# Entanglement Membrane in Chaotic Many-Body Systems

Tianci Zhou[1,*] and Adam Nahum[2,†]

[1]*Kavli Institute for Theoretical Physics, University of California, Santa Barbara, California 93106, USA*
[2]*Theoretical Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom*

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, describing a pairing of forward and backward Feynman trajectories, inspired by the structure emerging in random unitary circuits. This viewpoint provides a framework for applying ideas of coarse graining to dynamical quantities in chaotic systems. We apply the 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 nonrandom Floquet models. We also provide an efficient numerical algorithm for determining the entanglement line tension in $1 + 1$D.

## 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, suggest new aspects of universality [6–11,14]. In particular, they reveal 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" [7,8,11,19–34] of quantum information.

One way to motivate this membrane, which is relevant to the approach we take here, is via the multilayer structure of the quantum circuit (or the multisheet structure of the path integral, in continuum language) representing the observables in question. Algebraically, conventional correlation functions, such as an expectation value $\langle \mathcal{O}(t) \rangle$ 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 $U(t) \otimes U^*(t)$. 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

$$U^{(N)} \equiv U \otimes U^* \dots U \otimes U^* \tag{1}$$

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

The structure arising from this representation 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, i.e., a pairing of trajectories [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].

*tzhou@kitp.ucsb.edu
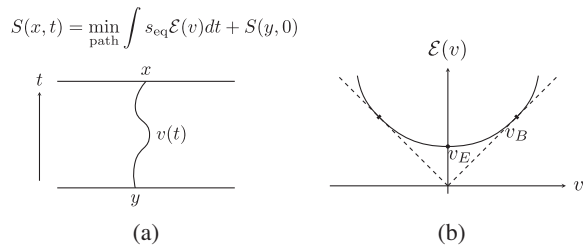†adam.nahum@physics.ox.ac.uk

FIG. 1. The membrane picture of entanglement and other dynamical quantities. (a) Example of a 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 $(\pm v_B, v_B)$.

The goal of this work is to show how to make sense of the pairing field $\sigma(x, t)$ and the membrane in nonrandom systems. Doing so allows the language of the renormalization group to be applied to chaotic dynamics.

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 pairings, and the membrane description, continue to make sense in systems more general and more realistic than random circuits, for the heuristic reason described in the next paragraph. In support of this expectation, there is numerical evidence that a consistent membrane tension can be defined in a generic spin chain [18] and in a dual-unitary circuit [35,36]. 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 [34,37], at leading order in the number of degrees of freedom; this derivation also begins to make a concrete connection between the entanglement membrane and the Ryu-Takayanagi and Hubeny-Rangamani-Takayanagi geometrical pictures for entanglement in AdS space [38–41]. Lattice toy models for the AdS-CFT correspondence, using random tensor networks, also exhibit domain walls between permutation degrees of freedom [42,43]. 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 shed light on various aspects of chaotic dynamics [6–17], and a natural question is how to generalize these calculations to nonrandom systems. Our aim here is to provide a formalism for making this generalization.

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 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 boundary condition 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, an exact local resolution of identity. The insertion includes states $|\sigma\rangle$ associated with pairings, but also a projector onto the complementary part of the Hilbert space, which we denote by $\perp$, containing nonpairing states. 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 nonpairing 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 contribution from any spin configuration that includes a $\perp$ vanishes after averaging, leaving a spin model for permutations only, 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 result is also true in a particular realization of the random circuit, *before* we average over the random unitaries), which 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.
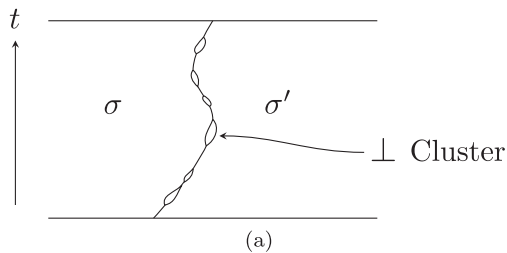
FIG. 2.   Thick membrane (schematic). Except in the special case of Haar-averaged unitaries, the membrane (a domain wall between permutations $\sigma$ and $\sigma'$) is thickened by clusters where the effective spin takes the value "$\perp$." In the multilayer tensor network for $U^{(N)}$, this value represents the propagation of states orthogonal to paired states. (In general, the interior of the thick membrane also contains paired states. For $N > 2$, these may be distinct from $\sigma$ or $\sigma'$ [11,14].)

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.

This description is loosely analogous to the renormalization of the mass of a quasiparticle due to interactions— here, the $\perp$ states dress the structure of the membrane, giving it a larger width (than in the averaged random circuit, where $\perp$ states are suppressed) and a modified line tension. This approach is an application of the renormalization group idea to find the coarse-grained quantities characterizing scrambling.

Before tackling translation-invariant systems, we analyze the case of a *fixed realization* of a random unitary circuit. This case 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 a 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 result leads 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), and give numerical evidence, 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. We first discuss Floquet models with a local unitary circuit structure, since these are the simplest case to visualize. However, the circuit structure is not required: We also apply the method to completely generic Floquet models, involving Hamiltonian evolution in continuous time without any circuit structure.

The numerical results show good convergence for a wide range of chaotic models, including prototypical spin models for quantum chaotic systems, such as Floquet Ising models with longitudinal and transverse fields (kicked-Ising models) [21,44,45].

The line tension $\mathcal{E}(v)$ satisfies some general constraints [18] which provide a highly nontrivial 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 [35,36] and the property of "dual unitarity" which, remarkably, allows a range of exact computations [36,46–49].

In addition to studying particular models, we discuss the general continuum theory for the membrane in $1 + 1$D. We argue that the entanglement membrane has special properties in dual-unitary circuits: Its equation of motion becomes a wave equation rather than being diffusive as it is in generic models.

## A. Organization of the paper

Section II introduces pairing degrees of freedom in the multilayer unitary circuit, leading to an effective spin model (Secs. II A–II C) in which a domain wall structure emerges (Sec. II D).

We first test our formalism in an analytically tractable setting, a fixed realization of a random circuit, in Sec. III. We show that large clusters of $\perp$ in the effective spin model are exponentially suppressed (Sec. III A) and use this result to obtain Kardar-Parisi-Zhang scaling of the Rényi entropy (Sec. III B).

We then develop a systematic formalism for resumming the effect of $\perp$ clusters for translationally invariant Floquet
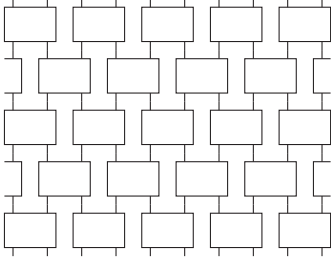
FIG. 3. The structure of the circuit with gates acting on nearest-neighbor sites. In the Floquet models we study, each gate is identical. In the random unitary case, each gate is independently sampled from the Haar ensemble.

circuits. Section IV describes technical preliminaries, and Sec. V applies the method to various models from the literature. Section VI discusses the coarse graining of the membrane, showing that two distinct universality classes arise, for a generic circuit and a self-dual circuit.

We discuss the application of this formalism to operator spreading in Sec. VII and to general Floquet models in continuous time, without any circuit structure, in Sec. VIII.

Section IX describes some open questions. Appendixes contain technical details about the spin model and further numerical results.

## II. DEFINING THE SPIN MODEL

For concreteness, let us focus on time evolution with a unitary circuit made of two-site gates, with the structure in Fig. 3. (The constructions below generalize straightforwardly to evolution without a circuit structure: We describe this generalization explicitly in Sec. VIII.) We use $U(t)$ or simply $U$ to denote the full many-body unitary represented by the entire circuit and the lowercase $u$ to stand for a local two-site gate.

At this point, the local gates $u$ are left arbitrary. In some sections, we restrict to Floquet models with both space and time translation invariance, but this restriction is not necessary. The circuit geometry guarantees that there is no propagation of information outside a strict light cone 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 $\text{Tr}\mathcal{O}(x_n, t_n)...\mathcal{O}(x_1, t_1)$ 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 (here, $U^*$ is the complex conjugate of $U$, taken in an arbitrary local basis, and not the Hermitian conjugate). In contrast, the quantities of interest to us here require traces of the $2N$-layer unitary circuit

$$U^{(N)} \equiv \underbrace{U \otimes U^*...U \otimes U^*}_{N \text{ copies of } U \text{ and } N \text{ copies of } U^*}, \qquad (2)$$

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

$$C(x, t) = \langle \mathcal{O}(x, t)\mathcal{O}'(0, 0)\mathcal{O}(x, t)\mathcal{O}'(0, 0)\rangle \qquad (3)$$

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

$$S_n = -\frac{1}{n - 1}\ln \text{tr}[\rho_A^n(t)], \qquad (4)$$

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 fact does 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 multilayer circuit has $N$ layers of $U$ and of $U^*$ and can be thought of as describing evolution of a "replicated," multicopy 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 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 show how a corresponding local "pairing field" $\sigma(x, t)$ can be defined by coarse graining over microscopic length and timescales.

### A. Introducing the effective spins

To motivate the structures we 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

$$\overline{u_{\text{Haar}}^{(N)}} = \sum_{\tau, \sigma \in S_N} \text{Wg}(\tau^{-1}\sigma; d)|\tau\tau\rangle\langle\sigma\sigma|, \qquad (5)$$

where Wg is the unitary Weingarten function [50–52]. Here, $\tau$ and $\sigma$ are elements of the permutation group $S_N$. These denote a pairing of the $u$ layers, labeled $1, 2, ..., N$, with the $u^*$ layers, labeled $\bar{1}, \bar{2}, ..., \bar{N}$. For example, $\sigma$

corresponds to pairing 1 with $\overline{\sigma(1)}$, 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 $\mathbb{I}$ and the transposition, denoted (12) in cycle notation. In our local basis, the corresponding states at a site are

$$\langle a_1, \overline{a_1}, a_2, \overline{a_2}|\mathbb{I}\rangle = \delta_{a_1\overline{a_1}}\delta_{a_2\overline{a_2}},$$
$$\langle a_1, \overline{a_1}, a_2, \overline{a_2}|(12)\rangle = \delta_{a_1\overline{a_2}}\delta_{a_2\overline{a_1}} \quad (6)$$

(note that we do not normalize these states).

In Ref. [14], we show 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 $\tau$ are 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 is to separate out the pairing (permutation) states $|\sigma\sigma\rangle$ in the multicopy Hilbert space for a pair of sites from states in the orthogonal complement of this space.

Regardless of the choice of unitary gate, we are able to decompose the multilayer unitary as

$$u^{(N)} = P_\parallel + \perp, \quad (7)$$

where $P_\parallel$ is a projection operator onto paired states, which as discussed below is also equal to $\overline{u_{\mathrm{Haar}}^{(N)}}$, and $\perp$ represents the projection of $u^{(N)}$ to the complement of the space of paired states. Equation (5) shows that in the Haar-averaged case the $\perp$ contributions vanish identically. More generally, they do not vanish, and their operator norm is not negligible. However, we 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.

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,

$$\overline{u_{\mathrm{Haar}}^{(N)}} \times \overline{u_{\mathrm{Haar}}^{(N)}} = \overline{u_{\mathrm{Haar}}^{(N)}}, \quad (8)$$

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

$$\overline{u_{\mathrm{Haar}}^{(N)}}|\sigma\sigma\rangle = |\sigma\sigma\rangle \quad (9)$$

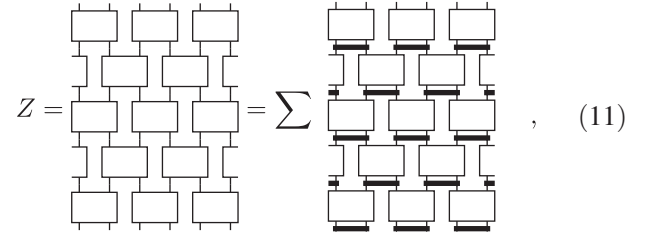for any permutation $\sigma$. This result shows that the Haar-averaged object, which is a Hermitian operator on the replicated space, is equal to $P_\parallel$, the orthogonal projector onto the subspace spanned by permutation states [53].

To separate out contributions from permutation states for a general circuit, we insert the resolution of the identity (on two physical sites)

$$I^{(N)} = P_\parallel + P_\perp \quad (10)$$

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 Eq. (11) below as a bar below each gate.

We view the quantity to be computed, such as 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 Eq. (10), this partition function is a sum of terms with all possible choices of projection operator inserted below each gate. Graphically,



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 property can be seen from the invariance of the Haar measure. For any two-site unitary $u$,

$$u^{(N)} \times P_\parallel = u^{(N)} \times \overline{u_{\mathrm{Haar}}^{(N)}} = \overline{u_{\mathrm{Haar}}^{(N)}} = P_\parallel. \quad (12)$$

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. (11) gives us a well-understood starting point, which we can go beyond by taking into account domains of $\perp$ insertions. Roughly speaking, this approach is 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 is 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 $P_\sigma$:

$$P_\parallel = \sum_{\sigma \in S_N} P_\sigma, \tag{13}$$

which are given by grouping together terms in Eq. (5) with a given $\sigma$:

$$P_\sigma = \left( \sum_\tau \mathrm{Wg}(\tau^{-1}\sigma; q^2) |\tau\tau\rangle \right) \langle\sigma\sigma|. \tag{14}$$

Using the properties of the Weingarten function, one can check that the states appearing in the brackets,

$$|(\sigma\sigma)^*\rangle = \sum_\tau \mathrm{Wg}(\tau^{-1}\sigma; q^2) |\tau\tau\rangle, \tag{15}$$

form the dual basis of $|\sigma\sigma\rangle$, i.e., $\langle(\sigma\sigma)^*|\tau\tau\rangle = \delta_{\tau,\sigma}$, so that the non-Hermitian operators $P_\sigma = |(\sigma\sigma)^*\rangle\langle\sigma\sigma|$ are projectors satisfying (see Appendix A 1)

$$P_\sigma P_{\sigma'} = \delta_{\sigma\sigma'} P_\sigma. \tag{16}$$

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 obstacle is not fundamental, 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 multilayer circuit, is now refined to

$$I^{(N)} = \sum_{\sigma \in S_N} P_\sigma + P_\perp. \tag{17}$$

Let us represent the possibilities graphically as


$$\tag{18}$$

Importantly, the blocks on the left, labeled by permutations, are *independent* of the choice of local unitary gate, because $P_\sigma$ again "absorbs" any unitary. For any gate $u$, we have $u^{(N)}|\sigma\sigma\rangle = |\sigma\sigma\rangle$, so

$$u^{(N)} P_\sigma = P_\sigma. \tag{19}$$

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

We now write $Z$ as a sum over spins $s$, associated with each block in the circuit, which take values in $S_N \cup \{\perp\}$:


$$\tag{20}$$

Throughout this paper, we use $x$, $t$ to label the bond between two lattice sites. The gates are located on half of the bonds.

Note that $\perp$ is simply a choice of label for one of the values that $s$ can take; we could also denote this state by "0" or anything else. At each location, $s$ runs over $N! + 1$ different values.

The weight of a spin configuration $\{s\}$ in this partition sum is given by contracting the network. In general, this weight is not simply a product of local terms. However, we argue that, for the boundary conditions of interest and for generic models, there is an effective notion of locality. We first describe some basic properties of the weights (Secs. II B and II C) and then specify to the case of interest, where the boundary conditions induce a domain wall structure in the spin configuration (Sec. II D).

### B. General properties of the spin model

First, consider the subpartition function without any $P_\perp$ insertions (i.e., with $s \neq \perp$ everywhere), which we denote $Z_{0\perp}$.

$Z_{0\perp}$ is a statistical model whose "spins" $\sigma$ are permutations [14]. The spins have local three-body interactions on "down-pointing triangles" (and no interactions on up-pointing triangles). To be specific, each down-pointing triangle such as


$$\tag{21}$$

has an associated weight


$$\tag{22}$$

From Eq. (14), the interaction triangle is (see Appendix A 1)

$$J(\sigma_b, \sigma_c; \sigma_a) = \sum_\tau \langle\sigma_b\sigma_c|\tau\tau\rangle \mathrm{Wg}(\tau^{-1}\sigma_a; q^2). \tag{23}$$

Equivalently, the triangle is the coefficient in the expansion

$$\langle \sigma_b \, \sigma_c | \, P_{\parallel}^{(N)} = \sum_{\sigma_a \in S_N} \overset{\sigma_b \quad \sigma_c}{\underset{\sigma_a}{\bigtriangledown}} \langle \sigma_a \, \sigma_a |, \qquad (24)$$

where the two arguments in the bras refer to the pair of spatial states.

This spin model is described in Ref. [14] for the random circuit. In that context, the spins $\sigma$ arise from the Haar average over the physical random unitary in Eq. (5). The three-body interaction appears after integrating out the "$\tau$" spins. Here, the sum over the $\tau$ spins appeared at the state level in defining the projection operator in Eq. (14).

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

$$\underset{\boxed{-}}{\boxed{+|+}} = 0, \qquad \underset{\boxed{+}}{\boxed{+|+}} = 1, \qquad \underset{\boxed{+}}{\boxed{+|-}} = K, \qquad (25)$$

with

$$K = \frac{q}{q^2 + 1}, \qquad (26)$$

which follows from Eq. (14) with the explicit expression

$$P_+ = |(++)^*\rangle\langle + + |$$
$$= \frac{1}{q^4 - 1} \left( | + +\rangle\langle + + | - \frac{1}{q^2} | - -\rangle\langle + + | \right) \quad (27)$$

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 Appendix A 2 for a further simplification). For example, as a generalization of the first equality in Eq. (25) above,

$$\overset{\sigma_b \quad \sigma_b}{\underset{\sigma_a}{\bigtriangledown}} \equiv \underset{\boxed{\sigma_a}}{\boxed{\sigma_b | \sigma_b}} = \delta_{\sigma_a, \sigma_b}. \qquad (28)$$

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

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 $|\sigma_b \sigma_b\rangle$:

$$\underset{\boxed{\perp}}{\boxed{\sigma_b | \sigma_b}} = 0. \qquad (29)$$

So again in the purity example above $\perp$ can be inserted only in the backward light cone emanating from the entanglement cut.

Let us discuss the locality of the spin interactions. We see 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 statement 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 result holds because surrounding a cluster of $\perp$ with permutation states dictates a definite way of contracting up the $\perp$ blocks, yielding a $c$ number. (We 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 is regained after coarse graining beyond this scale.

### C. Symmetry of the spin model

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

$$G_N = (S_N \times S_N) \rtimes \mathbb{Z}_2. \qquad (30)$$

This symmetry is a consequence of the fact that Eq. (2) involves multiple copies of the same unitary [14,43]. The importance of $S_N \times S_N$ symmetry has been emphasized in random tensor networks [43], where permutations labeling pairings also appear [42].

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 permutation gives the $S_N \times S_N$ subgroup of $G_N$, which acts on the spins via

$$\sigma \to g_L \sigma g_R^{-1}, \qquad \perp \to \perp. \qquad (31)$$

The $\mathbb{Z}_2$ subgroup of $G_N$ 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

$$\sigma \to \sigma^{-1}, \qquad \perp \to \perp. \qquad (32)$$

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

The simplest case is $N = 2$, when $G_N$ reduces to an Ising-like $\mathbb{Z}_2$ symmetry relating $+ = \mathbb{I}$ and $- = (12)$. In this case, the symmetry in Eq. (32) becomes trivial, since $\sigma = \sigma^{-1}$.

In the subpartition function $Z_{0\perp}$, the weights $J$ are invariant under $G$, because the Weingarten function and the overlaps in Eq. (14) depend only on the cycle structure of products of the form $\sigma\tau^{-1}$. This cycle structure is invariant under the above operations.

The $G_N$ symmetry is a symmetry of the bulk interactions for the spin model. It is, however, strongly broken by the boundary conditions we require.

### D. 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

$$\exp(-S_2) = \langle \ldots + + - - \ldots | U^{(2)}(t) | \psi^{(2)} \rangle, \quad (33)$$

where we label the two permutation states at a site, $\mathbb{I}$ and $(12)$, by $+$ and $-$, respectively. 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 subpartition 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. (25)], 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 result is illustrated in Fig. 4(top). Formally,

$$Z_{0\perp} = \sum_{\substack{\text{directed} \\ \text{paths}}} K^t = 2^t K^t, \quad (34)$$

where $K$, defined in Eq. (26), 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 showed above 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 appear only 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 $-$ on the right. In principle, this thick domain
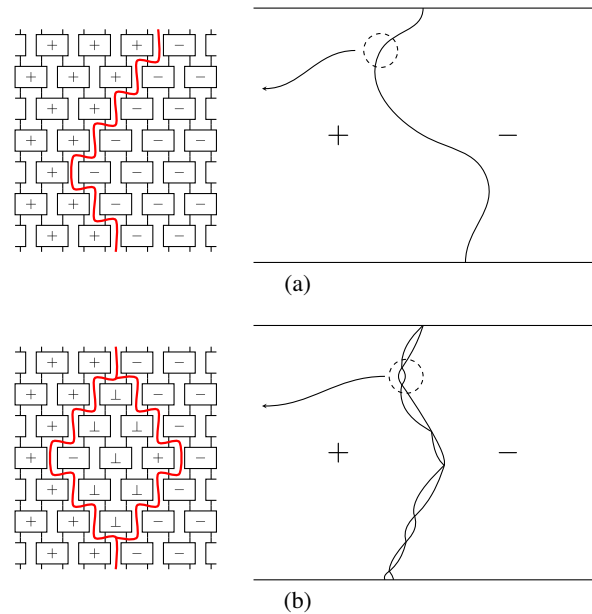


(a)

(b)

FIG. 4. Domain walls in (a) $Z_{0\perp}$ (or in the Haar-averaged circuit) and (b) $Z$, taking into account $\perp$ states. In (a), the width of the domain wall is 1. The width of the domain wall in (b) is argued to be of order 1.

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. (20) is not useful.

However, we argue that this filling 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 shown in Fig. 4. The domain wall has a nontrivial structure made up of clusters that can contain $\perp$, $+$, and $-$; these clusters are separated by the places in spacetime where the domain wall becomes "thin," which we 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 $\mathcal{E}(v)$ for this membrane.

We give analytical and numerical evidence for this picture in Secs. III and IV. In some cases, there may be a small parameter which makes the analytical calculation of renormalized line tension simple. We argue in Sec. III that such a simplification applies 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 see below how to extract it from simulations.

We now rewrite the partition function of the domain wall in a way which is useful when the width of the thick domain wall does not grow with $t$. [54] 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. (29)].

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 $-$ at position $x$ at the top boundary and at $y$ at the bottom boundary. This condition can be enforced using the dual states $|(\pm\pm)^*\rangle$ defined in Eq. (15):

$$Z(x, y; t) = \langle \ldots + + - - \ldots | U^{(2)}(t) | \ldots (++)^*(--)^* \ldots \rangle \tag{35}$$

[$x$ labels bonds of the lattice as in Eq. (20)]. In a translationally invariant system, the domain wall connecting $(x, t)$ and $(y, 0)$ [cf. the schematic picture in Fig. 1(a)] forms a path that is straight on scales of order $t \gg 1$, with coarse-grained velocity $v = (x - y)/t$. The free energy of the domain wall is proportional to the line tension $\mathcal{E}(v)$ for a path with this velocity: At leading order in $t$,

$$-\ln Z(x, y; t) \sim s_{\text{eq}} \mathcal{E}(v) t, \qquad v \equiv \frac{x - y}{t}. \tag{36}$$

This asymptotic scaling is one way to define $\mathcal{E}(v)$. However, it is more efficient to extract $\mathcal{E}(v)$ directly from $W$, as we discuss.

Let the number of time steps where the domain wall is thin be $M + 1$, and let their spacetime coordinates be

$$(y, 0), (x_1, t_1), \ldots, (x_{M-1}, t_{M-1}), (x, t). \tag{37}$$

The steps are of time duration and spatial extent

$$\Delta t_i = t_i - t_{i-1} \geq 1, \tag{38}$$

$$\Delta x_i = x_i - x_{i-1} \in \mathbb{Z}. \tag{39}$$

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 $M$:

$$Z(x, y; t) = \sum_{M \geq 1} \sum_{\substack{\{\Delta x\}, \{\Delta t\} \\ \sum \Delta x_i = x - y \\ \sum \Delta t_i = t}} \prod_{i=1}^{M} W(\Delta x_i, \Delta t_i). \tag{40}$$

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

The weight for a step of duration 1 is

$$W(\pm 1, 1) = K. \tag{41}$$

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

$$W_{0\perp}(x, t > 1) = 0. \tag{42}$$

The $\perp$ insertion produces nonzero 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 disappear under coarse graining. However, we suggest that in some fine-tuned cases negative weights are important (Sec. VI).

The rewriting in Eq. (40) is exact for the boundary conditions we choose above, but this rewriting is useful only 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 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 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. (40) to compute $W(x, t)$:

$$W(x, t) = Z(x, t) - \sum_{t' < t} \sum_{y} W(x - y, t - t') Z(y, t'). \tag{43}$$

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 $M = 1$ and from paths with $M > 1$ 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 $W(x, t)$ by starting with the weight for $t = 1$ in Eq. (41) and employing Eq. (43) 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 are limited to $t \le t_{\max}$ for some $t_{\max}$, so the usefulness of the algorithm depends on how rapidly the weights $W$ get smaller at larger $t$. We apply the algorithm to some nonrandom 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 of 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 modification 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. (41)]. For some choices of $u$, this modification may give better approximations for small $t_{\max}$ (see Sec. VIII). Here we 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.

### E. Weights for smallest nontrivial ⊥ cluster

In some limits, it is a good approximation to consider only the smallest possible ⊥ cluster, namely, a single isolated ⊥ 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 ⊥ is inserted.

Recall that

$$\frac{\boxed{+}\boxed{+}}{\boxed{\perp}} = \frac{\boxed{-}\boxed{-}}{\boxed{\perp}} = 0. \qquad (44)$$

as a consequence of unitarity [Eq. (29)]. Therefore, there are four possible local configurations for one isolated ⊥, up to $+ \leftrightarrow -$ symmetry (assuming the row above contains a single domain wall):



$$(45)$$

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. (49).

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 $t = 2$:

$$W(0, 2) = ①, \qquad W(\pm 2, 2) = ②. \qquad (46)$$

By contrast, ④ does not contribute to $W(x, 2)$, as it does not yield a thin domain wall at the bottom. This diagram appears 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 ①–④, recall that we can regard the gate $u$ as a quantum state for four $q$-state spins ("vectorization" of the operator). In tensor network language, this statement simply means that we regard all four of the legs sticking out of $u$ as physical spin indices. Let us label these legs $A$, $B$, $C$, and $D$ as

$$u = \begin{array}{cc} A & B \\ \boxed{\phantom{xxxx}} \\ C & D \end{array}. \qquad (47)$$

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 [55–59] and have been studied in the context of quantifying the "entangling power" of unitary gates [55,60]. 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 $\mathcal{P} = e^{-S_2}$:

$$\mathcal{P}_\parallel \equiv \mathcal{P}_{AC}, \qquad \mathcal{P}_\times \equiv \mathcal{P}_{AD}. \qquad (48)$$

All others are equivalent (e.g., $\mathcal{P}_{BD} = \mathcal{P}_{AC}$) or independent of $u$ (e.g., $\mathcal{P}_A = q^{-1}$ and $\mathcal{P}_{AB} = q^{-2}$). In terms of these, the weights in Eq. (45) are (recall ③ = ②)

$$① = \left(\frac{q^4}{q^4 - 1}\right)^2 \left(\mathcal{P}_\parallel + \frac{1}{q^6}\mathcal{P}_\times - \frac{2K}{q} - \frac{2K}{q^7}\right), \quad (49)$$

$$② = -\left(\frac{q^3}{q^4 - 1}\right)^2 \left(\mathcal{P}_\parallel + \frac{1}{q^2}\mathcal{P}_\times - \frac{2K}{q} - \frac{2K}{q^3}\right), \quad (50)$$

$$④ = \left(\frac{q^3}{q^4 - 1}\right)^2 \left(\mathcal{P}_\times + \frac{1}{q^2}\mathcal{P}_\parallel - \frac{2K}{q} - \frac{2K}{q^3}\right). \quad (51)$$

See Appendix A 4 for computations.

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

### 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 $q$. These models are a useful test ground for our approach, because we can quantify the suppression of large clusters of $\perp$ blocks analytically. This section provides analytical support for the thin domain wall conjecture: Readers keen to see visual evidence for the success of the method can skip ahead to the numerics in Sec. V.

It is important here to distinguish between the *Haar average* of $e^{-S_2}$ or of $U(t)^{(2)}$ 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 $U(t)^{(2)}$ over the ensemble.

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

Consider a particular spin configuration in the partition sum $Z$ [Eq. (20)]. As noted in Sec. II B, each cluster of $\perp$ blocks in the configuration contributes separately to the Boltzmann 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 $\pm$ 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 $\overline{\Omega_C} = 0$ because of cancellation between positive and negative values. This vanishing is just the statement that, after Haar averaging, $Z$ becomes equal to $Z_{0\perp}$, where no $\perp$ blocks appear.

However, we can ask what the magnitude of $\Omega_C$ is for a typical realization of the circuit. We can quantify this magnitude by computing $\overline{\Omega_C^2}$, which is not zero. We show below that (at least for large $q$, where the calculation is controlled) $\overline{\Omega_C^2}$ 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 quantitative 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 or directed polymer in a random medium (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 the use of the replica trick (Sec. III B).

### A. Exponential suppression of large $\perp$ clusters

Let us consider an isolated cluster $C$ of $\perp$ 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. Equation (44) shows that the weight of the spin configuration vanishes if any $\perp$ block has two spins of the same sign ($++$ or $--$) above it, so the $\perp$ cluster must lie on a domain wall between $+$ and $-$. We wish to compute the mean square weight of the cluster, $\overline{\Omega_C^2}$.

Haar averaging maps $\overline{\Omega_C^2}$ 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 Appendix A 3. The domain wall labeling 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 labeled by the permutation $\sigma_L^{-1}\sigma_R$.

If we Haar average the circuit $u^{(4)}$, we obtain spins $\sigma \in S_4$ with the interactions $J$ specified in Eq. (14). Here, the interactions are modified, because we fix the configurations of spins $s \in \{+, -, \perp\}$ before squaring and averaging. This fixing 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 $S_4$ spins are fixed to $\mathbb{I}$, 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 $S_2 \times S_2 \subset S_4$ are suppressed in the interior of the cluster: A triangle whose vertices are all equal and take one of the values $\mathbb{I}$, (12), (34), or (12)(34) has weight zero. In fact, we have
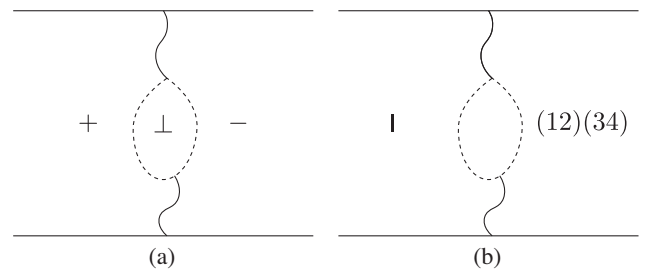


FIG. 5.  (a) A domain wall configuration with a $\perp$ cluster. The weight is nonzero only when the domain wall passes through the $\perp$ cluster. (b) The average of the square of (a) in terms of the $S_4$ spin model. The spin states outside the cluster are fixed.

$$\sigma \underset{\sigma}{\overset{\perp}{\bigtriangledown}} \sigma = \begin{cases} 1 & \text{for } \sigma \notin S_2 \times S_2, \\ 0 & \text{for } \sigma \in S_2 \times S_2, \end{cases} \quad (52)$$

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 in Ref. [14]):

$$(12)(34) = (14)(23) \times (24)(13). \quad (53)$$

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 value does not incur any "bulk" cost, because a triangle all of whose vertices are equal to (14)(23) has weight unity by Eq. (52).

The two doubled domain walls along the two boundaries of the cluster incur a weight $q^{-4}$ per time step. [61] Consequently, the mean squared cluster weight is of the order of $\overline{\Omega^2} \sim q^{-4t_{cluster}}$ or even smaller, where $t_{cluster}$ is the time duration of the cluster. (We do not include the cost of the "thin" domain wall sections above and below the cluster.) This result indicates a typical magnitude

$$\Omega \sim q^{-2t_{cluster}} \quad (54)$$

for a $\perp$ cluster of a given size in a typical large $q$ circuit. Note that this weight is much *smaller* than the weight $\sim q^{-t_{cluster}}$ 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 $t_{cluster}$ steps has a weight that is suppressed by *at least* $q^{-t_{cluster}}$ compared to the leading configurations in the same spacetime region.

It should be noted that $\perp$ clusters are suppressed by a cost in the exponent that scales not with their area in spacetime but instead with the length of their boundary. This scaling 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 [62]. 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 favors $\perp$ 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 compensation does not occur for the quantities considered here, however. Indeed,

for entanglement growth starting from a product state, the $t = 0$ boundary condition favors $\pm$ over $\perp$.

The results above strongly support the conjectures of the previous section. For example, there we define weights $W$ for "steps" of variable temporal duration that connect locations where a domain wall between $+$ and $-$ is thin. We conjecture that the contribution of long steps to the partition function is suppressed. The calculation above confirms this conjecture explicitly for a typical circuit realization at large $q$. (A realization of a random circuit is not translation invariant, but $W$ generalizes directly to this case.) We find that the contribution of long steps to the line tension is suppressed exponentially ($\sim q^{-\Delta t}$) in the step duration $\Delta t$. Note that, according to our definition in Sec. II D, a step of duration $\Delta t \gg 1$ does not need to include $\mathcal{O}(\Delta t) \perp$ blocks: For example, we can have a step with a single $\perp$ block at the top, which allows the domain wall to branch into three domain walls between $+$ and $-$, which merge again at the bottom of the step. Such configurations are also exponentially suppressed, simply because they contain extra domain walls, each contributing an $\mathcal{O}(q^{-\Delta t})$ factor to the cost [14].

Our results in Sec. V for translation-invariant Floquet spin-1/2 circuits are also compatible with exponential suppression of large steps (see Appendix B 5).

The results in this section 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 situation next, before moving on to deterministic models.

## B. KPZ and 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 $\overline{Z} = Z_{0\perp}$ (Sec. II). However, the averaged Rényi entanglement entropy $\overline{S_2}$ cannot be obtained from $-\log(\overline{Z})$ because of fluctuations.

Let us first recall the replica approach used previously. It involves computing the replicated partition function $\overline{Z^k}$ and extracting quantities like $\overline{S_2}$ from the $k \to 0$ limit. At large $q$, there is a simple picture for the replicated partition function $\overline{Z^k}$. 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 [63,64]. 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 have exactly the same mutually attractive interactions as the domain walls here. Therefore, we indirectly map the
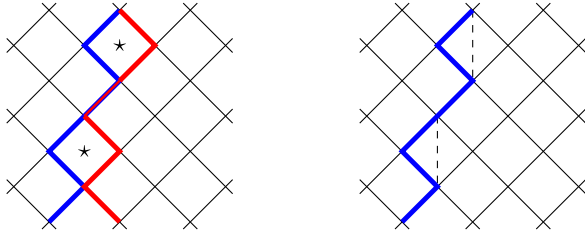
FIG. 6. Domain walls on a tilted square lattice. Left: In the replica approach, two domain walls have an attractive interaction when they both visit two sides of a plaquette (starred). Right: This replica treatment describes a single domain wall subject to a random potential on the vertical bonds. The domain wall in the figure is affected by the potentials on the two dashed bonds.

entanglement to the free energy of a single domain wall in a random potential that modifies the weights of steps. This mapping leads to KPZ scaling [63,64] 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 show 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. Furthermore, at large $q$, the local configuration labeled ① in Eq. (45) dominates over the other configurations in Eq. (45) (see Appendix A 5).

By Eq. (44), these $\perp$ insertions are restricted to lie *on* the domain wall. In the notation of Sec. II D, we have a nonzero weight

$$W(0, 2; u) \qquad (55)$$

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, the weight yields an uncorrelated random potential for our "polymer" in spacetime [65]. 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 ① in Eq. (45) (see Appendix A 5):

$$\overline{\boxed{\begin{array}{c} + \\ + \ \perp \ - \\ + \end{array}}} \simeq \frac{2K^4}{q^4}, \qquad (56)$$

where $K = q/(q^2 + 1)$ is defined in Eq. (26). This result 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.

## IV. FLOQUET SPIN-1/2 CHAINS: PRELIMINARIES

We now apply the formalism of Sec. II to models that are invariant under both space and time translations. In the brickwork circuit geometry, such a model is defined by a single two-site unitary $u$.

In Sec. II D, we define the "thick" domain wall that appears in the calculation of $Z(x, t)$. This definition reduces this partition function to one for a directed path with irreducible step weights $W(\Delta x, \Delta t)$ that take into account successively larger clusters of $\perp$. The line tension of this path can be extracted easily from the weights $W$.

In Sec. V, we present numerical data for several models using this scheme. In preparation for this presentation, we next describe how to obtain systematic approximations to $\mathcal{E}(v) = \mathcal{E}_2(v)$ and the butterfly velocity $v_B$ from the weights $W(\Delta x, \Delta t)$ (see Sec. IV A).

The numerical method for obtaining the weights $W$ themselves is described in Appendix B 1. We use exact diagonalization, which allows us to treat $\Delta t$ up to $t_{\max} = 8$. Together with the analytical results below, this formalism yields a straightforward algorithm which can be applied to any circuit. Section VIII gives an extension to dynamics that are not of circuit form.

### A. Extracting $\mathcal{E}(v)$ from $W$

Recall that the line tension function is encoded in the asymptotics of the partition function $Z(x, t)$ defined in Eq. (40):

$$Z(x, t) \sim \exp[-s_{\rm eq}\mathcal{E}(v)t], \qquad v = x/t. \qquad (57)$$

We use generating functions to determine $\mathcal{E}(v)$. Define generating functions for the partition function $Z$ and for the irreducible weights $W$:

$$\zeta(s, b) = \sum_{x,t} Z(x, t)e^{sx}b^t, \qquad (58)$$

$$\omega(s, b) = \sum_{x,t} W(x, t)e^{sx}b^t. \qquad (59)$$

(The variable $s$ in this section should not be confused with the spin variables elsewhere in the paper.) We can relate these generating functions using Eq. (40), which yields a geometric sum for $\zeta$:

$$\zeta(s, b) = \frac{\omega(s, b)}{1 - \omega(s, b)}. \qquad (60)$$

By truncating the sum in Eq. (59) at the order of $b^{t_{\max}}$, we obtain a polynomial approximation to $\omega(s, b)$. This approximation is physically motivated: It amounts to setting a maximum step length $t_{\max}$ in the directed walk of Eq. (40). We expect this approximation to improve systematically with $t_{\max}$ for chaotic models.

We then need the prescription for obtaining the asymptotics of $Z$, or, in other words, the line tension $\mathcal{E}(v)$, from the numerically accessible quantity $\omega(s, b)$, which can be done with the pole method. For a given $s$, let $b_0(s)$ be the smallest root (in absolute value) of the denominator $\omega(s, b) - 1$ in Eq. (60). Then, the desired relation is

$$\mathcal{E}(v) = \frac{1}{s_{\text{eq}}} \max_s [\ln |b_0(s)| + vs]. \tag{61}$$

This relation is explained in Appendix C. The root $b_0$ appearing here has a physical meaning (although we do not need it here): If we write

$$s_{\text{eq}} \Gamma(s) = \ln |b_0(s)|, \tag{62}$$

then $s_{\text{eq}} \Gamma(s)$, which is related to $s_{\text{eq}} \mathcal{E}(v)$ by a Legendre transformation, Eq. (61), 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 $t_{\max} = 1$, corresponding to the partition function without any $\perp$ blocks. We then need only

$$W(x, 0) = 0, \qquad W(\pm 1, 1) = K. \tag{63}$$

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

$$b_0(s) = \frac{1}{2K \cosh s}. \tag{64}$$

The quantity $\ln b_0(s) + vs$ in Eq. (61) is maximized by $s = \tanh^{-1} v$. Plugging this result in gives

$$s_{\text{eq}} \mathcal{E}(v) = \ln K^{-1} + \frac{1-v}{2} \ln \frac{1-v}{2} + \frac{1+v}{2} \ln \frac{1+v}{2}. \tag{65}$$

This equation is indeed the correct result for $Z_{0\perp}$ [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 $t = t_{\max}$ and neglect $W$ for larger times. Given a constant $s$, we then solve numerically for the smallest zero $b_0(s)$ of

$$1 - \sum_{x, t \leq t_{\max}} W(x, t) e^{sx} b_0^t = 0. \tag{66}$$

If we define [66]

$$v_s = \text{Re} \frac{\sum_{x, t \leq t_{\max}} W(x, t) e^{sx} b_0(s)^t \times x}{\sum_{x, t \leq t_{\max}} W(x, t) e^{sx} b_0(s)^t \times t}, \tag{67}$$

then $\mathcal{E}(v_s)$ is given by $s_{\text{eq}}^{-1} [\ln |b_0(s)| + v_s s]$. [We see this result by setting the $s$ derivative of the argument in Eq. (61) to zero and extracting the derivative of $\ln |b_0(s)|$ from Eq. (66).] We iterate over $s$ to construct the entire curve.

This approach also allows us to extract the butterfly speed $v_B$, as the point [67] 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 $\mathcal{E}(0)$.

It is worth noting that this generating function approach sums up an infinite number of domain wall configurations directly in the thermodynamic limit, eliminating a significant source of finite-$t$ effects. As an illustration, if (hypothetically) steps of size greater than $t_{\max}$ have zero weight, the present approach gives the *exact* result for $\mathcal{E}(v)$ already at time $t_{\max}$. In reality, the weights are not zero for large $t$, but the analytical results in Sec. III A suggest that the convergence in $t_{\max}$ is typically exponential.

By contrast, attempting to directly extract the exponential decay rate of $Z(vt, t)$ by a fit for $t \leq t_{\max}$ is subject to finite-$t$ effects that are generically only polynomially small in $t_{\max}$ (as one can see in the case where the domain wall is a simple random walk). Therefore, understanding the domain wall structure improves the computation of the line tension.

Reference [18] extracts the line tension associated with the von Neumann entropy for a chaotic Ising model directly from the operator entanglement of $U(t)$. That study indeed notes larger finite-$t$ effects than those found here. However, we do not address the von Neumann entropy (as opposed to the higher Rényi entropies) here.

## B. General parameterization

In our discussion of the numerical results, it is sometimes useful to use the following parameterization of the two-site gate $u$, which is an arbitrary SU(4) matrix [36,47,70,71]:

$$u = \begin{array}{c} \boxed{u_1} \ \boxed{u_2} \\ \boxed{u_{\text{sym}}} \\ \boxed{u_3} \ \boxed{u_4} \end{array} \tag{68}$$

where

$$u_i = \exp \left( -i \sum_{\alpha=x,y,z} h_\alpha^{(i)} \sigma^\alpha \right), \tag{69}$$

$$u_{\text{sym}} = \exp \left( -i \sum_{\alpha=x,y,z} J_\alpha \sigma^\alpha \otimes \sigma^\alpha \right). \tag{70}$$

The $u_i$ scramble the individual sites, and the two-site unitary $u_{\text{sym}}$ entangles the two sites.

We 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 $u \to (v \otimes v)^{-1} u (v \otimes v)$, where $v$ is any single site unitary, which 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

$$\mathbf{h}^{(i)} = \mathbf{h} \quad \text{for } i = 1, 2, 3, 4, \tag{71}$$

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

## V. APPLICATION TO FLOQUET CIRCUITS

We are finally ready to apply the numerical scheme in Sec. IV A to a variety of Floquet circuit models, described in the following subsections. In Sec. VIII, we show that the restriction to circuits is not necessary. But circuits are especially convenient for numerics, since they have a strict light-cone structure (propagation of information outside the light cone is not only exponentially suppressed, as in generic spin chains [19], but exactly zero). The space of circuit models is also already very rich.

For each model we study, we show the sequence of approximations to $\mathcal{E}(v)$, indexed by $t = 1, \ldots, t_{\max}$ with $t_{\max} = 8$; 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. (65). 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. V C, we 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 $v_B^{L/R}$. To visualize this check, we plot the boundary curve $\mathcal{E} = |v|$ with dotted lines. We also mark an estimate of $v_B$ from an independent numerical computation of the OTOC (the protocol is described in Appendix B 3). 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 $t_{\max}$ (indeed, we see that they can be disobeyed by the small-$t_{\max}$ approximations).

### A. Generic symmetric and asymmetric gates

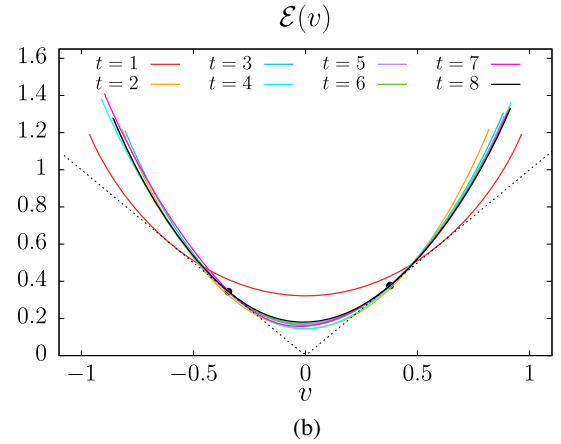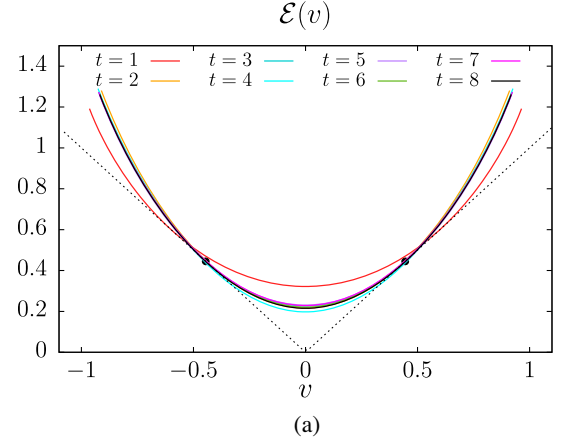First, we consider the following one-parameter family of reflection-symmetric gates [see Eqs. (68)–(71)]:



(a)



(b)

FIG. 7. Numerical computation of $\mathcal{E}(v)$ with $t_{\max}$ up to 8 for (a) the reflection-symmetric gate in Eq. (72) with $x = 0.8$; (b) a generic nonsymmetric gate (specified in Appendix B). Black dots mark estimates of $v_B^{L/R}$ from an independent calculation of the OTOC in Appendix B 3.

$$\mathbf{J} = x\frac{\pi}{4}(3, -4, 5)/\sqrt{50}, \qquad \mathbf{h} = \frac{\pi}{5}(1, 2, 3)/\sqrt{14}. \tag{72}$$

The parameter $x$ tunes the strength of the interaction between the two qubits. For small $x$, local scrambling takes a long time: We do not expect our algorithm to converge rapidly in that case, so we choose $x$ reasonably large. Otherwise, the numbers above are arbitrary and are chosen to ensure (i) that there is no fine-tuning in the sense of any of the coefficients $J_\alpha$ vanishing and (ii) that $u_i$ represents a rotation on the Bloch sphere by a significant angle (here, $\pi/5$).

Figure 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 $(v_B, v_B)$ as obtained from an independent calculation of $v_B$.

Next, Fig. 7(b) shows an example of a generic gate without the reflection-symmetry constraint. We simply pick a random gate from the Haar distribution on U(2) and use it to build a translation-invariant circuit. The matrix elements of this gate are given explicitly in Appendix (B2).

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 $v_B^L$.

## B. Chaotic Ising models

Next, we consider Floquet "kicked" Ising models [21,44,45] with the Floquet unitary

$$U_{\text{Floq}} = \exp(-i\tau H_z)\exp(-i\tau H_x) \qquad (73)$$

with a longitudinal field to spoil integrability:

$$H_x = \sum_{i=1}^{L} h_x \sigma_i^x, \qquad H_z = \sum_{i=1}^{L-1} \sigma_i^z \sigma_{i+1}^z + \sum_{i=1}^{L} h_z \sigma_i^z. \qquad (74)$$

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

$$u = \begin{array}{c} \boxed{u_z} \\ \boxed{u_x}\,\boxed{u_x} \\ \boxed{u_z} \end{array}, \qquad (75)$$

where

$$u_x = \exp(-i\tau h_x \sigma^x), \qquad (76)$$

$$u_z = \exp\{-i\tau[\sigma^z \otimes \sigma^z + ih_z(\sigma_i^z \otimes \mathbb{I} + \mathbb{I} \otimes \sigma^z)/2]\}. \qquad (77)$$

We expect that for generic values of the parameters $(\tau, h_x, h_z)$ 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 [35,36,46–48,72]:

$$\tau = \frac{\pi}{4}, \qquad h_x = 1, \qquad h_z \text{ generic}, \qquad (78)$$

where some quantities can be computed exactly. In particular, entanglement growth is maximal on this line. This fact is related to the property 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 [36,46,47], which is 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 [36,72].

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]:

$$\mathcal{E}(v) = 1. \qquad (79)$$

This case is an interesting test 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 case in Sec. VI.

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

$$\tau = 1, \qquad h_x = 0.9045, \qquad h_z = 0.8090. \qquad (80)$$

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

Kim and Huse study similar values of the fields [73] but with the smaller period $\tau = 0.8$ [44]. However, the Kim-Huse values $(\tau, h_x) \simeq (0.8, 0.90)$ happen to be quite close to the dual-unitary values, [74] which are $(\tau, h_x) \simeq (0.785, 1)$. This coincidence explains the strong entanglement growth in the Kim-Huse model. Our numerical results for this case (not shown) give 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 [75]

$$\tau = \frac{\pi}{4}, \qquad h_x = 1, \qquad h_z = 0.6. \qquad (81)$$

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. Empirically, we see only very small differences for the line tension function among different choices of $h_z$, including the special point $h_z = 1$ where the dynamics is Clifford as well as being dual unitary.

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.

Therefore, the domain wall structure appears to make sense even for dual-unitary gates, which represent a limiting case. However, in this case, negative signs in $W$ are important: Without these, it is impossible to have a flat $\mathcal{E}(v)$; see Sec. VI.

## C. Structure of W matrix

We now examine the temporal decay of the step weights $W$ defined in Sec. II D. In order to examine the importance
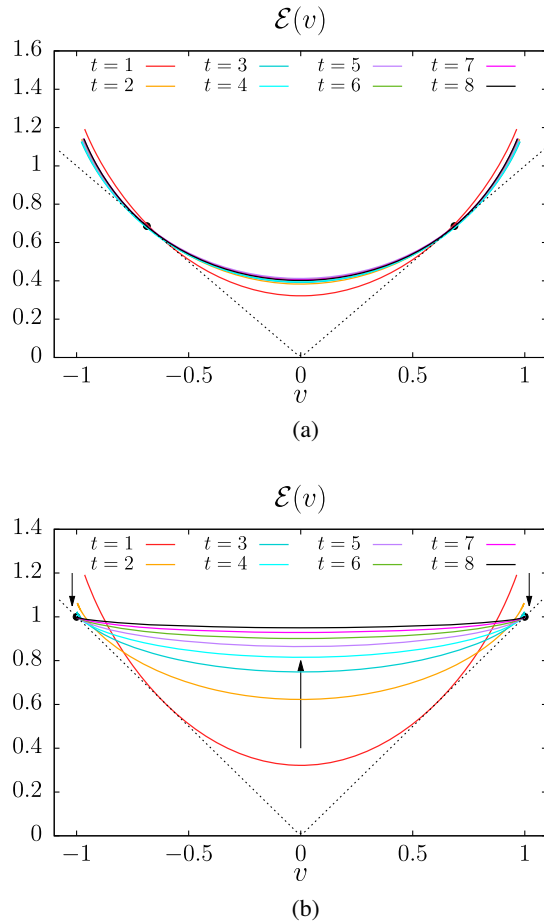
(a)



(a)



(b)



(b)

FIG. 8. Numerical calculations of $\mathcal{E}(v)$ for Floquet Ising models. (a) shows the parameters in Eq. (80). (b) shows the "dual-unitary" parameter values in Eq. (81). 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).
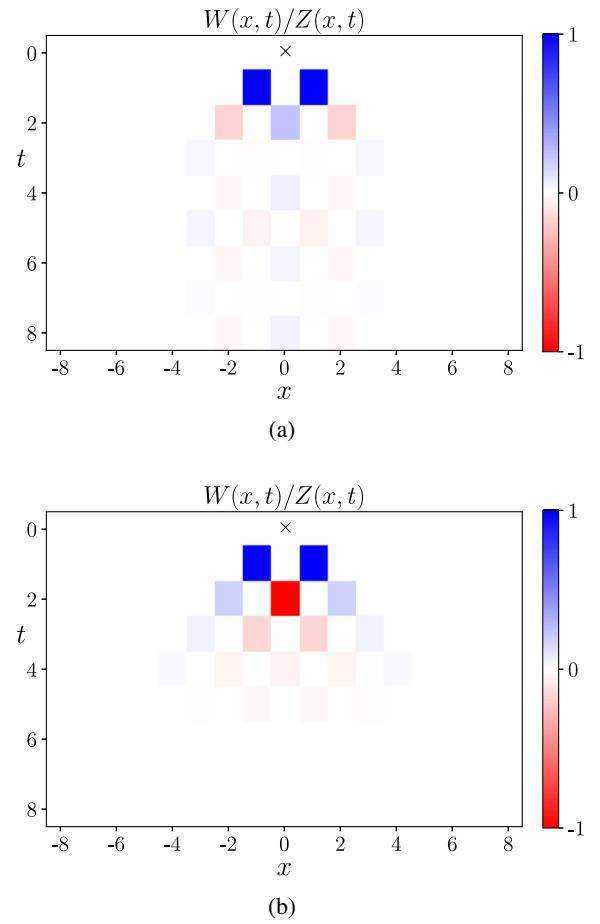
FIG. 9. The spacetime structure of $W(x,t)/Z(x,t)$ for (a) the generic reflection-symmetric gate in Eq. (71) [Fig. 7(a)] and (b) the maximally entangling dual-unitary gate specified in Eq. (81) [Fig. 8(b)]. The cross marks the coordinate (0,0). In our construction, only steps with even $x + t$ are possible.

of long steps, we consider the normalized quantity $W(x,t)/Z(x,t)$: This quantity 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 end points. 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. III A 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 $q^{-t}$.

Figure 9 shows heat maps of $W(x,t)/Z(x,t)$ for examples of translation-invariant Floquet models for spin-1/2. Figure 9(a) is for the generic reflection-symmetric gate in Eq. (71), whose line tension function is shown in Fig. 7(a), and Fig. 9(b) shows the dual-unitary gate of Eq. (81) 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. (80) (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. VI. Note that the weight at $x = 0$, $t = 2$ is positive for the gate in Fig. 9(a) and negative for the dual-unitary gate in Fig. 9(b). In the former case, this extra positive step weight increases $Z(0,t)$ compared to $Z_{0\perp}(0,t)$, leading to decreased $v_E^{(2)}$. In the latter case, the negative step weight decreases $Z(0,t)$ compared to $Z_{0\perp}(0,t)$, 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 the local gate is very weakly entangling, resulting in slower convergence of the algorithm. This example is the reflection-symmetric gate defined in Eq. (72) but with the smaller interaction constant $x = 0.5$ (in Sec. V A, we show results for $x = 0.8$).
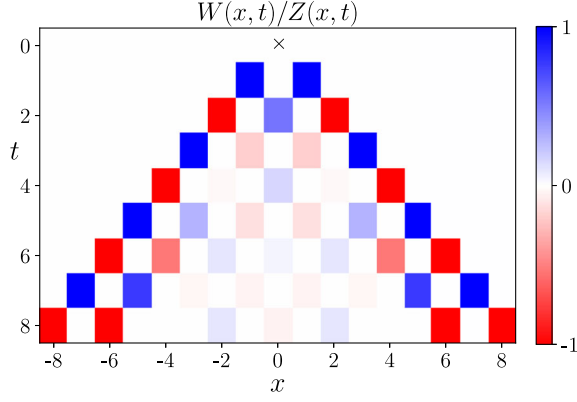
FIG. 10.  Weakly entangling gate with $x = 0.5$ [Eq. (72)]: Heat map showing the relative importance of long steps.

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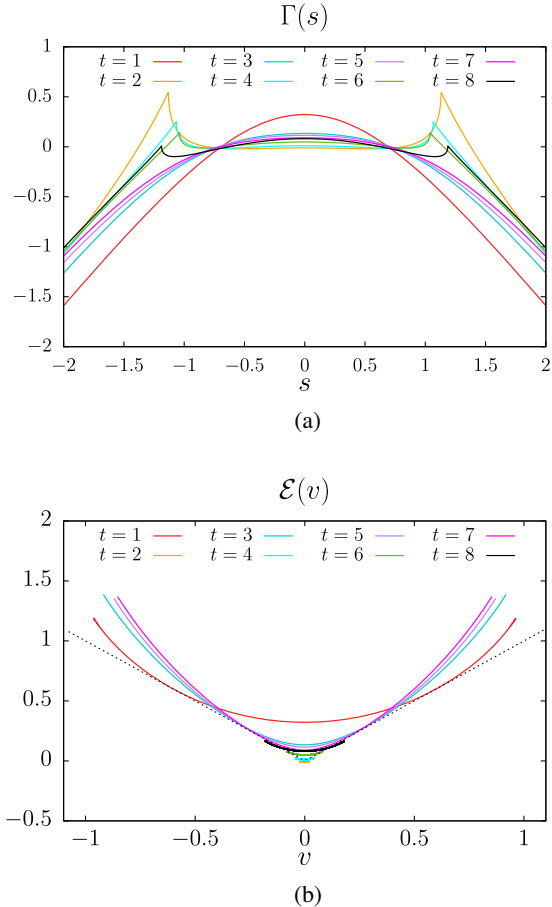


(a)



(b)

FIG. 11.  Weakly entangling gate with $x = 0.5$ [Eq. (72)]: (a) Successive approximations to $\Gamma(s)$ (see Sec. IV A). For some small $t$ values, $\Gamma(s)$ is negative around $s = 0$. (b) Approximations to $\mathcal{E}(v)$ obtained by Legendre transform of (a). Note the small $\mathcal{E}(0) = \Gamma(0)$.

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 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 result is consistent with slow convergence to a function that satisfies the general constraints, with a very small entanglement growth rate.

For a quantitative analysis of convergence rates in the various models, see Appendix B 5.

## VI. COARSE GRAINING: GENERIC MODELS VERSUS DUAL-UNITARY CIRCUITS

The thin domain wall conjecture implies that the entanglement membrane is well defined once we coarse grain sufficiently in space and time. The relevant length scale, the width of the domain wall, is microscopic, in the sense that it remains finite as the system size and the total time $t$ of the evolution go to infinity. In the models investigated here, which do not contain any small parameter, the width of the domain wall is an order of 1 multiple of the lattice spacing. In cases where the dynamics is tuned close to a special point, a large length scale may emerge.

In this section, we discuss this coarse graining in slightly more detail, in order to make a distinction between two universality classes.

Microscopically, the step weights $W(x, t)$ can be of either sign. But, we conjecture that, for generic models, the minus signs do not alter the universality class from that of a classical directed path. 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 $v < v_B$). We restate this conjecture in terms of the recursion relation for $Z(x, t)$ below.

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 $v$ that we condition on.) This fact implies, for example, that the scaling of $Z(x, t)$ is

$$Z(x, t) \sim t^{-1/2} e^{-s_{eq}\mathcal{E}(v)t}. \tag{82}$$

The power-law prefactor is universal and comes from the fact that both end points of the random path are fixed. For example, consider the case where $x \ll t$, so that the coarse-grained velocity is close to zero, and let the model be parity symmetric. Expanding in $v = x/t$, the above is then

$$\tilde{Z}(x,t) \sim t^{-1/2} \exp\left(-\frac{s_{\mathrm{eq}}\mathcal{E}''(0)x^2}{2t}\right) \quad (x \ll t), \quad (83)$$

where we factor out the extensive free energy of the path by defining [recall $v_E = \mathcal{E}(0)$]

$$Z(x,t) = e^{-s_{\mathrm{eq}}v_E t}\tilde{Z}(x,t). \quad (84)$$

Up to a normalization constant, Eq. (83) is the probability for a random walker with diffusion coefficient

$$D = \frac{1}{2s_{\mathrm{eq}}\mathcal{E}''(0)} \quad (85)$$

to be at position $x$ at time $t$. In Fig. 12(a), we show $\tilde{Z}(x,t)$ for a generic kicked-Ising gate: These numerical results are
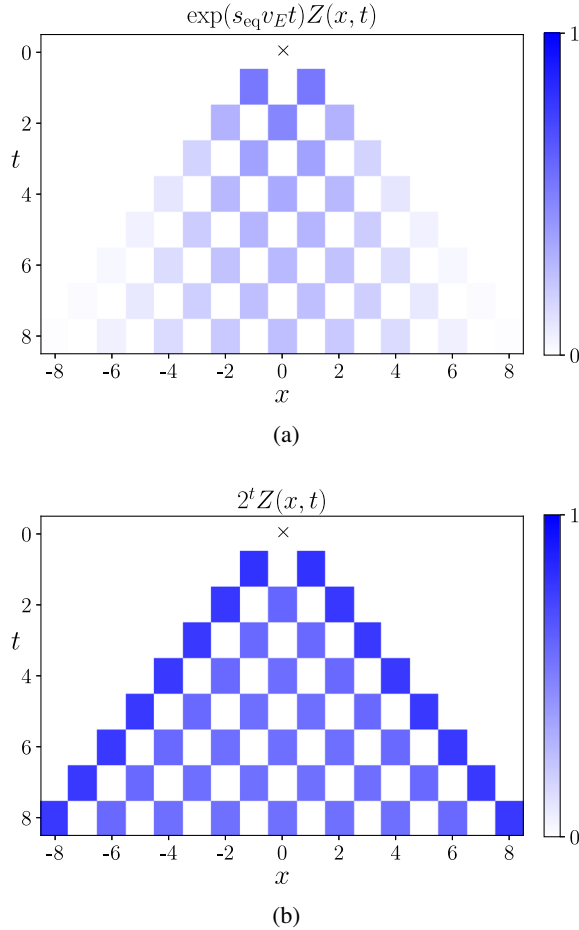
(a)

(b)

FIG. 12. Heat map of $\tilde{Z}(x,t) = \exp(s_{\mathrm{eq}}v_E t)Z(x,t)$. (a) Floquet Ising model as in Fig. 8(a). $\tilde{Z}(x,t)$ is consistent with the solution of a parabolic equation. For example, the decay at $x = 0$ is compatible with the $t^{-1/2}$ scaling in Eq. (83): This scaling is demonstrated in Appendix B 6. (b) Dual-unitary model as in Fig. 8(b). Unlike the previous model, $\tilde{Z}(x,t)$ is approximately constant in the interior of the light cone and zero outside.

in good agreement with the $t^{-1/2}$ scaling in Eq. (83) (see Appendix B 6).

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. The numerical results in Sec. VI A are consistent with the large-$t$ scaling

$$Z(x,t) \sim e^{-s_{\mathrm{eq}}t} = 2^{-t}, \quad (86)$$

with a trivial $x$ dependence inside the light cone (we neglect the even-odd effect at the lattice scale). This result is also consistent with the analytical calculation in Sec. VI A below.

Evidently, in the dual-unitary case, the membrane is not a classical random walk. Indeed, since for the dual-unitary model $\mathcal{E}(v)$ is constant (inside the light cone) and $\mathcal{E}''(v)$ vanishes, the diffusivity in Eq. (85) must diverge in the limit that a model becomes dual unitary.

The distinction between the two cases may be clearer if we think of the recursion relation as a discrete analog of a linear partial differential equation. In the limit of large $t$, we may write Eq. (43) as

$$Z(x,t) = \sum_{\Delta t = 1}^{\infty}\sum_{\Delta x} W(\Delta x, \Delta t)Z(x - \Delta x, t - \Delta t). \quad (87)$$

We would like to relate this to a continuum equation in the limit of large times. We cannot immediately perform a derivative expansion of $Z$: Since it is exponentially decaying, higher derivatives are not parametrically smaller than lower ones. For simplicity, consider first the regime $x \ll t$ discussed above (close to the $v = 0$ ray) and assume reflection symmetry [so $\mathcal{E}'(0) = 0$]. Then, $\tilde{Z}$ in Eq. (84) has a sensible derivative expansion. Let us define the following "average" for an arbitrary function $F$:

$$\langle F(y,\tau)\rangle_0 = \sum_{y,\tau} W(y,\tau)e^{s_{\mathrm{eq}}v_E \tau}F(y,\tau). \quad (88)$$

(We replace $\Delta x$, $\Delta t$ with $y$, $\tau$ to avoid clutter. The subscript on the average specifies the spacetime ray whose vicinity we are considering.) This equation satisfies $\langle 1\rangle_0 = 1$, by the relations in Sec. IV A. The derivative expansion of Eq. (87) yields

$$\langle\tau\rangle_0\partial_t\tilde{Z} = \frac{1}{2}\langle y^2\rangle_0\partial_x^2\tilde{Z}, \quad (89)$$

plus less-relevant terms. The thin domain wall assumption implies that the two coefficients shown are finite. While the weights $W$ defining the averages are not necessarily positive, the averages shown are nonzero in the absence of fine-tuning. The higher terms can then be dropped,

leading to the diffusive (random walk) scaling discussed above. We can also check using the relations between generating functions that

$$\langle \tau \rangle_0 = s_{\text{eq}} \mathcal{E}''(0) \langle y^2 \rangle_0, \qquad (90)$$

so that the diffusion constant in Eq. (89) agrees with that in Eq. (82). We emphasize that the above equation is valid only close to the $v = 0$ ray; see below for a more general equation. The information contained in Eq. (89) is already contained in $\mathcal{E}(v)$, and the spacetime picture is usually more convenient.

The universality class changes if $\langle \tau \rangle_0$ vanishes due to fine-tuning. It is clear from the definition in Eq. (88) that the vanishing of $\langle \tau \rangle_0$ requires negative step weights (which are absent in the large-$q$ limit described in Sec. III). When this vanishing occurs, we must take into account the next time derivative term, which was previously irrelevant:

$$\langle \tau^2 \rangle_0 \partial_t^2 \tilde{Z} = \langle y^2 \rangle_0 \partial_x^2 \tilde{Z}. \qquad (91)$$

This vanishing is what occurs in the dual-unitary model above. We see that the equation becomes a wave equation, consistent with the constancy of $\tilde{Z}$ inside the light cone (see Sec. VI A). We have confirmed that $\langle \tau \rangle_0$ is consistent with zero in the dual-unitary circuit: The finite $t$ approximation to $\langle \tau \rangle_0$ from the truncated $W(x, t)$ matrix decays exponentially in $t$, with a characteristic time of approximately 1.8 steps.

To simplify further, let us consider a toy model, which takes only a single additional nonzero $W$ element (beyond the ones present in $Z_{0\perp}$). We parameterize its elements in terms of a constant $\mu$:

$$W(\pm 1, 1) = K, \qquad W(0, 2) = -\mu K^2. \qquad (92)$$

When $\mu < 1$, we are in the diffusive class. However, when $\mu = 1$ (in fact, the dual-unitary model is numerically quite close to this simplified model), removing the factors of $K$ by defining $Z(x, t) = K^t \tilde{Z}(x, t)$ gives

$$\tilde{Z}(x, t) = \tilde{Z}(x + 1, t - 1) + \tilde{Z}(x - 1, t - 1) - \tilde{Z}(x, t - 2), \qquad (93)$$

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

Numerically, we find that the structure in the specific dual-unitary kicked-Ising model that we study above is similar: $2^t Z(x, t)$ is approximately constant within the light cone and zero outside. Figure 12 contrasts the spacetime pattern of $e^{s_{\text{eq}} v_E t} Z(x, t)$ for the generic kicked-Ising model and the dual-unitary kicked-Ising model.

In the derivations of the continuum equations [Eqs. (89) and (91)], we restrict to the vicinity of a particular ray.

For a more general equation, we may write $Z(x, t) = \exp[-S(x, t)]$ and expand in derivatives of $S(x, t)$. The resulting dynamical equation [18] also applies to the second Rényi entropy $S(x, t)$ of a time-evolving quantum state in a 1D system, for an entanglement cut at $x$. [76] As described in Appendix D, this expansion gives the equation

$$\frac{dS}{dt} = s_{\text{eq}} \left[ \Gamma \left( \frac{dS}{dx} \right) - \frac{1}{2} \Gamma'' \left( \frac{dS}{dx} \right) \frac{d^2 S}{dx^2} + \cdots \right], \qquad (94)$$

where $\Gamma$ is the Legendre transform of $\mathcal{E}$ (see Sec. IV A). The first term is the leading-order dynamics that captures the $\mathcal{O}(t)$ term in the entanglement. Here, $dS/dx$ and $dS/dt$ are generally of order 1 at late times, so we keep all powers of $dS/dx$, but higher derivatives are small and we can expand in them. The second term gives subleading corrections [18]. The random walk picture implies that the two coefficients are related. The above equation can also be written in terms of $\mathcal{E}(v)$, giving a form consistent with Eqs. (89) and (90) (see Appendix D).

## A. Partition function for dual-unitary models

Let us confirm by an explicit calculation that dual unitarity implies the above-mentioned approximate constancy of $Z(x, t)$ within the light cone [which is related to the wavelike structure of the recursion equation for $\tilde{Z}$ and which implies constancy of $\mathcal{E}(v)$ as a function of $v$].

This fact can be seen by considering a slightly different partition function, in which we place states $|\pm\rangle$ at the lower boundary, rather than the dual states $|(\pm\pm)^*\rangle$ used in Eq. (35):

$$Z_{\text{op}}(x, y, t) = \langle \ldots + - \ldots | U^{(2)}(t) | \ldots + - \ldots \rangle / q^{2L}. \qquad (95)$$

Here, $x$ and $y$ label the positions of the domain walls between $+$ and $-$ on the top and bottom boundaries, respectively. Figure 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. (35), $Z_{\text{op}}(x, y, t)$ does not force the end point of the domain wall to lie precisely at $y$ but instead penalizes configurations whose end point deviates from $y$ (via a negative power of $q$ in the weight). However, in analogy to the random circuit case, we expect that this difference leads only to an $\mathcal{O}(1)$ boundary term in the free energy (so long as $|x - y| \leq v_B t$, which in the dual-unitary circuit is a trivial restriction, since $v_B = 1$), so that

$$\ln Z_{\text{op}}(x, y, t) \sim \ln Z(x, y, t) \sim -\mathcal{E}(v) t \ln q. \qquad (96)$$

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 [36,46,49,72].
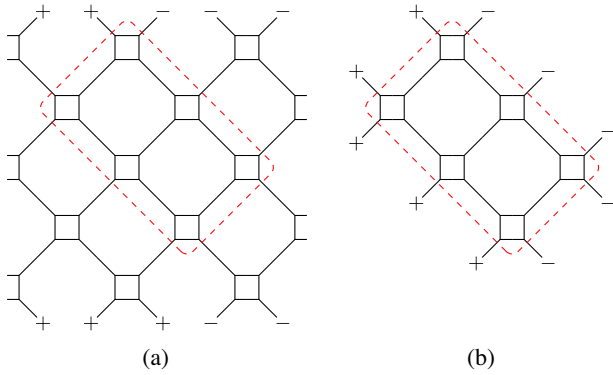
FIG. 13.   The partition function $Z_{op}(x, t)$ for a dual-unitary circuit. (a) $q^{2L} Z_{op}(x = 1, t = 5)$. By unitarity [the first identity in Eq. (98)], we can remove the unitaries outside the light cone, giving $q^{2t} Z_{op}(x = 1, t = 5)$ as expression (b). Then, Eq. (98) allows the remaining gates to be removed.

This result is easily seen by manipulations similar to those in Refs. [36,46,47]. Specifically (setting $y = 0$), we show that $Z_{op}(x, t) = q^{-t}$ for a dual-unitary circuit, so long as $x$ is within the light cone of the origin.

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

$$u^{(N)} = \begin{array}{c} A \quad B \\ \square \\ C \quad D \end{array}, \qquad u^{(N)}_{dual} = \begin{array}{c} B \quad C \\ \square \\ A \quad D \end{array}. \qquad (97)$$

We label the external legs of this tensor by $A$, $B$, $C$, and $D$. As a unitary, the legs $A$ and $B$ correspond to the row index and $C$ and $D$ to the column index. We rotate the tensor by 90° to obtain the dual tensor. Dual unitarity requires the dual tensor to be unitary (with $B$ and $C$ as row legs and $A$ and $D$ as column legs) as well. To contrast with the generic gate in Eq. (47), we draw the dual-unitary tensor as a (fourfold symmetric) square in Eq. (97). For $N = 2$, unitarity and dual unitarity can be used to propagate $++$ (or $--$) states from the top to bottom, right to left, and vice versa. Graphically, we have (the same holds for arbitrary $N$ with $+$ replaced by any $\sigma$)

$$\begin{array}{c} + \quad + \\ \square \\ {} \end{array} = \begin{array}{c} + \\ | \end{array} \begin{array}{c} + \\ | \end{array}, \qquad \begin{array}{c} + \\ \square \\ + \end{array} = \begin{array}{c} - \quad + \\ {} \\ - \quad + \end{array}. \qquad (98)$$

The first of these relations (unitarity) allows us to remove the unitary gates outside the light cones in Fig. 13(a). Keeping track of the $q$ factor, $q^{2t} Z_{op}(x, t)$ 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 $\langle + | - \rangle^t = q^t$. Hence, $q^{2t} Z_{op}(x, t) = q^t$, so that $Z_{op}(x, t) = e^{-s_{eq}t}$ within the light cone.

## VII. OPERATOR SPREADING

We focus 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 $e^{-S_2}$ but with different boundary conditions. In the Haar-averaged circuit, this sum 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:

$$\tilde{C}(x, t) = -\frac{1}{2} \langle [\mathcal{O}(x, t), \mathcal{O}'(0, 0)]^2 \rangle. \qquad (99)$$

We take the expectation value at infinite temperature, and, to simplify the boundary conditions, we average over local unitary rotations of the local operators $\mathcal{O} \to V \mathcal{O} V^{\dagger}$ and
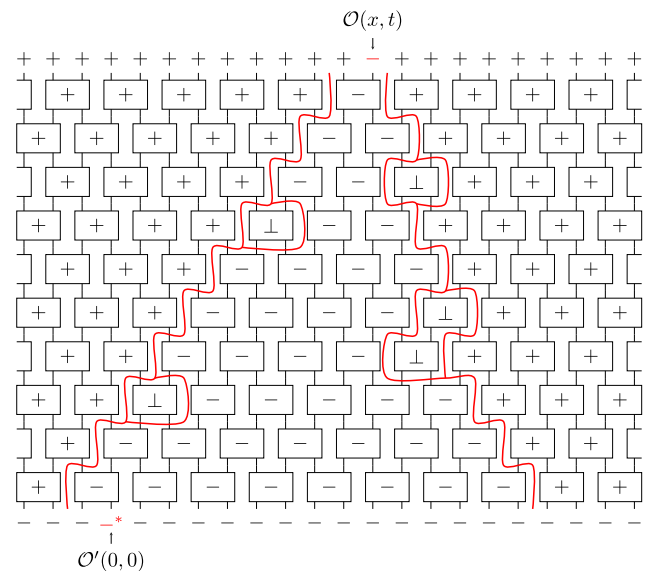


FIG. 14.   The membrane picture for the out-of-time-ordered commutator $\tilde{C}(x, t)$. The top boundary condition creates two domain walls. The bottom boundary condition favors $-$, providing an outward force on the lower end points of the two domain walls. The right domain wall thus moves with the optimal speed $v_B$. Here, the left wall is forced by the dual state $|-^*\rangle$ at the operator insertion point to travel with $v > v_B$.

$\mathcal{O}' \to V'\mathcal{O}'V'^{\dagger}$. Carrying out manipulations similar to Sec. VI B in Ref. [7], we rewrite $\tilde{C}(x,t)$ using the present notation (here, tr is a single-site trace):

$$\tilde{C}(x,t)$$
$$= \frac{\text{tr}(\mathcal{O}^2)\text{tr}(\mathcal{O}'^2)}{q^2-1}$$
$$\times \langle \ldots + + - + + \ldots | U^{(2)} | \ldots - - - ^* - - \ldots \rangle / q^{L-1}. \tag{100}$$

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 rate $\pm v_B$ (assuming reflection symmetry) [77].

However, the operator insertion at the initial time cannot lie in the $+$ region exterior to the $-$ domain, because this leads to a factor of $\langle + | -^* \rangle = 0$. If $|x| < v_B t$, then $\tilde{C}(x,t)$ is constant at leading order: $\tilde{C}(x,t) = q^{-2}\text{tr}\mathcal{O}^2\text{tr}\mathcal{O}'^2$. But, if $|x| > v_B$, the lower operator insertion "stretches" one of the domain walls, forcing it to move at a speed greater than $v_B$. This mechanism is illustrated for the left domain wall in Fig. 14. The additional free energy cost means that, outside the light cone, for $x/t > v_B$, the OTOC scales as [78]

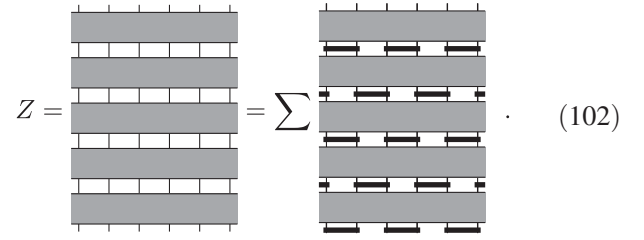$$\tilde{C}(x,t) \sim e^{-\mu(v)t}, \qquad \mu(v) \equiv s_{\text{eq}}[\mathcal{E}(v) - v]. \tag{101}$$

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 $v$ [79]. 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 [80].
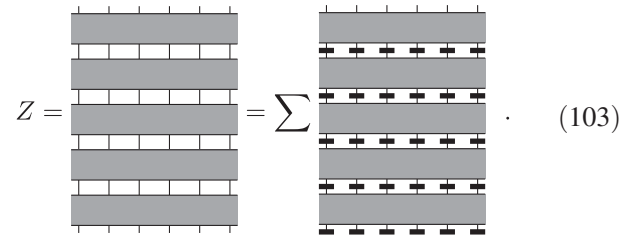
## VIII. BEYOND CIRCUIT MODELS

In Sec. II, we describe how to introduce the pairing field (the spin $s \in S_N \cup \{\perp\}$ in the effective partition function) in the context of models with a circuit structure. These models were a natural starting point, since the idea of inserting permutation states is inspired by the random unitary circuit. However, there is no need to restrict to circuits.

In this section, we study generic Floquet systems, i.e., models where the Hamiltonian is periodic in time but which do not have the form of a circuit. We are still free to insert resolutions of identity as in Eq. (11):



$$Z = \quad = \sum \quad . \tag{102}$$

In this graphical equation, the shaded region denotes the Floquet operator (evolution operator for one time period), which, in general, does not have an exact circuit decomposition. As in Eq. (11), the horizontal bars represent projectors onto either a permutation state or the $\perp$ subspace, these possibilities being summed over.

For models that are translationally invariant by one lattice spacing, an alternative that may appear at first sight more natural is to insert single-site, rather than two-site, projectors:



$$Z = \quad = \sum \quad . \tag{103}$$

Both rewritings of $Z$ are exact. In the following, we consider both schemes.

The partition functions are sums over spins $s_{x,t}$, taking $N! + 1$ values, associated with each horizontal bar. We focus on the case of the purity ($N = 2$), where $s_{x,t}$ can be $\pm$ or $\perp$. We choose a graphical notation that resembles that for the circuit, labeling the shaded region immediately above the horizontal bar with the bar's spin value; see Fig. 15. This figure shows possible spin configurations for a section of domain wall of finite width in the two schemes.

The numerical algorithm in Sec. IV A can be carried over with little modification. We first define the partition function $Z(x,t)$ the same way as in Eq. (35)—the difference here is that $Z(x,t)$ may be nonzero even when $x$ is arbitrarily large. We then solve for the irreducible step weights $W(x,t)$ recursively, using Eq. (43).

The approach is feasible, because the Floquet evolution has an emergent Lieb-Robinson light cone [81] which suppresses $Z(x,t)$ for large $x$. For practical computation of $Z(x,t)$, we must truncate the Floquet operator to a chain of finite length $L$, so at the end we must check that our results are converged in $L$. For a fixed $t$, the error introduced due to the finiteness of $L$ should be exponentially small at large $L$, but, of course, the required $L$ grows linearly with $t$.

We consider two examples of an Ising model with an oscillating transverse field $h_x(t)$:

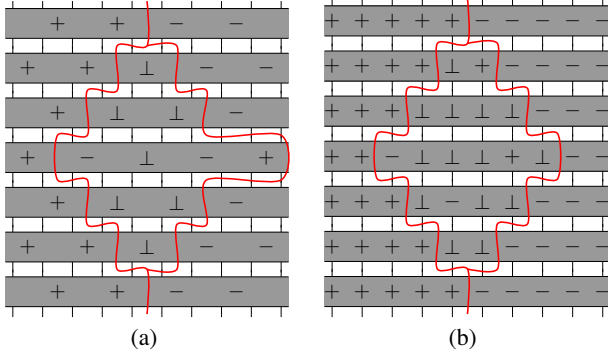(a)                                    (b)

FIG. 15. Domain wall (membrane) in Floquet evolution without a circuit structure. The spin variables $+$, $-$, and $\perp$ denote the three choices of projection operators inserted on the horizontal bars beneath them; see Eqs. (102) and (103). Figures show samples of thick domain walls for (a) two-site and (b) one-site insertion.

$$H(t) = \sum_{i=1}^{L-1} J\sigma_i^z\sigma_i^{z+1} + \sum_i h_x(t)\sigma_i^x + \sum_i h_z\sigma_i^z. \quad (104)$$

In the first example, $h_x(t)$ is a square wave with period $\tau = 1$, taking the values $h^{(1)}$ and $h^{(2)}$ for half the period each:

$$h_x^{(1,2)} = 0.9045 \pm w, \qquad w = 0.3,$$
$$J = 0.5, \qquad h_z = 0.8090. \quad (105)$$

The Floquet operator is then

$$U_{\text{Floq}} = \exp(-iH_1\tau/2)\exp(-iH_2\tau/2) \quad (106)$$

with $H_{1,2}$ equal to Eq. (104) with the appropriate value of $h_x$.

In the second example, $h_x(t)$ oscillates sinusoidally, with period $\tau = 1$:

$$h_x = 0.9045[1 + w\cos(2\pi t/\tau)], \quad (107)$$

the parameters $J$, $h_z$, and $w$ being the same as in Eq. (105). Neither of these two examples allows a circuit representation.

For the first system (square wave), we compare the one- and two-site insertion schemes. Figure 16 shows approximations to the line tension function, for increasing values of $t_{\max}$, in a system of $L = 12$ spins, for the two-site and one-site schemes, respectively. We position the domain wall at the top boundary so as to minimize finite-$L$ effects; see the numerical details in Appendix B 1. (Performing calculations for a given $L$ with multiple choices of the top boundary condition incurs an additional cost, but this cost is a subleading factor.) We compute up to $t_{\max} = 6$. We estimate the error due to the finiteness of $L$ by computing
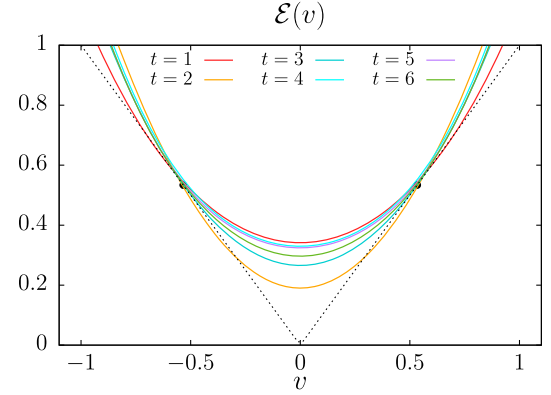


(a)
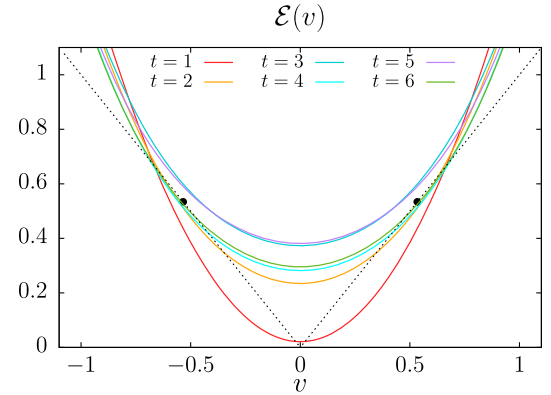


(b)

FIG. 16. Numerical calculations of $\mathcal{E}(v)$ for Floquet evolution with the square wave $h_x(t)$ [Eqs. (104) and (105)] for (a) two-site and (b) one-site schemes.

the difference in the $t = 5$ result for $\mathcal{E}(0)$ for $L = 10$ and for $L = 12$. We find a relative error of 0.2% for the two-site scheme and 1% for the one-site scheme.

For both schemes, the $t = 6$ curves in Fig. 16 are consistent with reasonable convergence to a legitimate line tension function, with a value of $v_B$ consistent with that computed directly from the OTOC.
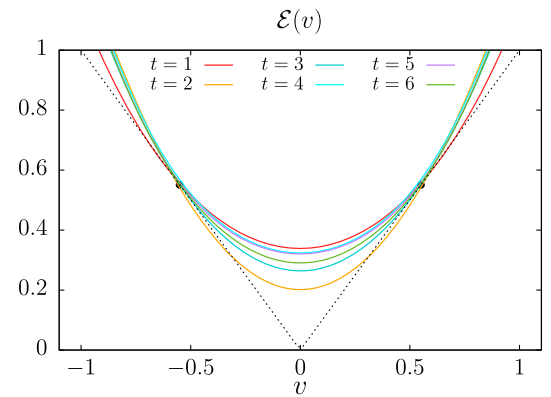


FIG. 17. Numerical calculations of $\mathcal{E}(v)$ for Floquet evolution with the sinusoidally varying $h_x(t)$ in Eq. (107) for the two-site scheme.

One counterintuitive feature of Fig. 16(a) is that the two-site scheme performs better than the one-site scheme for small $t$ (the one-site scheme is also more expensive computationally). The contrast is most obvious at $t = 1$, where the curve for the one-site scheme [Fig. 16(a)] strongly violates the constraints of the line tension. Despite this violation, the two curves for $t = 6$ are in remarkably good agreement, with 0.5% relative difference in the entanglement velocity estimates.

Finally, in Fig. 17, we show results for the sinusoidally oscillating field [Eq. (107)]. These are qualitatively similar to Fig. 16(a).

## IX. OUTLOOK

In random circuits, the pairing field and the membrane picture emerge in a simple way from the average over unitaries whenever we deal with dynamical quantities that require multiple forward and backward evolution operators (such as Rényi entropies or OTOCs). Here, we have shown how to make sense of the pairing field, which describes a pairing of trajectories in the multisheeted path integral, in realistic systems that do not involve randomness. This description opens the way to extending the various results that have been obtained in random circuits to a much broader class of more realistic models. More generally, by identifying appropriate effective fields, it provides a framework for applying the renormalization group to chaotic real-time dynamics.

We separated the local multicopy Hilbert space into paired states and states in their orthogonal complement, $\perp$. Dynamical quantities then became partition functions for an effective field $s(x, t) \in \{\perp\} \cup S_N$. 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 $1 + 1$D 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

Floquet models (whether circuits or not) 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]. Our approach can be extended to models with conservation laws by refining the resolution of the identity to specify information about the charge as well as the permutation: We will discuss this extension elsewhere. 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, can be qualitatively different from that of the von Neumann entropy in models with conservation laws [15,82–84].

In the present approach, the von Neumann entropy corresponds to a nontrivial replicalike 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, in addition to giving new hydrodynamic modes [85,86].

There is considerable freedom in the precise way in which the effective spins are introduced. We focused mainly on inserting a resolution of the identity for pairs of sites, since this convention was convenient for making contact with the Haar-random circuits studied previously. We also explored the use of single-site resolutions of the identity, which appeared to be less efficient for the model we studied. It would be useful to understand this better in order to optimize the efficiency of the scheme.

The basic idea of this paper is not restricted to $1 + 1$D: 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. The method may also be applied to dynamics without spatial structure, in which any qubit can interact with any other.

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 nonlocal in time play a role [11,33,87]. 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 consideration 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 [6,88,89].

The approach could also be used to study questions motivated by quantum information such as the emergence of the "design" property in random circuits with non-Haar-distributed gates [5,17,90–92].

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 [88,89,93–103].

The approach could also be adapted to study tensor networks that do not involve a time direction, for example, in the context of the AdS-CFT correspondence where random tensor networks have been used [42,43].

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

## APPENDIX A: PROPERTIES OF THE SPIN MODEL

### 1. Projectors onto permutation states

The operator $P_\sigma$ defined in Eq. (14) is a projection operator onto the $\sigma$ permutation state (it is a nonorthogonal projector, i.e., $P_\sigma^\dagger \neq P_\sigma$). Here, we prove the projector property in Eq. (16). According to the definition of $P_\sigma$ in terms of the Weingarten function and permutation states in Eq. (14), we have

$$P_{\sigma_1} P_{\sigma_2} = \sum_{\tau_1,\tau_2} \mathrm{Wg}(\tau_1^{-1}\sigma_1)|\tau_1\rangle\langle\sigma_1|\tau_2\rangle \mathrm{Wg}(\tau_2^{-1}\sigma_2)\langle\sigma_2|. \quad \text{(A1)}$$

We simplify the notation by writing $|\tau\rangle$ as a shorthand for the two-site state $|\tau\tau\rangle$, which emphasizes that the formulas here apply to an arbitrary Hilbert space, using the Weingarten functions for the appropriate value of the Hilbert space dimension (equal to $q^2$ here). We return to the previous notation after Eq. (A4). Note that states associated with distinct permutations are not orthogonal.

Unless $q^2 = 0, \pm 1, \pm 2, \ldots, \pm(N-1)$, the function $\mathrm{Wg}(\tau^{-1}\sigma)$, viewed as a matrix with $\tau$ and $\sigma$ as its indices, is the inverse of the Gramian matrix $\langle\sigma|\tau\rangle$ [50,51]:

$$\sum_\tau \langle\sigma_1|\tau\rangle \mathrm{Wg}(\tau^{-1}\sigma_2) = \delta_{\sigma_1,\sigma_2}. \quad \text{(A2)}$$

This equation means the states $|\sigma^*\rangle$ defined in Eq. (15) indeed form the dual basis:

$$\langle\sigma_1^*|\sigma_2\rangle = \delta_{\sigma_1,\sigma_2}. \quad \text{(A3)}$$

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

$$P_{\sigma_1} P_{\sigma_2} = \sum_{\tau_1,\tau_2} \mathrm{Wg}(\tau_1^{-1}\sigma_1)|\tau_1\rangle\delta_{\sigma_1,\sigma_2}\langle\sigma_2| = P_{\sigma_1}\delta_{\sigma_1\sigma_2}. \quad \text{(A4)}$$

Any four-layer unitary gate with the projector $P_\sigma$ attached at the bottom is equal to $P_\sigma$ [Eq. (19)]. In Eq. (18), we represent the projector as a tensor (box) labeled with a $\sigma$ inside. Given the expression in Eq. (14), it can be decomposed as

$$\boxed{\phantom{xx}\sigma\phantom{xx}} = \frac{\boxed{\sigma^*}}{\boxed{\sigma}\,\boxed{\sigma}}. \quad \text{(A5)}$$

where the ket $|(\sigma\sigma)^*\rangle = \sum_\tau \mathrm{Wg}(\tau^{-1}\sigma)|\tau\tau\rangle$ is the two-site dual state of $|\sigma\sigma\rangle$ and the bras at the bottom just give $\langle\sigma\sigma| = \langle\sigma| \otimes \langle\sigma|$ 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:

$$\frac{\boxed{\sigma_b}\,\boxed{\sigma_c}}{\boxed{\sigma_a}} \equiv \text{shaded part of} \left( \begin{array}{c} \boxed{\sigma_b}\ \boxed{\sigma_c} \\ \boxed{\sigma_a^*} \\ \end{array} \right) \quad \text{(A6)}$$

$$= \langle\sigma_b\sigma_c|(\sigma_a\sigma_a)^*\rangle, \quad \text{(A7)}$$

which yields the expression of $J(\sigma_b, \sigma_c; \sigma_a)$ in Eq. (23).

### 2. $N$ independence of the weights

The triangle weights

$$J(\sigma_b, \sigma_c; \sigma_a) = \frac{\boxed{\sigma_b}\,\boxed{\sigma_c}}{\boxed{\sigma_a}} = \overset{\sigma_b}{\underset{\sigma_a}{\bigtriangledown}}\overset{\sigma_c}{} \quad \text{(A8)}$$

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 multilayer circuit.

In some cases of interest to us, the spins $\sigma_a$, $\sigma_b$, and $\sigma_c$ 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 ($M$ "$U$" layers and $M$ "$U^*$" layers) can be paired arbitrarily, with the pairing labeled 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 $\sigma_c = (12)$:

$$\begin{array}{c} \mathbb{I} \qquad (12) \\ \bigtriangledown \\ \mathbb{I} \end{array} = K. \qquad (A9)$$

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

Algebraically, the triangle is the coefficient in the expansion (24):

$$\langle \sigma_b\, \sigma_c | \, P_{\parallel}^{(N)} = \sum_{\sigma_a \in S_N} \begin{array}{c} \sigma_b \qquad \sigma_c \\ \bigtriangledown \\ \sigma_a \end{array} \langle \sigma_a\, \sigma_a |, \qquad (A10)$$

(the two spins refer to the two spatial states), which is equivalent to the definition in Eq. (23):

$$\begin{array}{c} \sigma_b \qquad \sigma_c \\ \bigtriangledown \\ \sigma_a \end{array} = \sum_{\tau \in S_N} \langle \sigma_b\, \sigma_c | \tau\, \tau \rangle \mathrm{Wg}(\tau^{-1}\sigma_a; q^2), \qquad (A11)$$

as we see by using Eq. (A2) to extract a particular coefficient on the right-hand side of Eq. (A10).

Now consider the case where $\sigma_b, \sigma_c \in S_M$, where $S_M$ is the subgroup mentioned above. [104] The fact that the resulting weights are independent of $N$ is not immediately apparent from Eq. (A11) but can be seen from Eq. (24). When $\sigma_b, \sigma_c \in S_M$, we have

$$\langle \sigma_b \sigma_c | P_{\parallel}^{(N)} = \langle \sigma_b \sigma_c | (P_{\parallel}^{(M)} \otimes \mathbb{1}^{(N-M)}). \qquad (A12)$$

For example, we can see this independence 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 contraction 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 result shows that in Eq. (A12) the only nonvanishing terms on the right-hand side have $\sigma_a \in S_M$ as well. As a result, Eq. (A12) 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 result 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 *not* factorize, in general. That is, if we have

$$\sigma_a = \sigma_{a1}\sigma_{a2}, \qquad \sigma_b = \sigma_{b1}\sigma_{b2}, \qquad \sigma_c = \sigma_{c1}\sigma_{c2}, \qquad (A13)$$

where $\sigma_{b1} \in S_M$ and $\sigma_{b2} \in S_{N-M}$, etc., the triangle

$$\begin{array}{c} \sigma_b \qquad \sigma_c \\ \bigtriangledown \\ \sigma_a \end{array} \qquad (A14)$$

is, in general, not equal to

$$\begin{array}{c} \sigma_{b1} \qquad \sigma_{c1} \\ \bigtriangledown \\ \sigma_{a1} \end{array} \times \begin{array}{c} \sigma_{b2} \qquad \sigma_{c2} \\ \bigtriangledown \\ \sigma_{a2} \end{array} . \qquad (A15)$$

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

### 3. Suppression of large ⊥ clusters

We analyze the spin model that arises in computing the mean square weight $\overline{\Omega_C^2}$ of a ⊥ 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 $\Omega_C^2$. We 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 ⊥ region. Each block there is labeled as either $+$ or $-$, indicating a local insertion of $P_{\pm}$ before squaring. Recall that this projection operator "absorbs" the local unitary gate: Equivalently, in the squared system, a given $+$ block is (recall that each $P_+^{(2)}$ acts on two physical sites)

$$\overline{u^{(4)}}(P_+^{(2)} \otimes P_+^{(2)}) = P_+^{(2)} \otimes P_+^{(2)} \qquad (A16)$$

and similarly for $-$. This result defines an interaction triangle for the case where the lower spin in the triangle is from a plus block, in analogy to the triangle interaction defined in Sec. II B for the case without the $P_+^{(2)} \otimes P_+^{(2)}$ factor. However, because of the projection operator, the spin associated with the block (the lower spin of the triangle) is *fixed* to be $\mathbb{I}$. Similarly, the spin associated with a $-$ block is fixed to be $\sigma = (12)(34)$. Therefore, in computing $\Omega_C$, 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. (24) with $P_\perp^{(4)}$ replaced by $P_+^{(2)} \otimes P_+^{(2)}$; the $-$ weights are related to the following by symmetry]

$$
\overset{\sigma_b \overset{\displaystyle +}{\bigtriangledown} \sigma_c}{\underset{\mathbb{1}}{}} = q^{-8} \langle \sigma_b \, \sigma_c | \left( P_+^{(2)} \otimes P_+^{(2)} \right) | \mathbb{1} \, \mathbb{1} \rangle . \quad (A17)
$$

Using the expression for the projector in Eq. (27), this $+$ triangle is equal to

$$
\frac{q^8}{(q^4 - 1)^2} \sum_{\tau \in S_2} q^{-|\sigma_b \tau| - |\sigma_c \tau| - 2|\tau|} (-1)^{|\tau|} . \quad (A18)
$$

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

$$
\overset{\sigma_b \overset{\displaystyle +}{\bigtriangledown} \sigma_c}{\underset{\mathbb{1}}{}} \lesssim q^{-|\sigma_b| - |\sigma_c|} \quad (A19)
$$

and, similarly, the $-$ triangle can be bounded by

$$
\overset{\sigma_b \overset{\displaystyle -}{\bigtriangledown} \sigma_c}{\underset{(12)(34)}{}} \lesssim q^{-|\sigma_b (12)(34)| - |\sigma_c (12)(34)|} . \quad (A20)
$$

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 $\sigma_c$ terms are for the right edge. We thus interpret Eqs. (A19) and (A20) as costing $q^{-1}$ per outgoing domain wall on the two edges.

Inside the cluster, we have $\perp$ blocks. These lead to a $\perp$ interaction triangle defined by

$$
\sum_{\sigma_a \in S_4} \overset{\sigma_b \overset{\displaystyle \perp}{\bigtriangledown} \sigma_c}{\underset{\sigma_a}{}} \langle \sigma_a \, \sigma_a | \equiv \langle \sigma_b \, \sigma_c | \left( \overline{u^{(4)}} - \overline{u^{(2)}} \otimes \overline{u^{(2)}} \right)
$$

$$
\equiv \langle \sigma_b \, \sigma_c | \left( P_\parallel^{(4)} - P_\parallel^{(2)} \otimes P_\parallel^{(2)} \right) , \quad (A21)
$$

where we use the equivalence between the parallel projection and the Haar-averaged stack of unitaries.

The right-hand side of Eq. (A21) 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

$$
\overset{\sigma_b \overset{\displaystyle \perp}{\bigtriangledown} \sigma_c}{\underset{\sigma_a}{}} \begin{cases} \sim q^{-|\sigma_a^{-1} \sigma_b| - |\sigma_a^{-1} \sigma_c|} & \sigma_a \notin S_2 \otimes S_2 , \\ \lesssim q^{-|\sigma_a^{-1} \sigma_b| - |\sigma_a^{-1} \sigma_c| - 2} & \sigma_a \in S_2 \otimes S_2 . \end{cases}
$$

$$(A22)$$

For the first case, the $\perp$ triangle weight is exactly equal to an ordinary triangle

$$
\overset{\sigma_b \overset{\displaystyle \perp}{\bigtriangledown} \sigma_c}{\underset{\sigma_a}{}} = \overset{\sigma_b \bigtriangledown \sigma_c}{\underset{\sigma_a}{}} , \quad \sigma_a \notin S_2 \times S_2 . \quad (A23)
$$

We again interpret the corresponding expansion (the first line) in Eq. (A22) 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 the order of $q^{-|\sigma_a^{-1} \sigma_b| - |\sigma_a^{-1} \sigma_c|}$ cancels, which therefore incurs an additional cost of $q^{-2}$. In particular, if $\sigma_b$ and $\sigma_c$ are also in $S_2 \times S_2$, then the $\perp$ triangle vanishes exactly:

$$
\overset{\sigma_b \overset{\displaystyle \perp}{\bigtriangledown} \sigma_c}{\underset{\sigma_a}{}} = 0, \quad \sigma_a, \sigma_b, \sigma_c \in S_2 \times S_2 . \quad (A24)
$$

With this result in hand, we can estimate the cost of a $\perp$ cluster in Fig. 5. Because of the vanishing weight in Eq. (A24), each triangle inside the $\perp$ 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 $1/q$ 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 result 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. (A23). 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

$$
\overset{\mathbb{1} \overset{\displaystyle \perp}{\bigtriangledown} (12)(34)}{\underset{(13)(24)}{}} \quad (A25)
$$

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 Eqs. (A19) and (A20). In total, the weight is $q^{-4t}$ for the squared average of the $\perp$ cluster.

#### 4. Minimal ⊥ cluster and operator purity

In Eq. (45), we list four nontrivial configurations for an isolated ⊥ surrounded by + and −. Here, we compute their weights for a fixed unitary $u$ at the location of the ⊥. We see that they are related to the operator purity of $u$.

It is more transparent if we represent each + and − block graphically as in Eq. (A5):

$$
\begin{array}{c} \boxed{+}\ \boxed{-} \\ \boxed{+}\boxed{\perp}\boxed{-} \\ \boxed{\sigma_1}\ \boxed{\sigma_2} \end{array} \equiv \text{shaded part of} \left( \begin{array}{c} \text{diagram} \end{array} \right). \tag{A26}
$$

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 four-leg ⊥ tensor (red) is $u^{(4)}P_\perp$. We can rewrite it as $u^{(4)} - P_\parallel = u^{(4)} - P_+ - P_-$. It is contracted with states $\langle +|$ and $\langle -|$ on the top (blue) and the kets at the bottom from the green tensors:

$$
\begin{aligned}
&= \frac{Kq}{q^2-1}\left[|+\rangle - \frac{1}{q^3}|-\rangle\right], \\
&= \frac{K}{q^2-1}\left[|+\rangle - \frac{1}{q}|-\rangle\right], \\
&= \frac{K}{q^2-1}\left[|-\rangle - \frac{1}{q^3}|+\rangle\right], \\
&= \frac{K}{q^2-1}\left[|+\rangle - \frac{1}{q}|-\rangle\right].
\end{aligned} \tag{A27}
$$

Combining these, we have

$$
\begin{array}{c} \boxed{+}\ \boxed{-} \\ \boxed{+}\boxed{\perp}\boxed{-} \\ \boxed{+}\ \boxed{-} \end{array} = -2K^2 + K^2\left(\frac{q}{q^2-1}\right)^2 \times
$$

$$
\left( \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ + \quad - \end{array} - \frac{1}{q^3} \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ + \quad + \end{array} - \frac{1}{q^3} \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ - \quad - \end{array} + \frac{1}{q^6} \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ - \quad + \end{array} \right) \tag{A28}
$$

where the $\pm$ 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,

$$
\begin{array}{c} \boxed{+}\ \boxed{-} \\ \boxed{+}\boxed{\perp}\boxed{-} \\ \boxed{+}\ \boxed{+} \end{array} = \begin{array}{c} \boxed{+}\ \boxed{-} \\ \boxed{+}\boxed{\perp}\boxed{-} \\ \boxed{-}\ \boxed{-} \end{array} = -K^2 + K^2\frac{q}{(q^2-1)^2} \times
$$

$$
\left( \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ + \quad + \end{array} - \frac{1}{q} \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ + \quad - \end{array} - \frac{1}{q^3} \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ - \quad + \end{array} + \frac{1}{q^4} \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ - \quad - \end{array} \right), \tag{A29}
$$

$$
\begin{array}{c} \boxed{+}\ \boxed{-} \\ \boxed{+}\boxed{\perp}\boxed{-} \\ \boxed{-}\ \boxed{+} \end{array} = K^2\frac{1}{(q^2-1)^2} \times
$$

$$
\left( \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ - \quad + \end{array} - \frac{1}{q} \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ + \quad + \end{array} - \frac{1}{q} \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ - \quad - \end{array} + \frac{1}{q^2} \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ + \quad - \end{array} \right). \tag{A30}
$$

The blocks on the second line, representing different contractions of $u^{(4)}$, are the operator purities for different partitions. Two of them are actually independent of the unitary gate $u$:

$$
\begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ + \quad + \end{array} = \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ - \quad - \end{array} = q^3 \tag{A31}
$$

This independence is because $|+\rangle|+\rangle$ and $|-\rangle|-\rangle$ are invariant under the action of $u^{(4)}$. The other two terms are proportional to the operator purity of vertical and diagonal partitions:

$$
\begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ + \quad - \end{array} = q^4\mathcal{P}_{||}, \qquad \begin{array}{c} + \quad - \\ \boxed{u^{(4)}} \\ - \quad + \end{array} = q^4\mathcal{P}_{\times}. \tag{A32}
$$

Using these operator purities and simplifying, we obtain the expressions in Eq. (49).

#### 5. The size of the fluctuations

In Sec. III, we list four local configurations where an isolated ⊥ lies at the path of the domain wall. The average weight of each configuration is zero in the Haar ensemble. In this Appendix, we estimate their fluctuations, i.e., their mean-square values:

$$
\underbrace{\begin{array}{c}\boxed{+}\boxed{-}\\\boxed{+}\boxed{\perp}\boxed{-}\\\boxed{+}\boxed{-}\end{array}}_{\textcircled{1'}} \quad \underbrace{\begin{array}{c}\boxed{+}\boxed{-}\\\boxed{+}\boxed{\perp}\boxed{-}\\\boxed{+}\boxed{+}\end{array}}_{\textcircled{2'}} = \underbrace{\begin{array}{c}\boxed{+}\boxed{-}\\\boxed{+}\boxed{\perp}\boxed{-}\\\boxed{+}\boxed{+}\end{array}}_{\textcircled{3'}} \quad \underbrace{\begin{array}{c}\boxed{+}\boxed{-}\\\boxed{+}\boxed{\perp}\boxed{-}\\\boxed{-}\boxed{+}\end{array}}_{\textcircled{4'}} \tag{A33}
$$

$\textcircled{2'} = \textcircled{3'}$ by the symmetry under spatial reflection and $+/-$ exchange. We thus evaluate the other three.

The averages of these diagrams generate the $S_4$ model discussed in Appendix A 3. They can be expressed in terms of ⊥ and $\pm$ triangles:

$$
\begin{array}{c} \| \diagup (12)(34) \\ \| \diagdown \perp \diagup \\ \diagup \sigma_b \diagdown \diagup \sigma_c \diagdown (12)(34) \end{array} \tag{A34}
$$

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. (A33). Because of the doubling of layers (due to squaring), $+$ corresponds to (e.g.) $\sigma_b = \mathbb{1} \in S_4$, and $-$ corresponds to $\sigma_b = (12)(34)$. We can write these weights as

$$w_\perp(\sigma_b, \sigma_c) = \sum_{\sigma_a} w_\perp(\sigma_b, \sigma_c; \sigma_a), \qquad (A35)$$

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

$$w_\perp(\sigma_b, \sigma_c; \sigma_a) \qquad (A36)$$

$$= \sum_{\sigma_a} \overset{\mathbb{1}\diagup\!\!\!\overline{\triangledown}\diagdown^{(12)(34)}}{\underset{\sigma_a}{\perp}} \times {}^{\mathbb{1}}\!\diagup\!\overline{\triangledown}\diagdown^{\sigma_a}_{\sigma_b} \times {}^{\sigma_a}\!\diagup\!\overline{\triangledown}\diagdown^{(12)(34)}_{\sigma_c} . \qquad (A37)$$

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. (A19) and (A20), the $\sigma_b$ and $\sigma_c$ triangles have the leading-order expansion

$${}^{\mathbb{1}}\!\diagup\!\overline{\triangledown}\diagdown^{\sigma_a}_{\sigma_b} \lesssim q^{-|\sigma_b|-|\sigma_b^{-1}\sigma_a|} \le q^{-|\sigma_a|}, \qquad (A38)$$

$${}^{\sigma_a}\!\diagup\!\overline{\triangledown}\diagdown^{(12)(34)}_{\sigma_c} \lesssim q^{-|\sigma_c(12)(34)|-|\sigma_c^{-1}\sigma_a|} \le q^{-|\sigma_a(12)(34)|}. \qquad (A39)$$

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

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

$${}^{\mathbb{1}}\!\diagup\!\overline{\triangledown}\diagdown^{\sigma_a}_{\sigma_b} \times {}^{\sigma_a}\!\diagup\!\overline{\triangledown}\diagdown^{(12)(34)}_{\sigma_c} \le q^{-|\sigma_a(12)(34)|-|\sigma_a^{-1}|}$$

$$\le q^{-2}, \qquad (A40)$$

$${}^{\mathbb{1}}\!\diagup\!\overline{\triangledown}\diagdown^{(12)(34)}_{\underset{\sigma_a}{\perp}} \le q^{-8}, \qquad (A41)$$

where Eq. (A41) is based on the large $q$ expansion in Eq. (60) of Ref. [14].

Hence,

$$w_\perp(\sigma_b, \sigma_c; \sigma_a) \le q^{-10} \quad \text{for } \sigma_a \in S_2 \times S_2. \qquad (A42)$$

On the other hand, if $\sigma_a \notin S_2 \times S_2$, the $\perp$ triangle is equal to an ordinary triangle [Eq. (A23)]. There will be at 8 walls in the $\perp$ triangle, and we have

$${}^{\mathbb{1}}\!\diagup\!\overline{\triangledown}\diagdown^{\sigma_a}\;{}^{\sigma_a}\!\diagup\!\overline{\triangledown}\diagdown^{(12)(34)}_{\sigma_b\qquad\sigma_c} \le q^{-|\sigma_a(12)(34)|-|\sigma_a|} \qquad (A43)$$

$$\le q^{-4} \qquad (A44)$$

and

$${}^{\mathbb{1}}\!\diagup\!\overline{\triangledown}\diagdown^{(12)(34)}_{\underset{\sigma_a}{\perp}} \sim q^{-|\sigma_a(12)(34)|-|\sigma_a|} \le q^{-4}. \qquad (A45)$$

So

$$w_\perp(\sigma_b, \sigma_c; \sigma_a) \le q^{-8} \quad \text{for } \sigma_a \notin S_2 \times S_2. \qquad (A46)$$

These inequalities can be saturated *only* when

$$\sigma_b = \mathbb{I}, \qquad \sigma_c = (12)(34),$$
$$\sigma_a = (13)(24) \quad \text{or} \quad (14)(23). \qquad (A47)$$

Therefore, the largest terms in Eq. (A33) are

$$w_\perp[\mathbb{I}, (12)(34)] \sim w_\perp[\mathbb{I}, (12)(34); (13)(24)] \qquad (A48)$$

$$+ w_\perp[\mathbb{I}, (12)(34); (14)(23)] \qquad (A49)$$

$$\sim 2q^{-8}. \qquad (A50)$$

In other words, ①′ dominates and scales as

$$①′ \sim 2q^{-8}. \qquad (A51)$$

Note that we may also write

$$\overline{①′}^n = \overline{(Z_u - \bar{Z}_u)^n}, \qquad (A52)$$

where we use $P_\perp = 1 - P_\parallel$ to split ① into $Z_u - \overline{Z}_u$ and

$$Z_u = \boxed{\begin{smallmatrix} + & - \\ + & u & - \\ + & - \end{smallmatrix}} . \qquad (A53)$$

We may then use the result for $\overline{Z_u^n}$ in Sec. VI A of Ref. [14].

## APPENDIX B: FURTHER DETAILS OF NUMERICAL RESULTS

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

In the main text, we define 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 $(++)^*$ and $(--)^*$ states at the bottom; see Eq. (35)].

A naïve 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 contain only $|\pm\rangle$. We discuss a subset of these modified partition functions, which we call $Z_{op}(x, t)$, in the context of dual-unitary circuits in Eq. (43). 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 $Z_{op}$ is then proportional to the operator purity for this partition. By rewriting the states at the bottom boundary via

$$|(++)^*\rangle = \frac{1}{q^4 - 1}\left(|++\rangle - \frac{1}{q^2}|--\rangle\right),$$

$$|(--)^*\rangle = \frac{1}{q^4 - 1}\left(|--\rangle - \frac{1}{q^2}|++\rangle\right), \quad (B1)$$

we find that $Z(x, t)$ is a linear combination of these operator purities. Our numerical scheme then computes these operator purities and obtains $Z(x, t)$ as the linear combination. We can then recursively compute $W(x, t)$ through Eq. (43). This scheme enables us to compute up to $t_{max} = 8$ for a circuit.

For general Floquet evolutions that do not have a circuit representation, discussed in Sec. VIII, the *exact* computation of $Z(x, t)$ requires, in principle, infinitely many operator purities of the Floquet operator on an infinite chain. As argued in the text, in practice, we may truncate the Floquet operator to a chain of $L$ sites. To minimize finite size effects, we choose the entry and exit locations of the membrane (at the top and bottom, respectively) to be symmetric about the center of the chain. To compute the partition function $Z(x, t)$, in the one-site scheme we contract on the top with $\langle + + \cdots - -|$, where the last $+$ is at $x_1$, and we contract on the bottom with $|+^* +^* \cdots -^* -^*\rangle$, where the last $+^*$ is at site $x_2$ (i.e., the domain wall enters at the link to the right of $x_1$ and leaves at the link to the right of $x_2$). For the two-site scheme, we use two-site boundary states, $\langle \pm \pm |$ at the top and $|(\pm\pm)^*\rangle$ at the bottom, except that at the spatial boundaries we use one-site states when commensuration effects make it necessary. For both schemes, we take $x_1 = \lceil (L - x)/2 \rceil$ and $x_2 = \lceil (L + x)/2 \rceil$.

The computational cost is dominated by the largest $L$ we study. In addition to the cost of computing each purity, we must sum over of the order of $2^{L/2}$ or $2^L$ (for the two-and one-site insertions, respectively) different purities. We use the same $L$ for each $x$ and $t$ (in principle, one could reduce $L$ for the smaller values of $t$). Since we use different top boundary conditions for different $x$, there is an additional polynomial factor in the cost proportional to the number of distinct $x$ and $t$ values.

### 2. The random gate

The matrix elements of the random gate used in Fig. 7(b) are approximately

$$u = \begin{bmatrix} -0.10054 & 0.01426 & -0.53043 & 0.35559 \\ 0.48842 & 0.52408 & -0.44873 & -0.19577 \\ -0.22473 & 0.50727 & 0.11189 & 0.26009 \\ 0.77248 & 0.02169 & 0.26725 & 0.12918 \end{bmatrix}$$

(B2)

$$+i \begin{bmatrix} -0.04094 & 0.27291 & -0.05859 & 0.70873 \\ -0.28822 & 0.22707 & -0.15763 & -0.29600 \\ 0.11952 & -0.46156 & -0.61862 & 0.04465 \\ 0.07122 & -0.35814 & 0.14956 & 0.39873 \end{bmatrix}.$$

(B3)

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 [18,25,105] to estimate $v_B$ in the Floquet circuit. The Pauli weight is defined to be the fraction of the operator weight at a site on nonidentity operators. To minimize the number of distinct length scales affecting the data, we place $\sigma^x$ at the leftmost site 0 and probe its Pauli weight at the rightmost site $d = L - 1$:

$$W_{Pauli}(t, d) = -\frac{1}{8}\sum_{\alpha=x,y,z} tr\{[\sigma_0^x(t), \sigma_d^\alpha]^2\}/2^L$$

$$= \frac{3}{4} - \frac{1}{4}\sum_{\alpha=x,y,z} tr[\sigma_0^x(t)\sigma_d^\alpha\sigma_0^x(t)\sigma_d^\alpha]/2^L. \quad (B4)$$

According to the operator spreading picture, the Pauli strings in the time-evolved operator $\sigma_0^x(t)$ 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 $v_B t \gg d$, $\sigma_0^x(t)$ can be regarded as a random operator for the purpose of calculating the Pauli weight, which is then $\frac{3}{4}$.

Figure 18(a) shows the Pauli weight for the kicked-Ising model with the parameters in Eq. (80). Because of 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 are greater than zero. Following Ref. [18], we define $t_{arrival}$ as the time when the interpolated curve is at $\frac{3}{8}$. Fitting $d$ with $t_{arrival}$ gives us an estimation of $v_B$. Figure 18(b) shows the results of fitting for the generic symmetric gate and the fixed but randomly chosen gate (Sec. VA) and for the kicked-Ising model [Eq. (80)].
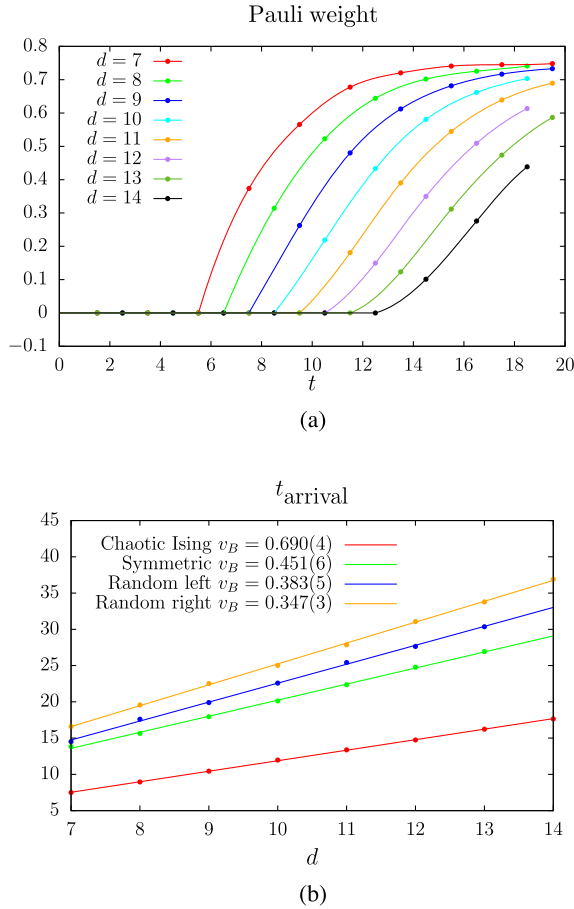
(a)



(b)

FIG. 18. (a) Pauli weight for the kicked-Ising model with parameters in Eq. (80). (b) Estimation of $v_B$ by fitting $t_{\text{arrival}}$ versus $d$. "Chaotic Ising" is the kicked-Ising model with the parameters in Eq. (80), while "Symmetric" and "Random" refer to the generic parity-symmetric gate with $x = 0.8$ [Eq. (72)] and the fixed random parity-asymmetric gate (see Sec. V A and Appendix B 2). For the asymmetric gate, we fit separately for the left and right butterfly velocities. For the dual-unitary model (not shown), $v_B = 1$ exactly.

## 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 selects out $v_E = \min_v \mathcal{E}(v)$ as the entanglement speed for $S_2$. We extract the approximate entanglement speed from our numerical calculation of $\mathcal{E}(v)$ and show its dependence on the truncation time $t$ in Fig. 19.

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 toward the analytical result $v_E = 1$. Curves for the Floquet Ising model, the generic reflection symmetric gate, and the fixed random gate behave reasonably well. Fluctuations are still visible for $t_{\text{max}}$ up to 8, and it is hard
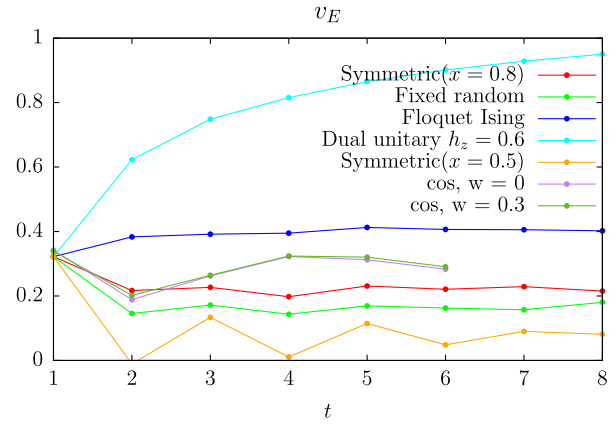


FIG. 19. The approximate entanglement velocity $v_E$ computed by truncating to times $\leq t$, for $t$ up to 8. The first four curves are for the four circuits studied in Sec. V [the first three are also the subject of Fig. 18(b) for $v_B$]. The fifth shows the generic symmetric gate [Eq. (72)] at a weaker coupling strength. The two curves with $w = 0$ and $w = 0.3$ are the entanglement velocities for the Floquet Hamiltonian with sinusoidal field in Eq. (107). Because of the absence of circuit structure, their starting values at $t = 1$ are not identical to that obtained from the random average, while the others are.

to determine the functional form for convergence directly from this plot. However, in the next subsection, we give evidence for exponential suppression of the irreducible step weights in these models at large $t$.

We also include in the plot a continuous time model with a static Hamiltonian [the limit $w = 0$ of the Floquet model with a sinusoidal field, Eq. (107)]. This model requires a separate theoretical analysis, and we include it only for comparison. Because of continuous time translation symmetry, it has a diffusive conserved energy density. This diffusive density is expected to suppress $v_E$ to zero for the second Rényi entropy [83,106,107] and is likely to introduce parametrically larger (in $t$) finite time effects [108].

## 5. The suppression of irreducible weights

In Sec. III A, we conjecture that $\perp$ clusters inside a thick domain wall are exponentially suppressed at large $t$, leading, in turn, to an exponential suppression of the irreducible step weights at large $t$. We show this suppression explicitly for a typical fixed realization of a random circuit at large $q$: There, long steps of size $t$ are suppressed by $\mathcal{O}(q^{-t})$ compared to the leading-order contributions to the domain wall free energy on this scale.

Here, we examine the nature of the decay in step weights at large $t$ numerically for translation-invariant spin-$1/2$ models. Quantitatively, we can compare $W(x, t)$ with $\exp(-s_{\text{eq}}v_E t)$, where the former represents the irreducible step weight for a single step of size $t$, while the latter represents the asymptotic scaling of the partition function $Z(0, t)$ for a domain wall of duration $t \gg 1$. Their ratio should decay exponentially according to our conjecture.
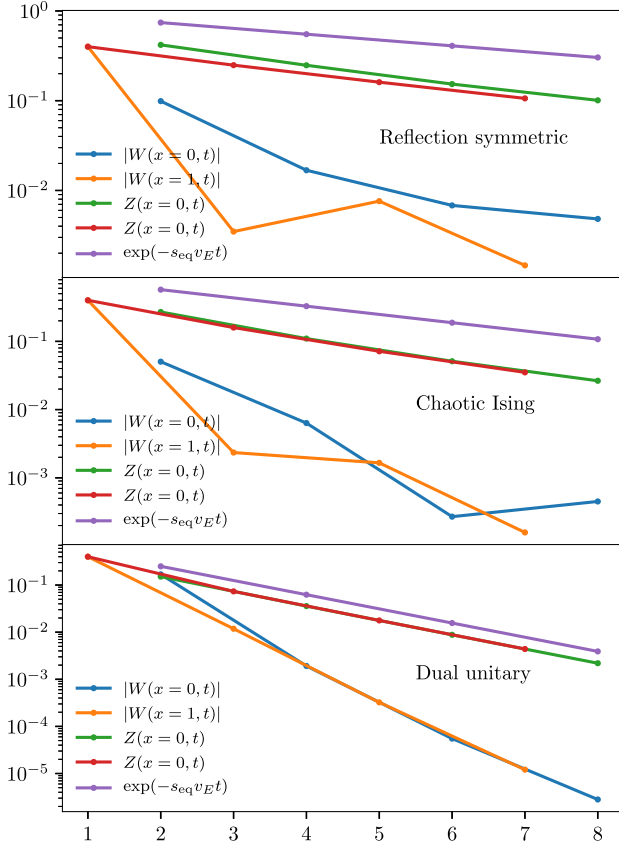
FIG. 20. The decay of $|W|$ and $Z$ at $x = 0, 1$ and $\exp(-s_{\text{eq}} v_E t)$ for (top) reflection symmetric gate in Eq. (72), (center) Floquet Ising model in Eq. (80), and (bottom) dual-unitary gate in Eq. (81).

In Fig. 20, we plot $W(x, t)$ along with both $Z(x, t)$ and $\exp(-s_{\text{eq}} v_E t)$ as a comparison. The best case is the dual-unitary circuit, where a faster exponential decay of $W(x, t)$ with respect to both $Z(x, t)$ and $\exp(-s_{\text{eq}} v_E t)$ is clearly visible. The Floquet Ising model follows a similar trend. Though this trend is less clear for the reflection symmetric gate at $x = 0$, that is consistent with the fact that this gate is
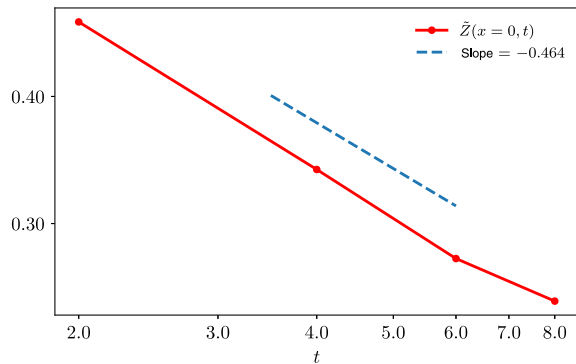


FIG. 21. The diffusive decay of $\tilde{Z}(x, t)$ at $x = 0$ for the chaotic Ising model in Eq. (80). Both axes are logarithmic.

more weakly entangling than the others (Fig. 19) and takes larger $t$ to reach convergence.

### 6. Diffusive character of the domain wall

In Sec. VI, we argue that for generic choices of the gate the wandering of the domain wall is diffusive at large scales. One consequence of this result is the $t^{-1/2}$ scaling in Eq. (83): $\tilde{Z}(0, t) \sim t^{-1/2}$. We state in the caption to Fig. 12(a) that the data are in good agreement with this scaling. This agreement is demonstrated in Fig. 21, which shows $\tilde{Z}(0, t)$ versus $t$ on a log-log plot.

## APPENDIX C: RELATION BETWEEN ROOT AND LINE TENSION

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

$$z(s, t) \equiv \sum_x Z(x, t) e^{sx}. \quad (C1)$$

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

$$z(s, t) \sim \exp[-s_{\text{eq}} \Gamma(s) t] \quad (C2)$$

with [18]

$$\Gamma(s) = \min_v \left[ \mathcal{E}(v) - \frac{vs}{s_{\text{eq}}} \right], \quad (C3)$$

so that $-s_{\text{eq}} \Gamma(s)$ and $s_{\text{eq}} \mathcal{E}(v)$ form a standard Legendre transform pair. On the other hand,

$$\sum_t z(s, t) b^t = \frac{\omega(s, b)}{1 - \omega(s, b)}. \quad (C4)$$

Assuming that the smallest root of $\omega(s, b)$ for a given $s$ is $b_0(s)$, then for $|b| \lesssim |b_0|$

$$\sum_t z(s, t) b^t \sim \frac{1}{1 - b/b_0}. \quad (C5)$$

Matching the expansions in $b$,

$$s_{\text{eq}} \Gamma(s) = \ln|b_0(s)|. \quad (C6)$$

Finally, inverting the Legendre transform yields the expression (61) 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. VI, we postulate that $Z(x, t)$ for the dual-unitary model is approximately constant in space, and so $z(s = 0, t) \sim t e^{-s_{\text{eq}} t}$. The factor of

$t$ in the front indicates a double root on the right-hand side of Eq. (C4) in the limit of large $t_{\max}$.

## APPENDIX D: THE CONTINUUM EQUATION FOR $\tilde{Z}$

For completeness, let us give a more general derivation of the continuum equation which makes contact with the equation in Ref. [18] for the dynamical evolution of the state entanglement. Writing $Z = \exp(-S)$, we assume that the solution is such that higher derivatives of $S(x,t)$ are small at late times (to be checked self-consistently) and expand in derivatives.

So far, we consider a domain wall which is tethered at both the top and the bottom. To compute the entanglement $S(x,t)$ of a state across an entanglement cut at position $x$, we should instead take the domain wall to be tethered at position $x$ at the top and free at the bottom; the initial condition $Z(x,0) = \exp[-S(x,0)]$ is set by the entanglement in the initial state. The recursion relation still applies.

To reduce clutter, let us write $(y,\tau)$ instead of $(\Delta x, \Delta t)$ in the recursion relation. In terms of $S(x,t)$, it reads

$$e^{-S(x,t)} = \sum_{y,\tau} W(y,\tau) e^{-S(x-y,t-\tau)}. \qquad \text{(D1)}$$

Expanding in the exponent gives

$$1 = \sum_{y,\tau} W(y,\tau) \exp\left(yS' + \tau\dot{S} - \frac{y^2}{2}S'' - \frac{\tau^2}{2}\ddot{S} - y\tau\dot{S}'\right). \qquad \text{(D2)}$$

Generically, $S$ is of order $t$, $S'$ and $\dot{S}$ are of order 1, and the higher derivatives are smaller. Therefore, the leading-order equation is simply

$$1 = \sum_{y,\tau} W(y,\tau) \exp\left(yS' + \tau\dot{S}\right). \qquad \text{(D3)}$$

Using the definitions in Sec. IV A, this equation is solved by

$$\dot{S} = s_{\text{eq}}\Gamma(S'), \qquad \text{(D4)}$$

so that $\Gamma$ is the growth rate, which depends on the derivative $S'$ of the entanglement as a function of the spatial location. We are interested in the subleading corrections to this equation. Let us define an $s$-dependent "average" of an arbitrary function $F$ of the step:

$$\langle F(y,\tau)\rangle_s \equiv \sum_{y,\tau} W(y,\tau) e^{sy+s_{\text{eq}}\Gamma(s)\tau} F(y,\tau). \qquad \text{(D5)}$$

From the relations for the generation functions,

$$\langle 1\rangle_s = 1. \qquad \text{(D6)}$$

We also define the associated velocity $v(s) = \langle y\rangle_s/\langle\tau\rangle_s$. By differentiating Eq. (D6) with respect to $s$, we have

$$v(s) = -s_{\text{eq}}\frac{d\Gamma(s)}{ds}. \qquad \text{(D7)}$$

This equation is the coarse-grained velocity of the membrane, since it also solves the Legendre transform [Eq. (61)] relating the growth rate $\Gamma[s]$ and the line tension $\mathcal{E}(v)$.

We may then write Eq. (D2) as

$$\left\langle e^{\tau[\dot{S}-s_{\text{eq}}\Gamma(S')]} e^{-(y^2/2)S''-(\tau^2/2)\ddot{S}-y\tau\dot{S}'+\cdots}\right\rangle_{S'}, \qquad \text{(D8)}$$

where the $\cdots$ are higher derivatives. The exponent is now written in terms of quantities that are small at late times and so may be expanded:

$$\dot{S} = s_{\text{eq}}\Gamma(S') + \frac{\langle y^2\rangle_{S'}S'' + \langle\tau^2\rangle_{S'}\ddot{S} + 2\langle y\tau\rangle_{S'}\dot{S}'}{2\langle\tau\rangle_{S'}}. \qquad \text{(D9)}$$

We may use the leading-order solution to replace the time derivatives on the right-hand side with space derivatives [cf. Eqs. (D4) and (D7)]: $\dot{S}' \simeq -v(S')S''$ and $\ddot{S} \simeq v(S')^2 S''$. Then,

$$\dot{S} = s_{\text{eq}}\Gamma(S') + \frac{\langle[y-v(S')\tau]^2\rangle}{2\langle\tau\rangle}S''. \qquad \text{(D10)}$$

By differentiating Eq. (D6) twice,

$$\frac{\langle[y-v(s)\tau]^2\rangle}{\langle\tau\rangle} = -\frac{d^2\Gamma(s)}{ds^2}, \qquad \text{(D11)}$$

so

$$\dot{S} = s_{\text{eq}}\left(\Gamma(S') - \frac{1}{2}\Gamma''(S')S''\right). \qquad \text{(D12)}$$

(The primes on $\Gamma$ indicate derivatives with respect to its argument.)

We may also write everything in terms of $\mathcal{E}(v)$. From the Legendre transform relationship between $\mathcal{E}(v)$ and $\Gamma(s)$ and Eq. (D7),

$$\mathcal{E}(v) = \Gamma[s] + \frac{v(s)s}{s_{\text{eq}}}, \qquad \text{(D13)}$$

we differentiate twice and obtain

$$\mathcal{E}'(v) = \frac{s}{s_{\text{eq}}}, \qquad \mathcal{E}''(v) = \frac{1}{s_{\text{eq}}v'(s)} = -\frac{1}{s_{\text{eq}}^2\Gamma''[s]}. \qquad \text{(D14)}$$

This result along with Eq. (D12) gives a diffusive equation

$$\dot{S} = s_{\text{eq}}\mathcal{E}(v) + \frac{1}{2s_{\text{eq}}\mathcal{E}''(v)}S'' \qquad \text{(D15)}$$

with diffusion constant $D = 1/[2s_{\text{eq}}\mathcal{E}(v)]$, consistent with Eq. (85).

## APPENDIX E: TRIANGLE WEIGHTS FOR TWO COMMUTING DOMAIN WALLS

For reference and use in Appendix A 5, here we list exact expressions for triangle weights $J[\mathbb{I}, (12)(34); \sigma_a]$. This expression corresponds to two incoming "commutating" elementary domain walls at the top of the triangle. (An

elementary domain wall is a transposition.) We also list 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.

There are in total six 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 $\sigma_a$:

(i) $\mathbb{I}$ and (12)(34).—

$$\underset{(12)(34)}{\overset{\mathbb{I}}{\triangledown}}\,^{(12)(34)} = \;\;\;\;\; = \frac{q^8 - q^6 - 12q^4 + 14q^2 + 22}{(q^4 - 9)(q^4 - 4)(q^2 + 1)} = \left[\;\;\;\;\;\right]^2 \left[1 + \frac{2}{q^6} + \mathcal{O}\!\left(\frac{1}{q^{12}}\right)\right]. \qquad (E1)$$

(ii) (12) and (34).—

$$\underset{(12)}{\overset{\mathbb{I}}{\triangledown}}\,^{(12)(34)} = \;\;\;\;\; = \frac{(q^8 - 12q^4 + 12)(q^2 - 1)}{(q^4 - 9)(q^4 - 4)(q^2 + 1)q^2} = \left[\;\;\;\;\;\right]^2 \left[1 - \frac{12}{q^8} + \mathcal{O}\!\left(\frac{1}{q^{12}}\right)\right]. \qquad (E2)$$

(iii) (13)(24) and (14)(23).—

$$\underset{(13)(24)}{\overset{\mathbb{I}}{\triangledown}}\,^{(12)(34)} = \;\;\;\;\; = \frac{q^2 - 1}{(q^4 - 9)(q^2 + 1)} = \left[\;\;\;\;\;\right]^2 \left[\frac{1}{q^2} + \mathcal{O}\!\left(\frac{1}{q^6}\right)\right]. \qquad (E3)$$

(iv) (13), (24), (14), (23), (1234), (1423), (1243), and (1342).—

$$\underset{(1234)}{\overset{\mathbb{I}}{\triangledown}}\,^{(12)(34)} = \;\;\;\;\; = -\frac{(q^4 + 6)(q^2 - 1)}{(q^4 - 9)(q^4 - 4)(q^2 + 1)q^2} = -\left[\;\;\;\;\;\right]^2 \left[\frac{1}{q^4} + \mathcal{O}\!\left(\frac{1}{q^8}\right)\right]. \qquad (E4)$$

(v) (123), (124), (132), (142), (234), (134), (243), and (143).—

$$\underset{(123)}{\overset{\mathbb{I}}{\triangledown}}\,^{(12)(34)} = \;\;\;\;\; = \frac{5(q^2 - 1)}{(q^4 - 9)(q^4 - 4)(q^2 + 1)} = \left[\;\;\;\;\;\right]^2 \left[\frac{5}{q^6} + \mathcal{O}\!\left(\frac{1}{q^{10}}\right)\right]. \qquad (E5)$$

(vi) (1324) and (1423).—

$$\mathbb{I}\underset{(1324)}{\overset{(12)(34)}{\bigtriangledown}} \;=\; \bigtriangledown\!\!\!\!\bigtriangledown \;=\; -\frac{3(q^2-1)}{(q^4-9)(q^2+1)q^2} \;=\; -\left[\;\bigtriangledown\;\right]^2\left[\frac{3}{q^4}+\mathcal{O}\left(\frac{1}{q^8}\right)\right]. \qquad (E6)$$

[1] L. P. Kadanoff and P. C. Martin, *Hydrodynamic Equations and Correlation Functions*, Ann. Phys. (N.Y.) **24**, 419 (1963).

[2] P. C. Hohenberg and B. I. Halperin, *Theory of Dynamic Critical Phenomena*, Rev. Mod. Phys. **49**, 435 (1977).

[3] R. Oliveira, O. C. O. Dahlsten, and M. B. Plenio, *Generic Entanglement Can Be Generated Efficiently*, Phys. Rev. Lett. **98**, 130502 (2007).

[4] A. Hamma, S. Santra, and P. Zanardi, *Quantum Entanglement in Random Physical States*, Phys. Rev. Lett. **109**, 040502 (2012).

[5] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, *Local Random Quantum Circuits Are Approximate Polynomial-Designs*, Commun. Math. Phys. **346**, 397 (2016).

[6] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, *Quantum Entanglement Growth under Random Unitary Dynamics*, Phys. Rev. X **7**, 031016 (2017).

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

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

[9] T. Rakovszky, F. Pollmann, and C. W. von Keyserlingk, *Diffusive Hydrodynamics of Out-of-Time-Ordered Correlators with Charge Conservation*, Phys. Rev. X **8**, 031058 (2018).

[10] V. Khemani, A. Vishwanath, and D. A. Huse, *Operator Spreading and the Emergence of Dissipative Hydrodynamics under Unitary Evolution with Conservation Laws*, Phys. Rev. X **8**, 031057 (2018).

[11] A. Chan, A. De Luca, and J. T. Chalker, *Solution of a Minimal Model for Many-Body Quantum Chaos*, Phys. Rev. X **8**, 041019 (2018).

[12] A. Chan, A. De Luca, and J. T. Chalker, *Spectral Statistics in Spatially Extended Chaotic Quantum Many-Body Systems*, Phys. Rev. Lett. **121**, 060601 (2018).

[13] D. A. Rowlands and A. Lamacraft, *Noisy Coupled Qubits: Operator Spreading and the Fredrickson-Andersen Model*, Phys. Rev. B **98**, 195125 (2018).

[14] T. Zhou and A. Nahum, *Emergent Statistical Mechanics of Entanglement in Random Unitary Circuits*, Phys. Rev. B **99**, 174205 (2019).

[15] T. Rakovszky, F. Pollmann, and C. W. von Keyserlingk, *Sub-ballistic Growth of Rényi Entropies due to Diffusion*, Phys. Rev. Lett. **122**, 250602 (2019).

[16] A. J. Friedman, A. Chan, A. De Luca, and J. T. Chalker, *Spectral Statistics and Many-Body Quantum Chaos with Conserved Charge*, Phys. Rev. Lett. **123**, 210603 (2019).

[17] N. Hunter-Jones, *Unitary Designs from Statistical Mechanics in Random Quantum Circuits*, arXiv:1905.12053.

[18] C. Jonay, D. A. Huse, and A. Nahum, *Coarse-Grained Dynamics of Operator and State Entanglement*, arXiv:1803.00089.

[19] E. H. Lieb and D. W. Robinson, *The Finite Group Velocity of Quantum Spin Systems*, in *Statistical Mechanics* (Springer, New York, 1972), pp. 425–431.

[20] P. Calabrese and J. Cardy, *Evolution of Entanglement Entropy in One-Dimensional Systems*, J. Stat. Mech. (2005) P04010.

[21] H. Kim and D. A. Huse, *Ballistic Spreading of Entanglement in a Diffusive Nonintegrable System*, Phys. Rev. Lett. **111**, 127205 (2013).

[22] A. Kitaev, *Hidden Correlations in the Hawking Radiation and Thermal Noise*, in Proceedings of the Fundamental Physics Prize Symposium, 2014, http://online.kitp.ucsb.edu/online/joint98/kitaev/.

[23] H. Liu and S. J. Suh, *Entanglement Tsunami: Universal Scaling in Holographic Thermalization*, Phys. Rev. Lett. **112**, 011601 (2014).

[24] S. H. Shenker and D. Stanford, *Black Holes and the Butterfly Effect*, J. High Energy Phys. 03 (2014) 067.

[25] D. A. Roberts, D. Stanford, and L. Susskind, *Localized Shocks*, J. High Energy Phys. 03 (2015) 051.

[26] J. Maldacena, S. H. Shenker, and D. Stanford, *A bound on Chaos*, J. High Energy Phys. 08 (2016) 106.

[27] D. A. Roberts and B. Swingle, *Lieb-Robinson Bound and the Butterfly Effect in Quantum Field Theories*, Phys. Rev. Lett. **117**, 091602 (2016).

[28] A. M. Kaufman, M. E. Tai, A. Lukin, M. Rispoli, R. Schittko, P. M. Preiss, and M. Greiner, *Quantum Thermalization through Entanglement in an Isolated Many-Body System*, Science **353**, 794 (2016).

[29] I. L. Aleiner, L. Faoro, and L. B. Ioffe, *Microscopic Model of Quantum Butterfly Effect: Out-of-Time-Order Correlators and Traveling Combustion Waves*, Ann. Phys. (Amsterdam) **375**, 378 (2016).

[30] A. A. Patel and S. Sachdev, *Quantum Chaos on a Critical Fermi Surface*, Proc. Natl. Acad. Sci. U.S.A. **114**, 1844 (2017).

[31] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, *Chaos in Quantum Channels*, J. High Energy Phys. 02 (2016) 004.

[32] M. Mezei, *On Entanglement Spreading from Holography,* J. High Energy Phys. 05 (2017) 064.

[33] P. Kos, M. Ljubotina, and T. Prosen, *Many-Body Quantum Chaos: Analytic Connection to Random Matrix Theory,* Phys. Rev. X **8,** 021062 (2018).

[34] M. Mezei and J. Virrueta, *Exploring the Membrane Theory of Entanglement Dynamics,* J. High Energy Phys. 02 (2020) 013.

[35] B. Bertini, P. Kos, and T. Prosen, *Entanglement Spreading in a Minimal Model of Maximal Many-Body Quantum Chaos,* Phys. Rev. X **9,** 021033 (2019).

[36] L. Piroli, B. Bertini, J. I. Cirac, and T. Prosen, *Exact Dynamics in Dual-Unitary Quantum Circuits,* Phys. Rev. B **101,** 094304 (2020).

[37] M. Mezei, *Membrane Theory of Entanglement Dynamics from Holography,* Phys. Rev. D **98,** 106025 (2018).

[38] S. Ryu and T. Takayanagi, *Holographic Derivation of Entanglement Entropy from the Anti–de Sitter Space/ Conformal Field Theory Correspondence,* Phys. Rev. Lett. **96,** 181602 (2006).

[39] V. E. Hubeny, M. Rangamani, and T. Takayanagi, *A Covariant Holographic Entanglement Entropy Proposal,* J. High Energy Phys. 07 (2007) 062.

[40] C. A. Agón and M. Mezei, *Bit Threads and the Membrane Theory of Entanglement Dynamics,* arXiv:1910.12909.

[41] M. Freedman and M. Headrick, *Bit Threads and Holographic Entanglement,* Commun. Math. Phys. **352,** 407 (2017).

[42] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, *Holographic Duality from Random Tensor Networks,* J. High Energy Phys. 11 (2016) 009.

[43] R. Vasseur, A. C. Potter, Y.-Z. You, and A. W. W. Ludwig, *Entanglement Transitions from Holographic Random Tensor Networks,* Phys. Rev. B **100,** 134203 (2019).

[44] H. Kim, T. N. Ikeda, and D. A. Huse, *Testing Whether All Eigenstates Obey the Eigenstate Thermalization Hypothesis,* Phys. Rev. E **90,** 052105 (2014).

[45] T. Prosen, *Chaos and Complexity of Quantum Motion,* J. Phys. A **40,** 7881 (2007).

[46] S. Gopalakrishnan and A. Lamacraft, *Unitary Circuits of Finite Depth and Infinite Width from Quantum Channels,* Phys. Rev. B **100,** 064309 (2019).

[47] B. Bertini, P. Kos, and T. Prosen, *Exact Correlation Functions for Dual-Unitary Lattice Models in $1 + 1$ Dimensions,* Phys. Rev. Lett. **123,** 210601 (2019).

[48] M. Akila, D. Waltner, B. Gutkin, and T. Guhr, *Particle-Time Duality in the Kicked Ising Spin Chain,* J. Phys. A **49,** 375101 (2016).

[49] B. Bertini, P. Kos, and T. Prosen, *Operator Entanglement in Local Quantum Circuits I: Maximally Chaotic Dual-Unitary Circuits,* SciPost Phys. **8,** 067 (2020).

[50] B. Collins and P. Śniady, *Integration with Respect to the Haar Measure on Unitary, Orthogonal and Symplectic Group,* Commun. Math. Phys. **264,** 773 (2006).

[51] Y. Gu, *Moments of Random Matrices and Weingarten Functions,* thesis, Queen's University, 2013.

[52] D. Weingarten, *Asymptotic Behavior of Group Integrals in the Limit of Infinite Rank,* J. Math. Phys. (N.Y.) **19,** 999 (1978).

[53] A. W. Harrow, *The Church of the Symmetric Subspace,* arXiv:1308.6595.

[54] As we discuss later, this property depends on the boundary conditions.

[55] P. Zanardi, *Entanglement of Quantum Evolutions,* Phys. Rev. A **63,** 040304 (2001).

[56] T. Prosen and I. Pižorn, *Operator Space Entanglement Entropy in a Transverse Ising Chain,* Phys. Rev. A **76,** 032316 (2007).

[57] I. Pižorn and T. Prosen, *Operator Space Entanglement Entropy in XY Spin Chains,* Phys. Rev. B **79,** 184416 (2009).

[58] J. Dubail, *Entanglement Scaling of Operators: A Conformal Field Theory Approach, with a Glimpse of Simulability of Long-Time Dynamics in $1 + 1$D,* J. Phys. A **50,** 234001 (2017).

[59] T. Zhou and D. J. Luitz, *Operator Entanglement Entropy of the Time Evolution Operator in Chaotic Systems,* Phys. Rev. B **95,** 094206 (2017).

[60] P. Zanardi, C. Zalka, and L. Faoro, *Entangling Power of Quantum Evolutions,* Phys. Rev. A **62,** 030301 (2000).

[61] At large $q$, an "elementary" domain wall, i.e., a transposition, costs $1/q$ per time step. Here, we have a total of four elementary domain walls [the right-hand side of Eq. (53)].

[62] J. Cardy, *Renormalisation Group Theory of Branching Potts Interfaces,* Nucl. Phys. **B565,** 506 (2000).

[63] D. A. Huse, C. L. Henley, and D. S. Fisher, *Huse, Henley, and Fisher Respond,* Phys. Rev. Lett. **55,** 2924 (1985).

[64] M. Kardar, G. Parisi, and Y.-C. Zhang, *Dynamic Scaling of Growing Interfaces,* Phys. Rev. Lett. **56,** 889 (1986).

[65] 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 reformulation 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].

[66] $W$ here is real by definition. In principle, the smallest root $b_0$ may be complex (in which case it has a complex conjugate partner $b_0^*$). We allow for this complexity above. However, from the physical interpretation of the quantities above, we expect that $b_0$ is real and positive at least for large $t_{max}$.

[67] If $u$ is not reflection symmetric, there are separate left and right butterfly velocities $v_B^L < 0$ and $v_B^R > 0$, respectively [18,68,69]. In this situation, $v_E = \min_v \mathcal{E}(v)$, $\mathcal{E}(|v_B^L|) = |v_B^L|$, and $\mathcal{E}(-|v_B^R|) = |v_B^R|$. Our convention here is that the support of the operator $\mathcal{O}(t)$ grows, with $t$, to the left at speed $|v_B^L|$ and to the right at $|v_B^R|$.

[68] C. Stahl, V. Khemani, and D. A. Huse, *Asymmetric Butterfly Velocities in Hamiltonian and Circuit Models,* arXiv:1812.05589.

[69] Y.-L. Zhang and V. Khemani, *Asymmetric Butterfly Velocities in 2-Local Hamiltonians,* arXiv:1909.03988.

[70] B. Kraus and J. I. Cirac, *Optimal Creation of Entanglement Using a Two-Qubit Gate,* Phys. Rev. A **63,** 062309 (2001).

[71] N. Khaneja and S. Glaser, *Cartan Decomposition of SU(2n̂), Constructive Controllability of Spin Systems*

and Universal Quantum Computing, arXiv:quant-ph/0010100.

[72] B. Bertini, P. Kos, and T. Prosen, *Exact Spectral Form Factor in a Minimal Model of Many-Body Quantum Chaos*, Phys. Rev. Lett. **121**, 264101 (2018).

[73] More precisely, $h_x = (\sqrt{5} + 5)/8 \simeq 0.9045$ and $h_x = (\sqrt{5} + 1)/4 \simeq 0.8090$. Here, we use the truncated decimal values.

[74] Writing the gate with $(\tau, h_x, h_z) = (0.8, 0.9045, 0.8090)$ in the representation in Eq. (68) gives $\mathbf{J} \simeq (0.8326, 0.8617, 0.0001)$. One branch of dual-unitary values is $\mathbf{J} = (\pi/4, \pi/4, J_z)$ with any $J_z$ and any $\mathbf{h}^{(i)}$.

[75] In terms of the general parameterization in Eq. (68) [35,36, 46,47,72], $\mathbf{J} = (0, \pi/4, \pi/4)$, $u_1 = u_2 = \exp(-ih_z\sigma_z/2)$ $\exp(-i\pi\sigma_x/4)$, and $u_3 = u_4 = \exp(-ih_z\sigma_z/2)$.

[76] This equation corresponds to a different lower boundary condition, where the domain wall end point at $t = 0$ is free and weighted with $Z(x, 0) = e^{-S(x,0)}$. The recursion relation is the same.

[77] The lower boundary condition contributes a weight proportional to $q^d = e^{s_{eq}d}$, 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 $\mathcal{E}$ this velocity minimizes $s_{eq}[-v_L + \mathcal{E}(v_L)]t$.

[78] V. Khemani, D. A. Huse, and A. Nahum, *Velocity-Dependent Lyapunov Exponents in Many-Body Quantum, Semiclassical, and Classical Chaos*, Phys. Rev. B **98**, 144304 (2018).

[79] A. Nahum, S. Roy, S. Vijay, and T. Zhou, *Chaos and the Time-Ordered Two-Point Function* (to be published).

[80] E. McCulloch and C. W. von Keyserlingk (private communication).

[81] E. H. Lieb and D. W. Robinson, *The Finite Group Velocity of Quantum Spin Systems*, Commun. Math. Phys. **28**, 251 (1972).

[82] Y. Huang, *Dynamics of Renyi Entanglement Entropy in Local Quantum Circuits with Charge Conservation*, arXiv:1902.00977.

[83] T. Zhou and A. W. W. Ludwig, *Diffusive Scaling of Rényi Entanglement Entropy*, Phys. Rev. Research **2**, 033020 (2020).

[84] M. Znidaric, *Entanglement Growth in Diffusive Systems*, Commun. Phys. **3**, 100 (2020).

[85] H. Liu and P. Glorioso, *Lectures on Non-equilibrium Effective Field Theories and Fluctuating Hydrodynamics*, arXiv:1805.09331.

[86] M. Rangamani, *Gravity & Hydrodynamics: Lectures on the Fluid-Gravity Correspondence*, Classical Quantum Gravity **26**, 224003 (2009).

[87] A. Chan, A. De Luca, and J. T. Chalker, *Spectral Statistics in Spatially Extended Chaotic Quantum Many-Body Systems*, Phys. Rev. Lett. **121**, 060601 (2018).

[88] Y. Li, X. Chen, and M. P. A. Fisher, *Measurement-Driven Entanglement Transition in Hybrid Quantum Circuits*, Phys. Rev. B **100**, 134306 (2019).

[89] M. J. Gullans and D. A. Huse, *Dynamical Purification Phase Transition Induced by Quantum Measurements*, arXiv:1905.05195.

[90] A. W. Harrow and R. A. Low, *Random Quantum Circuits Are Approximate 2-Designs*, Commun. Math. Phys. **291**, 257 (2009).

[91] D. A. Roberts and B. Yoshida, *Chaos and Complexity by Design*, J. High Energy Phys. 04 (2017) 121.

[92] J. Cotler, N. Hunter-Jones, J. Liu, and B. Yoshida, *Chaos, Complexity, and Random Matrices*, J. High Energy Phys. 11 (2017) 048.

[93] B. Skinner, J. Ruhman, and A. Nahum, *Measurement-Induced Phase Transitions in the Dynamics of Entanglement*, Phys. Rev. X **9**, 031009 (2019).

[94] Y. Li, X. Chen, and M. P. A. Fisher, *Quantum Zeno Effect and the Many-Body Entanglement Transition*, Phys. Rev. B **98**, 205136 (2018).

[95] A. Chan, R. M. Nandkishore, M. Pretko, and G. Smith, *Unitary-Projective Entanglement Dynamics*, Phys. Rev. B **99**, 224307 (2019).

[96] S. Choi, Y. Bao, X.-L. Qi, and E. Altman, *Quantum Error Correction in Scrambling Dynamics and Measurement-Induced Phase Transition*, Phys. Rev. Lett. **125**, 030505 (2020).

[97] M. Szyniszewski, A. Romito, and H. Schomerus, *Entanglement Transition from Variable-Strength Weak Measurements*, Phys. Rev. B **100**, 064204 (2019).

[98] X. Cao, A. Tilloy, and A. De Luca, *Entanglement in a Fermion Chain under Continuous Monitoring*, SciPost Phys. **7**, 24 (2019).

[99] M. J. Gullans and D. A. Huse, *Scalable Probes of Measurement-Induced Criticality*, arXiv:1910.00020.

[100] Q. Tang and W. Zhu, *Measurement-Induced Phase Transition: A Case Study in the Nonintegrable Model by Density-Matrix Renormalization Group Calculations*, Phys. Rev. Research **2**, 013022 (2020).

[101] Y. Bao, S. Choi, and E. Altman, *Theory of the Phase Transition in Random Unitary Circuits with Measurements*, Phys. Rev. B **101**, 104301 (2020).

[102] C.-M. Jian, Y.-Z. You, R. Vasseur, and A. W. W. Ludwig, *Measurement-Induced Criticality in Random Quantum Circuits*, Phys. Rev. B **101**, 104302 (2020).

[103] A. Zabalo, M. J. Gullans, J. H. Wilson, S. Gopalakrishnan, D. A. Huse, and J. H. Pixley, *Critical Properties of the Measurement-Induced Transition in Random Quantum Circuits*, Phys. Rev. B **101**, 060301 (2020).

[104] By the symmetry in Eq. (30), the following also applies to other subsets of states related to this one by left and right multiplication with fixed elements of $S_N$.

[105] M. Mezei and D. Stanford, *On Entanglement Spreading in Chaotic Systems*, J. High Energy Phys. 05 (2017) 065.

[106] T. Rakovszky, F. Pollmann, and C. W. Von Keyserlingk, *Sub-ballistic Growth of Rényi Entropies due to Diffusion*, Phys. Rev. Lett. **122**, 250602 (2019).

[107] Y. Huang, *Dynamics of Renyi Entanglement Entropy in Local Quantum Circuits with Charge Conservation*, arXiv:1902.00977.

[108] We expect that $\mathcal{E}(v)$ for $v \neq 0$ will still be nontrivial. The suppression of $v_E$ all the way to zero is due to rare Feynman trajectories in which a large spatial region is depleted of energy and cannot entangle. Since the dynamics of the energy density is diffusive, these empty regions can persist for a long time along a ray at $v = 0$, suppressing the cost of a membrane with $v = 0$, but not along a ray at nonzero $v$.