The entanglement membrane in chaotic many-body systems

Tianci Zhou\textsuperscript{1} and Adam Nahum\textsuperscript{2}

\textsuperscript{1}Kavli Institute for Theoretical Physics, University of California, Santa Barbara, CA 93106, USA
\textsuperscript{2}Theoretical Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom

(Dated: January 1, 2020)

In certain analytically-tractable quantum chaotic systems, the calculation of out-of-time-order correlation functions, entanglement entropies after a quench, and other related dynamical observables, reduces to an effective theory of an “entanglement membrane” in spacetime. These tractable systems involve an average over random local unitaries defining the dynamical evolution. We show here how to make sense of this membrane in more realistic models, which do not involve an average over random unitaries. Our approach relies on introducing effective pairing degrees of freedom in spacetime, inspired by the structure emerging in random unitary circuits. We apply this approach to some translationally invariant Floquet spin chains studied in the literature. We show that a consistent line tension may be defined for the entanglement membrane, and that there are qualitative differences in this tension between generic models and “dual-unitary” circuits. These results allow scaling pictures for out-of-time-order correlators and for entanglement to be taken over from random circuits to non-random Floquet models. We also provide an efficient numerical algorithm for determining the entanglement line tension in 1+1D.

I. INTRODUCTION

This paper is about universality in the dynamics of chaotic many-body systems. One familiar type of universality is encapsulated in hydrodynamics for conserved quantities and other slow modes [1, 2]. But random circuits [3–17], a family of tractable many-body systems, have suggested new aspects of universality [6–11, 14]. In particular they have revealed a generic structure associated with an emergent “membrane” in spacetime [6, 18]. The effective statistical mechanics of this membrane determine the production of entanglement after a quantum quench, the spreading of quantum operators, and other aspects of the “scrambling” [7, 8, 11, 19–33] of quantum information.

One way to motivate this membrane, which is relevant to the approach we take here, is via the multi-layer structure of the quantum circuit (or the multi-sheet structure of the path integral, in a continuum language) representing the observables in question. Algebraically, conventional correlation functions, such as an expectation value $\langle O(t) \rangle$ following a quench, involve a single copy of the unitary time evolution operator $U(t)$, and a single copy of its Hermitian conjugate $U^\dagger(t)$. Formally, we can represent them in terms of matrix elements of the operator $U(t) \otimes U^\ast(t)$. But the quantities mentioned above (for example the $N$th Rényi entropy, obtained by tracing the $N$th power of the reduced density matrix) require matrix elements of the multiply “replicated” operator

$$U^{(N)} \equiv U \otimes U^\ast \ldots U \otimes U^\ast$$

with $N > 1$ copies of $U$, and $N$ copies of $U^\ast$. In a path integral language, we require multiple forward and backward paths [20, 26].

The structure arising from this is most easily understood in random unitary circuits, which are chaotic models built from random unitary gates [6–8, 11, 14]. Averaging $U^{(N)}$ in Eq. 1 over the random unitaries introduces a degree of freedom $\sigma(x,t)$, at each location in spacetime, that labels a pairing between the set of forward replicas and the set of backward replicas [14]. Formally this pairing is a permutation, $\sigma \in S_N$, and quantities like the Rényi entropies and the out-of-time ordered correlator (OTOC) map to effective partition functions for $\sigma(x,t)$. In this setting, the membrane can be understood as a domain wall between different values of the pairing field $\sigma(x,t)$ [7, 11, 14].

The basic quantity characterizing the membrane is its coarse-grained “line tension” $\mathcal{E}(v)$ [18]. This line tension determines the system’s butterfly velocity or operator spreading speed and its entanglement growth rate. It can be thought of loosely as associating an entanglement cost with a given curve in spacetime (in general this tension depends on $N$). See Fig. 1 for a schematic picture. We expect that the membrane description continues to make sense in systems more general and more realistic than random circuits, for the heuristic reason described in the

\begin{equation}
S(x,t) = \min_{\text{path}} \int s_\nu \mathcal{E}(v) dt + S(y,0)
\end{equation}

Figure 1. The membrane picture of entanglement and other dynamical quantities. (a) Example of an entanglement membrane for evaluating Rényi entropy growth in a 1D quench. Minimizing the membrane tension gives the entanglement. The tension is a function of local velocity $v = \dot{x}$. (b) Schematic line tension function for a chaotic system with parity symmetry. $\mathcal{E}(v)$ is convex and tangential to $|v|$ at $(\pm v_B, v_B)$.
next paragraph. In support of this, there is numerical evidence that a consistent membrane tension can be defined in a generic spin chain [18], and in a dual-unitary circuit [34]. There are analytic computations in random Floquet circuits in the large Hilbert space dimension (large $q$) limit [11], showing that the domain wall structure makes sense in that limit. Finally, there is an analytical derivation of the form of the membrane tension in holographic systems [35], at leading order in the number of degrees of freedom; this also begins to make a concrete connection between the entanglement membrane, and the Ryu-Takayanagi/Hubeny-Rangamani-Takayanagi geometrical pictures for entanglement in AdS space [36–38]. (Lattice toy models for the AdS-CFT correspondence, using random tensor networks, also exhibit domain walls between permutation degrees of freedom [49?].) But at present there is no formalism for explaining how the pairing degrees of freedom emerge when there is no average over randomness, or how to compute the properties of the membrane in a generic system without randomness and away from any large-$N$-like limit. As a more general point, random circuits and related stochastic models have shed light on various aspects of chaotic dynamics [6–17], and a natural question is how to generalize these calculations to non-random systems. Our aim here is to provide a formalism for doing this.

Heuristically, the paired configurations mentioned above can be thought of as saddle points of the path integral for the replicated evolution operator (1). This path integral contains multiple forward and backward paths. In a continuum language, these are weighted by $e^{-iS}$ and $e^{+iS}$ respectively, where $S$ is the action; in the discrete setting, $e^{-iS}$ is replaced by a product of matrix elements of local unitaries. By pairing each forward path with a backward path we can cancel the phases in the exponent, so we might guess that paired configurations predominate. In the cases of interest to us, however, the boundary conditions force the pairing to differ in different regions of spacetime. This leads to saddle-point solutions with a nontrivial membrane structure.

In this work we show explicitly, for a large class of systems without conservation laws, how the pairing field $\sigma(x, t)$ and the entanglement membrane emerge, making the above heuristic picture precise. We emphasize that we do not require either a limit of large local Hilbert space dimension, or any kind of randomness. To demonstrate the utility of our approach, we compute the entanglement line tension (for $N = 2$, the case relevant to the OTOC and to the second Rényi entropy) for a variety of translationally-invariant Floquet spin-1/2 chains.

Formally, our approach is to work in the replicated Hilbert space (acted on by Eq. 1) and to insert, in each time slice, a local resolution of identity: this includes states $|\sigma\rangle$ associated with pairings, but also states in the complement of the Hilbert space, which we denote by $\perp$. Traces of $U^{(N)}$, yielding for example correlation functions or Rényi entropies, can then be written as partition functions for an effective “spin” $s(x, t)$ which can take the values either $s = \sigma$ (a permutation) or $s = \perp$.

Including these non-pairing states allows the structures that appear in the random circuit to be generalized to models without any random average. The Haar-averaged random circuit can be thought of as a special case: there, the contributions from configurations with $\perp$ vanish after averaging, and the membrane is a domain wall between permutations [7, 14]. In, say, a translationally-invariant system, the domain wall structure becomes more complex, with the appearance of the spin value $s = \perp$ along the domain wall. (In fact, this is also true in a particular realization of the random circuit, before we average over the random unitaries.) This leads to a thickened domain wall with a more complex structure. However we argue that in chaotic systems, the domain wall’s width remains of order 1 in terms of microscopic length scales. Therefore, on large scales, the membrane is still well-defined, but with a nontrivial, renormalized membrane tension function. We illustrate this situation in Fig. 2.

Our initial introduction of the pairing field is exact but formal, because the microscopic weight for a configuration $s(x, t)$ is in general complicated. To make progress we first argue that configurations with large unpaired regions (large regions of $\perp$) are exponentially suppressed. Then we show how to resum an infinite number of configurations, or “Feynman diagrams”, exactly, to give a simpler description of the entanglement membrane, in terms of a reduced set of parameters.

Before tackling translation-invariant systems, we analyze the case of a fixed realization of a random unitary circuit. This illustrates the basic mechanism that makes the approach possible — suppression of $\perp$ states by phase cancellation — in a tractable setting. Since the disorder realization is fixed, we cannot introduce the pairing degrees of freedom by Haar-averaging. Instead we apply the new method described here. Working at large but finite Hilbert local space dimension, we show that the membrane remains well-defined, and is thickened slightly by occasional appearances of $\perp$ states. This leads to
to a membrane that inhabits a disordered “potential” in spacetime. This picture reproduces our previous results on Kardar-Parisi-Zhang (KPZ) fluctuations in the Rényi entropy $S_2$ between different realizations of the random circuit [14]. But previously these results required the replica trick, which we are able to dispense with here.

Next we turn to Floquet dynamics of spin-$\frac{1}{2}$ chains, without any kind of randomness. In order to control the calculations, we introduce a systematic expansion in the maximal temporal extent of connected $\perp$ clusters in the effective spin model. A priori, this expansion does not have a small parameter. However we conjecture (based on the tractable case of the random circuit) that it is a convergent expansion in chaotic Floquet systems.

Using this observation, we construct an efficient numerical scheme for computing the membrane line tension $\mathcal{E}(v)$ using the expansion above: we truncate the maximum size of a connected $\perp$ cluster (but we do not require that these clusters are dilute), and we examine convergence as a function of the order at which we truncate.

The numerical results show good convergence for a wide range of chaotic circuits, including prototypical spin models for quantum chaotic systems, such as Floquet Ising models with longitudinal and transverse fields (kicked Ising models) [21, 39, 40].

The line tension $\mathcal{E}(v)$ satisfies some general constraints [18] which provide a check on our results. Recall that $\int dt \mathcal{E}(\dot{x}(t))$ is proportional to the “free energy”, in the scaling limit, of a membrane whose shape is parameterized by $x(t)$ (Fig. 1). $\mathcal{E}(v)$ is a convex function of velocity, which is tangential to the line $\mathcal{E} = \pm v$ at the point $(\pm v_B, v_B)$, where $v_B$ is the quantum butterfly velocity (for parity symmetric systems), see the schematic picture in Fig. 1(b). For the chaotic systems we study, our numerical approximations to $\mathcal{E}(v)$ indeed appear to converge a form satisfying the constraints above.

We also investigate a special case of the kicked Ising model that has maximal entanglement growth [34] and the property of “dual unitarity” which, remarkably, allows a range of exact computations [41–44]. We argue that the entanglement membrane has special properties in dual-unitary models: its equation of motion becomes hyperbolic, rather than being parabolic (diffusive) as it is in generic circuits.

II. DEFINING THE SPIN MODEL
We consider time evolution by a unitary circuit made of two-site gates, with the structure in Fig. 3. We will use $U(t)$ or simply $U$ to denote the full many-body unitary represented by the entire circuit, and the lower-case $u$ to stand for a local two-site gate.

At this point the local gates $u$ are left arbitrary. In some sections we will restrict to Floquet models with both space and time translation invariance, but this is not necessary. The circuit geometry guarantees that there is no propagation of information outside a strict lightcone with speed $v = 1$; the actual butterfly speed in these circuits is in general strictly smaller than 1.

Conventional time ordered correlation functions such as $\text{Tr} \mathcal{O}(x_n, t_n) \ldots \mathcal{O}(x_1, t_1)$ involve a ‘two-layer’ quantum circuit made of $U$ and $U^*$, in the sense that they can be obtained from the contraction of $U \otimes U^*$ by inserting operators in the intermediate time-steps. In contrast, the quantities of interest to us here require traces of the $2N$-layer unitary circuit

$$U^{(N)} = \prod_{\text{N Us and N U*s}} U \otimes U^*\ldots U \otimes U^*,$$

where $N > 1$. For example, the out-of-time order correlation function

$$C(x, t) = \langle \mathcal{O}(x, t)\mathcal{O}'(0, 0)\mathcal{O}(x, t)\mathcal{O}'(0, 0) \rangle$$

requires $N = 2$. The the Rényi entanglement entropy for subsystem $A$

$$S_n = -\frac{1}{n-1} \ln \text{tr}(\rho^n_A(t))$$

requires $N = n$ to write the $n$-th power of the time-evolved reduced density matrix.

Taking $N > 1$ leads to new structure. One consequence is that it multiplies the number of conservation laws, since formally each ‘copy’ of the system undergoes independent unitary evolution. This will not concern us here, however, since we study models without conservation laws. The additional structure we examine is instead associated with pairings between spacetime histories in the ‘worlds’ associated with different layers [14].

The multi-layer circuit has $N$ layers of $U$ and of $U^*$, and can be thought of as describing evolution of a ‘replicated’, multi-copy system. For a heuristic picture, we may think in terms of Feynman histories of this replicated system in our chosen local basis. The amplitude for a given Feynman history is given by a product of matrix elements of the local unitaries. Feynman histories that are ‘paired’, so that the $N$ histories occurring in the $U$ replicas are pairwise equal to the $N$ histories occurring in the $U^*$ replicas, avoid phase cancellation [14]. (This pairing can happen in $N!$ different ways.) We will argue that in quantum chaotic systems, the observables mentioned above are dominated by Feynman histories that have a well-defined pairing structure after coarse-graining. We will show how a corresponding local ‘pairing field’ $\sigma(x, t)$ can be defined by coarse-graining over microscopic length and time scales.

### A. Introducing the effective spins

To motivate the structures we will consider, we first recall what happens when the local unitaries are averaged over the Haar ensemble instead of being fixed.

For a given two-site gate $u$, acting on sites $i$ and $i+1$, the replicated gate $u^{(N)}$ is a unitary operator on $2N$ copies ($N$ layers of $u$, $N$ layers of $u^*$) of the two-site Hilbert space, which has dimension $d = q^2$. Its average (which is no longer unitary) is the operator

$$\overline{u^{(N)}_{\text{Haar}}} = \sum_{\tau, \sigma \in \mathcal{S}_N} \text{Wg}(\tau^{-1}\sigma; d) |\tau\rangle \langle \sigma\sigma|$$

where Wg is the unitary Weingarten function [45–47]. Here $\tau$ and $\sigma$ are elements of the permutation group $\mathcal{S}_N$. These label a pairing of the $u$ layers, labelled $1, 2, \ldots, N$, with the $u^*$ layers, labelled $1', 2', \ldots, N'$. For example $\sigma$ corresponds to pairing $1$ with $\sigma(1)$, etc. At a given spatial site, the state $|\sigma\rangle_i$ is a product of maximally entangled states between paired copies of the site, and the states above are for a pair of spatial sites: $|\sigma\sigma\rangle = |\sigma\rangle_i \otimes |\sigma\rangle_{i+1}$.

For example, for $N = 2$, there are two possible permutations: the identity $\mathbb{1}$ and the transposition $(12)$. In our local basis the corresponding states at a site are

$$\langle a_1, \overline{a_1}, a_2, \overline{a_2}|l\rangle = \delta_{a_2, \overline{a_2}} \delta_{a_1, \overline{a_1}},$$

$$\langle a_1, \overline{a_1}, a_2, \overline{a_2}|(12)\rangle = \delta_{a_2, \overline{a_2}} \delta_{a_1, \overline{a_1}}$$

(note that we do not normalize these states).

In Ref. 14 we showed how, starting with the above expression, the expression for $U^{(N)}$ could be reduced to a lattice magnet, with local interactions, for the permutations $\sigma$ (the permutations $\tau$ were integrated out in this mapping). At the end of this mapping, there is a single $\sigma$ spin associated with each local block in the circuit.

Now let us consider a general circuit with the brickwork geometry in Fig. 3. The basic idea will be to separate out the pairing states, $|\sigma\rangle$, in the multi-copy Hilbert space for a pair of sites, from states in orthogonal complement of this space. To simplify notation we will often write $|\sigma\rangle$ for the two-site state $|\sigma\sigma\rangle$.

Eq. 5 shows that in the Haar-averaged case contributions from the orthogonal complement vanish identically. More generally they do not vanish. However we will argue that the permutation states dominate in a certain sense, and that states in the orthogonal complement can be taken into account via a systematic procedure.

In fact, the expression for the Haar average of a local unitary in Eq. 5 can be regarded simply as the projection operator onto the subspace of permutation states. By the invariance of the Haar measure under (say) left multiplications,

$$\overline{u^{(N)}_{\text{Haar}}} \times \overline{u_{\text{Haar}}} = \overline{u_{\text{Haar}}}$$

showing that this is a projector. The state $|\sigma\rangle$ satisfies $u^{(N)}|\sigma\rangle = |\sigma\rangle$ for any choice of two-site unitary, so after
Haar averaging we also have
\[ u_{\text{Haar}}^{(N)}(\sigma) = |\sigma\rangle. \] (8)

This shows that the Haar averaged object, which is a Hermitian operator on the replicated space, is equal to \( P_\parallel \), the orthogonal projector onto the subspace spanned by permutation states [48].

To separate out contributions from permutation states for a general circuit, we will insert the resolution of the identity (on two physical sites)
\[ I^{(N)} = P_\parallel + P_\perp \] (9)
immediately before each two-site gate in the circuit. Here \( P_\perp \) is the projector onto the subspace complementary to that spanned by pairing states. This insertion is shown in (10) below as a bar below each gate.

We will view the quantity to be computed, for example the purity, as a partition function which we denote \( Z \). (In the purity example, the second Rényi entropy is the free energy associated with this partition function.) After using 9, this partition function is a sum of terms with all possible choices of projection operator inserted below each gate. Graphically,

\[ Z = \sum_{\parallel, \perp} \sum_{\text{choices of projection operators}} \] (10)

where the sum is over all possible assignments of \( \parallel, \perp \) to the projection operators, which are represented by horizontal bars. We leave the boundary conditions at the initial and final times unspecified at this point, since they depend on the observable to be computed.

Note that the operator \( P_\parallel \) can "absorb" an arbitrary two-site gate. This can be seen from the invariance of the Haar measure. For any 2-site unitary \( u \),
\[ u^{(N)} \times P_\parallel = u^{(N)} \times u_{\text{Haar}}^{(N)} = u_{\text{Haar}}^{(N)} = P_\parallel. \] (11)

As a result, the term in the partition function where all the insertions are \( P_\parallel \) is identical to the expression where all gates are replaced with Haar-averaged gates. Therefore the decomposition in Eq. 10 gives us a well-understood starting point, which we can go beyond by taking into account domains of \( \perp \) insertions. Roughly speaking, this approach will be useful if these domains do not form a connected cluster ‘percolating’ from the top to the bottom of the spacetime slab, but instead form finite clusters. Their effect can then be taken into account as a renormalization of the interactions for the permutation degrees of freedom that are present in the random circuit.

To make these permutations explicit we decompose \( P_\parallel \) using Eq. 5. In the random circuit it was useful to perform the sum over the \( \tau \) variables (cf. Eq. 5) explicitly, leaving a partition sum solely for the \( \sigma \) variables. Algebraically, that corresponds to splitting \( P_\parallel \) into projection operators \( P_\sigma \),
\[ P_\parallel = \sum_{\sigma \in S_N} P_\sigma, \] (12)

which are given by grouping together terms in Eq. 5 with a given \( \sigma \):
\[ P_\sigma = \left( \sum_{\tau} W_\text{g}(\tau^{-1}\sigma; q^2)|\tau\rangle\langle\tau| \right) \langle\sigma\rangle. \] (13)

Using the properties of the Weingarten function, one can check that the states appearing in the brackets
\[ \langle(\sigma\sigma)^* \rangle = \sum_{\tau} W_\text{g}(\tau^{-1}\sigma; q^2)|\tau\rangle\langle\tau| \] (14)
form the dual basis of \( |\sigma\rangle \), i.e. \( \langle(\sigma\sigma)^* |\tau\rangle = \delta_{\tau,\sigma} \), so that the non-Hermitian operators \( P_\sigma = |(\sigma\sigma)^*\rangle\langle\sigma\sigma| \) are projectors satisfying (App. A):
\[ P_\sigma P_\sigma' = \delta_{\sigma\sigma'} P_\sigma. \] (15)

Here we assume that \( q \) is generic, to avoid divergences in the Weingarten function, which can arise for small values of \( q \) when \( N \) is sufficiently large (for \( q = 2 \), such divergences can arise for \( N > 4 \)). This is not a fundamental obstacle, as we could always define a “spin” with a larger local Hilbert space dimension by grouping sites.

Our resolution of the identity, inserted below each block in the multi-layer circuit, is now refined to
\[ I^{(N)} = \sum_{\sigma \in S_N} P_\sigma + P_\perp. \] (16)

Let us represent the possibilities graphically as:

\[ P_\sigma = \sigma, \quad P_\perp = \ ] (17)

Importantly, the blocks on the left, labelled by permutations, are independent of the choice of local unitary gate, because \( P_\sigma \) again “absorbs” any unitary. For any gate \( u \), we have \( u^{(N)}|\sigma\rangle = |\sigma\rangle \), so
\[ u^{(N)} P_\sigma = P_\sigma. \] (18)

However the \( \perp \) blocks, which represent the tensor \( u^{(N)} P_\perp \), depend on the local unitaries defining the dynamics.

We have now written \( Z \) as a sum over spins \( s \), associated with each block in the circuit, which take values in
average over the physical random unitary in Eq. 5. The three-body interaction appears after integrating out the “τ” spins. Here the sum over the τ spins appeared at the state level in defining the projection operator in Eq. 13.

The simplest case is \( N = 2 \) [7]. There are then two possible permutations, \( \perp \) and (12), which we denote + and - . The weights are symmetric under exchange of + and - and under spatial reflection. They are

\[
\begin{align*}
\langle \perp \rangle &= 0, \\
\langle + \rangle &= 1, \\
\langle - \rangle &= K,
\end{align*}
\]

with

\[
K = \frac{q}{q^2+1},
\]

which follows from Eq. 13 with the explicit expression

\[
P_+ = \frac{1}{q^2-1} \left( \langle + \rangle - \frac{1}{q^2} \langle - \rangle \langle + \rangle \right)
\]

and symmetrically for \( P_- \). Note the vanishing of the first weight, which amounts to a hard constraint on the spin configurations. The spin configuration \( \{ s \} \) is in fact highly constrained for any \( N \): the underlying unitarity means that \( J \) vanishes for many configurations of the triangle. In other words, many terms in the sum over \( s \in S_N \) defining \( Z_{0\perp} \) are zero (see Ref. [17] and App. B for a further simplification). For example, as a generalization of the first equality in Eq. 23 above,

\[
\frac{\sigma_b \sigma_c \sigma_a}{\sigma_a \sigma_c \sigma_b} = \delta_{\sigma_a, \sigma_b}.
\]

This enforces a light cone structure. For example, for the purity calculation which we discuss below in Sec. II C, the boundary spins on the top are \( s_{x,t} = \perp \) for \( x < 0 \) and \( s_{x,t} = (12) \) for \( x \geq 0 \). This means that the spin configuration is only nontrivial within a backwards lightcone emanating from the entanglement cut at the top boundary [i.e., we must have \( s_{x',t+1} = \perp \) and \( s_{x',t} = (12) \)].

This light-cone structure is preserved when we include configurations with \( s = \perp \). If the two spins at the top of a triangle are the same permutation \( \sigma_b \), the lower spin cannot be \( \perp \), because by definition the projector onto \( \perp \) is orthogonal to \( |\sigma_a, \sigma_b\rangle \).

\[
\frac{\sigma_a \sigma_b \perp}{\sigma_a \perp \sigma_b} = 0.
\]

So again in the purity example above \( \perp \) can only be inserted in the backward lightcone emanating from the entanglement cut.

Let us discuss the locality of the spin interactions. We have seen that in the absence of \( \perp \) the Boltzmann weight factorizes into three-body terms representing interactions between three adjacent spins. Once we have large clusters
of \( \perp \) this is no longer true. However, the weight does factorize into a product of separate weights for each cluster of \( \perp \) (together with the local weights mentioned above): there is no interaction between disconnected clusters of \( \perp \). This is because surrounding a cluster of \( \perp \) with permutation states dictates a definite way of contracting up the \( \perp \) blocks, yielding a c-number. (We will give some explicit formulas below.) Therefore, if the partition function is dominated by configurations in which clusters of \( \perp \) have a finite typical size, locality will be regained after coarse-graining beyond this scale.

Finally, we note that the model we have defined has a global symmetry, with the symmetry group

\[
G = (S_N \times S_N) \times \mathbb{Z}_2.
\] (28)

This is a consequence of the fact that Eq. 2 involves multiple copies of the same unitary \([14, 49]\). The importance of \(S_N \times S_N\) symmetry has been emphasized in random tensor networks \([49]\), where permutations labelling pairings also appear \([?]\).

Graphically, the symmetry arises from the possibility of permuting the layers of the original unitary circuit without changing Eq. 2. We can permute the \(U\) layers among themselves with a permutation \(g_L \in S_N\), and we can also separately permute the \(U^*\) layers among themselves with another permutation \(g_R\). This gives the \(S_N \times S_N\) subgroup of \(G\), which acts on the spins via

\[
\sigma \rightarrow g_L \sigma g_R^{-1}, \quad \perp \rightarrow \perp. \quad (29)
\]

The \(\mathbb{Z}_2\) subgroup of \(G\) arises from the fact that the multilayer circuit is invariant if we exchange all the \(U\) layers with all the \(U^*\) layers, and also complex conjugate the circuit. The resulting symmetry acts by

\[
\sigma \rightarrow \sigma^{-1}, \quad \perp \rightarrow \perp. \quad (30)
\]

For both types of operation, the spin state \(\perp\) is invariant because these exchanges of layers preserve the \(\|\) and \(\perp\) subspaces of the two-site Hilbert space.

The simplest case is \(N = 2\), when \(G\) reduces to an Ising-like \(\mathbb{Z}_2\) symmetry relating \(+ = \|\) and \(- = (12)\). In this case the symmetry in Eq. 30 becomes trivial, since \(\sigma = \sigma^{-1}\).

In the sub-partition function \(Z_{0\perp}\) the weights \(J\) are invariant under \(G\) because the Weingarten function and the overlaps in Eq. 13 depend only on the cycle structure of products of the form \(\sigma \tau^{-1}\). This cycle structure is invariant under the above operations.

The \(G\) symmetry is a symmetry of the bulk interactions for the spin model. It will however be strongly broken by the boundary conditions we require.

\[\text{C. Domain wall structure}\]

As a useful illustrative example, let us consider the purity of a region \(A\), \(e^{-S_2(A)} = \text{tr}(\rho_A^2)\). We take \(A\) be a semi-infinite half-system to the left of the origin in an infinite chain. This quantity requires \(N = 2\). If we take the physical system to start off in the state \(|\psi\rangle\), and if we denote the replicated state in the four-copy Hilbert space by \(|\psi^{(2)}\rangle\), then

\[
\exp (-S_2) = \langle \ldots + + - - \ldots |U^{(2)}(t)|\psi^{(2)}\rangle,
\] (31)

where we have labelled the two permutation states at a site, \(\perp\) and \((12)\), by + and −. For simplicity we take the initial state to be a product state, \(|\psi\rangle = \otimes_x |e_x\rangle\).

Again let us first consider the sub-partition-function \(Z_{0\perp}\), which is equivalent to that for Haar-random unitaries treated in Ref. [7]. The spins have fixed boundary conditions at the top, enforcing a domain wall between + and −, and free boundary conditions at the bottom, because \(\langle +|e^{(2)}\rangle = \langle -|e^{(2)}\rangle = 1\). As a result of the fact that \(J(+, +; -) = J(−, −; +) = 0\) (the first diagram in Eq. 23) the only nonvanishing configurations in \(Z_{0\perp}\) contain a single directed domain wall emanating from the entanglement cut, separating an infinite domain of + on the left and an infinite domain of − on the right. This is illustrated in Fig. 4, top. Formally

\[
Z_{0\perp} = \sum_{\text{directed paths}} K^t = 2^t K^t,
\] (32)

where \(K\), defined in Eq. 24, is the weight for a single step in the directed walk.

Now we consider including the terms in the partition function with \(s = \perp\), in order to address models without a Haar average.
We have shown that the spins outside the backward light cone are equal and fixed by the final-time boundary condition (a consequence of causality) and that \( \perp \) can only appear within the backward light cone. Hence, we can define a ‘thick’ domain wall separating the connected infinite domain of \(+\) on the left from the connected infinite domain of \(\perp\) on the right. In principle, this thick domain wall could fill the entire backward light cone: for example the \(\perp\) blocks could form a large percolating cluster of \(O(t^2)\) size. If such configurations are dominant, then the representation of the partition function in Eq. 19 is not useful.

However, we will argue that this does not happen for a typical chaotic choice of the circuit \(U\), because of a phase cancellation effect which suppresses large \(\perp\) clusters. Instead the typical thickness of the domain wall remains of order \(1\) in \(t\) as \(t \to \infty\), so that in a typical configuration the domain wall has the schematic structure in shown in Fig. 4. The domain wall has a nontrivial structure made up of clusters that can contain \(\perp, +\) and \(\perp\); these clusters are separated by the places in spacetime where the domain wall becomes “thin”, which we will take to mean “of minimal width”. After coarse-graining we recover a membrane picture similar to that in the random circuit, and we can define a renormalized line tension for this membrane.

We will give analytical and numerical evidence for this picture in Secs. III, IV. In some cases, there may be a small parameter which makes the analytical calculation of renormalized line tension simple. We will argue in Sec. III that this is the case for typical choices of \(u\) at large local Hilbert space dimension \(q\). But even if there is no small parameter, the renormalized line tension can be well defined. We will see below how to extract it from simulations.

We now rewrite the partition function of the domain wall in a way which will be useful when the width of the thick domain wall does not grow with \(t\).\(^1\) As noted above, we then expect that a typical configuration has order \(t\) locations where the domain wall is thin. We define a directed path that is made of steps connecting these locations: i.e. steps of length \(t \geq 1\) connecting the “pinch points” in Fig. 4(b). The definition of one of these “irreducible steps” is that domain wall is thin at the beginning and end of an irreducible step, but nowhere in between.

The weight for a given domain wall configuration can then be written as a product of weights \(W(\Delta x, \Delta t)\) for each step of extent \(\Delta t\) in the time direction and \(\Delta x\) in the space direction. Factorization into such a product follows from the locality property mentioned in the previous subsection (in the paragraph following Eq. 27).

To simplify the formulas, we restrict for now to systems with space and time translation symmetries such that all the local gates are identical, but the generalization to other cases is direct (Sec. III).

Define a partition function \(Z(x, y; t)\) with a modified bottom boundary condition, such that there is a domain wall between \(+\) and \(\perp\) at position \(x\) at the top boundary, and at \(y\) at the bottom boundary. This can be enforced using the dual states \(|\pm\rangle\) defined in Eq. 14:

\[
Z(x, y; t) = \langle \cdots + + - - \cdots | U^{(2)}(t) | \cdots + + - - \cdots \rangle
\]

\((x\text{ labels bonds of the lattice as in Eq. 19.})\) Let the number of timesteps where the domain wall is thin be \(N + 1\), and let their spacetime coordinates be:

\[
(y, 0), \ (x_1, t_1), \ \ldots, \ (x_{N-1}, t_{N-1}), \ (x, t).
\]

The steps are of time duration and spatial extent

\[
\Delta t_i = t_i - t_{i-1} \geq 1,
\]

\[
\Delta x_i = x_i - x_{i-1} \in \mathbb{Z}.
\]

For our choice of lattice geometry, \(\Delta x + \Delta t\) is necessarily even. Defining the weight for an irreducible step to be \(W(\Delta x, \Delta t)\), the total partition function is given by summing over paths of all possible lengths \(N\):

\[
Z(x, y; t) = \sum_{N \geq 1} \sum_{(\Delta x_i, \Delta t_i)} \prod_{i=1}^{N} W(\Delta x_i, \Delta t_i).
\]

Because of translational invariance \(Z\) only depends on one spatial argument, so we will also write \(Z(x, t) \equiv Z(x, 0; t)\).

The weight for a step of duration \(1\) is

\[
W(\pm 1, 1) = K.
\]

The simplest case is for \(Z_{0\perp}\), when these are the only steps allowed:

\[
W_{0\perp}(x, t > 1) = 0.
\]

The \(\perp\) insertion produces non-zero values for \(W(x, t > 1)\), which may be either positive or negative. The presence of negative steps is not necessarily an obstacle to defining a coarse-grained line tension. If positive steps predominate then these negative weights will disappear under coarse-graining. However we will suggest that in some fine-tuned cases negative weights are important (Sec. V D).

The rewriting in Eq. 37 is exact for the boundary conditions we have chosen, but this rewriting will only be useful if two conditions are satisfied. First, that the irreducible step weights decay sufficiently fast for large \(\Delta t\); at a minimum we require that the ratio \(W(x, t)/Z(x, t)\) tends to zero at large \(t\). Second we require that the original partition function of interest, where the lower boundary condition is free, is simply related to \(Z(x, y; t)\).
To be more precise, we should distinguish between different kinds of usefulness. First, we conjecture that the above is useful for generic chaotic models for deriving a coarse-grained picture that is in the right universality class. Second we will argue that for some strongly chaotic models the above representation is also practically useful for numerical determination of quantities like the line tension \( \mathcal{E}(v) \) and the butterfly speed \( v_B \). Third, in a more restricted class of models with a large parameter, the above can be used to obtain these quantities analytically.

We will describe the numerical algorithm that follows naturally from the above representation in Sec. IV. Its starting point is to use the recursive form of Eq. 37 to compute \( W(x, t) \):

\[
W(x, t) = Z(x, t) - \sum_{t' < t} \sum_{y} W(x - y, t - t')Z(y, t').
\]  

(40)

This expression follows simply from splitting the partition function \( Z(x, t) \) for the directed path into the term \( W(x, t) \) from the single-step path with \( N = 1 \), and from paths with \( N > 1 \) whose final irreducible step has weight \( W(x - y, t - t') \). Since \( Z(x, t) \) can be computed numerically by taking appropriate traces of \( U(t)^{(N)} \), we can compute the irreducible step weights \( \mathcal{E}(v) \) by starting with the weight for \( t = 1 \) in Eq. 38 and employing Eq. 40 recursively to get the irreducible weights for larger \( t \). The weights can then be used to extract \( \mathcal{E}(v) \) and \( v_B \).

In practice we will be limited to \( t \leq t_{\text{max}} \) for some \( t_{\text{max}} \), so the usefulness of the algorithm will depend on how rapidly the weights \( W \) get smaller at larger \( t \). We apply the algorithm to some non-random but chaotic Floquet models in Sec. V, with encouraging results. The results support the universal picture described in this section, with a domain wall that is not microscopically thin but has an order 1 width.

The approach can be modified in many ways which do not change the basic structure but which might improve the practical usefulness of the algorithm for a given choice of model. For example, we can insert projection operators on single sites instead of on double sites, yielding a slightly modified spin model. This has the advantage that the weight of a step of duration \( \Delta t = 1 \) is no longer independent of the gate \( u \) (which it is in the present setup, see Eq. 38). It is clear that at least for some choices of \( u \) this will give better approximations for small \( t_{\text{max}} \). Nevertheless here we will stick with the geometry above, since it is convenient for making contact with the random circuit case, and is sufficient to illustrate the basic ideas.

### D. Weights for smallest nontrivial \( \perp \) cluster

In some limits it is a good approximation to consider only the smallest possible \( \perp \) cluster, namely a single isolated \( \perp \) block (Sec. III). We can easily write down the weights for such a minimal cluster. They depend on the singular value structure of the local unitary gate \( u \) where \( \perp \) is inserted.

Recall that

\[
\begin{align*}
\begin{array}{c}
\uparrow \\
\perp \\
\downarrow \\
\end{array}
&= 0.
\end{align*}
\]  

(41)

as a consequence of unitarity (Eq. 27). So locally, there are four configurations for one isolated \( \perp \), up to \( + \leftrightarrow - \) symmetry:

\[
\begin{array}{cccc}
\begin{array}{c}
\uparrow \\
\downarrow \\
\end{array} & \begin{array}{c}
\downarrow \\
\uparrow \\
\end{array} & \begin{array}{c}
\downarrow \\
\downarrow \\
\end{array} & \begin{array}{c}
\downarrow \\
\uparrow \\
\end{array}
\end{array}
\]  

(42)

In fact only three of these are independent, as \( 2 \) and \( 3 \) are equal for any \( u \), even if it is not reflection-symmetric. Explicit formulas are given below in Eq. 46.

For the translationally invariant case, the first three diagrams above give the irreducible domain wall weights \( W(x, t) \) defined in the previous subsection for \( t = 2 \):

\[
W(0, 2) = 1, \quad W(\pm 2, 2) = 2.
\]  

(43)

By contrast \( 4 \) does not contribute to \( W(x, 2) \), as it does not yield a thin domain wall at the bottom. This diagram will appear only in configurations contributing to \( W(x, t > 2) \). In the limit discussed in Sec. III, namely typical unitaries at large \( q \), only \( 1 \) is required at leading nontrivial order in \( 1/q \).

In order to express the values of \( 1 \) to \( 4 \), recall that we can regard the gate \( u \) as a quantum state for four \( q \)-state spins (“vectorisation” of the operator). In a tensor network language, this simply means that we regard all four of the legs sticking out of \( u \) as physical spin indices. Let us label these legs \( A, B, C, D \) as

\[
u = \begin{array}{c}
A \\
B \\
C \\
D \\
\end{array}.
\]  

(44)

Regarding \( u \) as a state makes it clear that (after normalizing this state) we can define the entanglement between any subset of \( \{A, B, C, D\} \) and the complement following the usual prescription for states. These are referred to as operator entanglements [50–54], and have been studied in the context of quantifying the “entangling power” of unitary gates [50, 55]. Unitarity implies that \( \{A, B\} \) is maximally entangled with \( \{C, D\} \), but other entanglements depend on the gate \( u \). For example, if \( u \) is close to the identity then \( \{A, C\} \) is weakly entangled with \( \{B, D\} \).

For \( N = 2 \) we require the purities \( P = e^{-S_2} \)

\[
P_{\{1\}} \equiv P_{AC}, \quad P_{\{X\}} \equiv P_{AD}.
\]  

(45)

All others are equivalent (e.g. \( P_{BD} = P_{AC} \)) or independent of \( u \) (e.g. \( P_A = q^{-1}, P_{AB} = q^{-2} \)). In terms of these,
the weights in Eq. 42 are (recall $3 = 2$):

\begin{align}
\text{(1)} &= \left( \frac{q^4}{q^4 - 1} \right)^2 \left( P_{11} + \frac{1}{q^6} P_X - \frac{2K}{q} - \frac{2K}{q^3} \right) \\
\text{(2)} &= \left( \frac{q^3}{q^4 - 1} \right)^2 \left( P_{11} + \frac{1}{q^3} P_X - \frac{2K}{q} - \frac{2K}{q^3} \right) \\
\text{(4)} &= \left( \frac{q^3}{q^4 - 1} \right)^2 \left( P_X + \frac{1}{q^2} P_{11} - \frac{2K}{q} - \frac{2K}{q^3} \right)
\end{align}

See App. D for computations.

We now describe an analytically tractable case where it is sufficient to consider dilute $\downarrow\uparrow$ insertions of the above form (Sec. III). Then in Sec. IV we consider Floquet models for spin-$1/2$s, where it is important to allow for larger clusters of $\downarrow\uparrow$.

III. A TRACTABLE LIMIT: LARGE LOCAL HILBERT SPACE DIMENSION

We now apply the above formalism to a typical realization of a random unitary circuit at large $q$. This is a useful test ground for our approach, because we can quantify the suppression of large clusters of $\downarrow\uparrow$ blocks analytically.

It is important here to distinguish between the Haar average of $e^{-S_2}$ or of $U(t)^{\otimes 2}$, and a typical individual realization drawn from the random circuit ensemble. The latter represents a specific chaotic time evolution, with much more structure than is captured by simply averaging $U(t)^{\otimes 2}$ over the ensemble.

If we average $U(t)^{\otimes 2}$, the weight of any configuration with a $\downarrow\uparrow$ is set to zero, leading to $Z_{0\downarrow\uparrow}$. But in a given realization of the circuit this is not the case, and the $\downarrow\uparrow$ insertions have an important effect. For example, they modify the growth rate of averaged entropy $S_2$, because this quantity is not equal to the simpler “annealed average” $-\ln e^{-S_2}$.

Consider a particular spin configuration in the partition sum $Z$ (Eq. 19). As noted in Sec. II B, each cluster of $\downarrow\uparrow$ blocks in the configuration contributes separately to the Bolzmann weight of the spin configuration. We denote the weight of a given cluster $C$ by $\Omega_C$. It depends on the geometry of the cluster, the spins $\uparrow\uparrow$ on its boundary, and the local unitaries in the spacetime region inside the cluster.

If we average $\Omega_C$ over the random unitaries inside the cluster region, then we obtain $\Omega_C^{\backprime\backprime} = 0$ because of cancellation between positive and negative values. This is just the statement that, after Haar averaging, $Z$ becomes equal to $Z_{0\downarrow\uparrow}$, where no $\downarrow\uparrow$s appear.

However, we can ask what the magnitude of $\Omega_C$ is for a typical realization of the circuit. We can quantify this by computing $\Omega_C^{\backprime\backprime}$, which is not zero. We show below that (at least for large $q$, where the calculation is controlled) $\Omega_C^{\backprime\backprime}$ is exponentially suppressed in the temporal duration of the cluster, with the exponential suppression getting stronger and stronger as $q$ is increased.

Therefore the Haar circuit at large $q$ is a setting where we can put the phase cancellation conjecture on a qualitative footing. We do this next (Sec. III A). The present approach to the random circuit also allows us to compute statistical fluctuations in the second Rényi entropy due to circuit randomness in Sec. III B. We recover the result that these are in the KPZ/DPRM universality class. Previously this result required the replica trick [14]. Here however we can make the mapping to a directed polymer at the level of an individual circuit, so the correspondence with known classical problems can be made without use of the replica trick (Sec. III B).

A. Exponential suppression of large $\downarrow\uparrow$ clusters

Let us consider an isolated cluster $C$ of $\downarrow\uparrow$ blocks. For simplicity we consider a cluster with trivial topology like that shown in Fig. 5(a), but a general cluster with holes inside can be treated similarly. Eq. 41 shows that the weight of the spin configuration vanishes if any $\downarrow\uparrow$ block has two spins of the same sign ($\uparrow\uparrow$ or $\downarrow\downarrow$) above it, so the $\downarrow\uparrow$ cluster must lie on a domain wall between $+$ and $\downarrow$. We wish to compute the mean square weight of the cluster, $\Omega_C^{\otimes 2}$.

Haar-averaging maps $\Omega_C^{\otimes 2}$ to a partition sum for spins $\sigma \in S_4$. The permutations are now in $S_4$ rather than $S_2$ because squaring the diagram doubles the number of layers in the replicated circuit. The spins outside the cluster are fixed, so simply provide a boundary condition for those inside.

This partition sum can be computed relatively simply when $q$ is large. We summarize the main features here and give a detailed analysis in App. C. The domain wall labelling convention we use is described in Ref. [14]: a domain wall with a domain of $\sigma_L$ to its left, and a domain of $\sigma_R$ to its right, is labelled by the permutation $\sigma_L^{-1}\sigma_R$.

If we Haar-average the circuit $U^{\otimes 4}$ we obtain spins $\sigma \in S_4$ with the interactions $J$ specified in Eq. 13. Here
the interactions are modified, because we fixed the configurations of spins $s \in \{+, -, \perp\}$ before squaring and averaging. This involves insertions of $P_+, P_-$ and $P_\perp$ that modify the triangle weights discussed in Sec. II B.

In the doubled problem there is a doubled $(12)(34)$ domain wall incoming at the entanglement cut, as illustrated in Fig. 5. In a $+$ domain the $S_3$ spins are fixed to $\|$ and in a $-$ domain they are fixed to $(12)(34)$. Inside the $\perp$ cluster the spins must be summed over. However, they are very restricted. Because of the modified triangle weights, spin values in the subgroup $S_2 \times S_2 \in S_4$ are suppressed in the interior of the cluster: a triangle whose vertices are all equal and take one of the values $\|$, $(12)$, $(34)$ or $(12)(34)$ has weight zero. In fact we have

$$
\sigma \quad \frac{\sigma}{\sigma} = \begin{cases} 1 & \text{for } \sigma \notin S_2 \times S_2, \\
0 & \text{for } \sigma \in S_2 \times S_2,
\end{cases}
$$

where the symbol on the left denotes the weight for a triangle whose lower spin is associated with a $\perp$ block.

In the large $q$ limit, the leading term involves the incoming doubled $(12)(34)$ domain wall splitting at the top of the cluster into a pair of doubled domain walls, according to the allowed multiplication (cf. Sec. VI B of [14]):

$$(12)(34) = (14)(23) \times (24)(13).$$

One doubled domain wall, say $(14)(23)$, travels along the left boundary of the cluster, and the other along the right. (Since $(14)(23)$ commutes with $(24)(13)$, either of the two doubled domain walls can be on the left.) In the interior of the cluster, the spin value is then $(14)(23)$ [or $(24)(13)$]. This does not incur any “bulk” cost, because a triangle all of whose vertices are equal to $(14)(23)$ has weight unity by Eq. 49.

The two doubled domain walls along the two boundaries of the cluster incur a weight $q^{-4}$ per time step. Consequently the mean squared cluster weight is of order $\Omega^2 \sim q^{-4t_{\text{cluster}}}$ or even smaller, where $t_{\text{cluster}}$ is the time duration of the cluster. (We have not included the cost of the “thin” domain wall sections above and below the cluster.) This indicates a typical magnitude

$$\Omega \sim q^{-2t_{\text{cluster}}}$$

for a $\perp$ cluster of a given size in a typical large $q$ circuit. Note that this weight is much smaller than the weight $\sim q^{-t_{\text{cluster}}}$ of a section of thin domain wall of the same time duration. Similar considerations apply to clusters with more complex topology, i.e. with holes inside.

Therefore, any configuration with a $\perp$ cluster lasting for $t_{\text{cluster}}$ steps has a weight that is suppressed by at least $q^{-t_{\text{cluster}}}$ compared to the leading configurations in the same spacetime region.

It should be noted that “thick” domain walls are suppressed by a cost in the exponent that scales not with the area in spacetime, but instead with the length of the boundary of the region contained in the thick domain wall. This is reminiscent of a domain wall between two different spin states in the ordered phase of the Potts model with $Q > 2$ states, where bubbles of other spin states can appear on the domain wall [56]. Here this scaling is a necessary consequence of unitarity. As a result of it, the effect of boundary conditions can be subtle. For example, a boundary condition at $t = 0$ that favours $\perp$ may be able to induce a large domain of $\perp$, because a boundary free energy of order $t$ can compensate the additional domain wall cost of order $t$ in the bulk. This is not the case for the quantities considered here, however. Indeed, for entanglement growth starting from a product state, the $t = 0$ boundary condition favours $\pm$ over $\perp$.

The results above strongly support the conjectures of the previous section. They also mean that, at large $q$, we can obtain the nontrivial fluctuations of the entanglement entropy in a random circuit by considering only single-$\perp$ clusters. We discuss this next, before moving on to deterministic models.

B. KPZ/DPRM scaling in the random circuit

Recall that $e^{-S_2} = Z$ with the appropriate (free) lower boundary condition. Averaging over Haar-random gates gives $Z = Z_{0\perp}$ (Sec. I). However, the averaged Rényi entanglement entropy $\bar{S}_2$ cannot be obtained from $-\log(Z)$, because of fluctuations.

Let us first recall the replica approach used previously. This involves computing the replicated partition function, $Z_k$, and extracting quantities like $\bar{S}_2$ from the $k \to 0$ limit. At large $q$ there is a simple picture for the replicated partition function $Z_k$. It describes $k$ domain walls that interact via a mutual attraction. In the microscopic $S_{2k}$ spin model, the attraction is due to an additional local spin configuration that is possible when two domain walls come into contact with each other [14].

This description can be identified with the replica description of a well-known classical problem, the directed polymer in a random medium [57, 58]. In that context, the free energy of the polymer can be calculated by replicating the system $k$ times and integrating out the disorder. For an appropriate choice of lattice model, the resulting $k$ polymers will have exactly the same mutually attractive interactions as the domain walls here. Therefore, we have indirectly mapped the entanglement to the free energy of a single domain wall in a random potential that modifies the weights of steps. This leads to KPZ scaling [57, 58] for the fluctuations of the entanglement.

In the present approach we consider a single domain wall, without replicas. However, to go beyond $Z_{0\perp}$, we must “dress” the domain wall with insertions of $\perp$. We
have shown above that clusters of $\perp$ are exponentially suppressed in their size. At large $q$ it suffices to consider only dilute, isolated $\perp$ clusters made up of a single block. Further, at large $q$ the local configuration labelled $\mathbb{1}$ in Eq. 42 dominates over the other configurations in Eq. 42 (App. E).

By Eq. 41, these $\perp$ insertions are restricted to lie on the domain wall. In the notation of Sec. II C, this gives us a nonzero weight

$$W(0, 2; u)$$

for a step with $\Delta x = 0$ and $\Delta t = 2$. The notation indicates that this weight depends on the local gate $u$, unlike the deterministic weights $W(\pm 1, 1)$. Since all the $u$s are independent, this yields an uncorrelated random potential for our “polymer” in spacetime.\(^3\) In an equivalent stochastic differential equation representation for $S_2(x, t)$, which is the KPZ equation, $W(0, 2; u)$ determines spatiotemporal noise.

The above picture, with random weights assigned to vertical steps, agrees with the replica approach: we just need to check the strength of the noise matches.

The average $\overline{W}(0, 2; u)$ vanishes. Its variance is given by the squared average of $\mathbb{1}$ in Eq. 42 (App. E):

$$\overline{W}(0, 2; u) \approx \frac{2K^4}{q^4}.$$  

This matches the result for the variance in Ref.[14].

We therefore confirm that a thin domain wall with dilute $\perp$ insertions reproduces the result of the replica calculation.

\(^3\) In more detail: since the weight $W(0, 2; u)$ can be either positive or negative, it is convenient to absorb it as a correction to the weight of a pair of length-1 steps. This yields a directed path made of length-1 steps, with a random potential associated with pairs of sites vertically above each other: see Fig. 6, Right. The Boltzmann weight is positive in this representation [14].
This relation is explained in Appendix F. The root $b_0$ appearing here has a physical meaning (although we will not need it here): if we write

$$s_{\text{eq}} \Gamma(s) = \ln |b_0(s)|,$$

(59)

then $s_{\text{eq}} \Gamma(s)$, which is related to $s_{\text{eq}} \mathcal{E}(v)$ by a Legendre transformation, Eq. 58, is the Rényi entropy growth rate in a state with $\partial S/\partial x = s$ [18].

As an example, consider the lowest-order approximation where we truncate $W$ at $t_{\text{max}} = 1$, corresponding to the partition function with $s \neq \pm$. We then need only

$$W(x,0) = 0, \quad W(\pm 1,1) = K.$$  

At this order, $\omega(s,b) = 2Kb \cosh s$, giving the root

$$b_0(s) = \frac{1}{2K \cosh s}. $$

(61)

The quantity $\ln b_0(s) + vs$ in Eq. 58 is maximized by $s = \tanh^{-1}v$. Plugging this in gives

$$s_{\text{eq}} \mathcal{E}(v) = \ln K^{-1} + \frac{1 - v}{2} \ln \frac{1 - v}{2} + \frac{1 + v}{2} \ln \frac{1 + v}{2}. $$

(62)

This is indeed the correct result for $Z_{0L}$ [14, 18] (and also the leading order result in the random circuit at large $q$).

In the numerical algorithm we compute $W(x,t)$ up to $t = t_{\text{max}}$ and neglect $W$ for larger times. Given a constant $s$, we then solve numerically for the smallest zero $b_0(s)$ of

$$1 - \sum_{x,t \leq t_{\text{max}}} W(x,t)e^{sx}b_0 = 0.$$  

(63)

If we define

$$v_s = \text{Re} \frac{\sum_{x,t \leq t_{\text{max}}} W(x,t)e^{sx}b_0(s)^t \times x}{\sum_{x,t \leq t_{\text{max}}} W(x,t)e^{sx}b_0(s)^t \times t},$$

(64)

then $\mathcal{E}(v_s)$ is given by $s_{\text{eq}}^{-1}(\ln |b_0(s)| + v_s s)$. (We see this by setting the $s$-derivative of the argument in Eq. 58 to zero, and extracting the derivative of $\ln |b_0(s)|$ from Eq. 63.) We iterate over $s$ to construct the entire curve.

This approach also allows us to extract the butterfly speed $v_B$, as the point where $\mathcal{E}(v_B) = v_B$. The “entanglement speed” $v_E$ for the second Rényi entropy, for a quench from the product state, is simply $\mathcal{E}(0)$.

### B. General parameterization

In our discussion of the numerical results, it will sometimes be useful to use the following parameterization of the two-site gate $u$, which is an arbitrary SU(4) matrix [42, 61, 62]:

$$u = \begin{pmatrix}
    u_1 & u_2 \\
    u_3 & u_4 \\
    \end{pmatrix}$$

(65)

where

$$u_i = \exp \left(-i \sum_{\alpha = x,y,z} h^{(i)}_\alpha \sigma^\alpha \right), $$

(66)

$$u_{\text{sym}} = \exp \left(-i \sum_{\alpha = x,y,z} J_\alpha \sigma^\alpha \otimes \sigma^\alpha \right). $$

(67)

The $u_i$s scramble the individual sites and the two-site unitary $u_{\text{sym}}$ entangles the two sites. We will mostly restrict to the reflection-symmetric case such that $u_1 = u_2$ and $u_3 = u_4$. Additionally, $Z(x,t)$ is unchanged by the transformation $u \rightarrow (v \otimes v)^{-1}u(v \otimes v)$, where $v$ is any single site unitary. This amounts to a trivial redefinition of the circuit, and one may check that the conjugate operators $v$ and $v^{-1}$ all cancel out in the weights in the partition function. Exploiting this symmetry, we can take

$$h^{(i)} = h \quad \text{for } i = 1,2,3,4,$$

(68)

so that the general symmetric gate is parameterized only by the two vectors $h$ and $J$.

### V. APPLICATION TO FLOQUET MODELS

We are finally ready to apply the numerical scheme in Sec. IV A to a variety of Floquet models, described in the following subsections. For each model we show the sequence of approximations to $\mathcal{E}(v)$, indexed by $t = 1, \ldots, t_{\text{max}}$ with $t_{\text{max}} = 8$: see for example Fig. 7. These successive approximations take weights $W$ for longer and longer steps into account.

The curve for $t = 1$ (in red) is independent of the gate defining the circuit: this lowest approximation matches the result of the random average computed in Eq. 62. It can be thought of as a baseline showing how far the Rényi entropy growth in a particular circuit deviates from the random average.

Our algorithm works if the curve converges sufficiently rapidly to the actual line tension function. In addition to showing the bare plots, we perform several other checks of convergence. In Sec. VC we will directly examine the
decay of $W(x,t)$ and $W(x,t)/Z(x,t)$ at large $t$. The assumption on the structure of the domain wall implies that the latter should be negligible at large $t$.

Additionally, we check that the line tension function $\mathcal{E}(v)$ satisfies several constraints. It is positive, convex, greater than or equal to $|v|$, and tangent to this line only at $v_B^{L/R}$. We therefore plot the boundary curve $\mathcal{E} = |v|$ with dotted lines. We also mark an estimate of $v_B^2$ from an independent numerical computation of the OTOC (the protocol is described in App. E3). We show this estimate as a pair of black dots.

Convergence to a form consistent with these constraints is a nontrivial test, because they are not built into the formulas for finite $t_{\text{max}}$ (indeed we will see that they can be disobeyed by the small–$t_{\text{max}}$ approximations).

![Figure 7](image1.png)  
**Figure 7.** Numerical computation of $\mathcal{E}(v)$ with $t_{\text{max}}$ up to 8 for (a) the reflection-symmetric gate in Eq. 69 with $x = 0.8$; (b) a generic non-symmetric gate (specified in Appendix. E). Black dots mark estimates of $v_B^{2}$ from an independent calculation of the OTOC in App. E3.

![Figure 8](image2.png)  
**Figure 8.** Numerical calculations of $\mathcal{E}(v)$ for Floquet Ising models. Panel (a) shows the parameters in Eq. 77. Panel (b) shows the “dual-unitary” parameter values in Eq. 78. The bottom arrow in case (b) illustrates the convergence to the expected limiting form $\mathcal{E}(v) = 1$. The intersections with the line $\mathcal{E}(v) = |v|$ are in good agreement with the known value $v_B = 1$ (top arrows).

A. Generic symmetric and asymmetric gates

First, we consider the following one-parameter family of reflection-symmetric gates (see Eqs. 65–68):

$$ J = x \frac{\pi}{4} (3, -4, 5)/\sqrt{50}, \quad h = \frac{\pi}{5} (1, 2, 3)/\sqrt{14}. \quad (69) $$

The parameter $x$ tunes the strength of the interaction between the two qubits. For small $x$, local scrambling will take a long time: we do not expect our algorithm to converge rapidly in that case, so we will choose $x$ reasonably large. Otherwise, the numbers above are arbitrary and were chosen to ensure (1) that there is no fine-tuning in the sense of any of the coefficients $J_\alpha$ vanishing, (2) that $u_\alpha$ represents a rotation on the Bloch sphere by a significant angle (here $\pi/5$).
Fig. 7(a) shows results for the gate with $x = 0.8$. The results are consistent with a relatively fast convergence of $\mathcal{E}(v)$. We see that this chosen gate has a smaller $v_B$ and a smaller $v_E^{(2)}$ than the annealed average for the random circuit.

Note that the results are in striking agreement with the general constraints listed above: the asymptotic $\mathcal{E}(v)$ is a convex curve that touches the line $\mathcal{E} = v$ at a single point, which is consistent with $(v_B, v_E)$ as obtained from an independent calculation of $v_B$.

Next, Fig. 7(a) shows an example of a generic gate without the reflection-symmetry constraint. We simply picked a random gate from the Haar distribution on $U(2)$ and used it to build a translation-invariant circuit. The matrix elements of this gate are given explicitly in Appendix E23.

Again, the algorithm appears to be working. However, the convergence is now to an asymmetric curve with $\mathcal{E}(v) \neq \mathcal{E}(-v)$. The touching point on the left marks $v_B^R$ and that on the right marks $v_E^L$.

**B. Chaotic Ising models**

Next we consider Floquet (or “kicked”) Ising models [21, 39, 40] with the Floquet unitary

$$U_{\text{Floq}} = \exp(-i\tau H_z) \exp(-i\tau H_x)$$

(70)

with a longitudinal field to spoil integrability:

$$H_z = \sum_{i=1}^L h_z \sigma_i^z, \quad H_x = \sum_{i=1}^{L-1} \sigma_i^x \sigma_{i+1}^x + \sum_{i=1}^L h_z \sigma_i^z.$$  

(71)

This can be written in the brickwork circuit form (Fig. 3) with the gate

$$u = \begin{array}{c|c|c}
  1 & 1 & 1 \\
  \hline \\
  u_x & u_z & u_z \end{array}$$

(72)

where:

$$u_x = \exp(-i\tau h_z \sigma^z),$$

(73)

$$u_z = \exp(-i\tau [\sigma^z \otimes \sigma^z + i h_z (\sigma_i^z \otimes 1 + 1 \otimes \sigma^z)]/2).$$

(74)

We expect that for generic values of the parameters $(\tau, h_z, h_x)$ this gate defines a chaotic model. In special limits, such as $h_z = 0$ or $h_x = 0$, the model becomes integrable: in those limits we do not expect our algorithm to succeed.

Interestingly, this model also has a “self-dual” line in parameter space [34, 41–43, 63]:

$$\tau = \frac{\pi}{4}, \quad h_x = 1, \quad h_z = 0.$$  

(75)

where some quantities can be computed exactly. In particular, entanglement growth is maximal on this line. This is related to the fact that for this choice of parameters, the tensor $u$ remains unitary if it is rotated so as to exchange the roles of space and time [41, 42]. This has been referred to as “dual unitarity”. Though this property is highly fine-tuned, the model is believed to remain chaotic for generic values of $h_z$ on this line [63].

The fact that $v_E = 1$ for the dual-unitary models implies $v_B = 1$ and that the line tension is flat as a function of velocity [18]:

$$\mathcal{E}(v) = 1.$$  

(76)

This is an interesting test case for our algorithm. First, the above form is very far from our perturbative starting point. Second, the flatness of $\mathcal{E}(v)$ implies that for dual-unitary models the large-scale properties of the domain wall are rather different from the generic case, with negative signs in $W$ playing an important role. We describe this in Sec. V D.

To begin with we check the algorithm for a presumably generic set of parameters within the kicked Ising class in Fig. 8(a):

$$\tau = 1, \quad h_x = 0.9045, \quad h_z = 0.8090.$$  

(77)

The algorithm appears to be working, and the curves converging to one that represents a gate more entangling than the random average.

Kim and Huse studied similar values of the fields, but with a smaller period $\tau = 0.8$ [39]. However, the Kim–Huse values $(\tau, h_x) \approx (0.8, 0.90)$ happen to be quite close to the dual-unitary values, which are $(\tau, h_x) \approx (0.785, 1)!$ This explains the strong entanglement growth in the Kim-Huse model. Our numerical results for this case (not shown) show a line tension function with a $v_B$ close to 1 and a larger $\mathcal{E}(v)$ than for $\tau = 1$.

We now check our algorithm for the dual-unitary case, taking

$$\tau = \frac{\pi}{4}, \quad h_x = 1, \quad h_z = 1.$$  

(78)

The value of $h_z$ is not important for the dual-unitarity condition, but we should have $\tau h_z \neq 0 \mod 2\pi$ to avoid the model being free.

Results are shown in Fig. 8(b). The curves seem to converge to the expected flat line $\mathcal{E}(v) = 1$. The touching points where $\mathcal{E}(v_B) = |v_B|$ seem to converge to the expected value $v_B = 1$ very fast.

厖

---

6 More precisely $h_x = (\sqrt{2} + 0.5)/8 \approx 0.9045$ and $h_x = (\sqrt{2} + 1)/4 \approx 0.8090$. Here we have used the truncated decimal values.

7 Writing the gate with $(\tau, h_x, h_z) = (0.8, 0.9045, 0.8090)$ in the representation in Eq. 65 gives $J \approx (0.8326, 0.8617, 0.0001)$. The dual-unitary values are $J = (\pi/4, \pi/4, 0)$ and any $h^{(i)}$.

8 In terms of the general parameterization in Eq. 65 [34, 41, 42, 63], $J = (0, \pi/4, \pi/4)$, $u_L = u_R = \exp(-i h_z \sigma_x/2) \exp(-i \pi \sigma_x/4)$, and $v_L = v_R = \exp(-i h_z \sigma_x/2)$.
We now examine the temporal decay of the step weights $W$ defined in Sec. IIIC. In order to examine the importance of long steps, we consider the normalized quantity $W(x,t)/Z(x,t)$: this compares the weight of a single long step, with displacement $x$ and duration $t$, with the total weight of all paths with the same two endpoints. This ratio should decay with $t$ if asymptotically long steps can be neglected. (Individually, both $W(x,t)$ and $Z(x,t)$ decay exponentially with $t$.) A slight extension of the reasoning in Sec. IIIA shows that in a typical realization of a random circuit at large $q$, the typical value of this ratio (for a given starting point of the domain wall in spacetime) decays like $q^{-t}$.

Fig. 9 shows heat maps of $W(x,t)/Z(x,t)$ for examples of translation-invariant Floquet models for spin-1/2s. Panel (a) is for the generic reflection-symmetric gate in Eq. 68, whose line tension function was shown in Fig. 7(a), and Panel (b) shows the dual-unitary gate of Eq. 78 and Fig. 8(b).

In both cases the weight is small at large $t$, though finite-time effects in this quantity seem to be large for the former gate at displacement $x = 0$. We also find decay of $W/Z$ for the generic asymmetric gate discussed in Sec. V A and for the kicked Ising model with the generic parameters in Eq. 77 (data not shown).

The domain wall picture appears to be well-defined both for the generic models and for the dual-unitary model. However there are key differences between the two cases which we discuss in Sec. VD. Note that the weight at $x = 0$, $t = 2$ is positive for the gate in Panel (a) and negative for the dual-unitary gate in Panel (b). In the former case this extra positive step weight increases $Z(0,t)$ compared to $Z_{0,1}(0,t)$, leading to decreased $\nu_E^{(2)}$. In the latter case the negative step weight decreases $Z(0,t)$ compared to $Z_{0,1}(0,t)$, helping the dual-unitary gate to attain the fastest possible decay of $Z(0,t)$, i.e. the most rapid possible entropy growth.

Let us also show an example where local gate is very weakly entangling, resulting in slower convergence of the algorithm. This is the reflection-symmetric gate defined in Eq. 69, but with the smaller interaction constant $x = 0.5$ (in Sec. VA we showed results for $x = 0.8$).

First we show the ratio $W/Z$ in Fig. 10. We see that there are large values close to the edges of the lightcone. However, the apparent decay inside the lightcone suggests that the results may still converge, just more slowly than in the cases examined above.

The slow convergence in this case is related to the fact that the gate is weakly entangling. First we examine the function $\Gamma(s)$ defined in Sec. IV A (whose physical meaning, for $|s| < \ln 2$, is as an entanglement entropy growth

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure9.jpg}
\caption{The space-time structure of $W(x,t)/Z(x,t)$ for (a) the generic reflection-symmetric gate in Eq. 68 [Fig. 7(a)] and (b) the maximally-entangling dual-unitary gate specified in Eq. 78 [Fig. 8(b)]. The cross marks the coordinate $(0,0)$. In our construction, only steps with even $x + t$ are possible.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure10.jpg}
\caption{Weakly entangling gate with $x = 0.5$ (Eq. 69): Heatmap showing relative importance of long steps.}
\end{figure}
rate in a state with a nonzero gradient in the entanglement \( S(x) \) across a cut at position \( x \), with \( \partial S/\partial x = s \). The approximations to \( \Gamma(s) \) for \( t = 2, 4, 6, 8 \) are shown in Fig. 11(a). Note that they are no longer concave functions over the whole range \( s \). However, this effect gets less severe with increasing \( t \). If we restrict to the concave region in performing the Legendre transform, we get a reasonable result for \( \mathcal{E}(v) \): see Fig. 11(b). This is consistent with slow convergence to a function that satisfies the general constraints, with a very small entanglement growth rate.

**D. Entanglement membrane and dual-unitarity**

Microscopically, the step weights \( W(x, t) \) can be of either sign. But we conjecture that, for generic models, the minus signs do not change the universality class. Heuristically, we can think of the step weights becoming positive after coarse-graining, so that the membrane is effectively a classical directed path with diffusive wandering, at least for \( v < v_B \).

This coarse-grained picture implies a strictly convex \( \mathcal{E}(v) \), with a positive second derivative that is related to the diffusivity of the path. [This diffusivity can depend on the coarse-grained velocity \( v \), which we condition on in order to define \( \mathcal{E}(v) \).] In more detail, the scaling of \( Z(x, t) \) is:

\[
Z(x, t) \sim t^{-1/2} e^{-s_{eq}\mathcal{E}(v)t}. \tag{79}
\]

The power-law prefactor is universal and comes from the fact that both endpoints of the random path are fixed. For example, consider for concreteness the case where \( x \ll t \) and the coarse-grained velocity is close to zero. Expanding in \( v = x/t \), the above is

\[
Z(x, t) \sim t^{-1/2} e^{-s_{eq}\mathcal{E}(0)x^2/2t} \quad (x \ll t). \tag{80}
\]

We may identify this with the probability for a random walker, with diffusion constant \( D = (2s_{eq}\mathcal{E}(0))^{-1} \), to be at position \( x \) at time \( t \).

The dual-unitary model behaves differently. The membrane and its line tension function still appear to make sense, but its wandering is not diffusive. For example, we can contrast the scaling of \( Z(x, t) \) with the formula above. Our results are consistent with the large \( t \) scaling

\[
Z(x, t) \sim e^{-s_{eq}t}, \tag{81}
\]

with a trivial \( x \)-dependence inside the lightcone (we neglect the even-odd effect at the lattice scale).

Evidently, in the dual-unitary case, the membrane is not a classical random walk. To understand this, let us write Eq. 40 (in the limit of large \( t \)) as

\[
Z(x, t) = \sum_{\Delta t = 1}^{\infty} \sum_{\Delta x} W(\Delta x, \Delta t) Z(x - \Delta x, t - \Delta t). \tag{82}
\]

This is a discrete analogue of a linear partial differential equation in \( x \) and \( t \).

We expect that in the generic case we have an equation of parabolic — diffusive — type, analogous to the continuum equation

\[
\partial_t Z = (D \partial_x^2 + \text{const.} + \ldots) Z \tag{83}
\]

(this is schematic). However, by fine-tuning \( W \), we can make the coefficient of the leading time-derivative term \( \partial_t Z \) vanish, so that the leading term is second order,

\[
\partial_t^2 Z = (c^2 \partial_x^2 + \text{const.} + \ldots) Z, \tag{84}
\]

giving a hyperbolic, wave-like equation instead. We conjecture that this is what generically happens in the dual-unitary case.

A toy model is to take only a single additional nonzero \( W \) element (beyond the ones present in \( Z_{0,1} \)), which we parameterize in terms of a constant \( \mu \) as:

\[
W(\pm 1, 1) = K, \quad W(0, 2) = -\mu K^2. \tag{85}
\]
To simplify the notation, let us remove the factors of $K$ by extracting a factor of $K^3$: we define $Z(x, t) = K^3 \tilde{Z}(x, t)$ and $W(x, t) = K^3 \tilde{W}(x, t)$.

When $\mu < 1$ we are in the parabolic class. However when $\mu = 1$ (in fact the dual unitary model is numerically quite close to this simplified model) Eq. 82 becomes

$$\tilde{Z}(x, t) = \tilde{Z}(x+1, t-1) + \tilde{Z}(x-1, t-1) - \tilde{Z}(x, t-2),$$

which is a discrete version of the wave equation, $(\partial_t^2 - \partial_x^2) \tilde{Z} = 0$. This gives a $\tilde{Z}(x, t)$ which is constant inside the lightcone and zero outside.

Numerically, we find that the structure in the specific dual unitary kicked Ising model that we studied above is similar: $2^t Z(x, t)$ is approximately constant within the lightcone, and zero outside. Fig. 12 contrasts the space-time pattern of $e^{-s_{eq} v_B t} Z(x, t)$ for the generic kicked-Ising model and the dual-unitary kicked Ising model.

The approximate constancy of $Z(x, t)$ within the lightcone, which is related to the hyperbolic structure, is imposed by dual-unitarity. This fact can be seen by considering a slightly different partition function:

$$Z_{\text{op}}(x, y, t) = \langle \cdots + + - |U^{(2)}(t)| + + - \rangle/q_{2L}. \quad (87)$$

Here $x$ and $y$ label the positions of the domain walls between $+$ and $-$ on the top and bottom boundaries respectively. Fig. 13(a) shows an example of $q_{2L} Z_{\text{op}}(x, t)$.

In contrast to the domain wall partition function $Z(x, t)$ defined in Eq. 33, $Z_{\text{op}}(x, y, t)$ does not precisely fix the end point of the domain wall to $y$, but instead penalizes configurations whose end point deviates from $y$. However, in analogy to the random circuit case, we expect that (so long as $|x - y| \leq v_B t$, which in the dual-unitary circuit is a trivial restriction since $v_B = 1$) this only leads to an $\mathcal{O}(1)$ boundary term in the free energy, so that

$$\ln Z_{\text{op}}(x, y, t) \sim \ln Z(x, y, t) \sim E(v)t. \quad (88)$$

Physically, the above partition function $Z_{\text{op}}$ simply computes the operator entanglement of the time evolution operator itself, for a certain partition of the “legs” of the time evolution operator. This operator entanglement is maximal in the dual unitary circuits, like various other measures of entanglement in these systems [41, 44, 63]. This is easily seen by manipulations similar to those in [41, 42]. Specifically (setting $y = 0$) we show that $Z_{\text{op}}(x, t) = q^{-x}$ for a dual-unitary circuit, so long as $x$ is within the light cone of the origin.

We represent the multi-layer two-site unitary evolution matrix as a four-leg tensor:

$$u^{(N)} = \begin{array}{c} A \\ B \\ C \\ D \end{array}, \quad u^{(N)}_{\text{dual}} = \begin{array}{c} B \\ C \\ A \\ D \end{array}. \quad (89)$$

We have labelled the external legs of this tensor by $A, B, C, D$. As a unitary, the legs $A, B$ correspond to

![Figure 13](image_url)

Figure 13. The partition function $Z_{\text{op}}(x, t)$ for a dual-unitary circuit. (a) $q_{2L} Z_{\text{op}}(x = 1, t = 5)$. By unitarity (the first identity in Eq. 90), we can remove the unitaries outside the lightcone, giving $q^2 Z_{\text{op}}(x = 1, t = 5)$ as expression (b). Then Eq. 90 allows the remaining gates to be removed.
VI. OPERATOR SPREADING

We have focused on the second Rényi entropy as a test case, but the present formalism can be applied to almost any observable expressible in terms of $U^{(N)}$. One of the simplest is the out-of-time order correlation function, which requires $N = 2$, and which on large scales we expect to reduce to a stochastic growth process [7, 8].

The OTOC can be written in terms of the same partition sum as for $e^{-s_{2}t}$, but with different boundary conditions. In the Haar-averaged circuit this reduces to a simple partition function for a pair of domain walls. Since we have now made sense of the domain wall in a generic, fixed circuit, this picture from the random circuit carries over. The modification is that the domain walls now have a nontrivial internal structure, including finite regions of $\perp$, with an associated microscopic timescale $t_{0}$. A typical configuration from the partition sum is illustrated in Fig. 14.

For a brief heuristic summary, we consider the OTOC in the form of a commutator squared of two traceless operators,

$$\tilde{C}(x, t) = -\frac{1}{2} \langle [O(x, t), O'(0, 0)]^2 \rangle. \quad (91)$$

We take the expectation value at infinite temperature, and to simplify the boundary conditions we average over local unitary rotations of the local operators $O \rightarrow VOV^{\dagger}$, $O' \rightarrow V'O'V^{\dagger}$. Carrying out manipulations similar to Sec. VI B of [7], we rewrite $\tilde{C}(x, t)$ using the present notation (here tr is a single-site trace)

$$\tilde{C}(x, t) = \frac{\text{tr}(O^2)\text{tr}(O'^2)}{q^2 - 1} \langle \cdots + + - + \cdots |U^{(2)}| \cdots - - * - \cdots \rangle / q^{L-1}. \quad (92)$$

The operator insertion at the top initiates a domain of $+$ in the bulk. Because of the $-$ boundary condition at the bottom, this domain prefers to expand at a speed greater than $v_{B}$. This is illustrated for the left domain wall in Fig. 14. The additional free energy cost means that, outside the lightcone, for $x/t > v_{B}$, the OTOC scales as [64]

$$\tilde{C}(x, t) \sim e^{-\mu(v)t}, \quad \mu(v) \equiv s_{eq}(E(v) - v). \quad (93)$$

Fig. 14. The membrane picture for the out-of-time ordered commutator $\tilde{C}(x, t)$. The top boundary condition creates two domain walls. The boundary condition favors $-$, providing an outward force on the lower endpoints of the two domain walls. The right domain wall thus moves with the optimal slope $v_{B}$. The left domain wall is forced by the dual state $[-*]$ at the operator insertion point to travel with $v > v_{B}$.

9. The lower boundary condition contributes a weight proportional to $q^{d} = e^{s_{eq} t}$, which tries to increase the width $d$ of the $-$ domain at the lower boundary. The boundary weight means that the optimal velocity $v_{L}$ for the left domain wall, minimizing the free energy, is $v_{B}$, since by the properties of $E$ this minimizes $s_{eq} (-v_{L} + E(v_{L})) t$. 

---

The first of these relations (unitarity) allows us to remove the unitary gates outside the lightcones in Fig. 13(a). Keeping track of the $q$ factor, $q^{2i} Z_{op}(x, t)$ equals the expression in Fig. 13(b). Next, dual-unitarity (the second relation) can be used to propagate the $+$ states to the right, so that all the unitary gates are removed and Fig. 13(b) is reduced to $\langle + | - \rangle = q^{t}$. Hence $q^{2i} Z_{op}(x, t) = q^{t}$, so that $Z_{op}(x, t) = e^{-s_{eq}t}$ within the lightcone.
In random models these results can also be obtained from a stochastic growth process for the operator cluster \([7, 8, 13]\). In this picture, \(\mu(v)\) is the large deviation function describing the probability for the boundary of the operator cluster to propagate at speed \(v\) [65]. From the above discussion, we expect that in generic models this stochastic picture applies after coarse-graining to timescales greater than \(t_0\).

Forthcoming work using the memory matrix formalism will give a complementary view on the operator spreading problem [66].

VII. OUTLOOK

In random circuits, the membrane picture emerges in a simple way from the average over unitaries whenever we deal with dynamical quantities, such as Rényi entropies or OTOCs, that require multiple forward and backward evolution operators. Here we have shown how to make sense of this membrane in realistic systems that do not involve randomness. This opens the way to extending the various results that have been obtained in random circuits to a much broader class of more realistic models.

We separated the local multi-copy Hilbert space into “paired” states, and states in their orthogonal complement, \(\perp\). Dynamical quantities then became partition functions for an effective field \(s(x, t) \in \{\perp\} \cup \mathcal{S}_N\). The state \(\perp\) of the effective spin here is reminiscent of the state \(S = 0\), representing a vacancy, in the Blume-Capel model (the classical Ising model with vacancies, where the spin takes the values \(S = 0, \pm 1\)). We have shown that in strongly chaotic Floquet systems the “vacancies” \(\perp\) dress the structure of domain walls between different pairing states, but leave the basic structure of the domain wall intact on large scales.

After using a typical realization of a random circuit as a test case, we applied this picture to Floquet systems, for spin-1/2, that have no small parameter at all and no randomness.

The method gave convincing results for the line tension function for \(N = 2\) (which determines both the second Rényi entropy growth rate and the butterfly velocity) from calculations on relatively small length scales. A simple but important ingredient was a resummation of the weights in the spin model into a single function \(W(x, t)\), describing the amplitude for the membrane — which in 1+1D can be viewed as a directed path — to take a step of temporal extent \(t\) and spatial displacement \(x\). Interestingly, there was a qualitative difference in the structure of \(W\) for generic circuits and for dual-unitary, maximally entangling circuits.

Many questions remain for the future. First, there are many kinds of extra structure that could be incorporated and which might be expected to have interesting effects.

We have focused on Floquet models without any conserved quantities. Any slow mode, for example a conserved \(U(1)\) charge, will interact with the dynamics of the membrane. (A single physical conserved charge gives rise to multiple conserved charges in the replicated Hilbert space.) Results on random circuits with conservation laws provide a starting point [9, 10]. One cautionary note is that it has recently been shown that the dynamics of the higher Rényi entropies, for which the present method is suitable, is qualitatively different from that of the von Neumann entropy [15, 67].

In the present approach, the von Neumann entropy corresponds to a nontrivial replica-like limit, which it would be interesting to understand better even in the absence of conserved quantities.

Also interesting are continuous spatiotemporal, as opposed to internal, symmetries, for example Lorentz invariance. Such symmetries will constrain the effective field theory of the membrane, as well as giving new hydrodynamic modes [68, 69].

There is considerable freedom in the precise way in which the effective spins are introduced. We chose to insert a resolution of the identity for pairs of sites, since this was convenient for making contact with the Haar-random circuits studied previously. We can also insert the resolutions of identity for single sites. This does not introduce any fundamental differences, but is likely to improve the convergence of our numerical scheme (especially for weakly-entangling gates) since for the same \(t_{\text{max}}\) we may define more independent \(W\) elements.

We may also consider dynamics which is not of circuit form. The most straightforward generalization is to Floquet systems that do not have a circuit representation. Effective spins can be introduced in the same way, again with an arbitrariness in how we choose to place them.

The basic idea of this paper is not restricted to 1+1D: the effective partition function can be introduced in the same way in any number of dimensions, and we may again argue for the coarse-grained picture in terms of a codimension-1 membrane dressed with \(\perp\) states. However, the additional computational cost means that in higher dimensions we are likely to need either special structure, or a small parameter, in order to obtain quantitative results.

For the observables studied here, the pairing between layers of the circuit is local in time. For some dynamical quantities, for example the spectral form factor, pairings that are non-local in time play a role [11, 33? ]. It would be interesting to extend the present approach to this setting.

Our method for defining the effective spin model might illuminate other interesting universal dynamical phenomena. First, we may consider how the description in terms of an entanglement membrane breaks down on approaching a many-body-localized phase. We hope to return to this elsewhere. There are other failure modes that could be investigated, for example the approach to a noninteracting or integrable point.

Within the more limited domain of random circuits, the effective spin model here may also give a way of thinking about entanglement in random Clifford circuits.
which is more generalizable (e.g. to higher dimensions) than the approach of mapping the evolution of stabilizers to an effective classical stochastic dynamics [6, 70, 71].

The effective spin model may also be extended to dynamics with measurement, or other types of interaction with an environment, for example to clarify the relationship between different universality classes of measurement-induced criticality [70–82].

For chaotic dynamics, further numerical tests and further development of the numerical scheme, extending the range of models to which it could be applied usefully, would also be worthwhile and might give new insights into the physics on the scale of the membrane thickness (which, in many natural situations where there is a small parameter at hand, may be much larger than the lattice spacing).

ACKNOWLEDGMENTS

We thank John Chalker, Tibor Rakovszky, and in particular Curt von Keyserlingk for very useful discussions. TZ was supported by a postdoctoral fellowship from the Gordon and Betty Moore Foundation, under the EPiQS initiative, Grant GBMF4304, at the Kavli Institute for Theoretical Physics. This research is supported in part by the National Science Foundation under Grant No. NSF PHY-1748958. We acknowledge support from the Center for Scientific Computing from the CNSI, MRL: an NSF MRSEC (DMR-1720256) and NSF CNS-1725797. AN was supported by a Royal Society University Research Fellowship.

Appendix A: The projection operator to permutation states

The operator $P_{\sigma}$ defined in Eq. 13 is a projection operator onto the $\sigma$ permutation state. Here we prove the projector property in Eq. 15. According to the definition of $P_{\sigma}$ in terms of the Weingarten function and permutation states in Eq. 13, we have

$$P_{\sigma_1}P_{\sigma_2} = \sum_{\tau_1, \tau_2} Wg(\tau_1^{-1}\sigma_1)|\tau_1\rangle\langle\sigma_1|\tau_2\rangle Wg(\tau_2^{-1}\sigma_2)|\sigma_2\rangle$$

(A1)

Here each state is on two sites: $|\tau\rangle = |\tau, \tau\rangle$ if we write out the state on each site. Note that states associated with distinct permutations are not orthogonal. Unless $q^2 = 0, \pm 1, \pm 2, \cdots, \pm (N - 1)$, the function $Wg(\tau^{-1}\sigma)$, viewed as a matrix with $\tau$ and $\sigma$ as its indices, is the inverse of the Gramian matrix $\langle \sigma | \tau \rangle$ [45, 46]:

$$\sum_{\tau} \langle \sigma_1 | \tau \rangle Wg(\tau^{-1}\sigma_2) = \delta_{\sigma_1, \sigma_2}.$$  

(A2)

This means the states $|\sigma^*\rangle$ defined in Eq. 14 indeed form the dual basis,

$$\langle \sigma_1^* | \sigma_2 \rangle = \delta_{\sigma_1, \sigma_2}.$$  

(A3)

So for $q^2 \geq N$, we have

$$P_{\sigma_1}P_{\sigma_2} = \sum_{\tau_1, \tau_2} Wg(\tau_1^{-1}\sigma_1)|\tau_1\rangle\langle \sigma_1|\tau_2\rangle Wg(\tau_2^{-1}\sigma_2)|\sigma_2\rangle = P_{\sigma_1}\delta_{\sigma_1, \sigma_2}.$$  

(A4)

Any four-layer unitary gate with the projector $P_{\sigma}$ attached at the bottom is equal to $P_{\sigma}$ (Eq. 18). In Eq. 17, we represent the projector as a tensor (box) labelled with a $\sigma$ inside. Given the expression in Eq. 13, it can be decomposed as

$$P_{\sigma} =$$

(A5)

where the ket $|\sigma^*\rangle = \sum_{\tau} Wg(\tau^{-1}\sigma)|\tau\rangle$ is the two-site dual state of $|\sigma\rangle$, and the bras at the bottom are just $\langle \sigma | \sigma \rangle$ on the two sites.

This structure helps us to clarify the definition of the triangle weights $J(\sigma_b, \sigma_c; \sigma_a)$ as the result of contracting the red part of the following tensor:

$$\sigma_a \sigma_c \equiv \text{shaded part of }$$

(A6)

which yields the expression of $J(\sigma_b, \sigma_c; \sigma_a)$ in Eq. 22.

Appendix B: $N$-independence of the weights

The triangle weights

$$J(\sigma_b, \sigma_c; \sigma_a) = \sigma_a$$

(B1)

define the interactions of the effective spin models arising from the random unitaries. The spin $\sigma$ is an element of $S_N$, where $N$ is the number of copies of $U$ and of $U^*$ in the multi-layer circuit.

In some cases of interest to us the spins $\sigma_a$, $\sigma_b$, $\sigma_c$ are all in the smaller permutation group $S_M$, with $M \leq N$, corresponding to permutations involving the indices $1, \ldots, M$ only. For the corresponding bras, the first $2M$ layers (the "U" layers and the "U*" layers) can be paired arbitrarily, with the pairing labelled by a permutation in $S_M$, and the remaining $2(N - M)$ layers are paired in the manner prescribed by the identity permutation. For example, consider the case $\sigma_a = \sigma_b = \mathbb{I}$ and $\sigma_c = (12)$:

$$\begin{array}{c}
\sigma_a \\
\sigma_b \\
\sigma_c \\
\mathbb{I}
\end{array}$$

(B2)

This weight is independent of $N$ (proved algebraically in [14]). This $N$-independence holds whenever all the spins are in the permutation group $S_M$, as conjectured in [14]
and proved by Hunter-Jones in Ref. [17]. Here we review this result. Algebraically, the triangle is the coefficient in the following expansion:

$$\langle \sigma_b, \sigma_c | P_{\parallel}^{(N)} \rangle = \sum_{\sigma_a \in S_N} \sum_{\tau \in S_N} \langle \sigma_b | \tau \sigma_a | \tau \sigma_c \rangle W_{\parallel}(\tau^{-1} \sigma_a; q^2),$$  \hspace{1cm} (B3)

where the two arguments in the bras refer to the pair of spatial states. This is equivalent to the definition in Eq. B4,

$$\sum_{\sigma_a \in S_N} \langle \sigma_b | \tau \sigma_c | \tau \sigma_a \rangle = \sum_{\sigma_a \in S_N} \langle \sigma_b | \tau \sigma_c | \tau \sigma_a \rangle W_{\parallel}(\tau^{-1} \sigma_a; q^2),$$  \hspace{1cm} (B4)

as we see by using Eq. A2 to extract a particular coefficient on the right hand side of Eq. B3.

Now consider the case where $\sigma_a, \sigma_c \in S_M$, where $S_M$ is the subgroup mentioned above. \hspace{1cm} \footnote{By the symmetry in Eq. 28, the following also applies to other subsets of states related to this one by left and right multiplication with fixed elements of $S_N$.} The fact that the resulting weights are independent of $N$ is not immediately apparent from Eq. B4, but can be seen from Eq. B3. When $\sigma_b, \sigma_c \in S_M$, we have

$$\langle \sigma_b, \sigma_c | P_{\parallel}^{(N)} \rangle = \langle \sigma_b, \sigma_c | (P_{\parallel}^{(M)} \otimes 1^{(N-M)}) \rangle.$$  \hspace{1cm} (B5)

For example, we can see this by writing the projection operator $P_{\parallel}^{(N)}$ as the average of a stacked random unitary. The unitaries in the layers with index greater than $M$ cancel, since the bra contracts them together in the same way on each site. This leaves the average of a stack acting only on the first $2M$ layers, which is equal to $P_{\parallel}^{(M)}$ acting on those layers. The identity acts on the remaining $2(N-M)$ layers.

This shows that in Eq. B5 the only nonvanishing terms on the right-hand-side have $\sigma_a \in S_M$ as well. As a result Eq. B5 becomes equivalent to the same equation for the $2M$-layer case (except that both sides are tensored with the “identity” bra for the remaining $2(N-M)$ layers). This shows that the triangle weights are the same as in the $2M$-layer case, i.e. independent of $N$.

Note however that if all of the spins are in $S_M \otimes S_{N-M} \subset S_N$, where the subgroups refer to the first $2M$ and the last $2(N-M)$ layers respectively, the weight does \textit{not} factorize in general. That is, if we have

$$\sigma_a = \sigma_{a1} \sigma_{a2}, \hspace{0.5cm} \sigma_b = \sigma_{b1} \sigma_{b2}, \hspace{0.5cm} \sigma_c = \sigma_{c1} \sigma_{c2},$$  \hspace{1cm} (B6)

where $\sigma_{a1} \in S_M$ and $\sigma_{b2} \in S_{N-M}$ etc., the triangle

$$\begin{array}{c}
\sigma_a \\
\sigma_b \\
\sigma_c
\end{array}$$  \hspace{1cm} (B7)

is in general not equal to

$$\begin{array}{c}
\sigma_{a1} \\
\sigma_{c1}
\end{array} \times \begin{array}{c}
\sigma_{b2} \\
\sigma_{c2}
\end{array}.$$  \hspace{1cm} (B8)

The two are equal if, for example, $\sigma_{b2} = \sigma_{c2}$.

\section*{Appendix C: Suppression of large $\perp$ clusters}

We analyze the spin model that arises in computing the mean square weight $\Omega_2^\perp$ of a $\perp$ cluster like that in Fig. 5. (The average is over Haar-random unitaries.) The spins take values in $S_4$, because in total we need four $U$ layers and four $U^*$ layers to write $\Omega_2^\perp$. We will see that there is a distinction between spin values in the $S_2 \times S_2$ subgroup generated by (12) and (34), and spin values outside this subgroup. At leading order, the former are forbidden inside the cluster.

We first look at the region outside the $\perp$ region. Each block there is labelled either as $+$ or $-$, indicating a local insertion of $P_{\perp}$ before squaring. Recall that this projection operator “absorbs” the local unitary gate: equivalently, in the squared system a given $+ \otimes -$ block is (recall that each $P_{\perp}^{(2)}$ acts on two physical sites):

$$\frac{u(q)}{2^4} (P_{\perp}^{(2)} \otimes P_{\perp}^{(2)}) = P_{\perp}^{(2)} \otimes P_{\perp}^{(2)}$$  \hspace{1cm} (C1)

and similarly for $-$. This defines an interaction triangle for the case where the lower spin in the triangle is from a $+$ block, in analogy to the triangle interaction defined in Sec. II B for the case without the $P_{\perp}^{(2)} \otimes P_{\perp}^{(2)}$ factor. However, because of the projection operator, the spin associated with the block (the lower spin of the triangle) is \textit{fixed} to be $\perp$. Similarly, the spin associated with a $-$ block is fixed to be $\sigma = (12)(34)$. Therefore in computing $\Omega_2^\perp$, the spins outside the cluster are not free to fluctuate. Their role is to provide a boundary condition for the spins inside the cluster.

Explicitly, the $+$ weights are (from Eq. B3 with $P_{\perp}^{(4)}$ replaced by $P_{\perp}^{(2)} \otimes P_{\perp}^{(2)}$; the $-$ weights are related to the following by symmetry)

$$\sum_{\sigma_a \in S_N} \langle \sigma_b, \sigma_c | (P_{\perp}^{(2)} \otimes P_{\perp}^{(2)}) | \parallel, \parallel \rangle = q^{-8} \langle \sigma_b, \sigma_c | (P_{\perp}^{(2)} \otimes P_{\perp}^{(2)}) | \parallel, \parallel \rangle.$$  \hspace{1cm} (C2)

Using the expression for the projector in Eq. 25, this $+$ triangle is equal to

$$\sum_{\tau \in S_2} q^{6(1/2 - |\sigma_b| - |\sigma_c|)}.$$  \hspace{1cm} (C3)

Since $|\sigma_b| + |\tau| \geq |\sigma_b|$, in the large $q$ expansion, the $+$ triangle is at most of order

$$\langle \sigma_b, \sigma_c | (P_{\perp}^{(2)} \otimes P_{\perp}^{(2)}) | \parallel, \parallel \rangle = q^{-8} \langle \sigma_b, \sigma_c | (P_{\perp}^{(2)} \otimes P_{\perp}^{(2)}) | \parallel, \parallel \rangle.$$  \hspace{1cm} (C4)
and similarly the $-$ triangle can be bounded by
\[
\sigma_b \sigma_c \lesssim q^{-|\sigma_b(12)(34)|-|\sigma_c(12)(34)|} \tag{C5}
\]
Here $|\sigma_b|$ and $|\sigma_b(12)(34)|$ can be understood as the numbers of elementary domain walls (i.e. number of transpositions) on the left edge of the $+$ and $-$ triangle respectively. And similarly the corresponding $\sigma$ terms are for the right edge. We thus interpret Eq. C4 and Eq. C5 as costing $q^{-1}$ per out-going domain wall on the two edges.

Inside the cluster we have $\perp$ blocks. These lead to a “$\perp$” interaction triangle defined by:
\[
\sum_{\sigma_a \in S_4} \frac{\sigma_b \sigma_c}{\sigma_a} \langle \sigma_a, \sigma_b \vert \sigma_c \rangle \equiv \langle \sigma_b, \sigma_c \vert \left( u^{(4)} - u^{(2)} \otimes u^{(2)} \right) \nonumber \\
\equiv \langle \sigma_b, \sigma_c \vert \left( P^{(4)} - P^{(2)} \otimes P^{(2)} \right) \tag{C6}
\]
where we have used the equivalence between the parallel projection and the Haar averaged stack of unitaries.

The right-hand-side of Eq. C6 gives the $\perp$ triangle weight as the difference of two terms. The first term is the ordinary triangle weight for $S_4$. The second term vanishes unless $\sigma_a \in S_2 \times S_2$. Depending on whether $\sigma_a \in S_2 \otimes S_2$, the large $q$ expansion yields
\[
\begin{aligned}
\sigma_b \sigma_c &\lesssim q^{-|\sigma_a^{-1} \sigma_b| - |\sigma_a^{-1} \sigma_c|} \quad \sigma_a \notin S_2 \otimes S_2 \\
&\lesssim q^{-|\sigma_a^{-1} \sigma_b| - |\sigma_a^{-1} \sigma_c| - 2} \quad \sigma_a \in S_2 \otimes S_2 
\end{aligned} \tag{C7}
\]
For the first case, the $\perp$ triangle weight is exactly equal to an ordinary triangle
\[
\sigma_b \sigma_c = \sigma_a \sigma_b \sigma_c , \quad \sigma_a \notin S_2 \times S_2 . \tag{C8}
\]
We again interpret the corresponding expansion (the first line) in Eq. C7 as a cost of $q^{-1}$ for each outgoing elementary domain wall. For the case of $\sigma_a \in S_2 \otimes S_2$, we can show using the large $q$ expansion techniques in Ref. 14 that the term of order $q^{-|\sigma_a^{-1} \sigma_b| - |\sigma_a^{-1} \sigma_c|}$ cancels, which therefore incurs an additional cost of $q^{-2}$. In particular, if $\sigma_b, \sigma_c$ are also in $S_2 \times S_2$, then the $\perp$ triangle vanishes exactly:
\[
\sigma_b \sigma_c = 0 , \quad \sigma_a, \sigma_b, \sigma_c \in S_2 \times S_2 . \tag{C9}
\]
With this in hand, we can estimate the cost of a $\perp$ cluster in Fig. 5. Because of the vanishing weight in Eq. C9, each triangle inside $\perp$ region must host at least one spin that is not in $S_2 \times S_2$. Therefore in each time slice that contains a $\perp$ there are at least four elementary domain walls. By the estimates above, each costs at least $\frac{1}{q}$ per time step, regardless of whether it is inside the $\perp$ region or on the boundary between the $\perp$ and $\pm$ regions. A $\perp$ cluster that persists for $t$ steps is therefore suppressed by a factor of $q^{-4t}$ or smaller. This completes the argument in Sec. III.

The cost of $q^{-4t}$ is achieved by the following configuration. We set the spins in the $\perp$ cluster to (13)(24). The triangle weights inside the $\perp$ cluster are then 1, by Eq. C8. There are two elementary domain walls making up the composite domain wall (13)(24) on the left boundary of the $\perp$ cluster and two elementary domain walls making up (14)(23) on the right boundary. They enter the $\perp$ cluster through the triangle
\[
\begin{array}{c}
\sum \frac{\sigma_b \sigma_c}{\sigma_a} \\
\sigma_a \in S_4 \\
\end{array} \tag{C10}
\]
which is equal to the ordinary triangle and costs $q^{-4}$. Meanwhile on the boundary, the four domain walls also cost $q^{-4}$, according to Eq. C4 and Eq. C5. In total, the weight is $q^{-4t}$ for the squared average of the $\perp$ cluster.

Appendix D: Minimal $\perp$ cluster and operator purity

In Eq. 42, we listed four non-trivial configurations for an isolated $\perp$ surrounded by $+$ and $-$. Here we compute their weights for a fixed unitary $u$ at the location of the $\perp$. We will see they are related to the operator purity of $u$.

It is more transparent if we represent each $+$ and $-$ block graphically as in Eq. A5:
\[
\equiv \text{shaded part of} \begin{pmatrix} 
\end{pmatrix} . \tag{D1}
\]
The weight on the left-hand-side is equal to the $c$-number given by contracting the colorful parts of the tensor on the right-hand-side. The 4-leg $\perp$ tensor (red) is $u^{(4)} P_\perp$. We can rewrite it as $u^{(4)} - P_\perp = u^{(4)} P_+ - P_-$. It is contracted with states $\langle + \vert$ and $\langle - \vert$ on the top (blue) and the kets at the bottom from the green tensors:
\[
\begin{array}{c}
\text{shaded part} \\
\text{of} \begin{pmatrix} 
\end{pmatrix} \\
\end{array} \tag{D2}
\]
We have:
\[
\begin{aligned}
\langle + \vert - \rangle &\equiv \frac{Kq}{q^2 - 1} \\
\langle + \vert - \rangle &\equiv \frac{K}{q^2 - 1} \\
\langle - \vert + \rangle &\equiv \frac{K}{q^2 - 1} \\
\langle - \vert + \rangle &\equiv \frac{K}{q^2 - 1}
\end{aligned}
\]
We have:
\[
\begin{aligned}
\langle + \vert - \rangle &\equiv \frac{Kq}{q^2 - 1} \\
\langle + \vert - \rangle &\equiv \frac{K}{q^2 - 1} \\
\langle - \vert + \rangle &\equiv \frac{K}{q^2 - 1} \\
\langle - \vert + \rangle &\equiv \frac{K}{q^2 - 1}
\end{aligned}
\]
Combining these, we have

\[
\begin{pmatrix}
\vdots \\
+ \\
+ \\
\vdots \\
\end{pmatrix}
= -2K^2 + K^2 \left( \frac{q}{q^2 - 1} \right)^2 \times
\begin{pmatrix}
+ \\
\vdots \\
+ \\
\vdots \\
\end{pmatrix}
\]

where the ± symbols around the \( u^{(4)} \) block in the second line indicate contraction with \( \langle \pm | \) (at the top) or \( | \pm \rangle \) (at the bottom). Similarly,

\[
\begin{pmatrix}
\vdots \\
+ \\
+ \\
\vdots \\
\end{pmatrix}
= -K^2 + \frac{q}{(q^2 - 1)^2} \times
\begin{pmatrix}
+ \\
\vdots \\
+ \\
\vdots \\
\end{pmatrix}
\]

The averages of these diagrams generate the \( S_4 \) model discussed in App. C. They can be expressed in terms of \( \perp \) and \( \pm \) triangles:

\[
\begin{pmatrix}
\vdots \\
\sigma_b \\
\vdots \\
\vdots \\
\end{pmatrix}
\]

where \( \sigma_b \) and \( \sigma_c \) label the types of the triangles and are determined by the spins \( \pm \) in the bottom two blocks of the configurations in Eq. E1. Because of the doubling of layers (due to squaring), \( + \) corresponds to (e.g.) \( \sigma_b = 1 \in S_4 \), and \( - \) corresponds to \( \sigma_b = (12)(34) \). We can write these weights as

\[
w_{\perp}(\sigma_b, \sigma_c) = \sum_{\sigma_a} w_{\perp}(\sigma_b, \sigma_c; \sigma_a)
\]

where the sum is over the spin \( \sigma_a \) at the center vertex and

\[
w_{\perp}(\sigma_b, \sigma_c; \sigma_a)
\]

The lower spins of the second and triangles in this product are fixed to be \( \sigma_b \) and \( \sigma_c \) respectively, so are not written explicitly. According to Eqs. C4 and C5, the \( \sigma_b \) and \( \sigma_c \) triangles have the leading order expansion

\[
\begin{pmatrix}
\vdots \\
\sigma_b \\
\vdots \\
\vdots \\
\end{pmatrix}
\leq q^{-|\sigma_b| - |\sigma_b^{-1}\sigma_a|} \leq q^{-|\sigma_a|}
\]

For \( \sigma_a \in S_2 \times S_2 \), we have \( \sigma_a = \sigma_a^{-1} \), and

\[
\begin{pmatrix}
\vdots \\
\sigma_a \\
\vdots \\
\vdots \\
\end{pmatrix}
\leq q^{-|\sigma_a(12)(34)| - |\sigma_a^{-1}\sigma_a|} \leq q^{-|\sigma_a(12)(34)|}
\]

where Eq. E9 is based on the large \( q \) expansion in Eq. (60) of Ref. [14].

Hence

\[
w_{\perp}(\sigma_b, \sigma_c; \sigma_a) \leq q^{-10} \text{ for } \sigma_a \in S_2 \times S_2.
\]
On the other hand, if $\sigma_a \notin S_2 \times S_2$, the $\perp$ triangle is equal to an ordinary triangle (Eq. C8). There will be at least four out-going domain walls in the $\perp$ triangle, and we have
\begin{align}
\int \sigma_b \sigma_c \sigma_a \sim \sigma_a^{(12)(34)} \leq q^{-|\sigma_a^{(12)(34)}|} \leq q^{-4}
(E11)
\end{align}
and
\begin{align}
\int \sigma_a^{(12)(34)} \sim q^{-|\sigma_a^{(12)(34)}|} \leq q^{-4}
(E13)
\end{align}
So
\begin{align}
w_{\perp}(\sigma_b, \sigma_c; \sigma_a) \leq q^{-8} \quad \text{for} \quad \sigma_a \notin S_2 \times S_2
(E14)
\end{align}
These inequalities can be saturated only when
\begin{align}
\sigma_b = 1, \quad \sigma_c = (12)(34) \quad \sigma_a = (13)(24) \quad \text{or} \quad (14)(23).
(E15)
\end{align}
Therefore, the largest terms in Eq. E1 are:
\begin{align}
w_{\perp}(1, (12)(34)) & \sim w_{\perp}(1, (12)(34); (13)(24)) \quad \text{for} \quad \sigma_a \notin S_2 \times S_2 \quad \text{(E16)}
& + w_{\perp}(1, (12)(34); (14)(23)) \quad \text{(E17)}
& \sim 2q^{-8}.
(E18)
\end{align}
In other words, $\Gamma$ dominates and scales as
\begin{align}
\Gamma \sim 2q^{-8}
(E19)
\end{align}
Note that we may also write
\begin{align}
\Gamma^\pi = (Z_u - Z_u^\perp)^\pi.
(E20)
\end{align}
where we have used $P_{\perp} = 1 - P_{\parallel}$ to split $\Gamma$ into $Z_u - Z_u^\perp$, and
\begin{align}
Z_u = \begin{array}{ccc}
\frac{1}{2} & \frac{1}{2} & 0 \\
\frac{1}{2} & \frac{1}{2} & 0 \\
0 & 0 & 1
\end{array}.
(E21)
\end{align}
We may then use the result for $Z_u^\perp$ in Sec. VI A of [14].

1. Protocol to compute $Z(x, t)$ and $W(x, t)$

In the main text we defined the partition function $Z(x, t)$ for a domain wall with fixed upper and lower end-points (fixed by contracting with + and − states at the top boundary and $+$ states at the bottom, see Eq. 33).

A na"ive way to evaluate $Z(x, t)$ numerically is by contracting $U^{(2)}$ with these boundary states as per its definition. However, it is numerically inefficient to store the full four-copy unitary $U^{(2)}$.

It is much easier to compute modified partition functions whose bottom boundary conditions only contain $|\pm\rangle$. We have discussed a subset of these modified partition functions, which we called $Z_{\text{op}}(x, t)$, in the context of dual-unitary circuits in Sec. 40. We can view the unitary circuit representing the single-copy time evolution operator $U(t)$ as a state on $2L$ spins, and view all the legs with the + boundary condition as a subsystem. The modified partition function $Z_{\text{op}}$ is then proportional to the operator purity for this partition. By rewriting
\begin{align}
|+\rangle^* = \frac{1}{q^4 - 1} \left(|+\rangle - \frac{1}{q^2} |\pm\rangle \right)
|\pm\rangle = \frac{1}{q^4 - 1} \left(|\pm\rangle - \frac{1}{q^2} |+\rangle \right)
(E22)
\end{align}
we find that $Z(x, t)$ is a linear combination of these operator purities. Our numerical scheme then computes these operator purities and obtains $Z(x, t)$ from a linear transformation. We can then recursively compute $W(x, t)$ through Eq. 40. This scheme enables us to compute up to $t_{\text{max}} = 8$.

2. The random gate

The matrix elements of the random gate used in Fig. 7(b) are approximately
\begin{align}
u = \begin{bmatrix}
-0.10054 & 0.01426 & -0.53043 & 0.35559 \\
0.48842 & 0.52408 & -0.44873 & -0.19577 \\
-0.22473 & 0.50727 & 0.11189 & 0.26009 \\
0.77248 & 0.02169 & 0.26725 & 0.12918 \\
\end{bmatrix}
(E23)
+ i \begin{bmatrix}
-0.04094 & 0.27291 & -0.05859 & 0.70873 \\
-0.28822 & 0.22707 & -0.15763 & -0.29060 \\
0.11952 & -0.46156 & -0.61862 & 0.04465 \\
0.07122 & -0.35814 & 0.14956 & 0.39873 \\
\end{bmatrix}
(E24)
\end{align}
This gate is less entangling than the Haar average but nevertheless the scheme for extracting $\mathcal{E}(v)$ seems to work well.

3. Protocol to estimate $v_B$

We use the Pauli weight [18, 25, 83] to estimate $v_B$ in the Floquet circuit. The Pauli weight is defined to be fraction the operator weight at a site on non-identity operators. To reduce the number of lengths affecting the data, we place $\sigma^z$ at the leftmost site 0 and probe its Pauli weight at the rightmost site $d = L - 1$:
\begin{align}
W_{\text{Pauli}}(t, d) = \frac{1}{8} \sum_{\alpha=x,y,z} \text{tr}((\sigma_0^\alpha(t), \sigma_0^z)^2)/2^L
& = \frac{3}{4} - \frac{3}{4} \sum_{\alpha=x,y,z} \text{tr}(\sigma_0^\alpha(t)\sigma_0^z\sigma_0^\alpha(t)\sigma_0^z)/2^L
(E25)
\end{align}
According to the operator spreading picture, the Pauli strings in the time-evolved operator $\sigma_0^{-}(t)$ spread to the right with mean velocity $v_B$. Before the Pauli strings reach the right boundary, i.e. when $v_B t \ll d$, the Pauli weight is close to zero. At long times when $v_B t \gg d$, $\sigma_0^{-}(t)$ can be regarded as a random operator, and the Pauli weight is $\frac{3}{4}$.

Fig. 15(a) shows the Pauli weight for the kicked Ising model with parameters in Eq. 77. Due to the circuit structure, the Pauli weights at consecutive integer time steps are equal. Hence we place the data at half integer points, and interpolate the curve (with cubic spline) when the data is greater than zero. Following Ref. [18], we define $t_{\text{arrival}}$ to the time when the interpolated curve is at $\frac{3}{4}$. Fitting $d$ with $t_{\text{arrival}}$ gives us an estimation of $v_B$. Fig. 15(b) shows the results of fitting for kicked Ising gates and the random gate.

![Pauli Weight](image1)

**Figure 15.** (a) Pauli weight for the kicked Ising model with parameters in Eq. 77. (b) Estimation of $v_B$ by fitting $t_{\text{arrival}}$ versus $d$. For the random gate we fit separately for the left and right butterfly velocities.

Figure 16. The approximate entanglement velocity $\mathcal{E}(0)$ computed by truncating to times $\leq t$, for $t$ up to 8.

### 4. The approximate entanglement velocity

The second Rényi entropy after a quench corresponds to a free boundary condition at the bottom for the domain wall. Therefore the system will select out $\min_v \mathcal{E}(v)$ as the entanglement velocity. We extract the approximate entanglement velocity from our numerical data and show its dependence on the truncation time $t$ in Fig. 16.

All the curves start off for $t = 1$ at the purity velocity of the random circuit. Differences between the gates then drive the curves to their respective entanglement velocities, either above or below the value at $t = 1$.

Notably, the curve for the dual-unitary gate shows a clear trend towards the analytical result $v_B = 1$. Curves for the Floquet Ising model, the generic reflection symmetric gate and the fixed random gate behave reasonably well, although up to $t_{\text{max}} = 8$ the fluctuations do not decay fast enough to allow us to understand the nature of the convergence.

### Appendix F: Relation between the root and line-tension function

We give details of the relation between $\mathcal{E}(v)$ and the root $b_0(s)$ described in Sec. IV A. Let

$$z(s,t) \equiv \sum_x Z(x,t)e^{sx}. \quad (F1)$$

This is the partition function with a fixed force on the endpoint rather than a fixed endpoint. Asymptotically, it should scale as

$$z(s,t) \sim \exp(-s_{eq}\Gamma(s) t) \quad (F2)$$

with [18]

$$\Gamma(s) = \min_v \left[ \mathcal{E}(v) - \frac{vB}{s_{eq}} \right], \quad (F3)$$
so that \(-s_{eq}\Gamma(s)\) and \(s_{eq}\mathcal{E}(v)\) form a standard Legendre transform pair. On the other hand
\[
\sum_t z(s, t)b^t = \frac{\omega(s, b)}{1 - \omega(s, b)}. \quad (F4)
\]
Assuming that the smallest root of \(\omega(s, b)\) for a given \(s\) is \(b_0(s)\), then for \(|b| \lesssim |b_0|\)
\[
\sum_t z(s, t)b^t \sim \frac{1}{1 - b/b_0}. \quad (F5)
\]
Matching the expansions in \(b\),
\[
s_{eq}\Gamma(s) = \ln |b_0(s)|. \quad (F6)
\]
Finally, inverting the Legendre transform yields the expression (58) in the main text for \(\mathcal{E}(v)\).

Generically \(b_0(s)\) is a single root. But in the dual-unitary case, \(b_0\) is a double root for \(s = 0\). In Sec. V D, we postulate that \(Z(x, t)\) for the dual-unitary model is approximately constant in space, and so \(z(s, t) \sim te^{-\kappa_0 t}\). The factor of \(t\) in the front indicates a double root on the right hand side of Eq. F4 in the limit of large \(t_{\text{max}}\).

**Appendix G: Triangle weights with two commutative domain walls**

For reference, here we list the exact expressions for \(J(l, (12)(34); \sigma_a)\) for all \(\sigma_a\) and their order relative to the weight of two separated domain walls. We use the notation from Ref. 14, drawing one line for each elementary domain wall crossing an edge of the triangle. An elementary domain wall is a transposition.

There are in total 6 classes. Within each class the configurations are related by the combination of reflection symmetry about the center axis, conjugation by (12), (34) or both.

We categorize them in terms of \(\sigma_a\):

1. I, (12)(34).
   \[
   \begin{align*}
   \begin{array}{c}
   \includegraphics[width=1cm]{triangle1.pdf}
   \end{array} & = \begin{array}{c}
   \includegraphics[width=1cm]{triangle2.pdf}
   \end{array} = \frac{q^8 - q^6 - 12q^4 + 14q^2 + 22}{(q^4 - 9)(q^4 - 4)(q^2 + 1)} \\
   & = \left[ \begin{array}{c}
   \includegraphics[width=1cm]{triangle3.pdf}
   \end{array} \right]^2 \left( 1 + \frac{2}{q^2} + O\left(\frac{1}{q^{12}}\right) \right) \\
   & \quad (G1)
   \end{align*}
   \]

2. (12)(34)
   \[
   \begin{align*}
   \begin{array}{c}
   \includegraphics[width=1cm]{triangle4.pdf}
   \end{array} & = \begin{array}{c}
   \includegraphics[width=1cm]{triangle5.pdf}
   \end{array} = \frac{(q^8 - 12q^4 + 12)(q^2 - 1)}{(q^4 - 9)(q^4 - 4)(q^2 + 1)q^2} \\
   & = \left[ \begin{array}{c}
   \includegraphics[width=1cm]{triangle6.pdf}
   \end{array} \right]^2 \left( 1 - \frac{12}{q^4} + O\left(\frac{1}{q^{12}}\right) \right) \\
   & \quad (G2)
   \end{align*}
   \]

3. (13)(24), (14)(23)
   \[
   \begin{align*}
   \begin{array}{c}
   \includegraphics[width=1cm]{triangle7.pdf}
   \end{array} & = \begin{array}{c}
   \includegraphics[width=1cm]{triangle8.pdf}
   \end{array} = \frac{q^2 - 1}{(q^4 - 9)(q^2 + 1)} \\
   & = \left[ \begin{array}{c}
   \includegraphics[width=1cm]{triangle9.pdf}
   \end{array} \right]^2 \left( \frac{1}{q^2} + O\left(\frac{1}{q^6}\right) \right) \\
   & \quad (G3)
   \end{align*}
   \]

4. (13), (24), (14), (23), (1234), (1423), (1243), (1324)
   \[
   \begin{align*}
   \begin{array}{c}
   \includegraphics[width=1cm]{triangle10.pdf}
   \end{array} & = \begin{array}{c}
   \includegraphics[width=1cm]{triangle11.pdf}
   \end{array} = -\frac{(q^4 + 6)(q^2 - 1)}{(q^4 - 9)(q^4 - 4)(q^2 + 1)q^2} \\
   & = -\left[ \begin{array}{c}
   \includegraphics[width=1cm]{triangle12.pdf}
   \end{array} \right]^2 \left( \frac{1}{q^4} + O\left(\frac{1}{q^8}\right) \right) \\
   & \quad (G4)
   \end{align*}
   \]

5. (123), (124), (132), (142), (234), (134), (243), (143)
   \[
   \begin{align*}
   \begin{array}{c}
   \includegraphics[width=1cm]{triangle13.pdf}
   \end{array} & = \begin{array}{c}
   \includegraphics[width=1cm]{triangle14.pdf}
   \end{array} = \frac{5(q^2 - 1)}{(q^4 - 9)(q^4 - 4)(q^2 + 1)} \\
   & = \left[ \begin{array}{c}
   \includegraphics[width=1cm]{triangle15.pdf}
   \end{array} \right]^2 \left( \frac{5}{q^6} + O\left(\frac{1}{q^{10}}\right) \right) \\
   & \quad (G5)
   \end{align*}
   \]

6. (1324), (1423)
   \[
   \begin{align*}
   \begin{array}{c}
   \includegraphics[width=1cm]{triangle16.pdf}
   \end{array} & = \begin{array}{c}
   \includegraphics[width=1cm]{triangle17.pdf}
   \end{array} = -\frac{3(q^2 - 1)}{(q^4 - 9)(q^2 + 1)q^2} \\
   & = -\left[ \begin{array}{c}
   \includegraphics[width=1cm]{triangle18.pdf}
   \end{array} \right]^2 \left( \frac{3}{q^4} + O\left(\frac{1}{q^8}\right) \right) \\
   & \quad (G6)
   \end{align*}
   \]
[1] Leo P. Kadanoff and Paul C. Martin. Hydrodynamic equations and correlation functions. *Annals of Physics*, 24:419–469, October 1963.

[2] Pierre C. Hohenberg and Bertrand I. Halperin. Theory of dynamic critical phenomena. *Reviews of Modern Physics*, 49(3):435, 1977.

[3] R. Oliveira, OCO Dahlsten, and MB Plenio. Generic entanglement can be generated efficiently. *Physical Review Letters*, 98(13):130502, 2007.

[4] Alioscia Hamma, Siddhartha Santra, and Paolo Zanardi. Quantum entanglement in random physical states. *Physical Review Letters*, 109(4):040502, 2012.

[5] Fernando G. S. L. Brandão, Aram W. Harrow, and Michael Horodecki. Local random quantum circuits are approximate polynomial designs. *Communications in Mathematical Physics*, 346(2):397–434, 2016.

[6] Adam Nahum, Jonathan Ruhman, Sagar Vijay, and Jeongwan Haah. Quantum Entanglement Growth under Random Unitary Dynamics. *Physical Review X*, 7(3):031016, July 2017.

[7] Adam Nahum, Sagar Vijay, and Jeongwan Haah. Operator spreading in random unitary circuits. *Physical Review X*, 8(2):021014, 2018.

[8] C. W. von Keyserlingk, Tibor Rakovszky, Frank Pollmann, and S. L. Sondhi. Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws. *Physical Review X*, 8(2):021013, April 2018.

[9] Tibor Rakovszky, Frank Pollmann, and C. W. von Keyserlingk. Diffusive Hydrodynamics of Out-of-Time-Ordered Correlators with Charge Conservation. *Physical Review X*, 8(3):031058, September 2018.

[10] Vedika Khemani, Ashvin Vishwanath, and David A. Huse. Operator Spreading and the Emergence of Dissipative Hydrodynamics under Unitary Evolution with Conservation Laws. *Physical Review X*, 8(3):031057, September 2018.

[11] Amos Chan, Andrea De Luca, and J. T. Chalker. Solution of a Minimal Model for Many-Body Quantum Chaos. *Physical Review X*, 8(4):041019, November 2018.

[12] Amos Chan, Andrea De Luca, and J. T. Chalker. Spectral statistics in spatially extended chaotic quantum many-body systems. *arXiv:1803.03841 [cond-mat, physics:hep-th, physics:quant-ph]*, March 2018.

[13] Daniel A. Rowlands and Austen Lamacraft. Noisy coupled qubits: Operator spreading and the fredrickson-keyserlingk. Diffusive Hydrodynamics under Unitary Evolution with Conservation Laws. *Physical Review X*, 8(2):021014, 2018.

[14] Aaron J. Friedman, Amos Chan, Andrea De Luca, and J. T. Chalker. Spectral statistics and many-body quantum chaos with conserved charge. *Physical Review Letters*, 123(21):210603, November 2019.

[15] Nicholas Hunter-Jones. Unitary designs from statistical mechanics in random quantum circuits. *arXiv:1905.12053 [cond-mat, physics:hep-th, physics:quant-ph]*, May 2019.
ics from holography. Physical Review D, 98(10):106025, November 2018.

[36] Shinsei Ryu and Tadashi Takayanagi. Holographic derivation of entanglement entropy from the anti–de sitter space/conformal field theory correspondence. Physical review letters, 96(18):181602, 2006.

[37] Veronika E Hubeny, Mukund Rangamani, and Tadashi Takayanagi. A covariant holographic entanglement entropy proposal. Journal of High Energy Physics, 2007(07):062, 2007.

[38] Cesar A Agón and Márk Mezei. Bit threads and the membrane theory of entanglement dynamics. arXiv preprint arXiv:1910.12909, 2019.

[39] Hyungwon Kim, Tatsuhiko N. Ikeda, and David A. Huse. Testing whether all eigenstates obey the eigenstate thermalization hypothesis. Physical Review E, 90(5):052105, November 2014.

[40] Tomáš Prosen. Chaos and complexity of quantum motion. Journal of Physics A: Mathematical and Theoretical, 40(28):7881, 2007.

[41] Sara Gopalakrishnan and Austen Lamacraft. Unitary circuits of finite depth and infinite width from quantum channels. Physical Review B, 100(6):064309, August 2019.

[42] Bruno Bertini, Pavel Kos, and Tomáš Prosen. Exact Correlation Functions for Dual-Unitary Lattice Models in $1+1$ Dimensions. Physical Review Letters, 123(21):210601, November 2019.

[43] M. Akila, D. Waltner, B. Gutkin, and T. Guhr. Particle-time duality in the kicked Ising spin chain. Journal of Physics A: Mathematical and Theoretical, 49(37):375101, August 2016.

[44] Bruno Bertini, Pavel Kos, and Tomáš Prosen. Operator entanglement in local quantum circuits i: Maximally chaotic dual-unitary circuits. arXiv preprint arXiv:1909.07407, 2019.

[45] Benoît Collins and Piotr Śniady. Integration with Respect to the Haar Measure on Unitary, Orthogonal and Symplectic Group. Communications in Mathematical Physics, 264(3):773–795, March 2006. bibtex: collins_integration_2006.

[46] Yinzheng Gu. Moments of Random Matrices and Weingarten Functions. Thesis, Queen’s University, September 2013.

[47] Don Weingarten. Asymptotic behavior of group integrals in the limit of infinite rank. Journal of Mathematical Physics, 19(5):999–1001, 1978.

[48] Aram W. Harrow. The Church of the Symmetric Subspace. arXiv:1308.6595 [quant-ph], August 2013. arXiv: 1308.6595.

[49] Romain Vasseur, Andrew C. Potter, Yi-Zhuang You, and Andreas W. W. Ludwig. Entanglement transitions from holographic random tensor networks. Physical Review B, 100(13):134203, October 2019.

[50] Paolo Zanardi. Entanglement of quantum evolutions. Phys. Rev. A, 63:040304, Mar 2001.

[51] Tomáš Prosen and Iztok Pizorn. Operator space entanglement entropy in a transverse Ising chain. Physical Review A, 76(3):032316, September 2007.

[52] Iztok Pizorn and Tomáš Prosen. Operator space entanglement entropy in $S^2 	imes S^2$ spin chains. Physical Review B, 79(18):184416, May 2009.

[53] J. Dubail. Entanglement scaling of operators: a conformal field theory approach, with a glimpse of simillability of long-time dynamics in 1+1d. Journal of Physics A: Mathematical and Theoretical, 50(23):234001, May 2017.

[54] Tianci Zhou and David J. Luitz. Operator entanglement entropy of the time evolution operator in chaotic systems. Physical Review B, 95(9), March 2017. arXiv: 1612.07327.

[55] Paolo Zanardi, Christof Zalka, and Lara Faoro. Entangling power of quantum evolutions. Physical Review A, 62(3):030301, August 2000.

[56] John Cardy. Renormalisation group theory of branching potts interfaces. Nuclear Physics B, 565(3):506–520, 2000.

[57] David A. Huse, Christopher L. Henley, and Daniel S. Fisher. Huse, henley, and fisher respond. Phys. Rev. Lett., 55:2924–2924, Dec 1985.

[58] Mehran Kardar, Giorgio Parisi, and Yi-Cheng Zhang. Asymmetric butterfly velocities in hamiltonian and circuit models. arXiv preprint arXiv:1812.05589, 2018.

[59] Yong-Liang Zhang and Vedika Khemani. Asymmetric butterfly velocities in 2-local hamiltonians. arXiv preprint arXiv:1909.05988, 2019.

[60] B Kraus and JI Cirac. Optimal creation of entanglement using a two-qubit gate. Physical Review A, 63(6):062309, 2001.

[61] Navi Khaneja and Steffen Glaser. Cartan Decomposition of SU(2n), Constructive Controllability of Spin systems and Universal Quantum Computing. arXiv:quant-ph/0010100, October 2000.

[62] Bruno Bertini, Pavel Kos, and Tomáš Prosen. Exact Spectral Form Factor in a Minimal Model of Many-Body Quantum Chaos. Physical Review Letters, 121(26):264101, December 2018.

[63] Vedika Khemani, David A. Huse, and Adam Nahum. Velocity-dependent Lyapunov exponents in many-body quantum, semi-classical, and classical chaos. arXiv:1803.05902 [cond-mat, physics:hep-th, physics:nl, physics:quant-ph], March 2018. arXiv: 1803.05902.

[64] Adam Nahum, Sthitadhi Roy, Sagar Vijay, and Tianci Zhou. Chaos and the time-ordered two-point function, forthcoming.

[65] Ewan McCulloch and C. W. von Keyserlingk. Forthcoming.

[66] Yichen Huang. Dynamics of Renyi entanglement entropy in local quantum circuits with charge conservation. arXiv:1902.00977 [cond-mat, physics:hep-th, physics:quant-ph], February 2019.

[67] Hong Liu and Paolo Glorioso. Lectures on nonequilibrium effective field theories and fluctuating hydrodynamics. page 008, July 2018.

[68] Mukund Rangamani. Gravity & Hydrodynamics: Lectures on the fluid-gravity correspondence. arXiv:0905.4352 [gr-qc, physics:hep-th, physics:gr], October 2009. arXiv: 0905.4352.

[69] Yaodong Li, Xiao Chen, and Matthew Fisher. Measurement-driven entanglement transition in hybrid quantum circuits. arXiv preprint arXiv:1901.08092, 2019.

[70] Michael J Gullans and David A Huse. Dynamical purification phase transition induced by quantum measurements. arXiv preprint arXiv:1905.05195, 2019.
[72] Brian Skinner, Jonathan Ruhman, and Adam Nahum. Measurement-induced phase transitions in the dynamics of entanglement. *Phys. Rev. X*, 9:031009, Jul 2019.

[73] Yaodong Li, Xiao Chen, and Matthew PA Fisher. Quantum zeno effect and the many-body entanglement transition. *Physical Review B*, 98(20):205136, 2018.

[74] Amos Chan, Rahul M. Nandkishore, Michael Pretko, and Graeme Smith. Unitary-projective entanglement dynamics. *Phys. Rev. B*, 99:224307, Jun 2019.

[75] Soonwon Choi, Yimu Bao, Xiao-Liang Qi, and Ehud Altman. Quantum error correction and entanglement phase transition in random unitary circuits with projective measurements. *arXiv preprint arXiv:1903.05124*, 2019.

[76] M. Szyniszewski, A. Romito, and H. Schomerus. Entanglement transition from variable-strength weak measurements. *Phys. Rev. B*, 100:064204, Aug 2019.

[77] Xiangyu Cao, Antoine Tilloy, and Andrea De Luca. Entanglement in a fermion chain under continuous monitoring. *SciPost Phys.*, 7:24, 2019.

[78] Michael J Gullans and David A Huse. Scalable probes of measurement-induced criticality. *arXiv preprint arXiv:1910.00020*, 2019.

[79] Qicheng Tang and W Zhu. Measurement-induced phase transition: A case study in the non-integrable model by density-matrix renormalization group calculations. *arXiv preprint arXiv:1908.11253*, 2019.

[80] Yimu Bao, Soonwon Choi, and Ehud Altman. Theory of the phase transition in random unitary circuits with measurements. *arXiv preprint arXiv:1908.04305*, 2019.

[81] Chao-Ming Jian, Yi-Zhuang You, Romain Vasseur, and Andreas WW Ludwig. Measurement-induced criticality in random quantum circuits. *arXiv preprint arXiv:1908.08051*, 2019.

[82] Aidan Zabalo, Michael J Gullans, Justin H Wilson, Sarang Gopalakrishnan, David A Huse, and JH Pixley. Critical properties of the measurement-induced transition in random quantum circuits. *arXiv preprint arXiv:1911.00008*, 2019.

[83] Márk Mezei and Douglas Stanford. On entanglement spreading in chaotic systems. *Journal of High Energy Physics*, 2017(5):65, 2017.