

Operator Spreading in Random Unitary Circuits

Adam Nahum,^{1,2} Sagar Vijay,² and Jeongwan Haah³

¹*Theoretical Physics, Oxford University, 1 Keble Road, Oxford OX1 3NP, United Kingdom*

²*Department of Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*

³*Station Q Quantum Architectures and Computation, Microsoft Research, One Microsoft Way, Redmond, Washington, USA*



(Received 1 September 2017; published 11 April 2018)

Random quantum circuits yield minimally structured models for chaotic quantum dynamics, which are able to capture, for example, universal properties of entanglement growth. We provide exact results and coarse-grained models for the spreading of operators by quantum circuits made of Haar-random unitaries. We study both $1 + 1\text{D}$ and higher dimensions and argue that the coarse-grained pictures carry over to operator spreading in generic many-body systems. In $1 + 1\text{D}$, we demonstrate that the out-of-time-order correlator (OTOC) satisfies a biased diffusion equation, which gives exact results for the spatial profile of the OTOC and determines the butterfly speed v_B . We find that in $1 + 1\text{D}$, the “front” of the OTOC broadens diffusively, with a width scaling in time as $t^{1/2}$. We address fluctuations in the OTOC between different realizations of the random circuit, arguing that they are negligible in comparison to the broadening of the front within a realization. Turning to higher dimensions, we show that the averaged OTOC can be understood exactly via a remarkable correspondence with a purely classical droplet growth problem. This implies that the width of the front of the averaged OTOC scales as $t^{1/3}$ in $2 + 1\text{D}$ and as $t^{0.240}$ in $3 + 1\text{D}$ (exponents of the Kardar-Parisi-Zhang universality class). We support our analytic argument with simulations in $2 + 1\text{D}$. We point out that, in two or higher spatial dimensions, the shape of the spreading operator at late times is affected by underlying lattice symmetries and, in general, is not spherical. However, when full spatial rotational symmetry is present in $2 + 1\text{D}$, our mapping implies an exact asymptotic form for the OTOC, in terms of the Tracy-Widom distribution. For an alternative perspective on the OTOC in $1 + 1\text{D}$, we map it to the partition function of an Ising-like statistical mechanics model. As a result of special structure arising from unitarity, this partition function reduces to a random walk calculation which can be performed exactly. We also use this mapping to give exact results for entanglement growth in $1 + 1\text{D}$ circuits.

DOI: [10.1103/PhysRevX.8.021014](https://doi.org/10.1103/PhysRevX.8.021014)

Subject Areas: Condensed Matter Physics,
Quantum Physics, Statistical Physics

I. INTRODUCTION

A key challenge for many-body physics is to identify universal properties of quantum dynamics and the approach to thermalization. Particularly important are universal results that hold for generic quantum systems. Examples of such universal properties include the existence of effective light cones for the propagation of quantum information [1] and the existence of universal scaling forms for the growth and saturation of the von Neumann entanglement entropy in $1 + 1\text{D}$ [2–15].

By definition, generic systems lack the structures (for example, large numbers of symmetries or conservation

laws) that allow for exact results in typical solvable many-body systems. Surprisingly, insight into generic systems can come from studying dynamics with even *less* structure than a generic Hamiltonian system, such as the dynamics generated by a random quantum circuit. Random circuit dynamics provide a minimally structured model with which real Hamiltonian dynamics can be compared [16–24]. Despite its simplicity, this model is able to capture universal scaling forms for entanglement growth both in $1 + 1\text{D}$ and in higher dimensions [10]. Random circuits are also toy models for information scrambling in black holes and other strongly coupled systems [16–25].

In this paper, we provide both exact results and coarse-grained descriptions for the spreading of quantum operators under random circuit dynamics, as measured by the “out-of-time-order correlator” (OTOC). The OTOC originally appeared in the study of quasiclassical approximations to superconductivity [26] and is closely related to the commutator norm that appears in Lieb-Robinson bounds [1],

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article’s title, journal citation, and DOI.

but it has been studied recently as a means of quantifying the scrambling of quantum information [27–30]. It has been argued that early-time exponential growth of the OTOC is a characteristic feature of chaotic quantum systems, and such growth has been obtained within the AdS/CFT correspondence and in a range of physical systems [31–38]. The OTOC has also been applied to characterize slow dynamics in the presence of disorder and in the many-body localized phase [39–46]. It has been calculated in 2D conformal field theories [47] and integrable chains [48], and studied numerically in nonintegrable 1D systems [49–51]. Following theoretical proposals [52–54], experiments addressing the OTOC have been conducted [55–57].

Random quantum circuits provide an ideal theoretical setting for the exact calculation of quantities such as the OTOC. While the behavior of the OTOC in a random circuit is interesting in its own right, we conjecture that the long-distance properties of the OTOC that we derive will also be applicable to deterministic dynamics. Therefore, we believe that these results will provide a useful starting point for understanding the generic spatial structure of spreading operators.

An operator \mathcal{O}_0 that is initially localized near the spatial origin (say, on a single site of a spin chain) will evolve under Heisenberg time evolution into a vastly more complicated operator $\mathcal{O}_0(t) = U^\dagger(t)\mathcal{O}_0U(t)$ that acts non-trivially on many sites. The “size” of $\mathcal{O}_0(t)$ is the size of the region in which $\mathcal{O}_0(t)$ fails to commute with a typical local operator Y_x at position x . This may be measured by

$$\mathcal{C}(x, t) \equiv \frac{1}{2} \text{Tr} \rho [\mathcal{O}_0(t), Y_x]^\dagger [\mathcal{O}_0(t), Y_x], \quad (1)$$

where the expectation value has been taken in an appropriate Gibbs state. (For our purposes, this will be taken to be the infinite-temperature Gibbs state ρ_∞ , which is the state to which random circuit dynamics equilibrate.) To make the connection with the OTOC, we may expand out the commutators in Eq. (1). For simplicity, let us assume, for the moment, that the operators \mathcal{O}_0 and Y_x are both Pauli-like operators squaring to the identity. We then have

$$\mathcal{C}(x, t) = 1 - \text{Tr} \rho_\infty \mathcal{O}_0(t) Y_x \mathcal{O}_0(t) Y_x. \quad (2)$$

The second term, in which the operators are not time ordered, is the OTOC. Other measures of the structure of an operator could be defined (for example, correlation functions involving higher powers of the commutator), but in this paper, we restrict ourselves to the OTOC.

At a given time t , the range of x where the commutator $\mathcal{C}(x, t)$ is significantly larger than zero gives a measure of the size of the operator. This region typically grows ballistically [58], even when local conserved quantities exhibit diffusive transport [31, 38, 49, 50, 59]. The immediate natural questions about $\mathcal{C}(x, t)$ include the following: What is the “butterfly” velocity v_B associated with this ballistic

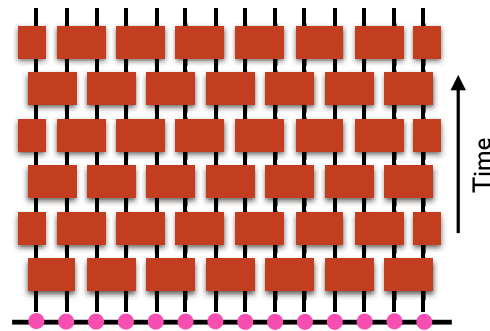


FIG. 1. Left: Random unitary circuit in 1 + 1D. Each brick represents an independently Haar-random unitary, acting on the Hilbert space of two adjacent “spins” of local Hilbert space dimension q .

growth? What is the spatial structure of $\mathcal{C}(x, t)$? Is there a “hydrodynamic” equation for $\mathcal{C}(x, t)$ at large time and distance scales? Are there important differences between 1 + 1D and higher dimensions? We will answer all of these questions for the case where the time evolution operator $U(t)$ is a circuit composed of Haar-random unitaries, as in Fig. 1.

We demonstrate that, both in 1D and in higher dimensions, operator spreading and the growth of the OTOC can be mapped to classical stochastic growth models. We show via an exact calculation that operator spreading in 1 + 1D can be understood in terms of an equation involving diffusion and drift. The front of the operator propagates at a finite velocity v_B . However, the front also *broadens* diffusively, so its width is proportional to \sqrt{t} (Fig. 2). We conjecture that this physics also occurs in generic (non-integrable) 1D systems undergoing deterministic Hamiltonian dynamics. For random circuit dynamics, we must also consider how the averaged correlator $\bar{\mathcal{C}}$ differs from the correlator \mathcal{C} within a given realization of the random circuit. We argue that fluctuations between realizations are small: Typical variations in the front position between

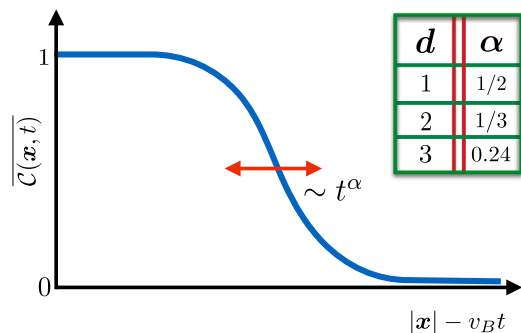


FIG. 2. Schematic behavior of the average OTOC: We find that the average OTOC $\bar{\mathcal{C}}(x, t)$ (where the average is over the local unitaries in the quantum circuit) has a front that broadens as t^α , with the indicated exponents in various spatial dimensions d .

different realizations are $O(t^{1/4})$, so they are negligible in comparison with the \sqrt{t} broadening of the front.

Turning to higher dimensions, we show by an exact mapping that there is a remarkable relationship with a classical droplet growth problem in the Kardar-Parisi-Zhang universality class [60]. (To avoid confusion, we note that this is *not* related to the connection between entanglement growth and KPZ introduced in Ref. [10].) We use this relationship to quantify the broadening of the front of a growing operator in a higher-dimensional random circuit. In $2 + 1D$, the front broadens like $t^{1/3}$ [60], and in $3 + 1D$, it broadens like $t^{0.240}$ [61,62]. In the two-dimensional case, and in the absence of lattice anisotropies, recent breakthroughs in the theory of interface growth [63–75] also translate to an exact form for the OTOC, in terms of the celebrated Tracy-Widom distribution (Fig. 3). The broadening of the front of the OTOC is summarized in Fig. 2. Again, we conjecture that these universal scaling forms extend to nonintegrable models with time-independent Hamiltonians, although we note that a previous calculation in a different setting has instead found a front that does not broaden with time and is governed by a traveling wave equation [31] (see also Refs. [37,38]). (A traveling wave equation arises from our mappings if we make a certain mean-field treatment, as in Appendix I. But this mean field is not valid in physical dimensionalities.)

In higher dimensions, the *shape* of the spreading operator [76] is also of interest. At first sight, one might expect the shape of the operator to be asymptotically spherical at late times. Instead, we argue that in systems with an underlying lattice, which have only discrete spatial symmetries, the spreading operator will not become spherical. Its asymptotic shape is determined by a model-specific velocity function $v_B(\hat{n})$, the speed of the front depending on the local normal vector \hat{n} . We verify this for random circuits by simulation in $2 + 1D$.

The results above are for random circuits composed of generic (Haar-random) unitary matrices. It is interesting to compare with random circuits composed of unitaries from the Clifford group, a discrete subgroup that leads to efficiently simulable dynamics [77,78]. In this case, the dynamics of the operator is much simpler [10], and randomness-induced fluctuations are much more severe.

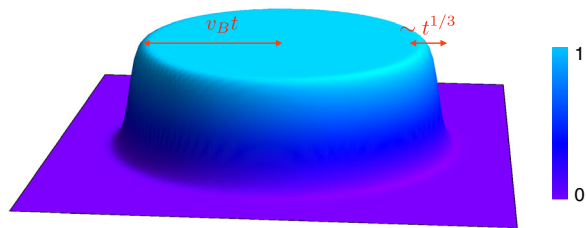


FIG. 3. Cartoon for the form proposed here for the OTOC in two spatial dimensions, when lattice anisotropy can be neglected. The functional form is given by the Tracy-Widom distribution F_2 .

But, remarkably, the results for the *averaged* OTOC \bar{C} coincide with the results for generic unitaries. This is a consequence of the fact that the Clifford group is a unitary 2-design [79].

In one dimension, we give a complementary exact calculation of the averaged OTOC, using a mapping to the partition function of a classical Ising model. These Ising degrees of freedom have a similar origin to those found in calculations in random tensor networks [80]. We show that the special structure arising from the unitarity of the quantum circuit means that this partition function is exactly calculable for any value of the local Hilbert space dimension.

Another important question is how the speed v_B associated with operator spreading relates to the speed v_E that can be associated with entanglement growth following a quench in 1D [2–4,6–8,10,11,81,82]. References [10,11] pointed out that, in general, v_E is smaller than v_B , unlike the situation in a 1D conformal field theory [3]. We extend this here, showing that arbitrarily small values of v_E/v_B can be achieved without any fine-tuning. (In a system with quenched, i.e., time-independent, disorder, it is even possible to have $v_B > 0$ but $v_E = 0$ [40].)

We also use the Ising mapping described above to give an exact calculation of the average entanglement purity [83–85] (the exponential of the second Renyi entropy) for a random circuit, complementing the scaling picture (in terms of a coarse-grained minimal cut) of Ref. [10].

II. OPERATOR DYNAMICS IN 1 + 1D

We begin by defining the random circuit dynamics, which we consider in $1 + 1D$, and by describing the “hydrodynamic” continuum picture we propose for the OTOC in $1 + 1D$. In Sec. IV, we give an alternative exact calculation of the OTOC, confirming and extending the results below.

A. Hydrodynamic equation for averaged OTOC

We consider time evolution by a quantum circuit on an infinite 1D spin chain, where each spin (qudit) has local Hilbert space dimension q . The structure of the quantum circuit is shown in Fig. 2(a). Two-site unitaries are applied to “even” bonds on even time steps and “odd” bonds on odd time steps (a “running bond” layout in the language of bricklaying). Each two-site unitary is drawn independently from the uniform distribution on the two-site unitary group $\mathbb{U}(q^2)$. Formally, our time evolution operator is $U(t) = U(t, t-1)U(t-1, t-2)\dots U(1, 0)$, where a single layer of the circuit is given by

$$U(t', t' - 1) = \begin{cases} \bigotimes_{x \in 2\mathbb{Z}} U_{x,x+1}(t', t' - 1) & \text{if } t' \text{ is even} \\ \bigotimes_{x \in 2\mathbb{Z}+1} U_{x,x+1}(t', t' - 1) & \text{if } t' \text{ is odd.} \end{cases} \quad (3)$$

Each two-site unitary $U_{x,x+1}(t', t' - 1)$ is Haar random and independent of all of the others.

Given an operator \mathcal{O} , we write $\mathcal{O}(t) = U(t)^\dagger \mathcal{O} U(t)$. We evaluate the following out-of-time-order correlator with respect to this time evolution:

$$\begin{aligned} \mathcal{C}(x, t) &\equiv \frac{1}{2} \text{Tr} \rho_\infty [X_0(t), Y_x]^\dagger [X_0(t), Y_x] \\ &= -\frac{1}{2} \text{Tr} \rho_\infty [X_0(t), Y_x]^2. \end{aligned} \quad (4)$$

Here, ρ_∞ is the infinite-temperature Gibbs state, i.e., the mixture of all possible spin configurations with equal weights. Note that X_0 is a Hermitian operator located at the origin of the spin chain, and Y_x is a Hermitian operator located at site x . We take both X and Y to be traceless and normalized such that $\text{Tr} X^2 = \text{Tr} Y^2 = q$. For example, if $q = 2$ (the spin-1/2 chain), we can take X and Y to be Pauli matrices at sites 0 and x , respectively.

Since the unitaries in the circuit are random, we must distinguish between averaged quantities (denoted by \mathbb{E}_U , or whenever unambiguous, by an overline $\overline{[\cdot \cdot]}$) and quantities within a given realization of randomness. However, we argue that fluctuations induced by the random circuit are small, meaning that the spatial profile of $\mathcal{C}(x, t)$ in a given realization of the circuit is, at large times, parametrically close to the average value $\overline{\mathcal{C}(x, t)}$.

Figure 2 is a schematic of the spatial profile we show for $\overline{\mathcal{C}(x, t)}$ at a fixed large time. The “size” of the operator is determined by a butterfly speed, which is

$$v_B(q) = \frac{q^2 - 1}{q^2 + 1}. \quad (5)$$

Within a region of size $\sim 2v_B(q)t$, the commutator $\mathcal{C}(x)$ has saturated to a value very close to unity. Note that for finite q , the butterfly velocity is smaller than the “naïve” speed limit of unity, which is set by the geometry of the quantum circuit, while in the limit $q \rightarrow \infty$, they coincide. The front of the operator, i.e. the region in which \mathcal{C} varies between 0 and 1, broadens diffusively. The width of the front is proportional to

$$\sigma(q, t) = \frac{2q}{q^2 + 1} \sqrt{t}. \quad (6)$$

More precisely, letting Φ denote the error function [the cumulative density function of the Gaussian distribution, $\Phi(y) = [1/(\sqrt{2\pi})] \int_{-\infty}^y e^{-x^2/2} dx$, which tends to zero for $y \ll 0$ and to 1 for $y \gg 0$], we have

$$\overline{\mathcal{C}(x, t)} \simeq \Phi\left(\frac{v_B t + x}{\sigma(t)}\right) \Phi\left(\frac{v_B t - x}{\sigma(t)}\right). \quad (7)$$

In Sec. IV, we see that Eq. (7) is the partition function of an Ising-like statistical mechanics problem, and we derive an exact formula on the lattice (without any continuum approximation):

$$\begin{aligned} \overline{\mathcal{C}(x, t)} &= (1 - \xi) g\left(t - 1, \frac{t - x - 3}{2}, p\right) g\left(t - 1, \frac{t + x - 3}{2}, p\right) \\ &\quad + \xi g\left(t - 1, \frac{t - x - 1}{2}, p\right) g\left(t - 1, \frac{t + x - 1}{2}, p\right), \end{aligned} \quad (8)$$

where

$$p = \frac{1}{q^2 + 1}, \quad \xi = \frac{q^4}{q^4 - 1},$$

and

$$g(n, a, p) = \sum_{k=0}^a \binom{n}{k} (1 - p)^{n-k} p^k.$$

Here, we show how Eq. (7) can be related to a continuum hydrodynamic equation, which is asymptotically accurate at large times. For x near the operator’s right-hand front, \mathcal{C} is related to a diffusing conserved density ρ :

$$\overline{\mathcal{C}(x, t)} = \int_x dx' \bar{\rho}(x', t), \quad (9)$$

$$\frac{\partial \bar{\rho}(x, t)}{\partial t} = v_B(q) \frac{\partial \bar{\rho}(x, t)}{\partial x} + D(q) \frac{\partial^2 \bar{\rho}(x, t)}{\partial x^2}. \quad (10)$$

We will explain the quantity ρ below.

To begin, focus on the spin-1/2 chain ($q = 2$). At time t , we may write the operator in the basis of products of Pauli matrices [11,22,23,58,82],

$$X_0(t) = \sum_{\mathcal{S}} a_{\mathcal{S}}(t) \mathcal{S}, \quad (11)$$

where $X_0(t=0)$ is a single-site operator. Here, the “string” \mathcal{S} can be any product of Pauli matrices on distinct sites. The number of strings in the sum generically grows exponentially with time (at the naive light-cone velocity, set by the geometry of the circuit). The \mathcal{S} are normalized as

$$\text{Tr} \rho_\infty \mathcal{S} \mathcal{S}' = \delta_{\mathcal{S} \mathcal{S}'}, \quad (12)$$

and X_0 is also normalized so $\text{Tr} \rho_\infty X_0^2 = 1$, implying

$$\sum_{\mathcal{S}} a_{\mathcal{S}}(t)^2 = 1. \quad (13)$$

It is also useful to introduce $\rho(x, t)$, the ‘‘fraction’’ of strings ending at x :

$$\rho(x, t) = \sum_{\substack{\text{strings } \mathcal{S} \\ \text{ending at } x}} a_{\mathcal{S}}(t)^2, \quad \sum_x \rho(x, t) = 1. \quad (14)$$

We observe that the OTOC is determined by $a_{\mathcal{S}}(t)^2$ as follows. Let Y_x be the Pauli matrix σ^y at site x . (This choice does not sacrifice generality because of the Haar randomness of the circuit.) Since distinct Pauli matrices anticommute, we see

$$\begin{aligned} [X_0(t), Y_x]^2 &= \left(\sum_{\mathcal{S}} a_{\mathcal{S}}(t) [\mathcal{S}, Y_x] \right)^2 \\ &= \left(\sum_{\mathcal{S}: \mathcal{S}_x = \sigma^y, \sigma^z} 2a_{\mathcal{S}}(t) \mathcal{S} Y_x \right)^2. \end{aligned} \quad (15)$$

Because of the orthonormality in Eq. (12), we have

$$C = -\frac{1}{2} \text{Tr} \rho_{\infty} [X_0(t), Y_x]^2 = \sum_{\mathcal{S}: \mathcal{S}_x = \sigma^y, \sigma^z} 2a_{\mathcal{S}}(t)^2. \quad (16)$$

This tells us that if we determine the evolution of $\overline{a_{\mathcal{S}}(t)^2}$, then the averaged OTOC is also determined. The dynamics of $\overline{a_{\mathcal{S}}(t)^2}$ turns out to be remarkably simple, as shown in Refs. [22,23]. It is best understood if we first consider a system of just two sites (rather than an infinite chain) over which a Haar-random unitary is applied at time t . It is straightforward to calculate (see Appendix A) that for arbitrary q ,

$$\overline{a_{\mathcal{S}'}(t+1)^2} = \sum_{\mathcal{S}} W_{\mathcal{S}'\mathcal{S}} \overline{a_{\mathcal{S}}(t)^2}, \quad (17)$$

where

$$W_{\mathcal{S}'\mathcal{S}} = \delta_{\mathcal{S}', I} \delta_{\mathcal{S}, I} + \frac{1}{q^4 - 1} (1 - \delta_{\mathcal{S}', I}) (1 - \delta_{\mathcal{S}, I}). \quad (18)$$

Note the following two features. First, the result is linear in $\overline{a_{\mathcal{S}}(t)^2}$. Second, \mathcal{S}' must be the identity if and only if \mathcal{S} is, but otherwise $\overline{a_{\mathcal{S}'}(t+1)^2}$ is a constant for all $\mathcal{S}' \neq I$. In other words, the random unitary introduces a (fictitious) Markov process on the probabilistic ensemble $\{(\mathcal{S}, \overline{a_{\mathcal{S}}(t)^2})\}$ of strings [22,23]. This Markov process describes a single string \mathcal{S} , which is stochastically updated over time. If \mathcal{S} is nontrivial, each update maps it to any nontrivial string, with uniform probability. The generalization to multiple spins is immediate: For each pair of spins that interact in a given time step, the stochastic update is applied to the corresponding two-site substring of \mathcal{S} . This Markov process will also be used in a higher-dimensional setting below, as it is not specific to the 1 + 1D setting. Note that the fictitious

stochastic dynamics, which involves a single evolving string, is entirely different from the stochastic dynamics of the operator $X_0(t)$ itself (which is a superposition of exponentially many strings).

Returning to the average of the OTOC, we realize that it only matters whether or not the string component of $X_0(t)$ at the site x is the identity. In the ensemble $\{(\mathcal{S}, \overline{a_{\mathcal{S}}(t)^2})\}$, the fraction

$$\mu(x, t) = \sum_{\mathcal{S}: \mathcal{S}_x \neq I} a_{\mathcal{S}}(t)^2 \quad (19)$$

of strings that occupy the site x may fail to commute with Y_x . There are $q^2 - 1$ possible nontrivial operators at the site, which are all equally probable in the ensemble of string components of $X_0(t)$. In the present case of $q = 2$ (see endnote for the general case [86]) this yields

$$\overline{C(x, t)} = \frac{q^2}{q^2 - 1} \overline{\mu(x, t)}. \quad (20)$$

In turn, the average occupation number $\mu(x, t)$ can be related to the end-point density ρ , assuming that x is far to the right of the left-hand front of the operator:

$$\overline{\mu(x, t)} = \mu_0 \sum_{x' \geq x} \overline{\rho(x', t)}, \quad \mu_0 = \frac{q^2 - 1}{q^2}. \quad (21)$$

The constant of proportionality μ_0 has been determined by assuming local equilibration of the structure of the strings [87]. Therefore,

$$\overline{C(x, t)} = \sum_{x' \geq x} \overline{\rho(x', t)}. \quad (22)$$

It is natural to conjecture that local equilibration of the strings, together with the exponentially large number of strings contributing to ρ , will make this identity asymptotically valid even without the average.

It remains to analyze the dynamics of $\overline{\rho(x, t)}$. The above Markov process implies a simple autonomous dynamics for $\overline{\rho}$:

$$\begin{aligned} \overline{\rho(t+1, x)} &= p \left[\overline{\rho(x, t)} + \overline{\rho(t, x+1)} \right], \\ \overline{\rho(t+1, x+1)} &= (1-p) \left[\overline{\rho(x, t)} + \overline{\rho(t, x+1)} \right], \end{aligned} \quad (23)$$

where

$$p = \frac{q^2 - 1}{q^4 - 1} = \frac{1}{q^2 + 1} \quad (24)$$

is calculated by counting the nonidentity two-site operators \mathcal{S} that have the identity at $x+1$, and the overline denotes averaging over unitaries applied up to a given time.

Recalling that unitaries are applied on even and odd bonds alternately, Eq. (23) gives a complete description of the dynamics of $\overline{\rho(x, t)}$. This is a lattice diffusion equation for the conserved density $\bar{\rho}$. Formally, $\bar{\rho}$ behaves like the probability density for a random walker who starts at the origin and who prefers to travel to the right since $p < \frac{1}{2}$. In the continuum (i.e., at long timescales), $\bar{\rho}$ satisfies a simple diffusion equation,

$$\partial_t \overline{\rho(x, t)} = v_B(q) \partial_x \overline{\rho(x, t)} + D(q) \partial_x^2 \overline{\rho(x, t)}, \quad (25)$$

whose drift and diffusion constants are determined in Appendix B:

$$v_B(q) = \frac{q^2 - 1}{q^2 + 1}, \quad D(q) = \frac{2q^2}{(q^2 + 1)^2}. \quad (26)$$

The peak in ρ corresponds to the front of the spreading operator $X_0(t)$. It travels at speed $v_B(q)$ and broadens as $\sigma(q, t)$ [Eq. (6)]. We emphasize that this fictitious random walker should not be thought of as the “end point” of the operator $X_0(t)$, which is a superposition of many strings with different end points.

From Eq. (22), or in the continuum

$$\overline{\mathcal{C}(x, t)} = \int_x dx' \overline{\rho(x', t)}, \quad (27)$$

we see that $\bar{\mathcal{C}}$ obeys the same equation as $\overline{\rho(x, t)}$ but with different boundary conditions,

$$\partial_t \overline{\mathcal{C}(x, t)} = v_B(q) \partial_x \overline{\mathcal{C}(x, t)} + D(q) \partial_x^2 \overline{\mathcal{C}(x, t)}. \quad (28)$$

Taking into account similar behavior at the left-hand front, we obtain Eq. (7).

Above, we make two (very natural) assumptions. One is that we can ignore the interaction between the left end and the right end, and the other is that the occupation density $\overline{\mu(x, t)}$ reaches its equilibrium value. In Sec. IV, we give an exact calculation of the averaged OTOC (including exact results for finite t and x , not necessarily large) without making any approximation.

B. Hydrodynamic description including fluctuations

Having determined the averaged OTOC, the key question is about the fluctuations between different realizations of the random circuits. From the point of view of exact results, this is a much harder question. (It is possible to obtain bounds in regions far from the front: We return to this in Sec. IV C.) However, we conjecture that the universal physics of fluctuations in $\rho(x, t)$ can be obtained by upgrading Eq. (25) to a *stochastic* diffusion equation for the random quantity $\rho(x, t)$. This description indicates that fluctuations are strongly suppressed at late times. Since the

diffusive broadening is present in a single realization (i.e., is not an artefact of disorder averaging), it is natural to conjecture that it will also be present in generic, non-random, 1D many-body systems.

The postulated stochastic diffusion equation may contain noise in both the diffusion constant and the drift, but we restrict to noise in the latter since it is more relevant in the renormalization group (RG) sense:

$$\partial_t \rho(x, t) = \partial_x (v_B + \eta(x, t)) \rho(x, t) + D \partial_x^2 \rho(x, t). \quad (29)$$

Here, $\eta(x, t)$ is white noise, uncorrelated in space and time.

The statistical properties of this equation are easy to analyze. In the absence of the noisy drift term, $\rho(x, t)$ forms a “wave packet” whose width grows like \sqrt{t} and whose center of mass is at $x_{\text{cm}} = v_B t$. When the noisy drift is turned on, it induces statistical fluctuations in x_{cm} whose magnitude scales with time as

$$\Delta x_{\text{cm}} \sim t^{1/4}. \quad (30)$$

A quick way to see this is to ask what the drift velocity has been in a given realization, *averaged* over the spacetime region visited by the wave packet. The wave packet visits a spacetime volume of order $\int^t dt' \sqrt{t'} \sim t^{3/2}$. Averaging the drift velocity $\eta(x, t)$ over this spacetime volume yields $v_{\text{av}} \sim t^{-3/4}$. The typical random displacement of the wave packet is thus of order $\Delta x_{\text{cm}} \sim v_{\text{av}} t \sim t^{1/4}$. A standard perturbative calculation in Appendix C reproduces this exponent 1/4, which also characterizes the spreading of directed waves in random media [88].

The quantity Δx_{cm} is parametrically smaller than \sqrt{t} , the width of the front of the averaged commutator. Therefore, this heuristic argument indicates that the front profile of the averaged OTOC also applies to the OTOC within a given instance of the random circuit. This is somewhat surprising. To see why, let us contrast the above Haar-random dynamics with Clifford dynamics for $q = 2$.

C. Comparison with Clifford circuit dynamics

The Clifford group is a discrete subgroup of the unitary group, defined by the property that any Pauli matrix is mapped to a product of Pauli matrices. When the quantum circuit consists of Clifford operators, an initial Pauli matrix remains a single string (rather than evolving into a superposition of exponentially many strings as for dynamics with generic unitaries), and at any given time, the end-point density $\rho(x, t)$ is localized on a single site.

However, uniformly random Clifford circuits have a crucial relationship with Haar-random circuits. Under a uniformly random Clifford update on a pair of sites, a nontrivial operator is mapped with equal probability to any of the nontrivial operators, and thus, the dynamics satisfies the master equation of the Markov process in Eq. (17) [10].

As a result, the *averaged* quantities, such as the average end-point density $\overline{\rho(x, t)}$, the average occupation number $\overline{\mu(x, t)}$, and, most importantly, the average OTOC $\overline{\mathcal{C}(x, t)}$, obey exactly the same dynamics as the Haar-random case. Formally, this is a consequence of the fact that random Clifford operators form a unitary 2-design [79]; see Appendix D for the definition of design and a proof for random Clifford operators. One may say that Clifford dynamics realizes the *a priori* fictitious Markov process in a physical system [89].

Despite the equivalence of averaged quantities, the quantities *within* a realization are entirely different. In the Clifford case, the end-point density ρ and the OTOC \mathcal{C} are strongly fluctuating, while we have argued that for generic unitaries, they are self-averaging (fluctuations are parametrically small).

III. HIGHER DIMENSIONS

We now address the structure of the out-of-time-order correlator $\mathcal{C}(x, t)$ in spatial dimensions greater than 1 by exploiting the relationship between the averaged OTOC and a fictitious classical Markov process (Sec. II A). We show that this process is a classical droplet growth problem whose universal physics can be understood in terms of the Kardar-Parisi-Zhang equation [60]. By taking appropriate averages, we then obtain exact universal exponents and scaling forms for the OTOC in a circuit composed of Haar-random unitary matrices. We conjecture that these scaling forms also apply to more realistic Hamiltonian dynamics in nonintegrable lattice models and field theories.

Somewhat surprisingly, we show that the “shape” of the spreading operator at late times does not become spherical, unless the microscopic dynamics has symmetry under continuous spatial rotations. In a lattice model, the spreading operator remembers forever that the lattice has only discrete point group symmetries. Our argument for this is not specific to random circuit dynamics. The point is simply that the butterfly velocity v_B , which sets the speed at which the operator’s front moves, generically depends on the front’s orientation, resulting in an anisotropic profile for the spreading operator at long times. Another surprising outcome, given previous work in the context of many-body perturbation theory including Ref. [31], is that, for the dynamics considered here, the averaged OTOC $\overline{\mathcal{C}}$ does not satisfy a local differential equation.

In 2 + 1D, when lattice anisotropy is absent (e.g., in an appropriate continuum model) or negligible, recent results in KPZ theory [63–75] yield the full functional form of $\overline{\mathcal{C}(x, t)}$ as a function of position and time. For an initially localized operator, this is expressed in terms of the Gaussian unitary ensemble (GUE) Tracy-Widom distribution [67,70] (which describes the extremal eigenvalue statistics for the GUE of Hermitian matrices [90,91]).

A. Higher dimensions: Setup and mapping to classical growth

We now describe the unitary dynamics for which we wish to study operator spreading and the OTOC. We choose a circuit where, in each time step, Haar-random two-site unitaries are applied to bonds of a d -dimensional cubic lattice in a manner that generalizes the 1 + 1D protocol. We describe the 2 + 1D case for concreteness; the generalization to higher dimensions is immediate. The periodicity of the circuit is four layers. Four successive layers cycle through the four columnar “dimer coverings” of the square lattice as shown schematically in Figs. 4 and 5, so the site at the origin interacts sequentially with its neighbors at $\mathbf{x} = (0, 1), (-1, 0), (0, -1), (1, 0)$.

We relate the dynamics of $\overline{\mathcal{C}}$ to a fictitious classical stochastic process for a growing cluster, described by the evolution of “occupation numbers”

$$n(\mathbf{x}) = 0 \quad \text{or} \quad 1 \quad (31)$$

at each site. These arise from the higher-dimensional generalization of the one-dimensional stochastic process for an evolving string \mathcal{S} described in Sec. II, in which we define the string’s occupation number $n(\mathbf{x})$ to be 1 if the string has support on site \mathbf{x} and zero otherwise. The effective stochastic process for these occupation numbers is simple (Appendix A). Consider two adjacent sites \mathbf{x} and \mathbf{y} that undergo a joint update in a given time step. If both sites are initially empty [$n(\mathbf{x}) = n(\mathbf{y}) = 0$], they remain so after the update. If at least one of the sites is initially occupied [$n(\mathbf{x}) = 1$ or $n(\mathbf{y}) = 1$ or both], then the configuration after the update can be $n(\mathbf{x}) = 1, n(\mathbf{y}) = 0$ with probability p , or $n(\mathbf{x}) = 0, n(\mathbf{y}) = 1$ with the same probability, or $n(\mathbf{x}) = n(\mathbf{y}) = 1$ with probability $1 - 2p$, where as before

$$p = \frac{1}{q^2 + 1}. \quad (32)$$

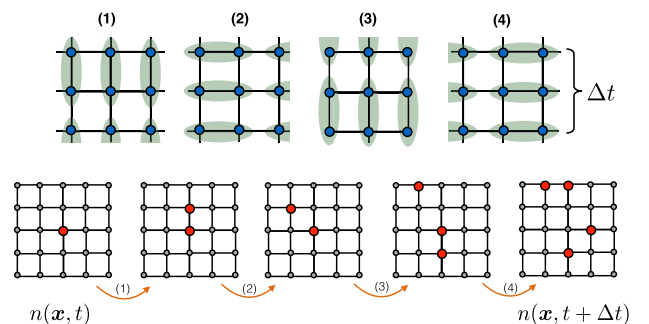


FIG. 4. Top row: 2 + 1D Haar-random quantum circuit. We consider unitary dynamics in which two-site Haar-random unitaries are applied on the bonds of a two-dimensional square lattice, in the columnar dimer configurations shown in (1)–(4). Bottom row: Allowed updates in the corresponding stochastic process.

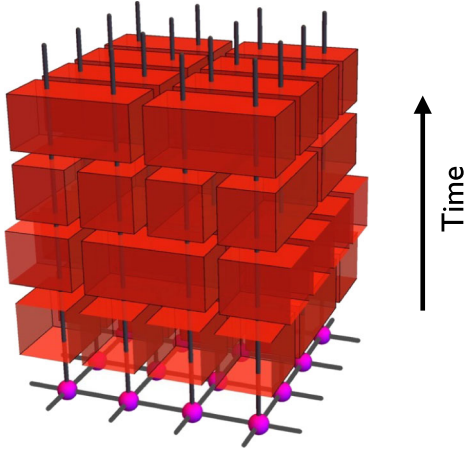


FIG. 5. Geometry of 2 + 1D Haar-random circuit.

If we consider the OTOC for a spreading operator that is initially localized at a single site, then the corresponding classical model is initialized with $n = 1$ at the origin and $n = 0$ everywhere else. A possible evolution of $n(\mathbf{x})$ in a single time step is shown in Fig. 4.

The Haar-averaged OTOC is related to the mean occupation number for this Markov process at time t by the relation

$$\overline{\mathcal{C}(\mathbf{x}, t)} = \frac{q^2}{q^2 - 1} \langle n(\mathbf{x}, t) \rangle_{\text{classical}} \quad (33)$$

as illustrated schematically in Fig. 6. The averages on the two sides of the above equation have different meanings. On the left, the bar denotes an average over realizations of a unitary circuit, and \mathcal{C} is a correlator for this quantum dynamics. On the right, the angle brackets denote an average in a classical stochastic process. The real number \mathcal{C} and the integer n are only related after averaging. (As we noted above, the fictitious Markov process can be

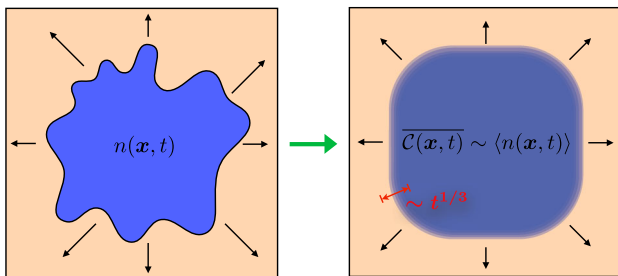


FIG. 6. Growth of a classical droplet and the OTOC: We relate the behavior of the OTOC (averaged over the unitaries in the circuit) to a classical stochastic process for the growth of a droplet in two spatial dimensions. A given configuration of the classical droplet is specified by a binary occupation number $n(\mathbf{x}, t)$ as shown on the left. Remarkably, the average droplet profile $\langle n(\mathbf{x}, t) \rangle$ precisely reproduces the averaged OTOC.

physically realized by random Clifford dynamics, whenever q is a prime power.)

B. Classical model in 2 + 1D: Analytical and numerical results

The seed at the origin grows to produce a cluster of linear size of order t . In the interior of this cluster, the state equilibrates rapidly to a state in which nearby sites are essentially uncorrelated, with average occupation $\langle n(\mathbf{x}, t) \rangle_{\text{classical}} = (q^2 - 1)/q^2$. In a given realization, there is an interface between the occupied and unoccupied regions which is sharp on length scales of the order of the lattice spacing. The evolution of the droplet is very similar to well-studied growth models such as the Eden model [92], and it reduces to the stochastic growth of this one-dimensional interface. The size of the occupied region grows linearly in time, with statistical fluctuations in the shape of the interface. (The average shape in our 2D model is not circular but has only fourfold rotational symmetry; we discuss this in Sec. III E.)

Typically, such interface growth processes are in the universality class of the KPZ equation [60]. Consider a section of the interface, with ξ a coordinate parallel to the interface and h its height in the perpendicular direction. The KPZ equation is

$$\partial_t h = c + \nu \partial_\xi^2 h + \frac{\lambda}{2} (\partial_\xi h)^2 + \zeta(x, t), \quad (34)$$

where ζ is uncorrelated spatiotemporal noise. The constant c contributes to the average normal growth rate for the interface, while the ν term describes diffusive smoothing of sharp features. Finally, the nonlinear λ term encodes the dependence of average growth rate on the slope. (In the present setting, lattice symmetry also allows a term $\propto \partial_\xi h$ if the ξ axis is at a generic angle to the lattice axes. This term may be removed by a coordinate boost.) The KPZ equation renormalizes to a nontrivial fixed point. One of its most basic properties is the fact that the fluctuations in the height at a given position ξ grow with time as t^β , with $\beta = 1/3$.

The KPZ description above is justified on RG grounds, assuming that the relevant length scales for the fluctuations of the classical interface are much larger than the lattice spacing. (See Ref. [93] for a review of lattice growth models.) We show numerically that KPZ scaling does indeed apply to the growing droplet, so long as the front is not locally parallel to the lattice axes, where—as a result of the discrete spacetime geometry we have chosen—an additional effect can come into play. (We expect this additional subtlety to be absent in models in continuous time.)

Let us write the shape of the droplet as a parametrized curve in polar coordinates, with $R(\theta)$ the radius at angle θ from the origin. [As mentioned above, the interface is sharp on an $O(1)$ length scale, and therefore, $R(\theta)$ is well defined

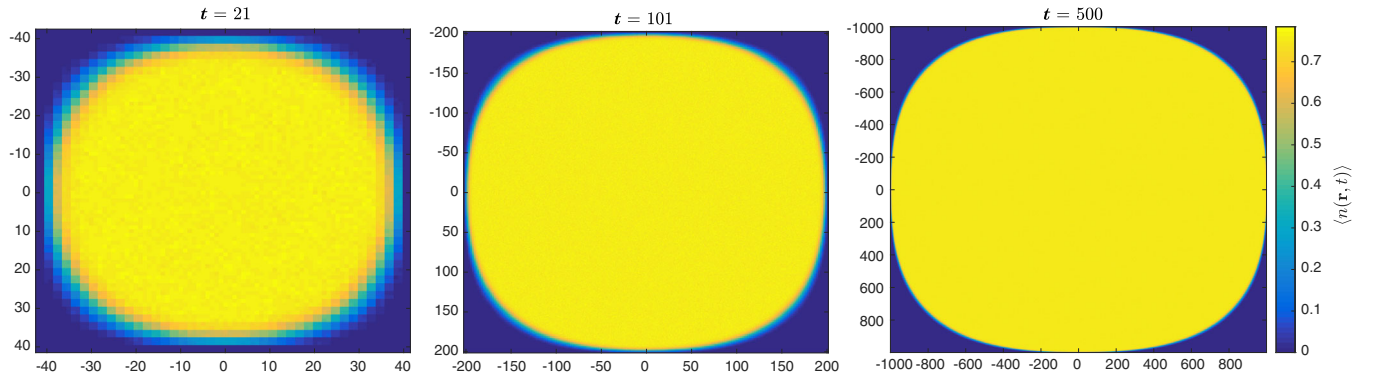


FIG. 7. Growth of a 2D cluster ($q = 2$): We determine the behavior of the averaged OTOC by simulating the stochastic growth of a two-dimensional cluster over $M = 2 \times 10^3$ realizations, with local updates applied at each time step, as described in the text. The average occupation number for the cluster $\langle n(\mathbf{x}, t) \rangle$ is shown for the indicated times in the evolution as it approaches its asymptotic shape.

up to an $O(1)$ uncertainty; this is sufficient since the properties we discuss below are on parametrically larger length scales when t is large.] From KPZ scaling, we would expect

$$\langle R(\theta, t) \rangle = r(\theta)t - A(\theta)t^\beta + \dots, \quad (35)$$

$$\sqrt{\langle R(\theta, t)^2 \rangle - \langle R(\theta, t) \rangle^2} = C(\theta)t^\beta + \dots, \quad (36)$$

with the exactly known exponent $\beta = 1/3$. We will discuss the nonuniversal function $r(\theta)$ below and in Sec. III E, and we will discuss more detailed universal properties in the next section.

We have examined the growth of the droplet for spin-1/2 degrees of freedom ($q = 2$) on the square lattice by tracking the average, evolving support of a cluster over $M = 2 \times 10^3$ realizations of the classical dynamics up to time $t = 1000$ (see Fig. 7). We store only the density $\langle n(\mathbf{x}, t) \rangle$, averaged over all M realizations, as a function of position and time, as this is the quantity with a direct interpretation in the quantum setting. We have also investigated smaller values of q : These do not have an interpretation in the quantum circuit, but in the classical model, decreasing q simply corresponds to increasing the probability p in the update. Taking $q < 2$ is a convenient way to explore two regimes (discussed below) that differ in the properties of a lattice-aligned front; both regimes could also be accessed at fixed q , say $q = 2$, by varying other parameters in the random circuit such as the circuit geometry.

At each time slice, the form of $\langle n(\mathbf{x}, t) \rangle$ is fitted, along cuts through lattice symmetry axes, to extract the cluster size and the width of the front [where $\langle n(\mathbf{x}, t) \rangle$ is appreciably different both from zero and from its $t \rightarrow \infty$ value]. We observe linear growth of the size, as expected. Note that

the fluctuations in the second equation of (35) imply that the width of the front region is expected to scale like $t^{1/3}$.

Figure 8 (top panel) shows the growing width of the front for cuts along the diagonal, $\theta = \pm\pi/4$. There, at the largest times we can access, the fitted exponent is $\beta = 0.3305 \pm 0.0269$, extracted from a fit to the blue data points in Fig. 8. As expected, this value is consistent with the KPZ value $\beta = 1/3$.

A slight surprise is that the behavior along the axis, e.g., at $\theta = 0$, is rather different: See Fig. 8 (bottom panel), which does not show KPZ growth. Generically, the only stable fixed point for the growth of a 1D interface is believed to be the KPZ fixed point. However, anomalous growth is possible in this model, for q greater than a critical value $q_c \lesssim 2$, when the direction of the front's local normal vector is fine-tuned to coincide with one of the axes, as occurs at $\theta = 0$. In this regime, a front with normal parallel to a lattice axis moves at a speed exactly equal to the naive light-cone speed, $v_B = 2$, and does not roughen. This is a known phenomenon in various lattice growth models in discrete time, which have synchronous parallel updates and can be understood by a relationship with directed percolation [94–98]: See Appendix E for an explanation. While interesting, this phenomenon is an artefact of the specific discrete spacetime geometry we have chosen, which could be eliminated by modifying this geometry, [99] and we certainly do not expect it to be relevant to continuous time dynamics. (It would be interesting to look for this effect in appropriate deterministic Floquet dynamics, however.) This phenomenon has an effect on the shape of the droplet, leading to flat “facets” on the front close to $\theta = 0, \pi/2, \pi$, and $3\pi/2$ (Sec. III E).

We now discuss the OTOC scaling that results from the KPZ mapping, neglecting effects of lattice anisotropy (which we will return to in Sec. III E).

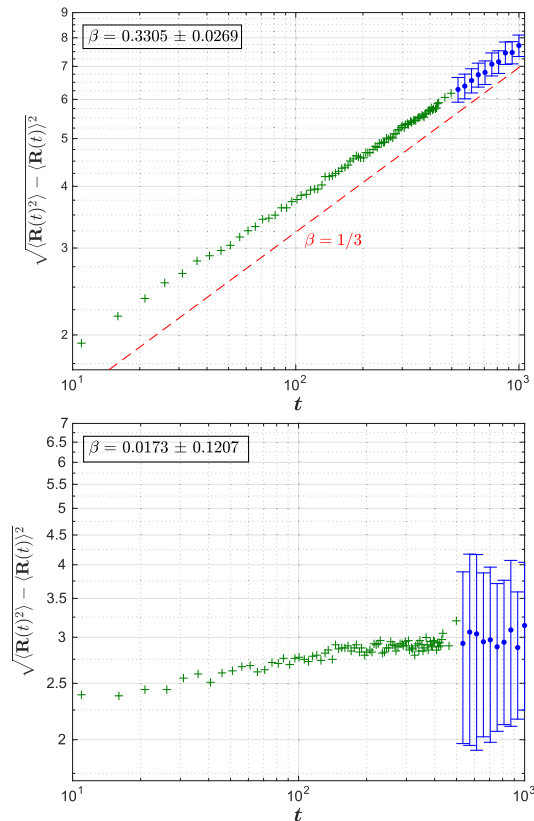


FIG. 8. Fluctuation exponent β : We fit the profile of the evolving droplet for $q = 2$ along the $\theta = \pm\pi/4$ directions (top panel) to extract the mean operator size and magnitude of the fluctuations about the mean. The fluctuations exhibit power-law growth with exponent $\beta = 0.3305 \pm 0.0269$, consistent with the KPZ value $\beta = 1/3$. When fitting the profile along $\theta = 0$ (bottom panel), we observe no appreciable growth of the fluctuations; we argue in Appendix E that this occurs for sufficiently large q when the front's local normal vector is precisely aligned with a lattice axis (as a result of the specific circuit geometry).

C. Scaling of the OTOC in 2+1D

We have already mentioned the basic consequence of KPZ growth, which is the $t^{1/3}$ broadening of the front. But, unusually for a nontrivial fixed point, not only the exact critical exponents but also certain exact scaling functions are known for the growth of an interface in 1D [63–75] (see Refs. [100–103] for reviews). We can now apply this information to the OTOC to obtain scaling functions that we propose are generic.

To simplify things, let us consider a case where lattice anisotropy is absent or very weak so that the spreading operator is circular and the OTOC depends only on a radial coordinate and time. Weak anisotropy could certainly be engineered in an appropriate random circuit. More importantly, we conjecture that the scaling form below captures universal scaling in realistic, rotationally invariant, many-body systems and field theories.

For the growth of a droplet, the probability distribution of the interface radius is given by the GUE Tracy-Widom distribution [67,70] (which has been observed experimentally in striking experiments on the growth of a turbulent domain in liquid crystals [104,105]). Following convention, we write

$$R(\theta, t) = v_B t + ct^{1/3}\chi(\theta, t), \quad (37)$$

where the nonuniversal constants v_B and c are of order one, and $\chi(\theta, t)$ is a random variable whose mean and variance are of order one at large times. Focusing on a fixed value of θ , the cumulative probability distribution of χ at a fixed time is t independent at large times and given by the Tracy-Widom distribution F_2 :

$$\Pr(\chi < s) = F_2(s). \quad (38)$$

Remarkably, this allows us to fix the full functional form of $\overline{\mathcal{C}(x, t)}$ in two dimensions, in the case where lattice anisotropy is absent. In polar coordinates (r, θ) , and in the continuum, Eq. (33) is

$$\overline{\mathcal{C}(r, \theta, t)} = \langle \Theta[R(\theta, t) - r] \rangle_{\text{classical}}, \quad (39)$$

where Θ is the Heavyside step function. The right-hand side is precisely the probability that χ is greater than $(r - v_B t)/ct^{1/3}$. We suppress the θ dependence since we are assuming rotational symmetry:

$$\overline{\mathcal{C}(r, t)} = 1 - F_2\left(\frac{r - v_B t}{ct^{1/3}}\right). \quad (40)$$

The form of $\overline{\mathcal{C}(r, t)}$ is shown in Fig. 9. The asymptotic behavior near the trailing edge, close to saturation (i.e., for $[v_B t - r]/ct^{1/3} \gg 1$), is [106,107]

$$\overline{\mathcal{C}(r, t)} = 1 - b_1 \frac{c^{1/8} t^{1/24}}{|r - v_B t|^{1/8}} \exp\left[\frac{(r - v_B t)^3}{12c^3 t}\right] + \dots, \quad (41)$$

where $b_1 = 2^{1/24} e^{\zeta'(-1)}$ with $\zeta'(-1) \approx -0.165$, the derivative of the Riemann zeta function. Near the leading edge, $[r - v_B t]/ct^{1/3} \gg 1$,

$$\overline{\mathcal{C}(r, t)} = \frac{c^{3/2} t^{1/2}}{16\pi(r - v_B t)^{3/2}} \exp\left[-\frac{4(r - v_B t)^{3/2}}{3c^{3/2} t^{1/2}}\right] + \dots. \quad (42)$$

The former asymptotic expansion of F_2 was achieved only recently [106,107].

One can also consider operator spreading with other initial conditions. For example, we can initialize an operator in a half-plane so that $\mathcal{C}(x, t)$ has a straight, rather than a circular, front. The scaling form for $\mathcal{C}(x, t)$ will then

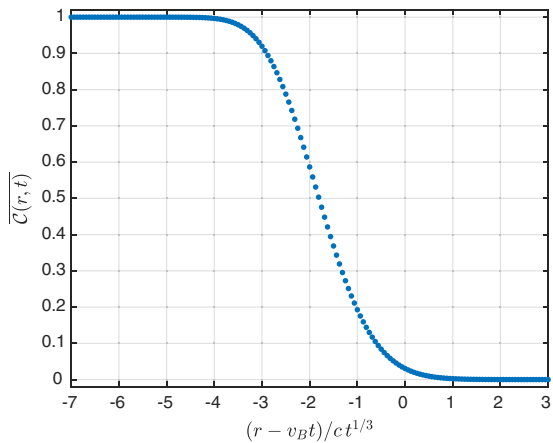


FIG. 9. The OTOC in $(2+1)D$: Plot of the front of the averaged OTOC $\overline{C}(r, t)$ in two spatial dimensions and in the absence of lattice anisotropy, as determined from the exact expression in terms of the Tracy-Widom distribution in the main text.

be given by the Tracy-Widom distribution of the Gaussian orthogonal ensemble, denoted F_1 . The objects F_1 and F_2 are of fundamental importance in a broad range of mathematical and physical problems, and it would be very interesting to see whether any of these connections shed light on operator growth.

It should be noted that the mapping to a classical growth process described above gives exact results for the averaged OTOC \overline{C} but does not immediately give access to the fluctuations in C between different realizations of the circuit. [These circuit-to-circuit fluctuations in the quantum expectation value should not be confused with the stochastic fluctuations of $n(\mathbf{x})$ in the auxiliary classical model.] A calculation of, say, $\overline{C^2} - \overline{C}^2$ would need to go beyond the mapping provided here for \overline{C} . Nevertheless, we conjecture that, by analogy with the 1D case described in Sec. II B, circuit-to-circuit fluctuations in C are parametrically small at late times whenever the front broadens with time.

D. Scaling of the OTOC in 3+1D and above

The basic features of the 3+1D case are very similar to those in 2+1D. The KPZ equation extends to an interface of arbitrary dimensionality [60]. For the 3+1D quantum problem, the dimensionality of the interface is two, and the critical exponent β relevant to the width is $\beta \simeq 0.240$ (Ref. [61] and references therein). The analogue of F_2 that yields the universal form of \overline{C} is not known analytically, but it has been determined numerically [103]. Numerical simulations for the 3+1D random circuit, along the lines of those above, would be feasible.

Dimensions equal to or higher than 4+1 are, of course, inaccessible experimentally, but they are nonetheless interesting because, in these high dimensionalities, the KPZ equation yields a phase transition as a function of the

strength of nonlinearity [60]. Both a rough phase, in which fluctuations grow as t^β with $\beta > 0$, and a smooth phase, where fluctuations remain of order one as $t \rightarrow \infty$, exist. It would be interesting to know whether both phases are accessible in appropriate many-body systems.

E. Shape of the operator at late times

It is interesting to consider the shape of the spreading operator at late times—i.e., the shape of the growing spatial region in which the OTOC $\overline{C}(x, t)$ has already saturated to its late-time value (to within an exponentially small correction). Rescaling distances by a factor of t^{-1} gives a droplet of $O(1)$ size, which we expect to reach a fixed asymptotic shape. [108] What is this shape?

At first glance, one might expect that, asymptotically, the front is a circle in two spatial dimensions and a sphere in higher dimensions. For example, this would be expected if the OTOC satisfied a local nonlinear differential equation in which derivatives higher than 2 could be neglected, as discrete lattice symmetries would ensure that such an equation had symmetry under continuous spatial rotations. Instead, we argue that, for many-body systems on the lattice, the shape of the operator is model dependent and retains information about the discrete symmetries of the lattice, even at arbitrarily late times. For the random circuit model, this follows immediately from the mapping to domain growth processes, for which anisotropy is a well-known feature [109–113]. Figure 10 shows the shape of the droplet in the present model for various values of q .

For concreteness, consider the 2D case (similar statements hold in higher dimensions). The asymptotic droplet shape is described by a radius $R(\theta) = h(\theta)t$ depending on the polar coordinate θ . Since the size of the operator is large at large times, the curvature of the front is parametrically small, except possibly at isolated θ values where $h(\theta)$ is not smooth. Away from such isolated points, the local *velocity* of the front, in the direction of its normal vector, can depend only on the orientation of this local normal vector. This dependence is captured by a velocity function $v_B(\phi)$, where $\phi = \phi(\theta)$ is the angle of the local normal vector to the x axis. *A priori* $v_B(\phi)$ is constrained only by lattice symmetry; for example, on the square lattice,

$$v_B(\phi) = v_0 + v_1 \cos 4\phi + v_2 \cos 8\phi + \dots \quad (43)$$

It is evident that the asymptotic shape cannot be a circle except when $v_B(\phi)$ is a constant function. Since the front of the operator advances by $v_B(\phi)dt$ in the normal direction \hat{n} , the distance between the front and the origin grows by $v_B(\hat{n})dt/\hat{n} \cdot \hat{r}$, which must be equal to $h(\hat{r})dt$. Expressing the normal vector in terms of h , one obtains

$$v_B(\phi(\theta)) = \frac{h(\theta)^2}{\sqrt{h(\theta)^2 + (\partial_\theta h(\theta))^2}}, \quad (44)$$

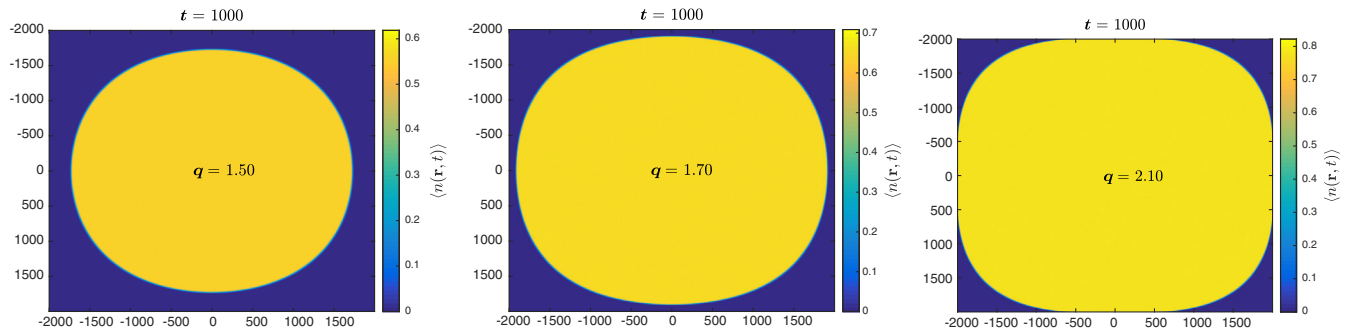


FIG. 10. “Faceting” of the cluster: Shown are the cluster shapes at fixed time $t = 10^3$, for the indicated values of q . When q is sufficiently large (third panel), the cluster develops facets along the $\theta = 0, \pi$ directions, where the normal growth speed is the maximum possible given the circuit geometry. The region shown is the naive light cone.

where $\phi(\theta)$ is the angle of the normal at polar position θ on the interface. This equation is solved by a geometrical construction described in Refs. [112,113]: $h(\theta) = \min_{\phi} \{ [v_B(\phi)] / [\cos(\phi - \theta)] \}$. When the effect of lattice anisotropy is weak (as is likely to be the case in many realistic situations when the relevant degrees of freedom are long-wavelength modes), we expect $v_B(\phi)$ to be a smooth, weakly varying function, and we may also solve for the shape perturbatively in $w(\phi) = v'_B(\phi)/v_B(\phi)$, as described in Appendix F. Restoring the time dependence, we obtain

$$R(\theta) = v_B(\theta)t \exp \left(-\frac{1}{2}w(\theta)^2 + \frac{1}{6}\partial_{\theta}w(\theta)^3 + \dots \right). \quad (45)$$

However, when $v_B(\phi)$ varies sufficiently strongly, the asymptotic shape $R(\theta)$ can include sharp corners or straight segments on the boundary: In this regime, the perturbative solution above is no longer appropriate.

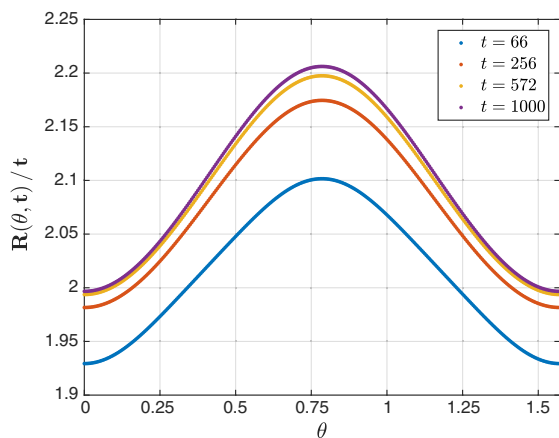


FIG. 11. Anisotropy in the cluster profile: Numerically determined anisotropy in the average shape of the 2D cluster $R(\theta, t)/t$, at the indicated times. The anisotropy in the cluster shape grows in time and appears to asymptote to a nontrivial steady-state shape.

For many-body systems in continuous time, we expect $v_B(\phi)$ to be analytic. In the present lattice model, $v_B(\phi)$ is analytic for $q < q_c$ ($q_c \lesssim 2$), while for $q > q_c$, this function is nonanalytic near $\phi = 0$ as a result of the anomalous behavior of a lattice-aligned front: $v_B(\phi) \simeq 2 + \text{const}|\phi|$ [98]. This leads to flat facets near $\theta = 0$ in the asymptotic shape [98]. This change in the surface morphology as q is varied is shown in Fig. 10.

For the random circuit model, it is straightforward to determine $v_B(\theta)$ in the extreme limit $q = \infty$, where growth becomes deterministic. The propagation of the front in this limit is similar to that of the “next-nearest-neighbor” deterministic Eden model introduced in Ref. [114], and it has the same nonanalytic angular dependence of the velocity [115] [114]:

$$v_B^{q=\infty}(\phi) = 2(|\cos \phi| + |\sin \phi|). \quad (46)$$

In this limit, the growing operator is simply a square.

Figure 11 shows the angular dependence of the radius for the 2 + 1D random circuit dynamics at $q = 2$, for several values of the time, showing a clear anisotropy. Note also that $R(\theta = 0) \rightarrow 2t$ at late times: The front has a small facet near $\theta = 0$ that moves at the naive light-cone speed for propagation along an axis direction. In the light of our 1D results where, for finite q , v_B is always less than the speed associated with the naive light cone, it is remarkable that, in a higher-dimensional circuit, it is possible for the OTOC front to propagate at the maximal speed in some directions. However, we emphasize that this effect relies on the specific discrete spacetime geometry.

F. Formal viewpoint

Before returning to 1D, we restate the higher-dimensional results of Sec. III A in a more formal language, which parallels our discussion in 1D. We introduce a density on clusters, C , where C is a collection of sites:

$$\rho(C) = \sum_{S: \text{supp}(S)=C} a_S^2, \quad \sum_C \rho(C) = 1. \quad (47)$$

Here, $\text{supp}(\mathcal{S})$ is the support of \mathcal{S} . After coarse graining, we can represent C by a closed surface of spherical topology, namely, the boundary of the coarse-grained cluster. Therefore, $\rho(C)$ is the natural analogue of the ‘‘end-point density’’ $\rho(x)$ in 1D. The surface growth picture implies that the effective dynamics of $\rho(C)$ are the dynamics of the probability distribution of a growing interface. Therefore, when this is KPZ, $\rho(C)$ satisfies the Fokker-Planck equation corresponding to the KPZ equation. We will discuss this further in a separate publication.

IV. EXACT CALCULATION OF OTOC IN THE SPACETIME PICTURE

We now give an analytical treatment of the OTOC from a spacetime point of view. This leads to connections with domain walls in an effective Ising model. Similar Ising degrees of freedom have appeared in work on random tensor networks [80]. Here, the effective Ising model looks complicated at first sight, but it turns out to be much simpler than those encountered in random (nonunitary) tensor networks because of the special structure arising from unitarity. We arrive at a problem of two nonintersecting directed walks, shown schematically in Fig. 12.

This spacetime picture may be much more generalizable than the dynamical point of view above. In Sec. V, we use it to calculate the entanglement purity. In the future, we hope that the tools introduced in this section will be generalizable to higher moments of the OTOC which capture fluctuations ($\overline{C^2}$, etc.), or higher powers of the commutator, or to a direct calculation of the von Neumann entropy. We note that the OTOC is not the only tool for characterizing operator spreading: The operator entanglement entropy of a spreading operator is also of interest [116,117]. Scaling forms for this quantity will be discussed in Ref. [118].

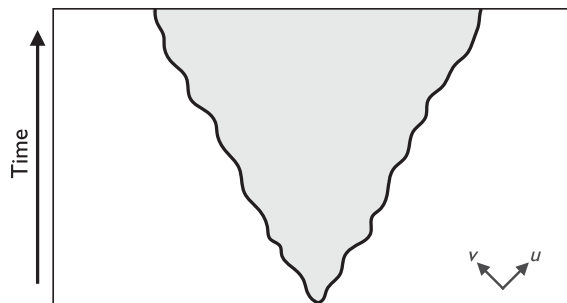


FIG. 12. The calculation of the OTOC in the spacetime picture leads to a partition function for two nonintersecting directed random walks (domain walls in an effective Ising model). These walks have a bulk energy cost [Eq. (69)] and a boundary potential [Eq. (70)] that biases them to a positive and negative velocity, respectively. Directions of null coordinates u, v used in the text are indicated.

Our exact result for the OTOC for arbitrary x and t (not necessarily large) is given in Eq. (8). There are no approximations in this equation. Approximating g by the cumulative density function $\Phi(y) = (1/\sqrt{2\pi}) \int_{-\infty}^y e^{-x^2/2} dx$ of the Gaussian distribution, we reproduce Eqs. (5)–(7) above. This approximation is valid when t is large.

Although the spin chain is spatially infinite in both directions, as is our quantum circuit, the time-evolved operator $U^\dagger(t)X_0U(t)$ is supported only on the interval $[-t, t-1]$ of length $2t$. Therefore, it suffices to consider an observable Y inserted in this interval, and our correlator becomes the trace of a $q^{2t} \times q^{2t}$ matrix. The infinite-temperature Gibbs state reduces to the identity matrix divided by the dimension q^{2t} . Expanding the commutator, we see

$$\begin{aligned} & -\frac{1}{2} \text{Tr} \rho_\infty ([U^\dagger(t)X_0U(t), Y_x])^2 \\ & = q^{-2t} (\text{Tr}[U^\dagger(t)X_0^2U(t)Y_x^2] - \text{Tr}[X_0(t)Y_x X_0(t)Y_x]) \\ & =: q^{-2t} \text{Tr}[U^\dagger(t)X_0^2U(t)Y_x^2] - F. \end{aligned} \quad (48)$$

The Haar average of the first term is easy to evaluate. The observable X_0^2 is conjugated by a unitary $U_{-1,0}(1,0)$, and after taking the Haar average, it becomes proportional to the identity. The constant of proportionality is fixed by the trace-preserving condition. By the normalization convention, $\text{Tr}X_0^2 = q = \text{Tr}Y_x^2$, and therefore, the Haar average of the first term is equal to $q^{-2t} \text{Tr}I = 1$. The second term F contains all the complexity.

Observe that the local unitaries form a square lattice that is rotated by 45° . It is thus natural to introduce null coordinates as

$$u := (t+x+1)/2, \quad v := (t-x+1)/2, \quad (49)$$

$$t = u+v-1, \quad x = u-v. \quad (50)$$

Because of the cyclic property of the trace, the only unitaries in the circuit that could affect the correlator are those in the intersection (a rectangle) of the future light cone of X_0 and the past light cone of Y_x . From now on, let us use ℓ_u and ℓ_v to denote the linear sizes of this intersection along the u and v directions, respectively. There are $\ell_u \ell_v$ local unitaries contained in the intersection of the light cones.

A. Reduction to Ising spins

For each local unitary U , the expression F contains two U 's and two U^\dagger 's. We will see that averaging over the local unitaries allows us to express F as a partition function for a set of classical Ising spins. To see why such Ising spins arise, consider the standard expression for the Haar average of a single unitary matrix in $\mathbb{U}(n)$:

$$\begin{aligned} \mathbb{E}_{U \in \mathbb{U}(n)} U_{a'a} U_{b'b}^* U_{c'c} U_{d'd}^* & \langle \uparrow | \downarrow \rangle = \frac{1}{q}. \quad (58) \\ &= \frac{1}{n^2 - 1} \left[\delta_{a'b'} \delta_{c'd'} \times \delta_{ab} \delta_{cd} + \delta_{a'd'} \delta_{b'c'} \times \delta_{ad} \delta_{bc} \right. \\ & \quad \left. - \frac{1}{n} (\delta_{a'b'} \delta_{c'd'} \times \delta_{ad} \delta_{bc} + \delta_{a'd'} \delta_{b'c'} \times \delta_{ab} \delta_{cd}) \right]. \quad (51) \end{aligned}$$

(See Appendix G for a self-contained derivation of this formula.) It is convenient to regard the above expression as a matrix whose rows are labeled by the multi-index (a', b', c', d') and whose columns are labeled by (a, b, c, d) . Note that two types of contractions appear for the unprimed indices, namely, $\delta_{ab} \delta_{cd}$ and $\delta_{ad} \delta_{bc}$, and similarly for the primed ones. Correspondingly, in bra-ket notation, the above matrix can be written in terms of two vectors, which we denote $|I_\uparrow\rangle$ and $|I_\downarrow\rangle$ (the reason for the notation will become clear below):

$$\begin{aligned} \overline{U \otimes U^* \otimes U \otimes U^*} & \\ &= \frac{n^2}{n^2 - 1} \left((|I_\uparrow\rangle\langle I_\uparrow| + |I_\downarrow\rangle\langle I_\downarrow|) - \frac{1}{n} (|I_\uparrow\rangle\langle I_\downarrow| + |I_\downarrow\rangle\langle I_\uparrow|) \right). \quad (52) \end{aligned}$$

In the natural basis, these vectors are

$$\langle abcd | I_\uparrow \rangle = \frac{1}{n} \delta_{ab} \delta_{cd}, \quad \langle abcd | I_\downarrow \rangle = \frac{1}{n} \delta_{ad} \delta_{cb}. \quad (53)$$

With this definition, we may write Eq. (52) as

$$\overline{U \otimes U^* \otimes U \otimes U^*} = \frac{n^2}{n^2 - 1} \sum_{s, s' = \uparrow, \downarrow} w(s, s') |I_s\rangle\langle I_{s'}|, \quad (54)$$

with

$$w(s, s') = \begin{cases} 1 & \text{if } s = s' \\ -\frac{1}{n} & \text{if } s \neq s'. \end{cases} \quad (55)$$

We see that the unitary may be associated with a *pair* of classical Ising “spins” s and s' .

For the application of interest to us, the unitaries are two-site unitaries acting on the q^2 -dimensional space associated with spins i and $i + 1$. In this case, it is easy to see that

$$|I_\uparrow\rangle = |\uparrow\rangle_i |\uparrow\rangle_{i+1}, \quad |I_\downarrow\rangle = |\downarrow\rangle_i |\downarrow\rangle_{i+1}, \quad (56)$$

with

$$\langle \alpha\beta\gamma\delta | \uparrow \rangle = \frac{1}{q} \delta_{\alpha\beta} \delta_{\gamma\delta}, \quad \langle \alpha\beta\gamma\delta | \downarrow \rangle = \frac{1}{q} \delta_{\alpha\delta} \delta_{\gamma\beta}, \quad (57)$$

and now α, \dots, δ run over the q basis vectors associated with a given spin. The vectors $|\uparrow\rangle$ and $|\downarrow\rangle$ have norm 1 and satisfy

When we consider F , the spins arising from each unitary in the circuit will form an interacting network. The interactions between spins from the same unitary will be given by $w(s, s')$, while the interactions between spins from different unitaries will arise from the inner products of kets $|\uparrow, \downarrow\rangle_i$ associated with a given spin. As a result of the second line of Eq. (55), the weight of an Ising configuration may be negative. However, we will see that the weights simplify when we integrate out half of the spins. This leads to a simpler Ising partition function, which can be treated exactly in terms of domain walls.

As a final piece of notation, we generalize the definition of $|I_\uparrow\rangle$ and $|I_\downarrow\rangle$. Given an operator \mathcal{O}_{ab} on the n -dimensional space, we define n^4 -dimensional vectors $|\mathcal{O}_\uparrow\rangle$ and $|\mathcal{O}_\downarrow\rangle$ via

$$\langle abcd | \mathcal{O}_\uparrow \rangle = \frac{\mathcal{O}_{ab} \mathcal{O}_{cd}}{\text{Tr} \mathcal{O} \mathcal{O}^\dagger}, \quad \langle abcd | \mathcal{O}_\downarrow \rangle = \frac{\mathcal{O}_{ad} \mathcal{O}_{cb}}{\text{Tr} \mathcal{O} \mathcal{O}^\dagger}. \quad (59)$$

Choosing \mathcal{O} to be the identity gives the vectors $|I_{\uparrow, \downarrow}\rangle$.

Before we evaluate F for arbitrary ℓ_u and ℓ_v , let us consider the simplest case, where $\ell_u = \ell_v = 1$.

$$\begin{aligned} \bar{F} &= q^{-2} \mathbb{E}_U \text{Tr} U X U^\dagger Y U X U^\dagger Y \\ &= q^{-2} \mathbb{E}_U \sum_{i_1, \dots, i_8=1}^{q^2} U_{i_1 i_2} X'_{i_2 i_3} U_{i_4 i_3}^* Y'_{i_4 i_5} U_{i_5 i_6} X'_{i_6 i_7} U_{i_8 i_7}^* Y'_{i_8 i_1} \\ &= q^{-2} \mathbb{E}_U \sum_{i_1, \dots, i_8=1}^{q^2} Y'_{i_8 i_1} Y'_{i_4 i_5} U_{i_1 i_2} U_{i_4 i_3}^* U_{i_5 i_6} U_{i_6 i_7}^* X'_{i_2 i_3} X'_{i_6 i_7} \\ &= q^2 \langle (I \otimes Y)_\downarrow | \overline{U \otimes U^* \otimes U \otimes U^*} | (I \otimes X)_\uparrow \rangle. \quad (60) \end{aligned}$$

In the second line, $X' = I \otimes X$ and $Y' = I \otimes Y$. The third line is a trivial rearrangement of the second, and the fourth employs the formal correspondence between matrices and normalized vectors introduced above.

The Haar average of the tensor product of four unitaries is given by Eq. (52) with $n = q^2$. To complete the evaluation of \bar{F} , we observe that $\langle (I \otimes X)_\uparrow | (I_{q^2})_\uparrow \rangle = q^{-4} (\text{Tr} X)^2 (\text{Tr} I_q)^2 = 0$, $\langle (I \otimes X)_\uparrow | (I_{q^2})_\downarrow \rangle = q^{-4} (\text{Tr} X^2) (\text{Tr} I_q) = q^{-2}$, $\langle (I \otimes Y)_\downarrow | (I_{q^2})_\uparrow \rangle = q^{-4} (\text{Tr} Y^2) (\text{Tr} I_q) = q^{-2}$, and $\langle (I \otimes Y)_\downarrow | (I_{q^2})_\downarrow \rangle = q^{-4} (\text{Tr} Y)^2 (\text{Tr} I_q)^2 = 0$. This gives $\bar{F} = -1/(q^4 - 1)$.

When $\ell_u, \ell_v > 1$, we map the layout of local unitaries to a partition function for the spins s, s' in Eq. (54). To facilitate the mapping, we decompose the input bra $\langle I_{s'} |$ and output ket $|I_s\rangle$ into separate “legs” corresponding to the two physical spins, as in Eq. (56),

$$\overline{U \otimes U^* \otimes U \otimes U^*} = \frac{q^4}{q^4 - 1} \sum_{s, s' = \uparrow, \downarrow} w(s, s') |s\rangle |s'\rangle \langle s'| \langle s|. \quad (61)$$

Similarly, the vectors encountered above for the case $\ell_u = \ell_v = 1$ can be decomposed as $|(IX)_\uparrow\rangle = |\uparrow\rangle |X_\uparrow\rangle$ and $|(IY)_\downarrow\rangle = |\downarrow\rangle |Y_\downarrow\rangle$, which satisfy

$$\begin{aligned} \langle \uparrow | Y_\downarrow \rangle &= q^{-2} \text{Tr} Y^2 = \frac{1}{q}, & \langle \downarrow | Y_\downarrow \rangle &= q^{-2} (\text{Tr} Y)^2 = 0, \\ \langle \uparrow | X_\uparrow \rangle &= q^{-2} (\text{Tr} X)^2 = 0, & \langle \downarrow | X_\uparrow \rangle &= q^{-2} \text{Tr} X^2 = \frac{1}{q}. \end{aligned} \quad (62)$$

The expression Eq. (60) is now depicted as in Fig. 13.

It is now clear that for general ℓ_u, ℓ_v , we may regard the array of unitaries as a tensor network composed of tensors of the form (61). The boundaries of this tensor network—i.e., the external legs of the array of $\ell_u \times \ell_v$ unitaries—involve inner products with fixed vectors. Two of the boundary legs are dressed with $q|X_\uparrow\rangle$ and $q|Y_\downarrow\rangle$; see Fig. 13. Apart from these, the external legs on the top boundary are dressed by states $q|\downarrow\rangle$, while those on the bottom boundary are dressed with $q|\uparrow\rangle$. In addition, \bar{F} includes an overall dimension factor q^{-2t} coming from the infinite-temperature Gibbs state. For convenience, we absorb the overall dimension factor q^{-2t} into the vectors on the lower boundaries; these vectors are taken to be normalized, whereas the boundary bras in the top boundaries have norm q .

We may now interpret \bar{F} as a partition function for the Ising spins $s_{u,v}$ and $s'_{u,v}$, which, according to Eq. (61), are associated with the unitary at position (u, v) . These spins take the values \uparrow, \downarrow . The weight associated with the “bond” between $s_{u,v}$ and $s'_{u,v}$ comes from the single-unitary Haar average and is $[q^4/(q^4 - 1)]$ if $s_{u,v} = s'_{u,v}$, and $[(-q^2)/(q^4 - 1)]$ if $s_{u,v} \neq s'_{u,v}$. The leg of the tensor network connecting the unitary at (u, v) to that on (for

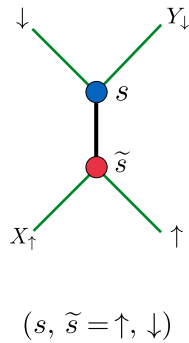


FIG. 13. Elementary tensor for the computation of \bar{F} . The boundary conditions in the top-left figure are for $\ell_u = \ell_v = 1$.

$$\begin{array}{c} \bullet s \\ | \\ \bullet \tilde{s} \end{array} = \frac{q^4}{q^4 - 1} w(s, \tilde{s}) = \begin{cases} \frac{q^4}{q^4 - 1} & s = \tilde{s} \\ -\frac{q^2}{q^4 - 1} & s \neq \tilde{s} \end{cases}$$

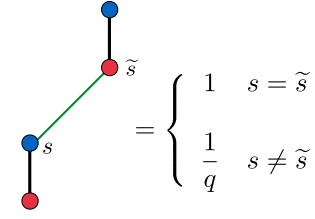


FIG. 14. Weights due to the interaction between adjacent Ising variables arising from the same unitary (top diagram) and from unitaries at adjacent time steps (bottom diagram). After integrating out the “bra” Ising variable s' , we obtain the weights shown in Fig. 15.

example) its lower right at $(u, v - 1)$ yields an interaction between $s'_{u,v}$ and $s_{u,v-1}$ that comes simply from the inner product $\langle s'_{u,v-1} | s_{u,v} \rangle$. This gives weight 1 if $s'_{u,v} = s_{u,v-1}$ and weight $(1/q)$ if $s'_{u,v} \neq s_{u,v-1}$. See Fig. 14.

Thus, we have mapped the Haar average of the out-of-time correlator to a partition function for Ising degrees of freedom on a honeycomb lattice (with the q dependence residing in the interactions on the bonds). At first sight, this may appear to be a formidable problem. Note, in particular, that some configurations have negative weight. However, a simplification is possible, as a result of the unitarity of the underlying dynamics. A hint that a simplification is possible comes from the fact that the expression for F , Eq. (48), becomes trivial if one of the operators X_0 and Y_x is the identity operator. In the Ising language, this corresponds only to a slight change of boundary conditions.

The simplification is effected by integrating out the bra variable $s'_{u,v}$ from each unitary. This generates a three-spin interaction among the ket variables $s_{u,v}$, $s_{u-1,v}$, and $s_{u,v-1}$. The calculation is straightforward and yields the table of weights in Fig. 15. For example, if $s_{u,v} = s_{u,v-1} = s_{u-1,v}$,

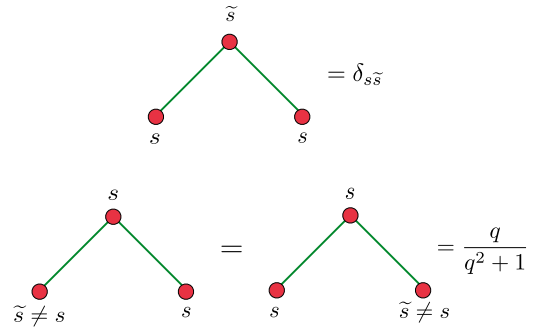


FIG. 15. Weights for the three-body interaction that arises after integrating out half of the Ising variables (the bra variables).

then the weight is $[q^4/(q^4 - 1)] \cdot 1 \cdot 1 + [(-q^2)/(q^4 - 1)] \cdot (1/q) \cdot (1/q) = 1$.

We are left with an effective Ising model for the s spins only. One s spin is associated with each unitary in the network. The fact that the weight in Fig. 15 is zero for two of the configurations means that only a very restricted subset of Ising configurations are allowed. We show that these can be summed exactly by viewing the configurations in terms of domain walls.

Let us specify the new boundary conditions. The rules of Fig. 15 apply along the bottom boundary (with the lower spins of Fig. 15 being fixed-up spins) because of our normalization convention for the boundary kets, except for the site where the observable ket $|X_\uparrow\rangle$ is dangling.

$$\begin{array}{c} \uparrow \\ \diagup \quad \diagdown \\ \uparrow_{\text{fixed}} \quad X_\uparrow \end{array} = \frac{-1}{q^4 - 1}, \quad (63)$$

$$\begin{array}{c} \downarrow \\ \diagup \quad \diagdown \\ \uparrow_{\text{fixed}} \quad X_\uparrow \end{array} = \frac{q^2}{q^4 - 1}. \quad (64)$$

The top boundary bras, which have norm q , follow the rule that

$$\begin{array}{c} \downarrow_{\text{fixed}} \\ | \\ | \\ | \\ \uparrow \end{array} = 1, \quad \begin{array}{c} \downarrow_{\text{fixed}} \\ | \\ | \\ | \\ \downarrow \end{array} = q, \quad \begin{array}{c} Y_\downarrow \\ | \\ | \\ | \\ \uparrow \end{array} = 1, \quad \begin{array}{c} Y_\downarrow \\ | \\ | \\ | \\ \downarrow \end{array} = 0. \quad (65)$$

B. Partition function for two directed paths

Now, the problem is reduced to a partition function of Ising variables with the three-body interaction and the boundary interaction. We first simplify the partition function by relating it to one with modified boundary conditions as follows. We denote the weight of a given configuration by $W_{X,Y}(s)$, where the subscripts indicate the dependence on the boundary conditions induced by the operators X and Y . Because of the last rule in Eq. (65), the spin at the site where Y_x is attached—null coordinate (ℓ_u, ℓ_v) —has to be $s_{\ell_u, \ell_v} = \uparrow$. As a result, we can replace Y_\downarrow with $\downarrow_{\text{fixed}}$, which, according to Eq. (65), gives the same weight when $s_{\ell_u, \ell_v} = \uparrow$. Let us denote the weight of a configuration s under this modified top boundary condition by $W_X(s)$, dropping the subscript Y . We may then write the desired quantity $\bar{F} = \sum_s W_{X,Y}(s)$ as

$$\bar{F} = \sum_{s: s_{\ell_u, \ell_v} = \uparrow} W_X(s) = \sum_s W_X(s) - \sum_{s: s_{\ell_u, \ell_v} = \downarrow} W_X(s). \quad (66)$$

We claim that the first term is equal to 1, and thus

$$-\frac{1}{2} \text{Tr} \rho_\infty [\overline{U(t)X_0U(t)^\dagger}, Y_x]^2 = 1 - \bar{F} \quad (67)$$

$$= \sum_{s: s_{\ell_u, \ell_v} = \downarrow} W_X(s). \quad (68)$$

The claim can be shown in two ways. First, the out-of-time correlator $1 - \bar{F}$ must vanish if the operator Y is replaced by the identity. The boundary vector $|Y_\downarrow\rangle$ then becomes $|\downarrow\rangle$, and the partition function for \bar{F} becomes precisely $\sum_s W_X(s)$. Therefore, $1 - \sum_s W_X(s) = 0$. The other way to show the claim is by directly integrating out the Ising variables inductively, starting from the top line with respect to the all- \downarrow boundary condition along the top boundary. This is a nontrivial consistency check on our reduction.

Now, we focus on $W_X(s)$, with the variable at null coordinate (ℓ_u, ℓ_v) fixed to be \downarrow . If the bottom variable where X is attached is \uparrow , then the second rule in Fig. 15, together with the boundary condition along the bottom boundary, dictates that all the bulk variables should be \uparrow . This cannot be fulfilled for the spin at (ℓ_u, ℓ_v) , implying that the weight is zero.

Hence, we have fixed two Ising variables at the bottom and top boundaries to be \downarrow , where the observables X_0 and Y_x are attached. Let us consider domain walls instead of spins. The key point is the first rule in Fig. 15, which leads to the domain walls being directed, drastically simplifying the partition function. If we follow a domain wall from the top to the bottom, it should always go down left or down right. This implies that there are two nonintersecting domain walls extending from the bottom to the top boundary. The starting vertices of the right and left domain walls have null coordinates $(1,0)$ and $(0,1)$, respectively.

The domain wall has a weight

$$\left(\frac{q}{q^2 + 1} \right)^{t-1} \quad (69)$$

from the three-spin interactions in Fig. 15, excluding that involving the spin at null coordinate $(1,1)$. The domain wall can fluctuate freely in the bulk without changing this weight; see Fig. 16.

However, the end position of the domain wall at the top boundary *does* affect the weight, and we will need to count the number of domain walls for each ending position. For the right domain wall, the end position can be at a vertex $(t-v, v)$ in null coordinates for some $v = 0, 1, \dots, \ell_v - 1$. Likewise, the end position of the left domain wall can be $(u, t-u)$ for some $u = 0, 1, \dots, \ell_u - 1$. The weight from the top boundary interaction is then

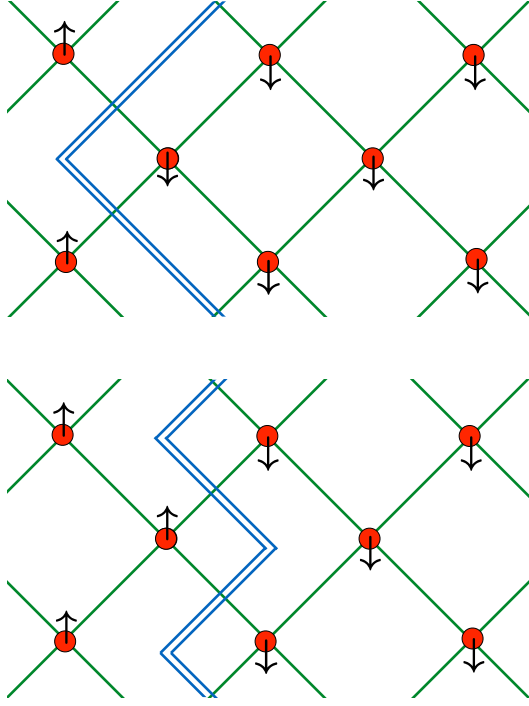


FIG. 16. A section of one of the domain walls in the bulk (double line). The two configurations shown have equal weight.

$$q^{2t-2u-2v} \quad (70)$$

from Eq. (65).

It remains to count the number of domain walls, given their end positions. The right domain wall connects $(1,0)$ to $(t-v, v)$, while the left domain wall connects $(0,1)$ to $(u, t-u)$, with the constraint that they must not intersect. To handle the constraint, we use a reflection trick. Regard the domain walls as random walkers. The right random walker randomly chooses between up-left and up-right edges, to proceed from $A = (1,0)$ to $B = (t-v, v)$. Similarly, the left random walker proceeds from $C = (0,1)$ to $D = (u, t-u)$. Any pair of paths $A \rightarrow B$ and $C \rightarrow D$ that has a common point can be viewed as a pair of paths $A \rightarrow D$ and $C \rightarrow B$. Conversely, any pair of paths $A \rightarrow D$ and $C \rightarrow B$, which must meet at a point, can be viewed as a pair of paths $A \rightarrow B$ and $C \rightarrow D$ with a common point. Therefore, the number of pairs of paths from $A \rightarrow B$ and $C \rightarrow D$ without intersection is the number of all unrestricted pairs from $A \rightarrow B$ and $C \rightarrow D$, minus the number of all unrestricted pairs from $A \rightarrow D$ and $C \rightarrow B$. The number of our domain wall configurations is therefore

$$\binom{t-1}{v} \binom{t-1}{u} - \binom{t-1}{v-1} \binom{t-1}{u-1}, \quad (71)$$

where the second factor vanishes when $u = 0$ or $v = 0$.

Finally, we combine the results above:

$$\begin{aligned} \sum_{\substack{s: \\ s_{\ell_u, \ell_v} = \downarrow}} W_X(s) &= \underbrace{\frac{q^2}{q^4-1}}_{\text{Eq. (64)}} \underbrace{\left(\frac{q}{q^2+1}\right)^{2t-2}}_{\text{Eq. (69)}} \sum_{u=0}^{\ell_u-1} \sum_{v=0}^{\ell_v-1} \underbrace{q^{2t-2u-2v}}_{\text{Eq. (70)}} \\ &\times \underbrace{\left[\binom{t-1}{v} \binom{t-1}{u} - \binom{t-1}{v-1} \binom{t-1}{u-1} \right]}_{\text{Eq. (71)}}. \end{aligned} \quad (72)$$

This correctly reproduces the answer $q^4/(q^4-1)$ when $\ell_u = \ell_v = 1$. This result can be conveniently rewritten as

$$\begin{aligned} \overline{\mathcal{C}(t, x)} &= \frac{(1-p)^2}{1-2p} g(t-1, \ell_v-1) g(t-1, \ell_u-1) \\ &\quad - \frac{p^2}{1-2p} g(t-1, \ell_v-2) g(t-1, \ell_u-2), \end{aligned} \quad (73)$$

where

$$t = \ell_u + \ell_v - 1, \quad x = \ell_u - \ell_v, \quad p := \frac{1}{q^2+1}, \quad (74)$$

$$g(n, a) := \sum_{k=0}^a \binom{n}{k} (1-p)^{n-k} p^k. \quad (75)$$

Further simplification is possible since $g(t-1, a) \simeq g(t-1, a-1)$ for large t .

$$\overline{\mathcal{C}(t, x)} \simeq g(t-1, \ell_v-1, p) g(t-1, \ell_u-1, p) \quad (76)$$

$$\simeq \Phi\left(\frac{v_B t + x}{\sigma}\right) \Phi\left(\frac{v_B t - x}{\sigma}\right), \quad (77)$$

where

$$v_B = \frac{q^2-1}{q^2+1}, \quad \sigma = \frac{2q\sqrt{t}}{q^2+1}, \quad \Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^y e^{-x^2/2} dx. \quad (78)$$

C. Bounds on fluctuations

Here, we estimate the fluctuation of $\mathcal{C}(t, x)$ due to the randomness of the unitaries. One might wish to calculate this fluctuation directly, using a similar technique that we employ for the average of $\mathcal{C}(t, x)$, but we found the exact computation unwieldy as it involves high powers of unitaries. (See Ref. [84].) Nevertheless, we can argue that the fluctuations are negligible in two regimes.

Since the random variable $\mathcal{C}(t, x)$ takes values between 0 and 2, the variance is upper bounded by $2\bar{\mathcal{C}}$. Therefore, the standard deviation is upper bounded by

$$\begin{aligned} \sqrt{\overline{\mathcal{C}(t, x)^2} - \overline{\mathcal{C}(t, x)}^2} &\leq \sqrt{2\bar{\mathcal{C}}(t, x)} \\ &\simeq O(1) \exp\left(-\frac{1}{4} \left(\frac{|x| - v_B t}{\sigma}\right)^2\right). \end{aligned} \quad (79)$$

This bound is valid for any t, x , but it is only meaningful when $|x| \gg v_B t$. This basically says that there is almost no “leakage” of operators beyond the light cone defined by v_B . (In passing, we note that one can also use Markov inequality $\Pr[X \geq a] \leq a^{-1} \mathbb{E}X$, which holds for any positive random variable X and a positive number a to have a probability tail bound.)

In the opposite regime where $|x| \ll v_B t$, we have shown that the average $\overline{\mathcal{C}(t, x)}$ is almost 1; the discrepancy is upper bounded by $O(1) \exp(-(v_B t - |x|)^2/2\sigma^2)$. Thus, in this regime, the fluctuation is basically given by

$$\mathbb{E}_U F^2(U) = N^{-2} \mathbb{E}_U (\text{Tr} U X U^\dagger Y U X U^\dagger Y)^2 \geq 0, \quad (80)$$

where F is defined in Eq. (48) and N is the dimension of the Hilbert space of spins on which $U X U^\dagger Y$ is supported. Here, U includes all the local unitaries in the evolution quantum circuit.

To estimate the fluctuation, we consider a slightly different system in which $2ct$ spins form a ring, where c is some absolute constant that depends on q only. If $c > 1$, this does not modify the dynamics at all since the evolved operator $U X_0 U^\dagger$ is supported on $2t$ spins. For $c < 1$, while we do not insist that this allows us to compute the fluctuation rigorously, we anticipate that qualitative conclusions from this modified setting carry over to the original open chain system.

In Appendix H, we show that if $c = M_q^{-1} = \tilde{O}(q^{-2})$, then for all $|x| < ct$,

$$\mathbb{E}_U F^2(U) \leq 11q^{-4ct}. \quad (81)$$

In other words, deep in the light cone, there is a region in spacetime bounded by a nonzero speed, where the fluctuation of $\mathcal{C}(t, x)$ is suppressed exponentially in t . It is likely that this is only a bound, rather than a tight estimate of fluctuation. Equation (81) is proved using previous results on approximate unitary designs [24], as well as estimates for $\mathbb{E}_U F^2(U)$ when U is truly Haar random [119].

V. ENTANGLEMENT GROWTH

Entanglement can be quantified in various ways, but perhaps the simplest measure is the entanglement purity

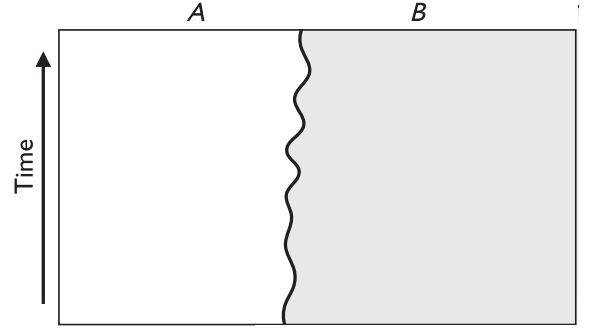


FIG. 17. The average entanglement purity $\bar{\mathcal{P}}$ maps to the partition function for a directed random walk. At the top boundary, the walk terminates at the position of the entanglement cut between subsystems A and B . At the bottom boundary, the end point is free.

$\mathcal{P} = \text{Tr}(\text{Tr}_{A^c} |\psi\rangle\langle\psi|)^2 \leq 1$, where A is some region. A pure state $|\psi\rangle$ on $A \cup A^c$ is entangled if and only if the purity is not equal to 1. The logarithm of the entanglement purity is the Renyi-2 entropy

$$S_2(A) = -\log \mathcal{P}. \quad (82)$$

In this section, we calculate exactly the average purity of “half” of the infinite chain, for arbitrary t , under the evolution protocol in Sec. II, with an initial product state. Previously, a bound on the saturation time for $q = 2$ was obtained [23]. Also, the evolution of average entanglement purity has been calculated using other techniques under a protocol where random two-site unitaries are applied to random bonds [83] and other protocols [84].

Here, we show that the averaged purity $\bar{\mathcal{P}}$ can be mapped to the partition function for a directed walk, shown schematically in Fig. 17. Our calculation technique below will be very similar to the OTOC calculation in the previous section; the difference is only in the boundary conditions.

A. Random walk picture

Let A be the left half of our chain, and B be the right half. The initial pure density matrix is $\rho(t=0) = \cdots \otimes P_{-1} \otimes P_0 \otimes P_1 \otimes \cdots$, where P_i is the projector $|i\rangle\langle i|$ onto the initial state at site i . If U is the full time-evolution unitary consisting of local unitaries, the entanglement purity \mathcal{P} across the cut between A and B is

$$\begin{aligned} \overline{\mathcal{P}(t)} &= \sum_{a, a'=1}^{q^{|A|}} \sum_{b, b'=1}^{q^{|B|}} \overline{\langle ab | U \rho(0) U^\dagger | a' b' \rangle \langle a' b' | U \rho(0) U^\dagger | ab \rangle} \\ &= q^{|A|+|B|} \langle \uparrow^{|A|} \downarrow^{|B|} | \overline{U \otimes U^* \otimes U \otimes U^*} | \cdots (P_0)_\uparrow (P_1)_\uparrow \cdots \rangle. \end{aligned} \quad (83)$$

The notation $|\uparrow\rangle, |\downarrow\rangle$ is the same as in Sec. IV A. For a one-dimensional projector P on q -dimensional space, the q^4 -dimensional vector $|P_\uparrow\rangle$ satisfies

$$\langle\uparrow|P_\uparrow\rangle = \frac{1}{q} = \langle\downarrow|P_\uparrow\rangle. \quad (84)$$

The expression for the purity can be thought of as a partition function for classical Ising spins as in Sec. IV A. There are two Ising spins associated with each local unitary; see Eq. (54). Because of Eq. (84), for any configuration of the Ising spins, the weight factor from the bottom boundary is $q^{-|A|-|B|}$, which cancels the factor $q^{|A|+|B|}$ in front of Eq. (83). Hence, the average purity is simply the sum of weights from the domain wall in the bulk (e.g., see Fig. 16).

In Sec. IV A, we first integrated out the bra Ising variables s' , but here we find it simpler to integrate out the ket Ising variables s . The transition rules of Fig. 15 are now upside down but otherwise the same. Then, we have a single domain wall starting from the top boundary to reach the bottom. Any domain wall has length exactly t , giving rise to weight $[q/(q^2 + 1)]^t$. The domain wall can choose between left-down or right-down moves as it proceeds from the top, and therefore, there are 2^t domain walls. We conclude that

$$\overline{\mathcal{P}(t)} = \left(\frac{2q}{q^2 + 1}\right)^t. \quad (85)$$

Note that the factor $2q/(q^2 + 1)$ is directly related to the average “entangling power” of Ref. [120]. We may define the “purity speed” as

$$\overline{\mathcal{P}(t)} \equiv q^{-v_P t}, \quad v_P = \log_q \frac{q^2 + 1}{2q}. \quad (86)$$

This quantity gives a bound on the growth rate of the second Renyi entropy:

$$\overline{S_2(\rho(t)_A)} = -\log_q \overline{\mathcal{P}(t)} \geq -\log_q \overline{\mathcal{P}(t)} = v_P t. \quad (87)$$

The inequality is because the function $f(x) = -\log x$ is convex. Note that this expression bounds the growth rate of S_2 but does not fix it. The distribution of S_2 fluctuates in a window of small size compared to its mean [10,121], but since S_2 appears in the exponential in q^{-S_2} , this does not rule out the possibility that this quantity is affected by rare, anomalously small values of S_2 , making it very different from $q^{-\overline{S_2}}$.

The von Neumann entropy S_{vN} is always greater than or equal to S_2 , so the growth rate v_E of S_{vN} is also bounded by v_P :

$$v_E \geq v_P = \log_q \frac{q^2 + 1}{2q} = 1 - \frac{\log 2}{\log q} + O\left(\frac{1}{q^2 \log q}\right), \quad (88)$$

where the expansion is for large q .

In Ref. [10], we argued that the universal fluctuations of the entanglement in random circuit dynamics may be understood in terms of a coarse-grained minimal cut, of random shape, through the random circuit. This picture may be contrasted with the domain wall calculation of the averaged purity, which reduces to a statistical mechanics problem without quenched randomness. This is reminiscent of the difference between a quenched and an annealed average in the statistical mechanics of disordered systems [122]. A direct exact calculation of $\overline{S_2}$ (not to mention $\overline{S_{vN}}$, or of the fluctuations in the entropy) for finite [123] q would be much more difficult than the calculation above, as a replica-like limit [122] would be required to handle the logarithm. However, structure arising from unitarity might make this calculation tractable. This is an interesting task for the future.

The scaling limit of the representation obtained in this section, where we take length scales and timescales to be large and of the same order, yields a “deterministic” domain wall configuration. This is simply a vertical line for the infinite geometry considered here. [124] We expect that extending the calculation to higher dimensions will give, in the scaling limit, a formula for $-\log \overline{\mathcal{P}}$ as the “energy” of a minimal surface (representing the Ising domain wall), which has a deterministic coarse-grained geometry, obtained from an effective, elastic, energy minimization problem. This is precisely the scaling picture proposed in Ref. [10] for the growth of entanglement in higher-dimensional systems.

B. Nonuniversality of the ratio v_E/v_B

In Ref. [10] (see also Ref. [11]), we showed that the speed v_E associated with entanglement growth is, in general, smaller than the operator growth speed v_B , and we gave explicit models displaying a ratio $v_E/v_B < 1$. In these models, [125] this ratio happened to be $1/2$. Values close to $1/2$ were also found numerically in Refs. [11,51]. These results might lead one to wonder whether this value is, in some sense, generic. Here, we show that it is not generic by constructing random circuit dynamics involving interactions of a large but finite range, which give *arbitrarily small* values of v_E/v_B without any fine-tuning. The construction uses random unitaries made up of “staircases” of length $O(q^2)$, which are made up of smaller random unitaries (Fig. 18). When q is large, we obtain a ratio v_E/v_B that is at most of order $1/q^2$. (In a deterministic spin chain with quenched spatial disorder, it is even possible to have $v_E/v_B = 0$ [40], but here we insist on *statistical*

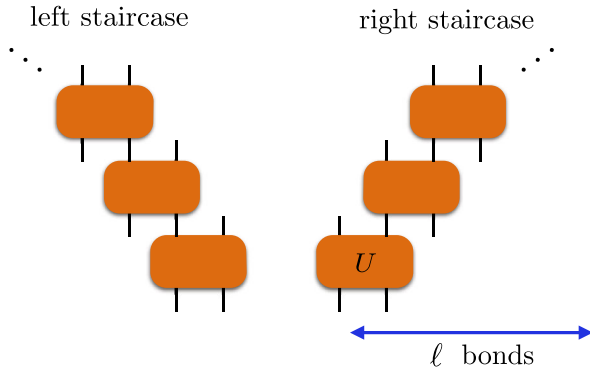


FIG. 18. Random circuit built from “staircase” unitaries: We use “left” and “right” staircases—built from random two-site unitary operators as shown, and extending over ℓ bonds—as the building blocks for a random quantum circuit in which the ratio of the entanglement and butterfly velocities v_E/v_B may be made arbitrarily small.

translational invariance: i.e., we insist that the probability distribution for the circuit is invariant under translations.)

Consider quantum circuit dynamics in which “staircase unitaries” are applied at random locations and at random times in a Poissonian fashion. A staircase is a collection of two-site unitaries arranged as in Fig. 18. Left- and right-oriented staircases are applied with equal probability. The staircase acts on ℓ bonds, and we take ℓ large but finite, satisfying $\ell/q^2 \gg 1$. Let r be the rate at which staircases are dropped at a given location.

A single staircase can increase the entanglement across a given bond by at most two units, implying $v_E \leq 2r\ell$. On the other hand, a single staircase can move the end point of an operator a long way when $\ell \gtrsim q^2 \gg 1$. The random walk picture of Sec. II A shows that in the limit of large ℓ/q^2 , a staircase advances the front of the OTOC by an average distance of about $q^2/2$. This involves an average over the two staircase orientations, only one of which is effective in advancing the front a long distance. The large value is because, when q is large, the small value of $p = 1/(q^2 + 1)$ [Eq. (24)] means the random walker can “run” a long way up a rightward-oriented staircase before falling off. The previous case implies $v_B \simeq q^2 r \ell / 2$ at leading order in ℓ . This yields a ratio $v_E/v_B \lesssim 4/q^2$ in this regime, which can be made arbitrarily small by taking q (and hence ℓ) to be large.

VI. OUTLOOK

We have argued that universal scaling forms for the out-of-time-order correlator can be obtained by using mappings to paradigmatic problems in classical statistical mechanics. In one dimension, we gave an extremely simple hydrodynamic picture in terms of diffusion. In higher dimensions, we gave a mapping to classical surface growth and the KPZ equation [126]: The OTOC is given by the

averaged density for a growth process in the KPZ universality class.

These mappings were derived exactly for random unitary circuits, which are natural “least structured” models for chaotic quantum dynamics in situations where conserved quantities do not play an important role. We have conjectured that the universal scaling forms found here also apply to OTOCs at asymptotically late times in generic, nonintegrable many-body systems and quantum field theories. It will be interesting to test this conjecture in other situations where calculations are possible.

This picture differs from that obtained in a number of previous calculations using many-body perturbation theory [31,33,37,38], and it will be interesting to understand the reasons for these differences. Reference [31] found an operator front that did not broaden in time, whereas here we find a broadening front in all dimensions below $4 + 1$. Additionally, in Ref. [31], the OTOC was found to obey a local, nonlinear, traveling wave equation, which is unlike what we found for random circuits. In 1D, we obtained a *linear* hydrodynamic equation, while in higher dimensions, $\bar{C}(x, t)$ in a random circuit is not governed by a local differential equation at all, contrary to standard lore about OTOCs.

Interestingly, a mean field approximation to the classical growth process would yield a local differential (or rather difference) equation for the OTOC of traveling wave form. This is discussed in Appendix I. However, the mean field approximation is not valid in physical dimensionalities.

Assuming that the results here do indeed apply to realistic systems with Hamiltonians that are fixed in time, it will be interesting to consider extensions of the present coarse-grained pictures which take conserved quantities into account.

We have also given exact results for entanglement growth in $1 + 1$ D that support the scaling ideas put forward in Ref. [10], as discussed in Sec. V. In this picture (in any D), entanglement growth is determined by a minimal surface in spacetime, whose geometry becomes well defined [127] in an appropriate scaling limit and is determined essentially by an elastic minimization problem. Furthermore, it was argued in that paper and in Ref. [11] that, generically, $v_E < v_B$, where v_E is the speed characterizing the growth of entanglement. Here, we have shown that it is possible to have arbitrarily small v_E/v_B in a random quantum circuit.

It turns out that the effective Ising partition functions for calculating the OTOC and the purity have interesting structure, making them drastically simpler than they appeared to be at first sight and much simpler than the analogous partition function for a nonunitary tensor network [80]. It would be very interesting to explore whether similar simplifications occur when the averaging involves higher powers of the unitary circuit. If so, this would permit calculations of, say, modified versions of the OTOC involving higher powers of the commutator (here, we have

considered only the square). Similar techniques would also allow a calculation of fluctuations in the OTOC due to randomness in the circuit. We have conjectured that these fluctuations are small whenever the front broadens with time, but we have not attempted a direct calculation, except via a heuristic argument in 1D. [128] Even more interesting would be a direct calculation of the von Neumann entropy, which would have to use a replica limit to handle the logarithm.

OTOCs involving higher powers of the commutator are important for comparison with the Lieb-Robinson bound. The OTOC considered here can be thought of as the squared Frobenius norm of the commutator divided by the Hilbert space dimension, whereas the Lieb-Robinson bound is on the operator norm of the commutator. Our results do not determine the operator norm: they yield a lower bound on the average squared operator norm, but not a nontrivial upper bound because of the large dimension factor. The exact relation of the two quantities is yet to be determined.

In addition to exploring implications for realistic many-body systems, interesting questions remain that are specific to the random circuit context. (Note that, at the most basic level, our results show that operator growth saturates the naive causal light cone of the quantum circuit as $q \rightarrow \infty$ but not for finite q .) The randomness in the circuit necessarily implies statistical fluctuations in all observables including $\mathcal{C}(x, t)$. We have argued that these statistical fluctuations are (perhaps counterintuitively) a subleading effect at late times. We have shown this in regimes far from the front of the OTOC by giving inequalities, and we have given a heuristic argument for it in the region near the front. This argument was based on a phenomenological extension of the hydrodynamic equation for $\overline{\mathcal{C}(x, t)}$ in the 1D case to allow for statistical fluctuations in $\mathcal{C}(x, t)$ [Eq. (29)]. It would be desirable to give a microscopic derivation of Eq. (29). (For the entanglement entropy, statistical fluctuations were investigated in Ref. [10].) It also remains to characterize the classical growth problem in Sec. III more fully, for example, by obtaining the nonuniversal constants via an approximate analytic treatment.

The KPZ equation is connected to a remarkable array of topics in classical statistical mechanics [129], including the directed polymer in a random medium [130] and one-dimensional hydrodynamics [131], and it has beautiful experimental applications [104,105]. Through the Tracy-Widom distribution [90], it is also connected to random matrix theory and an array of combinatorial problems (for example, the longest increasing subsequence problem and the statistics of random permutations [132,133]). It will be interesting to explore which members of this array can shed light on operator growth.

ACKNOWLEDGMENTS

We thank D. Huse, C. Jonay, D. Chowdhury, J. Ruhman, and H. Spohn for useful discussions. We are grateful to

T. Veness and J. Chalker for helpful discussions regarding droplet shape. J.H. thanks M.B. Hastings for useful discussions. A.N. was supported by the Gordon and Betty Moore Foundation under the EPiQS initiative (Grant No. GBMF4303) and by EPSRC Grant No. EP/N028678/1. S. V. is supported by the DOE Office of Basic Energy Sciences, Division of Materials Sciences and Engineering under Award No. de-sc0010526 and, for part of this work, by the National Science Foundation under Grant No. NSF PHY11-25915.

Note added.—Recently, we became aware of related work [134]. We also note forthcoming numerical and analytical work on operator spreading [118].

APPENDIX A: EVOLVING DISTRIBUTION ON OPERATOR STRINGS

In this appendix, we give a more detailed explanation of the relationship between the dynamics of the coefficients a_S and a Markov process [22,23] and of the derivation of the diffusion picture. We consider Haar-random, local unitary dynamics. In an N -site system with a q -dimensional Hilbert space at each site, a Hermitian operator that has evolved under the unitary circuit $\mathcal{O}(t) = U(t)^\dagger \mathcal{O} U(t)$ may be expanded in a basis of $SU(q^N)$ generators $\{S\}$ as

$$\mathcal{O}(t) = \sum_S a_S(t) S. \quad (\text{A1})$$

Our normalization convention is $\text{Tr}(SS') = q^N \delta_{SS'}$, so $a_S(t) = q^{-N} \text{Tr}(\mathcal{O}(t)S)$. The squared coefficient $a_S(t)^2$ evolves as

$$\begin{aligned} a_S(t)^2 &= q^{-2N} \sum_{S', S''} a_{S'}(t-1) a_{S''}(t-1) \\ &\quad \times \text{Tr}[\mathcal{U} S' \mathcal{U}^\dagger S] \text{Tr}[\mathcal{U} S'' \mathcal{U}^\dagger S] \\ &= q^{-2N} \sum_{S', S''} a_{S'}(t-1) a_{S''}(t-1) \\ &\quad \times \prod_r \text{tr}[U_r S'_r U_r^\dagger S_r] \text{tr}[U_r S''_r U_r^\dagger S_r], \quad (\text{A2}) \end{aligned}$$

where \mathcal{U} is a layer of m -site unitaries that were applied at time $t-1$. In the second line, we have written $\mathcal{U} = \prod_r U_r$, where \mathbf{r} is the coordinate of disjoint, m -site clusters on which the unitary $U_r \in U(q^m)$ acts, and we have also decomposed $S = \prod_r S_r$ as a product of basis elements acting on these m -site clusters. These operators are normalized according to $\text{tr}[S_r S'_r] = q^m \delta_{S_r, S'_r}$ and $\text{tr}[S_r] = q^m \delta_{S_r, 1}$. The Haar average of the above expression is given by

$$\begin{aligned} & \overline{\text{tr}[U_r S_r' U_r^\dagger S_r] \text{tr}[U_r S_r'' U_r^\dagger S_r]} \\ &= \frac{\delta_{S_r', S_r''}}{1 - q^{-2m}} \{q^{2m} \delta_{S_r, 1} \delta_{S_r', 1} + 1 - \delta_{S_r', 1} - \delta_{S_r, 1}\}. \end{aligned} \quad (\text{A3})$$

Thus, the Haar-averaged $a_S(t)^2$ evolves *linearly*,

$$\begin{aligned} \overline{a_S(t)^2} &= \frac{1}{q^{2N}} \sum_{S', S''} a_{S'}(t-1) a_{S''}(t-1) \\ &\quad \times \prod_r \frac{\delta_{S_r', S_r''} (q^{2m} \delta_{S_r, 1} \delta_{S_r', 1} + 1 - \delta_{S_r', 1} - \delta_{S_r, 1})}{1 - q^{-2m}} \\ &= \sum_{S'} W_{SS'} a_{S'}(t-1)^2 \end{aligned} \quad (\text{A4})$$

with the real, symmetric matrix

$$W_{SS'} = \prod_r \left[\delta_{S_r, 1} \delta_{S_r', 1} + \frac{(1 - \delta_{S_r, 1})(1 - \delta_{S_r', 1})}{q^{2m} - 1} \right]. \quad (\text{A5})$$

Averaging again over the unitaries applied in the previous time steps gives an equation for $P_S(t) \equiv \overline{a_S^2(t)}$,

$$P_S(t) = \sum_{S'} W_{SS'} P_{S'}(t-1), \quad (\text{A6})$$

which is formally a master equation for a fictitious Markov process [22,23]; at a given time, there is a single string \mathcal{S} that is updated stochastically in each time step, via local updates involving a cluster of m sites. From the form of $W_{SS'}$, we see that the local update on m sites is performed by replacing a nontrivial generator on the cluster *randomly* by any one of the $q^{2m} - 1$ nontrivial generators. We emphasize that this fictitious Markov process is not the true unitary dynamics of the operator $\mathcal{O}(t)$.

This fictitious classical stochastic process dramatically simplifies through the following observations. We focus here on one spatial dimension with updates on bonds. First, observe that the matrix elements $W_{SS'}$ only depend on the support of the generators \mathcal{S} and \mathcal{S}' , so Eq. (A6) gives rise to a simpler Markov process for the binary occupation number $n(x)$, which is 1 if the corresponding generator has support at site x and 0 otherwise (if \mathcal{S} acts as the identity at x). Formally, the probability distribution of the occupation numbers is given by

$$P[\{n\}; t] = \sum_{\mathcal{S}'} \overline{a_{\mathcal{S}'}(t)^2}, \quad (\text{A7})$$

where the prime indicates that the sum is only over strings \mathcal{S} that are compatible with the configuration $n(x)$. Further, the end point of the string observes an *autonomous* Markovian dynamics. Since $m = 2$, updates involving the end point either include the site to the right of it, which is empty, or that to the left, which may be empty or full.

The dynamical rule above implies that the probabilities for the position of the end point after the update are independent of whether the leftward site was initially occupied or empty. Formally, the probability distribution for the position of the end point in this fictitious dynamics is

$$p_{\text{end}}(x; t) = \sum_{\substack{n \text{ with} \\ \text{endpt at } x}} P[\{n\}; t] = \sum_{\mathcal{S} \text{ ends at } x} \overline{a_{\mathcal{S}}(t)^2}, \quad (\text{A8})$$

which is precisely $\overline{\rho(x, t)}$, as defined in Sec II. Therefore, for an end point at x or $x + 1$, a single update applied to the sites x and $x + 1$ leaves the end point at x with probability $p = (q^2 - 1)/(q^4 - 1) = 1/q^2 + 1$ and at $x + 1$ with probability $1 - p$. This establishes the claim in Sec. II for the evolution of $\overline{\rho(x, t)}$ in a single time step.

APPENDIX B: VELOCITY AND DIFFUSION CONSTANT FOR LATTICE DIFFUSION EQUATION

Our layout of the evolution operator is such that the local unitaries alternate between even and odd bonds. In other words, a bond at time step t is either at the left or right of the bond at $t - 1$. Thus, it suffices to count the left and right moves to specify the position of the right end bond of $X_0(t)$. As described in the main text, the probability of a left move is p . Let $u \geq 0$ be the number of right moves, and let $v \geq 0$ be the number of left moves. We have $u + v = t$, and $u - v$ (or $u - v \pm 1$) is the spatial coordinate of the right end point. Therefore, the probability distribution of the position of the right end bond is

$$f(u, v) = \binom{u+v}{u} (1-p)^u p^v. \quad (\text{B1})$$

This is correctly normalized since $\sum_{u+v=t} f(u, v) = (1-p+p)^t = 1$. Then, the probability that a site x is left to the right end of $X_0(t)$ is

$$\sum_{u+v=t, u-v \geq x} f(u, v) = \sum_{v=0}^{(t-x)/2} \binom{t}{v} (1-p)^{t-v} p^v \quad (\text{B2})$$

$$\simeq \Phi\left(\frac{v_B t - x}{\sigma}\right), \quad (\text{B3})$$

where Φ is the cumulative density function of the normal distribution, and

$$v_B = \frac{q^2 - 1}{q^2 + 1}, \quad \sigma = \frac{2q\sqrt{t}}{q^2 + 1}. \quad (\text{B4})$$

APPENDIX C: NOISY DIFFUSION EQUATION

Starting with Eq. (29), without loss of generality we rescale space so $D = 1$, and set $v = 0$ by going to the moving frame. Let ρ_0 be the solution without noise, $\rho_0 = (4\pi t)^{-1/2} e^{-x^2/4t}$. In terms of the Green's function,

$$\begin{aligned} \rho(x, t) - \rho_0(x, t) &= \int_{x', t'} G(x - x', t - t') \partial_x \eta(x', t') \rho(x', t') \\ &= \int_{x', t'} G'(x - x', t - t') \eta(x', t') \rho(x', t'). \end{aligned} \quad (\text{C1})$$

The center-of-mass position of the wave packet within a given realization is $x_{\text{cm}} = \int_x x \rho(x, t)$, so, if $\overline{x_{\text{cm}}}$ is the center-of-mass position averaged over realizations,

$$\begin{aligned} x_{\text{cm}} - \overline{x_{\text{cm}}} &= \int_{x, x', t'} x G'(x - x', t - t') \eta(x', t') \rho(x', t') \\ &= - \int_{x, x', t'} G(x - x', t - t') \eta(x', t') \rho(x', t'), \\ (x_{\text{cm}} - \overline{x_{\text{cm}}})^2 &= \int_{\substack{\tilde{x}, \tilde{x}' \\ \tilde{t}, \tilde{t}'}} G(x - x', t - t') G(\tilde{x} - \tilde{x}', t - \tilde{t}') \\ &\quad \times \eta(x', t') \eta(\tilde{x}', \tilde{t}') \rho(x', t') \rho(\tilde{x}', \tilde{t}'). \end{aligned}$$

Averaging over the noise with $\langle \eta(x, t) \eta(x', t') \rangle = \lambda \delta(x - x') \delta(t - t')$, we get

$$\overline{[x_{\text{cm}} - \overline{x_{\text{cm}}}]^2} = \lambda \int_{\substack{\tilde{x}, \tilde{x}' \\ \tilde{t}, \tilde{t}'}} G(x - x', t - t') G(\tilde{x} - \tilde{x}', t - \tilde{t}') \rho(x', t')^2. \quad (\text{C2})$$

For the leading-order scaling, we replace ρ with ρ_0 on the right-hand side. Then, dimensional analysis applied to the integral gives

$$\sqrt{\overline{[x_{\text{cm}} - \overline{x_{\text{cm}}}]^2}} \propto \lambda^{1/2} t^{1/4} + \dots \quad (\text{C3})$$

This is a statistical variation in x_{cm} of order $t^{1/4}$, in agreement with the heuristic argument and with Ref. [88]. This variation is small compared to the width of ρ , indicating that $\rho - \rho_0 \ll \rho$ at late times. The typical size of $\partial_x \rho$ near the peak is $O(1/t)$, so $x_{\text{cm}} - \overline{x_{\text{cm}}} \sim t^{1/4}$ corresponds to $\rho - \rho_0 \sim t^{-3/4}$, as compared with $\rho \sim t^{-1/2}$. The approximation above is therefore self-consistent.

APPENDIX D: RANDOM CLIFFORD OPERATORS

Here, we review that the left- and right-invariant probability distribution over the Clifford group on n q -dimensional qudits is a unitary 2-design when q is a

prime number. In other words, for a qudit of prime power dimension q^n , the unitary group $U(q^n)$ has a finite subgroup that is a unitary 2-design, and there is a linear operator basis that remains closed under conjugations by this subgroup. This is a well-known result [79], but we include it here for the reader's convenience.

To define the Clifford group, we first need the Pauli group. Define $X = \sum_{j=0}^{q-1} |j+1 \bmod q\rangle \langle j|$ and $Z = \sum_{j=0}^{q-1} e^{2\pi i j/q} |j\rangle \langle j|$. Then, the Pauli group is the subgroup of $U(q^n)$ generated by matrices $X_1, Z_1, \dots, X_n, Z_n$, where

$$\begin{aligned} X_j &= I_q^{\otimes(j-1)} \otimes X \otimes I_q^{\otimes(n-j)}, \\ Z_j &= I_q^{\otimes(j-1)} \otimes Z \otimes I_q^{\otimes(n-j)}. \end{aligned} \quad (\text{D1})$$

The Clifford group is defined to be the normalizer of the Pauli group in $U(q^n)$. The Pauli group quotiented out by its center $\langle \omega = e^{2\pi i/q} \rangle$ is Abelian since $XZX^\dagger = \omega^{-1}Z$, and it is isomorphic to the additive group \mathbb{Z}_q^{2n} . We define P_v for $v \in \mathbb{Z}_q^{2n}$ to be an element of the Pauli group (Pauli operator) as

$$P_v = X_1^{v_1} X_2^{v_2} \dots X_n^{v_n} Z_1^{v_{n+1}} Z_2^{v_{n+2}} \dots Z_n^{v_{2n}}. \quad (\text{D2})$$

The center of the Pauli group is also contained in the center of $U(q^n)$, and therefore, the conjugation action by the Clifford group on the Pauli group induces an action on \mathbb{Z}_q^{2n} . It turns out that this group \mathcal{S} of action consists precisely of those that preserve the symplectic form

$$\lambda_n = \begin{pmatrix} 0 & -I_n \\ I_n & 0 \end{pmatrix} \quad (\text{D3})$$

over \mathbb{Z}_q .

A probability distribution ν of unitary matrices forming a 2-design means that

$$\mathbb{E}_{U \sim \nu} U \otimes U^* \otimes U \otimes U^* = \mathbb{E}_{U \sim \mu} U \otimes U^* \otimes U \otimes U^*, \quad (\text{D4})$$

where μ is the Haar probability distribution over $U(q^n)$, and U^* is the complex conjugate of U . Tautologically, the Haar distribution is a 2-design. This is equivalent to

$$\mathbb{E}_{U \sim \nu} U O U^\dagger \otimes U O' U^\dagger = \mathbb{E}_{U \sim \mu} U O U^\dagger \otimes U O' U^\dagger \quad (\text{D5})$$

for any $q^n \times q^n$ matrices O and O' . Since Pauli operators (the elements of the Pauli group defined above) generate the full operator algebra over the complex numbers, it is enough to have Eq. (D5), with O and O' being Pauli operators.

Let ν be the left-invariant (hence right-invariant) probability distribution over the Clifford group. This is the uniform distribution over the finite Clifford group.

Consider a \mathbb{C} -linear map Π_ν on the set of operators defined by

$$\Pi_\nu: O \otimes O' \mapsto \mathbb{E}_{U \sim \nu} U O U^\dagger \otimes U O' U^\dagger. \quad (\text{D6})$$

Since ν is a left-invariant distribution over a group of unitaries, Π_ν is a projector (which is Hermitian under the Hilbert-Schmidt inner product). Since the Clifford group includes the Pauli group, we have, for arbitrary $a, b \in \mathbb{Z}_q^{2n}$,

$$\begin{aligned} \Pi_\nu(P_a \otimes P_b^\dagger) &= \sum_{x,y \in \mathbb{Z}_q^{2n}} \eta_{x,y}^{a,b} P_x \otimes P_y^\dagger \quad (\eta_{x,y} \in \mathbb{C}, \text{Pauli basis expansion}) \\ &= P_c^{\otimes 2} \Pi_\nu(P_a \otimes P_b^\dagger) (P_c^{\otimes 2})^{-1} \quad (\text{for any } c \in \mathbb{Z}_q^{2n} \text{ by the left invariance of } \nu) \\ &= \sum_{x,y \in \mathbb{Z}_q^{2n}} \eta_{x,y}^{a,b} \omega^{c^T \lambda_n(x-y)} P_x \otimes P_y^\dagger \quad (\text{commutation relation among Pauli operators}) \\ &= \sum_{x \in \mathbb{Z}_q^{2n}} \eta_{x,-x}^{a,b} P_x \otimes P_x^\dagger \quad (\det \lambda_n = 1, \text{ and } c \text{ was arbitrary}). \end{aligned}$$

The use of inverse P_b^\dagger here instead of P_b is for notational convenience later.

Now, observe that for any nonzero $x, y \in \mathbb{Z}_q^{2n}$, there exists a symplectic transformation $S \in \mathcal{S}$ such that $y = Sx$. For this step, it is essential that q is prime. By the right invariance of ν by S , we see $\Pi_\nu(P_x \otimes P_x^\dagger) = \Pi_\nu(P_y \otimes P_y^\dagger)$. This implies that

$$\Pi_\nu(P_a \otimes P_b) = \eta_{0,0}^{a,b} I + \eta^{a,b} \sum_{x \in \mathbb{Z}_q^{2n} \setminus \{0\}} P_x \otimes P_x^\dagger. \quad (\text{D7})$$

We claim that this is a linear combination of the identity operator and the swap operator $\mathcal{F} = \sum_{u,v=0}^{q-1} |u\rangle\langle v| \otimes |v\rangle\langle u|$. This is easily verified once we expand \mathcal{F} in the Pauli operator basis; using $\text{Tr}(\mathcal{F}O \otimes O') = \text{Tr}(OO')$ for any $q^n \times q^n$ matrices O and O' , we see that $\mathcal{F} \propto \sum_{x \in \mathbb{Z}_q^{2n}} P_x \otimes P_x^\dagger$.

The identity operator and the swap operator commute with $U \otimes U$, where $U \in U(q^n)$. This implies that $\Pi_\nu(P_a \otimes P_b^\dagger)$ commutes with $U \otimes U$ and hence is equal to $\Pi_{\mu_{\text{Haar}}} \circ \Pi_\nu(P_a \otimes P_b^\dagger)$. By the right invariance of the Haar distribution μ_{Haar} , we conclude that Eq. (D5) is proven.

When q is not prime, any probability distribution over the Clifford group fails to be a unitary 2-design. Let $n = 1$. Since the image of Π_μ is a linear combination of the identity and the swap, we must have (see Appendix G below)

$$\Pi_\mu(O \otimes O') = \sum_{s=\pm 1} \frac{\text{Tr}(O)\text{Tr}(O') + s\text{Tr}(OO')}{q(q+s)} \frac{I + s\mathcal{F}}{2}. \quad (\text{D8})$$

When $q = 6$, there are nonidentity Pauli operators P and Q such that $P^2 = I$ and $Q^3 = I$. From Eq. (D8), we have $\Pi_\mu(P \otimes P^\dagger) = \Pi_\mu(Q \otimes Q^\dagger) \neq 0$. However, $\Pi_\nu(P \otimes P^\dagger)$ is a linear combination of Pauli operators, each of which squares to identity, whereas $\Pi_\nu(Q \otimes Q^\dagger)$ is a linear

combination of those that cube to identity, so they cannot be equal.

APPENDIX E: ANOMALOUS BEHAVIOR OF THE FRONT FOR $\phi = 0$

Above, we noted that for sufficiently large $p, p > p_c$, the lattice growth process that we consider has anomalous behavior when the front is oriented parallel to a lattice plane. This is a known phenomenon in various lattice growth models in discrete time, which have synchronous parallel updates, and is well understood in terms of directed percolation [94–98].

In the regime $p > p_c$, the lattice-aligned ($\phi = 0$) front has a speed $v_B(\phi = 0) = 2$, which is precisely the maximum possible speed allowed by causality. In this regime, the front is pinned to the “light front” and is not rough (i.e., the width is of order one). (Exactly at p_c , the aligned front is logarithmically rough [97].) For our lattice model, it appears that $p_c \lesssim 2$.

This phenomenon is easily understood via a correspondence with directed percolation [96]. First, consider a straight, lattice-aligned front in the trivial deterministic limit $p = 0$ ($q = \infty$). Apart from possibly on the first time step, this flat front advances by two lattice spacings every period: The front keeps pace with the “light cone,” which is the line $x = 2t$. Let $\tilde{n}(y, t) = 0, 1$ denote the occupation numbers of the column of sites at the light cone: $\tilde{n}(y, t)$ is the occupation number of the site at position $(2t, y)$ at time t . When $p = 0$, we have $\tilde{n}(y, t) = 1$. We are interested in the density $\langle \tilde{n} \rangle$ (averaged over y) at late times when p is nonzero. If this density remains finite, that means the front has an $O(1)$ width and is attached to the light cone. If it instead tends to zero, the front detaches from the light cone, and we expect to recover standard KPZ roughening. Note that, in order to determine \tilde{n} at time $t + 1$, it is sufficient to know only \tilde{n} at time t . The dynamics of the occupation numbers $\tilde{n}(y, t)$ are as follows. Under a horizontal dimer

update (which advances the lightfront), each occupied y has a chance $(1 - p)$ of becoming unoccupied. Under a vertical update, pairs of adjacent y undergo the pairwise update described in the main text. This process allows occupied sites to “reproduce.” This is therefore a birth-death process of the directed percolation type [122]. When p is large, the death rate is small and the reproduction rate is large, and the process is in an “active” phase with $\langle \tilde{n} \rangle > 0$, while when p is small, the population of occupied sites dies out.

APPENDIX F: SHAPE OF A SPREADING DROPLET FOR WEAKLY VARYING $v(\phi)$

Consider an asymptotic front shape described by the parametrized curve $(\theta, r_t(\theta))$ in polar coordinates, which grows simply by rescaling: $r_t(\theta) = t \times r(\theta)$. Let $\phi(\theta)$ be the angle of the front’s normal (to the x axis) at polar position θ . The radial growth rate is $\dot{r}_t(\theta) = v_B(\phi(\theta))/\cos[\phi(\theta) - \theta]$. Since the curve grows by rescaling, we have $\partial_\theta[\dot{r}_t(\theta)/r_t(\theta)] = 0$. Note that

$$\partial_\theta \ln r(\theta) = -\tan[\phi(\theta) - \theta]. \quad (\text{F1})$$

Combining these gives [112]

$$(\tan[\phi(\theta) - \theta] + w(\phi(\theta)))\phi'(\theta) = 0. \quad (\text{F2})$$

Therefore, at a location where $r(\theta)$ is smooth, we either have $\phi'(\theta) = 0$, i.e., a straight segment, or

$$\tan[\phi(\theta) - \theta] = -w(\phi(\theta)). \quad (\text{F3})$$

If the solution is smooth everywhere, then the above equation must be satisfied everywhere. (Such solutions exist for sufficiently weakly varying v_B .) It is straightforward to solve this equation in powers of w :

$$-\tan[\phi(\theta) - \theta] = w(\theta) - \frac{1}{2}\partial_\theta w(\theta)^2 + \frac{1}{6}\partial_\theta^2 w(\theta)^3 + \dots \quad (\text{F4})$$

We find that the rhs involves only total derivatives of periodic functions. [Just from looking at Eq. (F3), this is, at first sight, surprising since it emerges from various cancellations.] Therefore, integrating the right-hand side according to Eq. (F1) gives a periodic $r(\theta)$.

For a formal explanation for why the expansion of $\tan(\phi - \theta)$ contains only total derivatives of periodic functions, consider a flow in the space of functions $v_B(\phi)$ that interpolates between the function of interest and the trivial function $v_B(\phi) = \text{const}$. Let $v_1(\phi)$ and $v_2(\phi)$ be two functions that are infinitesimally close on this flow, and let $\phi_1(\theta)$ and $\phi_2(\theta)$ be the corresponding solutions. Assuming that $\phi_1(\theta)$ is periodic and corresponds to a periodic $r(\theta)$, we show that this property is inherited by $\phi_2(\theta)$ to order $\phi_2 - \phi_1$. Using Eqs. (F1) and (F3), we obtain

$$\tan[\phi_1(\theta) - \theta] - \tan[\phi_2(\theta) - \theta] \quad (\text{F5})$$

$$= \partial_\theta[\ln v_2(\phi_1(\theta)) - \ln v_1(\phi_1(\theta))]. \quad (\text{F6})$$

As required, the rhs is indeed the total derivative of a periodic function (note that ϕ_2 does not appear on the rhs). Integrating along the flow then establishes the property for general $v_B(\phi)$ at the formal level—i.e., assuming that the solution evolves smoothly during the flow.

APPENDIX G: HAAR AVERAGE FORMULA

Here, we review a standard formula for the average of matrix elements of the unitary matrix with respect to the Haar probability measure μ on $U(N)$. (See, e.g., Ref. [84].) Let us abbreviate $\int_{U(N)} d\mu(U)$ as \mathbb{E}_U . We are going to prove that

$$\mathbb{E}_U U|a\rangle\langle b|U^\dagger \otimes U|c\rangle\langle d|U^\dagger \quad (\text{G1})$$

$$= \sum_{s=\pm} \frac{I + sF}{2N(N + s1)} (\delta_{ab}\delta_{cd} + s\delta_{cb}\delta_{ad}), \quad (\text{G2})$$

where F is the swap operator on $(\mathbb{C}^N)^{\otimes 2}$. Evaluating a particular matrix element, we have

$$\begin{aligned} & \mathbb{E}_U U_{a'a} U_{b'b}^* U_{c'c} U_{d'd}^* \\ &= \frac{1}{N^2 - 1} \left[\delta_{a'b'} \delta_{c'd'} \delta_{ab} \delta_{cd} + \delta_{a'd'} \delta_{b'c'} \delta_{ad} \delta_{bc} \right. \\ & \quad \left. - \frac{1}{N} (\delta_{ab} \delta_{cd} \delta_{a'd'} \delta_{b'c'} + \delta_{a'b'} \delta_{c'd'} \delta_{ad} \delta_{bc}) \right]. \quad (\text{G3}) \end{aligned}$$

Proof of Eq. (G2).—The average is a matrix on $\mathcal{H} = (\mathbb{C}^N)^{\otimes 2}$ that commutes with every $U^{\otimes 2}$. Hence, the average is block diagonal in the basis where the representation of $U(d)$ is block diagonal. The irreps appearing in \mathcal{H} are the symmetric subspace and the antisymmetric subspace. In each irrep, the average must be proportional to the identity I_\pm by Schur’s lemma, and we need to evaluate the trace in order to determine the constant of proportionality. The projection onto the (anti)symmetric subspace is $(I \pm F)/2$, where F is the swap operator: $F|ac\rangle = |ca\rangle$. So the trace is

$$\begin{aligned} & \frac{1}{2} \text{Tr}[U|a\rangle\langle b|U^\dagger \otimes U|c\rangle\langle d|U^\dagger \pm U|c\rangle\langle b|U^\dagger \otimes U|a\rangle\langle d|U^\dagger] \\ &= \frac{1}{2} [\delta_{ab}\delta_{cd} \pm \delta_{cb}\delta_{ad}]. \quad (\text{G4}) \end{aligned}$$

This must be equal to $C_\pm \text{Tr}(I_\pm) = C_\pm N(N \pm 1)/2$. Therefore, the average is equal to $\sum_{s=\pm} C_s (I + sF)/2$. \square

APPENDIX H: PROOF OF EQ. (81)

Let N be the Hilbert space dimension of n q -dimensional qudits: $N = q^n$. For any $N \times N$ unitary U , denote by $U^{\otimes t,t}$

the tensor product $(U \otimes U^*)^{\otimes t}$, where U^* is the complex conjugate of U . Let μ be the Haar probability distribution on $U(N)$, and define, for any probability distribution ν on $U(N)$, a real number

$$g(\nu, t) = \left\| \underbrace{\mathbb{E}_{U \sim \nu} U^{\otimes t, t}}_{\Pi_\nu} - \underbrace{\mathbb{E}_{U \sim \mu} U^{\otimes t, t}}_{\Pi_\mu} \right\|_\infty. \quad (\text{H1})$$

Here, $\|\cdot\|_\infty$ denotes the maximum singular value. Because of the left and right invariance of μ , it follows that $\Pi_\mu^2 = \Pi_\mu = \Pi_\nu \Pi_\mu = \Pi_\mu \Pi_\nu$. (Π_ν is not, in general, a projector.) Therefore,

$$g(\nu^{*m}, t) = \|\Pi_\nu^m - \Pi_\mu\|_\infty = \|(\Pi_\nu - \Pi_\mu)^m\|_\infty = g(\nu, t)^m, \quad (\text{H2})$$

where ν^{*m} is the m -fold convolution of ν ; i.e., ν^{*m} is the distribution of the product $U_1 U_2 \dots U_m$ when every U_i obeys distribution ν .

Now, let ν be the distribution on $U(N)$ obtained by applying one layer of even-bond local Haar-random unitaries $[U(q^2)]$ and then one layer of odd-bond local Haar-random unitaries. Brandao-Harrow-Horodecki's result [24] implies that

$$g(\nu, t) \leq \exp(-1/M_{t,q}), \quad (\text{H3})$$

$$M_{t,q} = 4250 \lceil \log_q(4t) \rceil^2 q^2 t^5 t^{3.1/\log q}. \quad (\text{H4})$$

Their theorem does not directly cover this, but they have lemmas that are good enough for our purpose; Eq. (48) of Ref. [24] is what we actually need.

Consider $f(U) = N^{-2}(\text{Tr} U X U^\dagger Y U X U^\dagger Y)^2 \geq 0$, where all the matrices U, X, Y are $N \times N$. Here, $f(U)$ can be thought of as $\langle \tilde{X} | U^{\otimes 4,4} | \tilde{Y} \rangle$ for some vectors $|\tilde{X}\rangle$ and $|\tilde{Y}\rangle$. Assume $\text{Tr}(X) = \text{Tr}(Y) = 0$, but $\text{Tr}(X^2) = \text{Tr}(Y^2) = N$. Then, the Euclidean norms of $|\tilde{X}\rangle$ and $|\tilde{Y}\rangle$ are both N^2 . Normalizing so that $|X\rangle := |\tilde{X}\rangle/N^2$ and $|Y\rangle := |\tilde{Y}\rangle/N^2$, we can write $f(U) = N^2 \langle X | U^{\otimes 4,4} | Y \rangle$.

By Eqs. (H2) and (H3), we have

$$|\mathbb{E}_{U \sim \nu^{*m}} f(U) - \mathbb{E}_{U \sim \mu} f(U)| \leq e^{-m/M_{4,q}} q^{2n}. \quad (\text{H5})$$

If $|\mathbb{E}_{U \sim \mu} f(U)| \leq q^{-cn}$, then $\mathbb{E}_{U \sim \nu^{*m}} f(U) \leq 2q^{-cn}$ whenever $m/n \geq (c+2)M_{4,q} \log q$.

Hastings' Schwinger-Dyson trick [119] gives

$$\mathbb{E}_{U \sim \mu} f(U) \leq 10N^{-2} = 10q^{-2n}. \quad (\text{H6})$$

Therefore, whenever $m/n \geq 4M_{4,q} \log q$, we have

$$\mathbb{E}_{U \sim \nu^{*m}} f(U) \leq 11q^{-2n}. \quad (\text{H7})$$

APPENDIX I: MEAN FIELD APPROXIMATION

Reference [31] argued, on the basis of Keldysh perturbation theory, that in various circumstances, the out-of-time-order correlator would satisfy a traveling wave equation such as the Fisher-KPP equation (the details of this equation depend on the physical system). An example is the Fisher-KPP equation itself:

$$\partial_t \mathcal{C} = D \nabla^2 \mathcal{C} + \lambda \mathcal{C}(1 - \mathcal{C}). \quad (\text{I1})$$

The key feature is the λ term, which means that if \mathcal{C} is ‘‘seeded’’ with a small nonzero value, it will increase to a value close to 1 on a timescale of order λ^{-1} (and then saturate). This equation has stable solutions describing a front propagating with a speed $v_B = 2\sqrt{D\lambda}$. This front does not broaden.

This phenomenology is very different from the picture that we have obtained from the random circuit and the mapping to classical growth processes. Recall that in 1D, we related \mathcal{C} to a homogeneous (linear) equation, and in higher dimensions, we found that \mathcal{C} was not governed by a partial differential equation.

The purpose of this appendix is to show that a traveling wave picture can emerge from our mappings if we make a certain mean field approximation. This mean field approximation is not valid in the systems we have studied—it is an uncontrolled approximation, which does not capture the true behavior at either short or large times. However, in variant models, a small parameter could be present that justifies the mean field approximation up to some finite but large timescale. In this situation, we expect that the mean field will nevertheless break down at asymptotically long times, with the front eventually roughening in the manner discussed in the text.

Recall that for the random circuit, we have

$$\bar{\mathcal{C}}(x, t) = \frac{q^2}{q^2 - 1} \langle n(x, t) \rangle, \quad (\text{I2})$$

where $n(x, t)$ is the occupation number in the fictitious classical cluster growth problem. Let us consider the joint probability distribution $P(\{n\}; t)$ for this occupation number. This distribution involves nontrivial correlations between sites, which are crucial for capturing the correct asymptotic behavior. Nevertheless, let us explore the mean field approximation in which we pretend all sites are independent, $P(\{n\}; t) = \prod_x P_x(n(x); t)$, with

$$P_x(n(x); t) = [1 - \langle n(x, t) \rangle] \delta_{n(x), 0} + \langle n(x, t) \rangle \delta_{n(x), 1}. \quad (\text{I3})$$

For simplicity, consider a model on the hypercubic lattice in d dimensions (with coordination number $z = 2d$) in which unitaries (‘‘updates’’) are applied to bonds in a Poissonian fashion at rate $\Gamma/2$ per bond. This continuous-time protocol does not change the basic point, but it simplifies the equations. Write $m = \langle n \rangle$. Note that if we update a bond

that contains at least one fictitious particle, the subsequent (conditionally) averaged density on that bond is $1 - p$. This implies

$$m(x, t + \Delta t) = (1 - z\Gamma\Delta t)m(x, t) + \Gamma\Delta t(1 - p) \sum_{y \in x} \langle (1 - \delta_{n(x),0} \delta_{n(y),0}) \rangle, \quad (I4)$$

where the first term is the probability that site x does not receive an update in the interval Δt . Making the mean field approximation, $\langle \delta_{n(x),0} \delta_{n(y),0} \rangle$ factorizes into $(1 - m(x))(1 - m(y))$, so

$$\partial_t m(x, t) = \Gamma \sum_{y \in x} (-pm(x, t) + (1 - p)m(y, t) \quad (I5)$$

$$- (1 - p)m(x, t)m(y, t)). \quad (I6)$$

The first term on the right is a “death rate,” the second term is the spreading, and the third term is a correction to overcounting in the second term. An analogous equation could be written down for the regular circuit considered in the main text, but we would have to use discrete time.

Equation (I5) is a lattice traveling wave equation. This is most apparent if we make a formal expansion in the lattice spacing a to second order (which is valid, given the approximations already made, if the solution is slowly varying). Recalling $p = 1/(q^2 + 1)$ and Eq. (I2),

$$\Gamma^{-1} \partial_t \bar{c}(x, t) = a^2((1 - p) - (1 - 2p)\bar{c})\nabla^2 \bar{c} \quad (I7)$$

$$+ 2d(1 - 2p)\bar{c}(1 - \bar{c}). \quad (I8)$$

This differs from the Fisher-KPP equation only in the C dependence of the diffusion constant, and we expect similar properties.

Above, the mean field limit is an unjustified formal approximation. We could, of course, construct random circuit models in which the (lattice) mean field approximation is quantitatively accurate up to a large time, for example, by using long-range interactions or a large coordination number to reduce the effect of correlations. However, at long times, in physical dimensionalities, we expect the front to roughen, so the mean field traveling wave picture breaks down. In (unphysically) high dimensions, the mean field may be valid even at late times (recall that the phase diagram of the KPZ equation allows for a nonroughening phase in high dimensions, as discussed in the text).

[1] 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.

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

[3] P. Calabrese and J. Cardy, *Quantum Quenches in 1 + 1-Dimensional Conformal Field Theories*, *J. Stat. Mech.* (2016) 064003.

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

[5] L. Zhang, H. Kim, and D. A. Huse, *Thermalization of Entanglement*, *Phys. Rev. E* **91**, 062128 (2015).

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

[7] C. T. Asplund, A. Bernamonti, F. Galli, and T. Hartman, *Entanglement Scrambling in 2D Conformal Field Theory*, *J. High Energy Phys.* **09** (2015) 110.

[8] H. Casini, H. Liu, and M. Mezei, *Spread of Entanglement and Causality*, *J. High Energy Phys.* **07** (2016) 077.

[9] W. W. Ho and D. A. Abanin, *Entanglement Dynamics in Quantum Many-Body Systems*, *Phys. Rev. B* **95**, 094302 (2017).

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

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

[12] M. Serbyn, Z. Papić, and D. A. Abanin, *Universal Slow Growth of Entanglement in Interacting Strongly Disordered Systems*, *Phys. Rev. Lett.* **110**, 260601 (2013).

[13] D. A. Huse, R. Nandkishore, and V. Oganesyan, *Phenomenology of Fully Many-Body-Localized Systems*, *Phys. Rev. B* **90**, 174202 (2014).

[14] R. Vosk and E. Altman, *Many-Body Localization in One Dimension as a Dynamical Renormalization Group Fixed Point*, *Phys. Rev. Lett.* **110**, 067204 (2013).

[15] 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).

[16] P. Hayden and J. Preskill, *Black Holes as Mirrors: Quantum Information in Random Subsystems*, *J. High Energy Phys.* **09** (2007) 120.

[17] Y. Sekino and L. Susskind, *Fast Scramblers*, *J. High Energy Phys.* **10** (2008) 065.

[18] W. Brown and O. Fawzi, *Scrambling Speed of Random Quantum Circuits*, arXiv:1210.6644.

[19] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, *Towards the Fast Scrambling Conjecture*, *J. High Energy Phys.* **04** (2013) 022.

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

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

[22] O. C. O. Dahlsten, R. Oliveira, and M. B. Plenio, *The Emergence of Typical Entanglement in Two-Party Random Processes*, *J. Phys. A* **40**, 8081 (2007).

- [23] M. Žnidarič, *Exact Convergence Times for Generation of Random Bipartite Entanglement*, *Phys. Rev. A* **78**, 032324 (2008).
- [24] F. G. S. L. Brandao, A. W. Harrow, and M. Horodecki, *Local Random Quantum Circuits Are Approximate Polynomial-Designs*, *Commun. Math. Phys.* **346**, 397 (2016).
- [25] S. H. Shenker and D. Stanford, *Stringy Effects in Scrambling*, *J. High Energy Phys.* **05** (2015) 132.
- [26] A. I. Larkin and Yu. N. Ovchinnikov, *Quasiclassical Method in the Theory of Superconductivity*, *Sov. Phys. JETP* **28**, 1200 (1969).
- [27] A. Kitaev, *Hidden Correlations in the Hawking Radiation and Thermal Noise*, in *The Fundamental Physics Prize Symposium* (2014).
- [28] S. H. Shenker and D. Stanford, *Black Holes and the Butterfly Effect*, *J. High Energy Phys.* **03** (2014) 067.
- [29] S. H. Shenker and D. Stanford, *Multiple Shocks*, *J. High Energy Phys.* **12** (2014) 046.
- [30] J. Maldacena, S. H. Shenker, and D. Stanford, *A Bound on Chaos*, *J. High Energy Phys.* **08** (2016) 106.
- [31] 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).
- [32] D. Stanford, *Many-Body Chaos at Weak Coupling*, *J. High Energy Phys.* **10** (2016) 009.
- [33] A. A. Patel and S. Sachdev, *Quantum Chaos on a Critical Fermi Surface*, *Proc. Natl. Acad. Sci. U.S.A.* **114**, 1844 (2017).
- [34] S. Banerjee and E. Altman, *Solvable Model for a Dynamical Quantum Phase Transition from Fast to Slow Scrambling*, *Phys. Rev. B* **95**, 134302 (2017).
- [35] Y. Gu, X.-L. Qi, and D. Stanford, *Local Criticality, Diffusion and Chaos in Generalized Sachdev-Ye-Kitaev Models*, *J. High Energy Phys.* **05** (2017) 125.
- [36] D. A. Roberts and B. Swingle, *Lieb-Robinson Bound and the Butterfly Effect in Quantum Field Theories*, *Phys. Rev. Lett.* **117**, 091602 (2016).
- [37] D. Chowdhury and B. Swingle, *Onset of Many-Body Chaos in the $O(N)$ Model*, *Phys. Rev. D* **96**, 065005 (2017).
- [38] A. A. Patel, D. Chowdhury, S. Sachdev, and B. Swingle, *Quantum Butterfly Effect in Weakly Interacting Diffusive Metals*, *Phys. Rev. X* **7**, 031047 (2017).
- [39] B. Swingle and D. Chowdhury, *Slow Scrambling in Disordered Quantum Systems*, *Phys. Rev. B* **95**, 060201 (2017).
- [40] A. Nahum, J. Ruhman, and D. Huse, *Dynamics of Entanglement and Transport in 1D Systems with Quenched Randomness*, [arXiv:1705.10364](https://arxiv.org/abs/1705.10364).
- [41] Y. Huang, Y.-L. Zhang, and X. Chen, *Out-of-Time-Ordered Correlators in Many-Body Localized Systems*, *Ann. Phys. (Berlin)* **529**, 1600318 (2017).
- [42] X. Chen, T. Zhou, D. A. Huse, and E. Fradkin, *Out-of-Time-Order Correlations in Many-Body Localized and Thermal Phases*, *Ann. Phys. (Berlin)* **529**, 1600332 (2017).
- [43] Y. Chen, *Universal Logarithmic Scrambling in Many Body Localization*, [arXiv:1608.02765](https://arxiv.org/abs/1608.02765).
- [44] R.-Q. He and Z.-Y. Lu, *Characterizing Many-Body Localization by Out-of-Time-Ordered Correlation*, *Phys. Rev. B* **95**, 054201 (2017).
- [45] R. Fan, P. Zhang, H. Shen, and H. Zhai, *Out-of-Time-Order Correlation for Many-Body Localization*, *Science bulletin* **62**, 707 (2017).
- [46] A. Chapman and A. Miyake, *Classical Simulation of Quantum Circuits by Dynamical Localization: Analytic Results for Pauli-Observable Propagation in Time-Dependent Disorder*, [arXiv:1704.04405](https://arxiv.org/abs/1704.04405).
- [47] D. A. Roberts and D. Stanford, *Diagnosing Chaos Using Four-Point Functions in Two-Dimensional Conformal Field Theory*, *Phys. Rev. Lett.* **115**, 131603 (2015).
- [48] B. Dóra and R. Moessner, *Out-of-Time-Ordered Density Correlators in Luttinger Liquids*, *Phys. Rev. Lett.* **119**, 026802 (2017).
- [49] A. Bohrdt, C. B. Mendl, M. Endres, and M. Knap, *Scrambling and Thermalization in a Diffusive Quantum Many-Body System*, *New J. Phys.* **19**, 063001 (2017).
- [50] D. J. Luitz and Y. B. Lev, *Information Propagation in Isolated Quantum Systems*, *Phys. Rev. B* **96**, 020406 (2017).
- [51] E. Leviatan, F. Pollmann, J. H. Bardarson, and E. Altman, *Quantum Thermalization Dynamics with Matrix-Product States*, [arXiv:1702.08894](https://arxiv.org/abs/1702.08894).
- [52] G. Zhu, M. Hafezi, and T. Grover, *Measurement of Many-Body Chaos Using a Quantum Clock*, *Phys. Rev. A* **94**, 062329 (2016).
- [53] B. Swingle, G. Bentsen, M. Schleier-Smith, and P. Hayden, *Measuring the Scrambling of Quantum Information*, *Phys. Rev. A* **94**, 040302 (2016).
- [54] N. Y. Yao, F. Grusdt, B. Swingle, M. D. Lukin, D. M. Stamper-Kurn, J. E. Moore, and E. A. Demler, *Interferometric Approach to Probing Fast Scrambling*, [arXiv:1607.01801](https://arxiv.org/abs/1607.01801).
- [55] M. Gärtner, J. G. Bohnet, A. Safavi-Naini, M. L. Wall, J. J. Bollinger, and A. M. Rey, *Measuring Out-of-Time-Order Correlations and Multiple Quantum Spectra in a Trapped Ion Quantum Magnet*, [arXiv:1608.08938](https://arxiv.org/abs/1608.08938).
- [56] J. Li, R. Fan, H. Wang, B. Ye, B. Zeng, H. Zhai, X. Peng, and J. Du, *Measuring Out-of-Time-Order Correlators on a Nuclear Magnetic Resonance Quantum Simulator*, *Phys. Rev. X* **7**, 031011 (2017).
- [57] K. X. Wei, C. Ramanathan, and P. Cappellaro, *Exploring Localization in Nuclear Spin Chains*, *Phys. Rev. Lett.* **120**, 070501 (2018).
- [58] D. A. Roberts, D. Stanford, and L. Susskind, *Localized Shocks*, *J. High Energy Phys.* **03** (2015) 051.
- [59] Strongly disordered Hamiltonians in 1 + 1D provide counterexamples to this ballistic spreading.
- [60] M. Kardar, G. Parisi, and Y.-C. Zhang, *Dynamic Scaling of Growing Interfaces*, *Phys. Rev. Lett.* **56**, 889 (1986).
- [61] A. Pagnani and G. Parisi, *Numerical Estimate of the Kardar-Parisi-Zhang Universality Class in $(2 + 1)$ Dimensions*, *Phys. Rev. E* **92**, 010101 (2015).
- [62] The phase diagram of the KPZ equation in higher dimensions [60] indicates that in 4 + 1D and above, two distinct universality classes may be possible for operator growth in a random circuit, one with a growing front width and one without.
- [63] K. Johansson, *Shape Fluctuations and Random Matrices*, *Commun. Math. Phys.* **209**, 437 (2000).

- [64] M. Prähofer and H. Spohn, *Universal Distributions for Growth Processes in 1 + 1 Dimensions and Random Matrices*, *Phys. Rev. Lett.* **84**, 4882 (2000).
- [65] P. Calabrese, P. Le Doussal, and A. Rosso, *Free-Energy Distribution of the Directed Polymer at High Temperature*, *Europhys. Lett.* **90**, 20002 (2010).
- [66] V. Dotsenko, *Bethe Ansatz Derivation of the Tracy-Widom Distribution for One-Dimensional Directed Polymers*, *Europhys. Lett.* **90**, 20003 (2010).
- [67] T. Sasamoto and H. Spohn, *One-Dimensional Kardar-Parisi-Zhang Equation: An Exact Solution and Its Universality*, *Phys. Rev. Lett.* **104**, 230602 (2010).
- [68] T. Sasamoto and H. Spohn, *Exact Height Distributions for the KPZ Equation with Narrow Wedge Initial Condition*, *Nucl. Phys.* **B834**, 523 (2010).
- [69] T. Sasamoto and H. Spohn, *The Crossover Regime for the Weakly Asymmetric Simple Exclusion Process*, *J. Stat. Phys.* **140**, 209 (2010).
- [70] G. Amir, I. Corwin, and J. Quastel, *Probability Distribution of the Free Energy of the Continuum Directed Random Polymer in 1 + 1 Dimensions*, *Commun. Pure Appl. Math.* **64**, 466 (2011).
- [71] P. Calabrese and P. Le Doussal, *Exact Solution for the Kardar-Parisi-Zhang Equation with Flat Initial Conditions*, *Phys. Rev. Lett.* **106**, 250603 (2011).
- [72] S. Prolhac and H. Spohn, *Height Distribution of the Kardar-Parisi-Zhang Equation with Sharp-Wedge Initial Condition: Numerical Evaluations*, *Phys. Rev. E* **84**, 011119 (2011).
- [73] P. Le Doussal and P. Calabrese, *The KPZ Equation with Flat Initial Condition and the Directed Polymer with One Free End*, *J. Stat. Mech.* (2012) P06001.
- [74] T. Imamura and T. Sasamoto, *Exact Solution for the Stationary Kardar-Parisi-Zhang Equation*, *Phys. Rev. Lett.* **108**, 190603 (2012).
- [75] T. Imamura and T. Sasamoto, *Stationary Correlations for the 1D KPZ Equation*, *J. Stat. Phys.* **150**, 908 (2013).
- [76] In other words, the shape of the spatial region in which C has saturated. Here, we can neglect the broadening of the front since, at late times, the length scale associated with this broadening is parametrically smaller than the size of the operator.
- [77] D. Gottesman, *The Heisenberg Representation of Quantum Computers*, [arXiv:quant-ph/9807006](https://arxiv.org/abs/quant-ph/9807006).
- [78] J. Gütschow, S. Uphoff, R. F. Werner, and Z. Zimborás, *Time Asymptotics and Entanglement Generation of Clifford Quantum Cellular Automata*, *J. Math. Phys.* (N.Y.) **51**, 015203 (2010).
- [79] D. P. DiVincenzo, D. W. Leung, and B. M. Terhal, *Quantum Data Hiding*, *IEEE Trans. Inf. Theory* **48**, 580 (2002).
- [80] 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.
- [81] P. Calabrese and J. Cardy, *Entanglement Entropy and Conformal Field Theory*, *J. Phys. A* **42**, 504005 (2009).
- [82] W. W. Ho and D. A. Abanin, *Entanglement Dynamics in Quantum Many-Body Systems*, *Phys. Rev. B* **95**, 094302 (2017).
- [83] A. Hamma, S. Santra, and P. Zanardi, *Quantum Entanglement in Random Physical States*, *Phys. Rev. Lett.* **109**, 040502 (2012).
- [84] A. Hamma, S. Santra, and P. Zanardi, *Ensembles of Physical States and Random Quantum Circuits on Graphs*, *Phys. Rev. A* **86**, 052324 (2012).
- [85] P. Zanardi, *Local random quantum circuits: Ensemble completely positive maps and swap algebras*, *J. Math. Phys.* (N.Y.) **55**, 082204 (2014).
- [86] For general q , one has to start with an operator basis that obeys our normalization condition in Eq. (12). It is easy to construct such a basis. Define $X = \sum_{k \in \mathbb{Z}_q} |k+1\rangle\langle k|$ and $Z = \sum_{k \in \mathbb{Z}_q} e^{2\pi i k/q} |k\rangle\langle k|$. Then, the discrete group generated by these two matrices contains exactly q^2 elements up to unimportant phase factors. These are not Hermitian, but no problem arises if one considers $|a_S|^2$. Over Haar-random unitaries, one easily obtains $\overline{C(x,t)} = [q^2/(q^2-1)]\overline{\mu(x,t)}$.
- [87] To find μ_0 , make the ansatz that each $\mu(x)$ is independent from $\mu(x')$ for $x \neq x'$. Under this ansatz, the probability that a pair of sites is partially or fully occupied is $1 - (1 - \mu_0)^2$, and such an occupied pair evolves to fill a given site of the pair with probability $1 - p$. Therefore, setting $\mu_0 = (1 - p)(2\mu_0 - \mu_0^2) = [(q^2 - 1)/q^2]$ yields the stationary state.
- [88] L. Saul, M. Kardar, and N. Read, *Directed Waves in Random Media*, *Phys. Rev. A* **45**, 8859 (1992).
- [89] All the statements here hold for any prime power q such as $q = 2, 3, 4, 5, 7, 8, 9, 11, 13, 16, \dots$
- [90] C. A. Tracy and H. Widom, *Level-Spacing Distributions and the Airy Kernel*, *Commun. Math. Phys.* **159**, 151 (1994).
- [91] P. J. Forrester, *The Spectrum Edge of Random Matrix Ensembles*, *Nucl. Phys.* **B402**, 709 (1993).
- [92] M. Eden, *A Two-Dimensional Growth Process*, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 4: Contributions to Biology and Problems of Medicine* (University of California Press, Berkeley, California, 1961), pp. 223–239.
- [93] A.-L. Barabási and H. E. Stanley, *Fractal Concepts in Surface Growth* (Cambridge University Press, Cambridge, England, 1995).
- [94] D. Richardson, *Random Growth in a Tessellation*, in *Mathematical Proceedings of the Cambridge Philosophical Society*, Vol. 74 (Cambridge University Press, Cambridge, England, 1973), pp. 515–528.
- [95] R. Durrett and T. M. Liggett, *The Shape of the Limit Set in Richardson's Growth Model*, *Ann. Probab.* **9**, 186 (1981).
- [96] R. Savit and R. Ziff, *Morphology of a Class of Kinetic Growth Models*, *Phys. Rev. Lett.* **55**, 2515 (1985).
- [97] J. Kertész and D. E. Wolf, *Anomalous Roughening in Growth Processes*, *Phys. Rev. Lett.* **62**, 2571 (1989).
- [98] J. Krug, J. Kertész, and D. E. Wolf, *Growth Shapes and Directed Percolation*, *Europhys. Lett.* **12**, 113 (1990).
- [99] The effect disappears for smaller q . For example, for $q = 1.4$, we see clear KPZ growth both at $\theta = 0$ (fitted exponent value $\beta = 0.3223 \pm 0.0199$) and at $\theta = \pi/4$ ($\beta = 0.3304 \pm 0.0149$).

- [100] T. Sasamoto and H. Spohn, *The 1 + 1-Dimensional Kardar-Parisi-Zhang Equation and Its Universality Class*, *J. Stat. Mech.* (2010) P11013.
- [101] T. Kriecherbauer and J. Krug, *A Pedestrian's View on Interacting Particle Systems, KPZ Universality and Random Matrices*, *J. Phys. A* **43**, 403001 (2010).
- [102] I. Corwin, *The Kardar-Parisi-Zhang Equation and Universality Class*, *Random Matrices Theory Applications* **01**, 1130001 (2012).
- [103] T. Halpin-Healy and K. A. Takeuchi, *A KPZ Cocktail-Shaken, Not Stirred...*, *J. Stat. Phys.* **160**, 794 (2015).
- [104] K. A. Takeuchi, M. Sano, T. Sasamoto, and H. Spohn, *Growing Interfaces Uncover Universal Fluctuations Behind Scale Invariance*, *Sci. Rep.* **1**, 34 (2011).
- [105] K. A. Takeuchi and M. Sano, *Evidence for Geometry-Dependent Universal Fluctuations of the Kardar-Parisi-Zhang Interfaces in Liquid-Crystal Turbulence*, *J. Stat. Phys.* **147**, 853 (2012).
- [106] P. Deift, A. Its, and I. Krasovsky, *Asymptotics of the Airy-Kernel Determinant*, *Commun. Math. Phys.* **278**, 643 (2008).
- [107] J. Baik, R. Buckingham, and J. DiFranco, *Asymptotics of Tracy-Widom Distributions and the Total Integral of a Painlevé II Function*, *Commun. Math. Phys.* **280**, 463 (2008).
- [108] In this scaling limit, the width of the front is negligible.
- [109] P. Freche, D. Stauffer, and H. E. Stanley, *Surface Structure and Anisotropy of Eden Clusters*, *J. Phys. A* **18**, L1163 (1985).
- [110] R. Hirsch and D. E. Wolf, *Anisotropy and Scaling of Eden Clusters in Two and Three Dimensions*, *J. Phys. A* **19**, L251 (1986).
- [111] P. Meakin, *Noise-Reduced and Anisotropy-Enhanced Eden and Screened-Growth Models*, *Phys. Rev. A* **38**, 418 (1988).
- [112] D. E. Wolf, *Wulff Construction and Anisotropic Surface Properties of Two-Dimensional Eden Clusters*, *J. Phys. A* **20**, 1251 (1987).
- [113] J. Krug, H. Spohn, and C. Godrèche, *Solids Far from Equilibrium* (Cambridge University Press, Cambridge, England, 1991).
- [114] J. Krug and H. Spohn, *Universality Classes for Deterministic Surface Growth*, *Phys. Rev. A* **38**, 4271 (1988).
- [115] This is seen by following the evolution of a front through the four layers comprising the time step.
- [116] T. Prosen, *Chaos and Complexity of Quantum Motion*, *J. Phys. A* **40**, 7881 (2007).
- [117] T. Prosen and I. Pižorn, *Operator Space Entanglement Entropy in a Transverse Ising Chain*, *Phys. Rev. A* **76**, 032316 (2007).
- [118] C. Jonay, D. A. Huse, and A. Nahum, *Coarse-Grained Dynamics of Operator and State Entanglement*, arXiv:1803.00089.
- [119] M. B. Hastings, *Random Unitaries Give Quantum Expanders*, *Phys. Rev. A* **76**, 032315 (2007).
- [120] P. Zanardi, C. Zalka, and L. Faoro, *Entangling Power of Quantum Evolutions*, *Phys. Rev. A* **62**, 030301 (2000).
- [121] Reference [10] argued that the width of the distribution scales as $t^{1/3}$. The mean value is of order t .
- [122] John Cardy, *Scaling and Renormalization in Statistical Physics* (Cambridge University Press, Cambridge, England, 1996), Vol. 5.
- [123] In the limit $q \rightarrow \infty$, it is easy to show that $v_E = v_P$ in the present model.
- [124] This is because the \sqrt{t} fluctuations in the transverse position of the domain wall are negligible compared to t ; compare Ref. [10], where the minimal cut configuration is also deterministic in the scaling limit.
- [125] This was determined analytically for a certain large q model, distinct from that here, and numerically for various circuits composed of Clifford gates.
- [126] Reference [10] obtained a connection between entanglement growth in 1 + 1D random unitary circuits and the KPZ equation. To avoid confusion, we emphasize that the connection with KPZ discussed here is physically entirely distinct from that one and is not even in the same number of spatial dimensions.
- [127] But model dependent above 1 + 1D.
- [128] Circuit-to-circuit fluctuations in the OTOC should not be confused with KPZ fluctuations in the auxiliary classical model.
- [129] T. Halpin-Healy and Y-C. Zhang, *Kinetic Roughening Phenomena, Stochastic Growth, Directed Polymers and All That. Aspects of Multidisciplinary Statistical Mechanics*, *Phys. Rep.* **254**, 215 (1995).
- [130] D. A. Huse and C. L. Henley, *Pinning and Roughening of Domain Walls in Ising Systems Due to Random Impurities*, *Phys. Rev. Lett.* **54**, 2708 (1985).
- [131] D. Forster, D. R. Nelson, and M. J. Stephen, *Large-Distance and Long-Time Properties of a Randomly Stirred Fluid*, *Phys. Rev. A* **16**, 732 (1977).
- [132] J. Baik, P. Deift, and K. Johansson, *On the Distribution of the Length of the Longest Increasing Subsequence of Random Permutations*, *J. Am. Math. Soc.* **12**, 1119 (1999).
- [133] C. A. Tracy and H. Widom, *On the Distributions of the Lengths of the Longest Monotone Subsequences in Random Words*, *Probab. Theory Relat. Fields* **119**, 350 (2001).
- [134] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, *Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws*, preceding article, *Phys. Rev. X* **8**, 021013 (2018).