ldots, N\}} 3^{||\Omega_{S'}||}$, which yields

$$
\langle a_S^2(t + 1) \rangle = \sum_{\Omega_{S'} \subset \{1, \ldots, N\}} \frac{1}{D(\Omega_{S'} | \Omega_S \cap 1)} \sum_{k=0}^{N-1} \sum_{m=0}^{3^k} \binom{|\Omega_S| - |\Omega_S \cap \{1\}|}{m} \binom{N-1 - |\Omega_S| + |\Omega_S \cap \{1\}|}{k-m} W(|\Omega_S|, k, m) h_t(\Omega_{S'}, |\Omega_{S'} \cap 1|).
$$

(S32)

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

$$
\langle a_S^2(t + 1) \rangle = \sum_{k=0}^{N-1} 3^k \binom{|\Omega_S| - |\Omega_S \cap \{1\}|}{m} \binom{N-1 - |\Omega_S| + |\Omega_S \cap \{1\}|}{k-m} W(|\Omega_S|, k, m) h_t(k, 0) \frac{D(k, 0)}{D(k)}
$$

$$
+ \sum_{k=1}^{N} 3^k \binom{|\Omega_S| - |\Omega_S \cap \{1\}|}{m} \binom{N-1 + |\Omega_S \cap \{1\}| - |\Omega_S|}{k-m-1 + |\Omega_S \cap \{1\}|} W(|\Omega_S|, k, m) h_t(k, 1) \frac{D(k)}{D(k-1)}.
$$

(S33)

Here, the first binomial in each bracket counts the number of ways one can choose the part of $\Omega_{S'}$ that is overlapping with $\Omega_S$ and the second binomial counts the number of ways to choose the non-overlapping part of $\Omega_{S'}$. It is clear at this point that the right-hand-side is a function of $w = |\Omega_S|$ and $w_1 = |\Omega_S \cap 1|$. Thus, replacing $\langle a_S^2(t + 1) \rangle$ by Eq. (S30) and simplifying gives

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

(S34)

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

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

(S35)
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

\[
W(w, w', v) = \left( \frac{1}{3} \right)^{w+w'} \sum_{k=0}^{v} \left( \frac{v}{k} \right) \left( \frac{k}{3} \right)^{N-k-(w+w'-2v)} \sum_{t=0, t \neq k}^{k} \left( \frac{k}{t} \right) \sin^2((2l - k)g') \] 

(1)

\[
\times \left[ \sin^2((2l - k)g') \right]^{w+w'-2v} + \delta_{2l,k} \delta_{w+w'-2v,0} \left( \frac{k}{3} \right)^{w+w'-2v} \right]. 
\]

(S43)

One may verify that \( \sum_i R_{i,j} = 1 \) 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

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

(S37)

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

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

(S38)

Note that \( h(0,0) \) does not actually participate in the dynamics since \( R(0,0,w',w'_1) = W(0,w',0) = \delta_{w',0} \).

### 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 \( g^2 \).

The factors of cosine and sine appearing in Eq. (S36), up to \( g^2 \), are given by

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

(S39)

We can now perform the sums over \( l \) and \( k \) appearing in Eq. (S36). We find

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

(S40)

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

\[
R(w, 0, w', 0) = \delta_{w,w'} 3^w W(w, w', w') + \delta_{w,w'+1} 3^w (N-w'-1) W(w, w', w') 
\]

(S41)

\[
R(w, 1, w', 0) = \delta_{w,w'+1} 3^w W(w, w', w') + O(g^4), 
\]

(S42)

\[
R(w, 0, w', 1) = \delta_{w,w'+1} 3^w W(w, w', w') + O(g^4), 
\]

(S43)

\[
R(w, 1, w', 1) = \delta_{w,w'} 3^w W(w, w', w') + \delta_{w,w'+1} 3^w (N-w') W(w, w', w') 
\]

(S44)

\[
+ \delta_{w,w'+1} 3^w (w'-1) W(w, w', w'-1) + O(g^4) 
\]
Writing out the master equation, Eq. (S34), within the \( g^2 \) approximation, we have

\[
\frac{h_{t+1}(w,0) - h_t(w,0)}{g^2} = \frac{2w}{9N}h_t(w+1,1) + \frac{2w}{9N}(1 - 3N + 2w)h_t(w,0) \\
+ \frac{2(N - w)}{3N}(w - 1)h_t(w-1,0) + \frac{2w(w+1)}{9N}h_t(w+1,0),
\]

\[
\frac{h_{t+1}(w,1) - h_t(w,1)}{g^2} = \frac{2(w - 1)}{3N}h_t(w-1,0) + \frac{2w}{9N}(1 - 3N + 2w)h_t(w,1) \\
+ \frac{2(w - 1)}{3N}(N - w + 1)h_t(w-1,1) + \frac{2w^2}{9N}h_t(w+1,1).
\]

We can define \( h_t(w) \equiv h_t(w,0) + h_t(w,1) \) and by adding both of Eqs. (S45) and (S46) we get a closed equation for \( h_t(w) \)

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

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

\[
\partial_t h(w,t) = -\partial_w(D_1(w)h(w,t)) + \partial^2_w(D_2(w)h(w,t))
\]

where we rescaled time \( g^2 t \to t \). The drift and diffusion coefficients are

\[
D_1(w) = \frac{2(4 + w + 3Nw - 4w^2)}{9N},
\]

\[
D_2(w) = \frac{-3 + 3N(w - 1) + 7w - 2w^2}{9N}.
\]

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

\[
\partial_t h(\phi,t) = -\partial_{\phi}\left(\frac{2}{3} \left( \phi - \frac{4}{3} \phi^2 \right) h(\phi,t) \right) + \partial^2_{\phi}\left( \frac{\phi}{3N} - \frac{2}{9N} \phi^2 \right) h(\phi,t)
\]

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

\section*{IV. Additional Details on the Time-Evolution of \( h(w) \)}

In this section, we provide additional numerical and analytical details regarding the probability weight distribution. In Fig. S1, we plot snapshots of \( h(w) \) at different times, computed numerically using the exact master equation. The initial distribution expands quickly and becomes extremely broad at intermediate times, before finally reaching the steady-state, which as we show below, is, to a good approximation, a Gaussian centered at \( w = 3N/4 \) with a width \( \Delta w/N \propto 1/\sqrt{N} \).

\subsection*{S1. Stationary Solution for \( h(w) \)}

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

\[
-\partial_{\phi}\left[D_1(\phi)h(\phi)\right] + \partial^2_{\phi}\left[D_2(\phi)h(\phi)\right] = 0,
\]

where

\[
D_1(\phi) = \frac{2}{3} \phi \left( 1 - \frac{4}{3} \phi \right), \quad D_2(\phi) = \frac{\phi}{3N} \left( 1 - \frac{2}{3} \phi \right).
\]

Integrating out Eq. S52 we obtain

\[
- D_1(\phi)h(\phi) + \partial_{\phi}\left[D_2(\phi)h(\phi)\right] = C.
\]
Equation (S54) can be rewritten as

\[
\partial_\phi h(\phi) = \left( \frac{D_1(\phi) - \partial_\phi D_2(\phi)}{D_2(\phi)} \right) h(\phi) + \frac{C}{D_2(\phi)}. \tag{S55}
\]

Solution of (S55) is straightforward:

\[
h(\phi) = \text{const} \times e^{J(\phi)} \int_0^\phi \frac{d\phi'}{D_2(\phi')} e^{-J(\phi')}, \tag{S56}
\]

\[J(\phi) = \int d\phi \frac{D_1 - \partial_\phi D_2}{D_2} = 4N\phi - \log \phi + (3N - 1) \log (3 - 2\phi).\]

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

\[
h(\phi) = \text{const} \times \frac{e^{NS(\phi)}}{(3 - 2\phi)^2} \int_0^\phi d\phi' e^{-NS(\phi')}, \tag{S57}
\]

where

\[S(\phi) = 4\phi + 3 \log (3 - 2\phi). \tag{S58}\]

In the limit \(N \to \infty\) the main contribution in the integral (S57) comes from the vicinity of the boundary point \(\phi = 0\). Expanding \(S(\phi)\) in Taylor series in powers \(\phi\): \(S(\phi) \approx S(0) + 2\phi\) and substituting it inside of the integrand in Eq. (S57) results in

\[h(\phi) \sim e^{NS(\phi)} \left[ 1 - e^{-2N\phi} \right]. \tag{S59}\]

Expression Eq. (S59) can be further simplified since \(e^{NS(\phi)}\) is strongly peaked in the vicinity of \(\phi_0 = 3/4\) which is the extremum of \(S(\phi)\): \(S(\phi) \approx S(\phi_0) + \frac{S''(\phi_0)}{2} (\phi - \phi_0)^2 + \ldots\), that gives

\[h(\phi) \sim \frac{e^{-8N(\phi-3/4)^2}}{\phi(3 - 2\phi)} \left[ 1 - e^{-2N\phi} \right]. \tag{S60}\]