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 $\mathrm{tr}(\mathcal{S}) = q^N\delta_{\mathcal{S},1}, \mathrm{tr}(\mathcal{S}\mathcal{S}') = q^N\delta_{\mathcal{S}\mathcal{S}'}$. 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 $a_{\mathcal{S}}(t) = q^{-N}\,\mathrm{tr}(\mathcal{O}(t)\mathcal{S})$, we can write $a^2_{\mathcal{S}}(t)$ in terms of the coefficients at the previous time step

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

Thus, we want to evaluate the quantity

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

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

Using properties of trace, we can write

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

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

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

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 $\text{tr}(\mathcal{I}_3(\mathcal{S}\otimes\mathcal{S}))$ 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]

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

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

From this, it follows that

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

where we used $\text{tr}(\mathcal{S}_r)=q\delta_{\mathcal{S}_r,1}$ and $\text{tr}(\mathcal{S}_r\mathcal{S}'_r)=q\delta_{\mathcal{S}_r,\mathcal{S}'_r}$. 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

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

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

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 $U_{\text{II}}\otimes U_{\text{II}}$ on Eq. (S12), giving us $\mathcal{I}_2$.

2. Then, we need to apply the layer of single-site Haar unitaries $U_{\text{I}}\otimes U_{\text{I}}$, 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

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

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

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

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

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 appearing 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 $U_{\rm II}\otimes U_{\rm II}$ and $U_{\rm I}\otimes U_{\rm I}$ to a single string of this form. Note that the result of applying $U_{\rm II}\otimes U_{\rm II}$, $U_{\rm I}\otimes U_{\rm I}$, 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 $U_{\rm II}$ and can be absorbed into $U_{\rm I}$, 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 $U_{\rm II}$, we can use the fact that $[X_iX_j, Z_iZ_j]=0$. We get

$$(U_{\rm II}\otimes U_{\rm II})(\prod_{i\in\Lambda}X_i\bar{X}_i\prod_{j\in\Sigma}Z_j\bar{Z}_j)(U_{\rm II}\otimes U_{\rm II})^\dagger=\prod_{i\in\Lambda}\left[(X_i\cos(Q_\Lambda)+Y_i\sin(Q_\Lambda))\big(\bar{X}_i\cos\big(\bar{Q}_\Lambda\big)+\bar{Y}_i\sin\big(\bar{Q}_\Lambda\big)\big)\right]\prod_{j\in\Sigma}Z_j\bar{Z}_j \tag{S16}$$

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 $\big\langle(U_{\rm I}\otimes U_{\rm I})A(U_{\rm I}\otimes U_{\rm I})^\dagger\big\rangle$ 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

$$\left\langle(U_{\rm II}\otimes U_{\rm II})(\prod_{i\in\Lambda}X_i\bar{X}_i\prod_{j\in\Sigma}Z_j\bar{Z}_j)(U_{\rm II}\otimes U_{\rm II})^\dagger\right\rangle=\prod_{i\in\Lambda}V_i\left\langle\cos^{|\Lambda|}(R_\Lambda)\prod_{j\in\Sigma}Z_j\bar{Z}_j\right\rangle \tag{S17}$$

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

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

$$\mathcal{I}_3=\delta_{\mathcal{S}',\mathcal{S}''}\sum_{\Omega_{\mathcal{S}'}\subset\{1,2,\cdots,N\}}\left(\prod_{j\notin\Omega_{\mathcal{S}'}}\delta_{\mathcal{S}'_j,1}\right)\left(\prod_{i\in\Omega_{\mathcal{S}'}}\frac{1-\delta_{\mathcal{S}'_i,1}}{3}\right)\sum_{\Lambda\subset\Omega_{\mathcal{S}'}}2^{|\Lambda|}\left(\prod_{m\in\Lambda}V_m\right)\left\langle\cos^{|\Lambda|}(R_\Lambda)\prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda}Z_n\bar{Z}_n\right\rangle. \tag{S18}$$

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 $\cos^k(x)=\frac{1}{2^k}\sum_{n=0}^k\binom{k}{n}\cos[(2n-k)x]$,

$$\left\langle\cos^{|\Lambda|}(R_\Lambda)\prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda}Z_n\bar{Z}_n\right\rangle=\frac{1}{2^{|\Lambda|}}\sum_{l=0}^{|\Lambda|}\binom{|\Lambda|}{l}\left\langle\cos((2l-|\Lambda|)R_\Lambda)\prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda}Z_n\bar{Z}_n\right\rangle. \tag{S19}$$

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 $D_j,\quad j\in\Omega_{\mathcal{S}'}\setminus\Lambda$ so that $R_\Lambda=R_{\Lambda\cup\{j\}}+D_j$. We then use the trig identity

$$\cos((2l-|\Lambda|)R_\Lambda)=\cos\big((2l-|\Lambda|)R_{\Lambda\cup\{j\}}\big)\cos((2l-|\Lambda|)D_j)-\sin\big((2l-|\Lambda|)R_{\Lambda\cup\{j\}}\big)\sin((2l-|\Lambda|)D_j). \tag{S20}$$

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

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

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

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

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

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

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

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

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

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

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 $\operatorname{tr}(V_i(\mathcal{S}_i\otimes\mathcal{S}_i)) = \frac{4}{3}(1-\delta_{\mathcal{S}_i,1})$, gives

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

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

$$W=\frac{1}{2^{2N}}\left(\frac{1}{3}\right)^{|\Omega_{\mathcal{S}'}|}\sum_{\Lambda\subset\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}}\left(\frac{4}{3}\right)^{|\Lambda|}\Bigg[\sum_{l=0,2l\neq|\Lambda|}^{|\Lambda|}\binom{|\Lambda|}{l}\left(4\cos^2((2l-|\Lambda|)g')\right)^{N-|\Omega_{\mathcal{S}}\cup\Omega_{\mathcal{S}'}|}$$
$$\times\left(\frac{4}{3}\sin^2((2l-|\Lambda|)g')\right)^{|\Omega_{\mathcal{S}}\setminus\Omega_{\mathcal{S}'}|}\times\left(\frac{4}{3}\cos^2((2l-|\Lambda|)g')\right)^{|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|-|\Lambda|}\times\left(4\sin^2((2l-|\Lambda|)g')\right)^{|\Omega_{\mathcal{S}'}\setminus\Omega_{\mathcal{S}}|}$$
$$+\delta_{2l,|\Lambda|}\binom{|\Lambda|}{|\Lambda|/2}\prod_{t\notin\Omega_{\mathcal{S}'}}(4\delta_{\mathcal{S}_t,1})\prod_{n\in\Omega_{\mathcal{S}'}\setminus\Lambda}\left(\frac{4}{3}(1-\delta_{\mathcal{S}_n,1})\right)\Bigg].\tag{S28}$$

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)

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

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

$$\begin{aligned}W_{\mathcal{S},\mathcal{S}'}&=W(|\Omega_{\mathcal{S}}|,|\Omega_{\mathcal{S}'}|,|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|)\\&=\left(\frac{1}{3}\right)^{|\Omega_{\mathcal{S}'}|+|\Omega_{\mathcal{S}}|}\sum_{k=0}^{|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|}\binom{|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|}{k}\Bigg[\sum_{l=0,2l\neq k}^{k}\binom{k}{l}\left[\cos^2((2l-k)g')\right]^{N-k-(|\Omega_{\mathcal{S}}|+|\Omega_{\mathcal{S}'}|-2|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|)}\\&\quad\times\left[\sin^2((2l-k)g')\right]^{|\Omega_{\mathcal{S}}|+|\Omega_{\mathcal{S}'}|-2|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|}+\delta_{2l,k}\delta_{|\Omega_{\mathcal{S}}|+|\Omega_{\mathcal{S}'}|-2|\Omega_{\mathcal{S}}\cap\Omega_{\mathcal{S}'}|,0}\binom{k}{k/2}\Bigg],\end{aligned}\tag{S30}$$

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),

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

We now claim that for these initial conditions, the probabilities $\left\langle a_{\mathcal{S}}^2(t)\right\rangle$ 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,

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

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

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

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

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

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

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

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

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

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

$$\mathcal{R}(w,w_1,w',w_1') = 3^w \sum_{m=\max\{0,w+w'-N+1-w_1-w_1'\}}^{\min\{w-w_1,w'-w_1'\}} \binom{w'-w_1'}{m}\binom{N-1-w'+w_1'}{w-w_1-m} W(w,w',m+w_1w_1'), \tag{S37}$$

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

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

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

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

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

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

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

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

The higher order terms will scale at most like $O(g^4N^2/N^2) = O(g^4)$ 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

$$\begin{aligned} \mathcal{R}(w, 0, w', 0) = {} & \delta_{w,w'}3^wW(w, w', w') + \delta_{w,w'+1}3^w(N - w' - 1)W(w, w', w') \\ & + \delta_{w,w'-1}3^ww'W(w, w', w' - 1) + O(g^4), \end{aligned} \tag{S43}$$

$$\mathcal{R}(w, 1, w', 0) = \delta_{w,w'+1}3^wW(w, w', w') + O(g^4), \tag{S44}$$

$$\mathcal{R}(w, 0, w', 1) = \delta_{w,w'-1}3^wW(w, w', w' - 1) + O(g^4), \tag{S45}$$

$$\begin{aligned} \mathcal{R}(w, 1, w', 1) = {} & \delta_{w,w'}3^wW(w, w', w') + \delta_{w,w'+1}3^w(N - w')W(w, w', w') \\ & + \delta_{w,w'-1}3^w(w' - 1)W(w, w', w' - 1) + O(g^4). \end{aligned} \tag{S46}$$

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:

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

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

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^2t 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 $r > 1$

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

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

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

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

where we used $\mathrm{tr}(\rho_\infty \mathcal{S}\mathcal{S}') = \delta_{\mathcal{S}\mathcal{S}'}$ 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

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

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

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

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

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

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

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

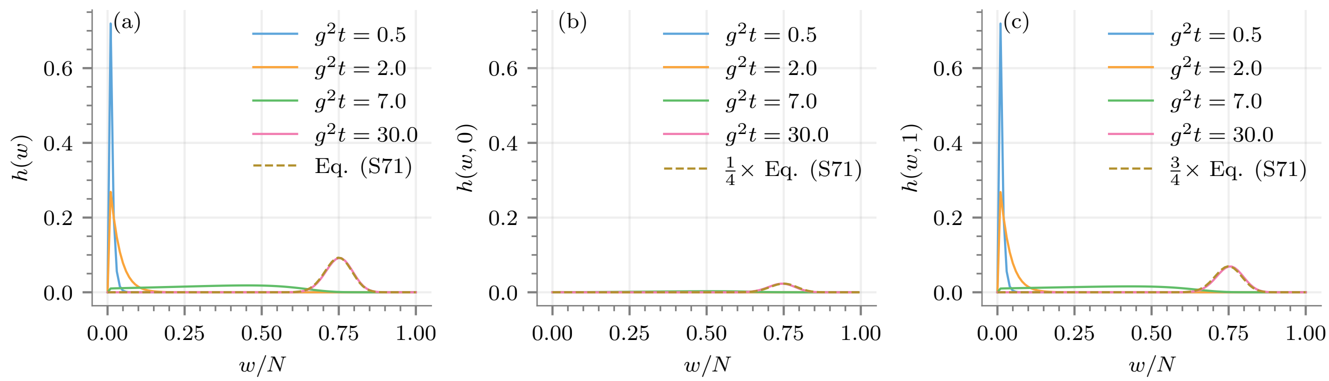

FIG. S1: (a) Snapshots of the numerically computed total probability weight distribution $h(w) = h(w,0) + h(w,1)$ for $g = 0.1$ and $N = 100$, together with the analytical expression of the steady-state from Eq. (S71), which essentially agrees with the $g^2t = 30$ numerics. (b) The same plot for $h(w,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

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

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 $U = U_{\rm I} U_{\rm II} U_{\rm I}$. 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).