Locality, Quantum Fluctuations, and Scrambling

Shenglong Xu[1] and Brian Swingle[2]

[1]*Condensed Matter Theory Center and Department of Physics, University of Maryland,*
College Park, Maryland 20742, USA
[2]*Condensed Matter Theory Center, Maryland Center for Fundamental Physics, Joint Center for Quantum*
Information and Computer Science, and Department of Physics, University of Maryland,
College Park, Maryland 20742, USA

Thermalization of chaotic quantum many-body systems under unitary time evolution is related to the growth in complexity of initially simple Heisenberg operators. Operator growth is a manifestation of information scrambling and can be diagnosed by out-of-time-order correlators (OTOCs). However, the behavior of OTOCs of local operators in generic chaotic local Hamiltonians remains poorly understood, with some semiclassical and large-N models exhibiting exponential growth of OTOCs and a sharp chaos wave front and other random circuit models showing a diffusively broadened wave front. In this paper, we propose a unified physical picture for scrambling in chaotic local Hamiltonians. We construct a random time-dependent Hamiltonian model featuring a large-N limit where the OTOC obeys a Fisher-Kolmogorov-Petrovsky-Piskunov (FKPP) type equation and exhibits exponential growth and a sharp wave front. We show that quantum fluctuations manifest as noise (distinct from the randomness of the couplings in the underlying Hamiltonian) in the FKPP equation and that the noise-averaged OTOC exhibits a crossover to a diffusively broadened wave front. At small N, we demonstrate that operator growth dynamics, averaged over the random couplings, can be efficiently simulated for all time using matrix product state techniques. To show that time-dependent randomness is not essential to our conclusions, we push our previous matrix product operator methods to very large size and show that data for a time-independent Hamiltonian model are also consistent with a diffusively broadened wave front.

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

I. INTRODUCTION

Information scrambling describes a process whereby information about the initial condition of a unitarily evolving system spreads over the entire system, becoming inaccessible to any local measurement [1–4]. Because it describes an effective loss of memory, scrambling is relevant for understanding quantum thermalization (e.g., Refs. [5–8]), i.e., the emergence of irreversibility from unitary time evolution, and is also tied to the black hole information problem. Scrambling is also closely related to the dynamics of initially simple Heisenberg operators, with the growth in size and complexity of these operators probing the spreading of quantum information [9–16].

Given two local operators W_0 and V_r at positions 0 and r, the out-of-time-order correlator (OTOC),

$$F(r, t) = \langle W_0^\dagger(t) V_r^\dagger W_0(t) V_r \rangle, \tag{1}$$

provides one way to quantify scrambling by probing how the Heisenberg operator $W_0(t)$ grows with time. Although scrambling can also be usefully characterized in entropic terms, OTOCs are more directly measurable, with early experiments having already been carried out in a variety of platforms [17–27]. A closely related quantity is the squared commutator between V and W, defined as

$$C(r, t) = \langle [W_0(t), V_r]^\dagger [W_0(t), V_r] \rangle = 2[1 - \text{Re}(F)]. \tag{2}$$

The physical picture is that under Heisenberg dynamics, the operator W_0 expands and eventually fails to commute with V_r, as manifested by the growth of $C(r, t)$ from zero. For chaotic local Hamiltonians, $W_0(t)$ is expected to expand ballistically, with speed called the butterfly velocity, so the OTOC exhibits a causal light-cone-like structure in

space-time. The squared commutator remains small outside the light cone and grows rapidly as the boundary of the light cone is crossed. Inside the light cone, $C(r, t)$ saturates for chaotic systems regardless of the specific form of operators W and V.

A particularly interesting question concerns the specific growth form of $C(r, t)$ near the wave front of the light cone. In some models, $C(r, t)$ grows exponentially with time, a phenomenon proposed as a quantum analog of the classical butterfly effect, the exponential divergence of initially nearby trajectories. This observation has led to an emphasis on probing the footprint of quantum chaos at an intermediate timescale, especially with a view towards defining the notion of the quantum Lyapunov exponent. A well-defined Lyapunov exponent λ_L, i.e., purely exponential growth of $C(r, t)$, plus the ballistic growth of the OTOC, implies that the wave front is sharp, and sharp wave fronts have been identified in a broad class of holographic or large-N models, including the O(N) model [28], the diffusive metal [29], and the coupled Sachdev-Ye-Kitaev (SYK) model [30–32]. On the other hand, although significant efforts have been made [33–36], a clear signature of purely exponential growth of the OTOC in more physical systems with finite on-site degrees of freedom (d.o.f.) is absent, and there are some counterexamples in random circuit models [9,10,37–39].

To reconcile the many different scenarios, in a recent paper [11], we proposed a universal form for the early growth region of the squared commutator,

$$C(r, t) \sim \exp\left(-\lambda_p (x/v_B - t)^{1+p}/t^p\right), \qquad (3)$$

assuming that there is a well-defined butterfly velocity v_B (a different ansatz is needed for localized systems [40–43]), which was further studied in [44]. The shape of the wave front is controlled by a single parameter p, denoted as the broadening exponent, associated with the growth rate λ_p. For large-N or holographic models, $p = 0$, and the corresponding λ_p is the Lyapunov exponent. However, an exact calculation in a Haar random brickwork circuit model gives $p = 1$ in one dimension, indicating a diffusive broadening of the wave front. Saddle-point analysis shows $p = \frac{1}{2}$ for general noninteracting systems with translational invariance [11,36,44]. Large-scale matrix product state (MPS) simulations using the time-dependent variational principle [45] and matrix product operator (MPO) simulations [11] also give strong evidence of wave-front broadening for chaotic local Hamiltonian systems.

In this work, we make two contributions to understanding the early growth region behavior of the OTOC. First, to understand the intriguing differences between the large-N models and the Haar random brickwork circuit models, we design and analyze a new random circuit model, denoted as the Brownian coupled cluster model (BCC). BCC, as an extension of the single cluster version

[46,47], describes the dynamics of clusters of N spins connected in a one-dimensional array (or more generally, connected according to any graph), similar to coupled SYK cluster models but with the couplings random in both space and time. We show that in the large-N limit, BCC is similar to other large-N models and has a well-defined Lyapunov exponent, but the finite-N correction qualitatively changes the broadening exponent from $p = 0$ to $p = 1$ in one dimension. We find that finite-N corrections are actually quite dramatic, with the broadening of the wave front characterized by a diffusion constant that scales as $1/\log^3 N$ at large N. We also find that there is a finite region in space-time where the wave front remains sharp, indicating strong finite-size effects on the broadening exponent.

With this new point of view, our second contribution is to push our numerical matrix product operator simulations of operator growth in a local Hamiltonian Ising system to include 200 spins in the wave front and up to 250 in units of the nearest-neighbor Ising coupling. By directly analyzing the way contours of constant C deviate in space-time, we find that the broadening exponent indeed converges to $p = 1$ in the large space-time limit. Therefore, we conclude that diffusive broadening of the wave front is generic for one-dimensional chaotic systems.

In more detail, our analysis of the BCC proceeds by focusing on operator dynamics, suitably averaged over the random couplings in the Hamiltonian. Any operator may be expanded in a complete basis of operators, with the expansion coefficients called operator amplitudes and with the square of the amplitudes forming a probability distribution, the operator probability distribution. This procedure is completely analogous to expanding a time-evolving wave function in a complete basis of states. The starting point of the analysis is the derivation of an equation of motion for the circuit-averaged operator probability distribution of a Heisenberg operator. The effect of averaging over the couplings in the quantum Hamiltonian is to give a closed stochastic equation for the operator probability distribution; physically, the operator amplitudes evolve via unitary time evolution for each choice of couplings, and the averaging dephases this dynamics to yield a master equation for the operator probability distribution. One point should be emphasized: The randomness of the couplings in the Hamiltonian, which we sometimes call "disorder," is physically distinct from the quantum randomness manifested in the operator probability distribution. The latter will, in a certain limit, be instantiated as a random process, which we call "noise." One of the key assertions of this paper is that the disorder average is a technical convenience, while the noise average contains essential physics of quantum fluctuations.

Starting from the master equation for the operator probability distribution, the analysis proceeds from two limits. In the large-N limit, a mean-field-like treatment of the operator distribution becomes exact, and the operator

dynamics can be translated into a closed nonlinear partial differential equation for the operator weight, a measure of the size of the operator that is linearly related to the OTOC. The resulting dynamical equation is similar to the Fisher-Kolmogorov-Petrovsky-Piskunov (FKPP) equation [48,49], which occurs, for example, in studies of combustion waves, invasive species, and quantum chromodynamics, among others; it was recently introduced in the context of scrambling to describe the growth of OTOCs [28,50,51]. The key physical effects embodied by FKPP-type equations are unstable exponential growth, diffusion, and eventual saturation; together, these effects lead to traveling-wave solutions with a sharp wave front, and they describe the spreading of local Heisenberg operators. The leading finite-N correction results in a stochastic partial differential equation, a noisy FKPP-like equation in which the noise is multiplicative and $1/N$ suppressed. Drawing from the noisy FKPP literature [52,53], we argue that the noisy FKPP-like equation in the BCC has a diffusively broadened wave front after averaging over noise. These analytical arguments are verified by direct numerical integration of the large-but-finite-N BCC stochastic equation. It should be emphasized again that the noise in the noisy FKPP-like equation represents quantum fluctuations not different instances of the microscopic couplings.

In the small-N limit, a different analytical treatment shows that the OTOC exhibits the same diffusive broadening as in the Haar random brickwork circuit model. Moreover, we show that by representing the operator probability distribution as a "stochastic" matrix product state, it is possible to numerically solve the master equation for the time dynamics. Thanks to the dephasing provided by the disorder average over couplings, one can show that the late-time operator probability distribution has low "correlation or entanglement" when viewed as a matrix product state. We further find that the operator probability distribution never has high correlation or entanglement, so matrix product state techniques can accurately capture the operator dynamics for all times. A modest bond dimension of $\chi = 32$ is already sufficient to converge the dynamics for 200 sites for all time.

Finally, taking these lessons from the BCC, especially the crucial role of noise, meaning quantum fluctuations, we argue that the diffusive broadening of the operator growth wave front is generic in one dimension. This idea has been previously conjectured based on work with random circuit models [9,10,37–39]. One piece of evidence is direct numerical simulation of the time-independent Hamiltonian dynamics of Heisenberg operators in a system of 200 spins for a very long time. A new analysis of the space-time contours of a constant squared commutator conclusively demonstrates diffusive broadening of the wave front at the largest sizes. Another piece of evidence is the prevalence of noiseless FKPP-like equations describing OTOC dynamics in large-N or weakly coupled models, including linearized

FKPP-like equations obtained in resummed perturbation theory [28] and fully nonlinear FKPP-like equations obtained from self-consistent Keldysh treatments [50]. We argue that, starting from these known results, quantum fluctuations should invariably be described by adding noise, specifically multiplicative noise of the type found in the BCC. Hence, these models will also suffer similar dramatic finite-N effects resulting in diffusively broadened wave fronts.

The remainder of the paper is organized as follows. Section II describes, in detail, the notion of operator dynamics used throughout the paper and their relations to OTOCs. Section III introduces and analyzes the BCC, both at small and large N. Section IV discusses implications of the results for generic Hamiltonian systems. We conclude with an outlook, including the effects of conserved quantities and going beyond one dimension, as well as open questions.

II. GENERAL FORMALISM OF OPERATOR DYNAMICS AND ITS RELATION TO THE OUT-OF-TIME-ORDER CORRELATOR

Consider a generic quantum system consisting of L sites with N spin-1/2's per site. The dynamics of the system is governed by a local unitary circuit $U(t)$. In the Pauli basis, a Heisenberg operator $W(t)$ takes the form

$$W(t) = \sum c(\mathcal{S})\mathcal{S}, \qquad (4)$$

where $\mathcal{S}$ is a product of Pauli operators with length NL and the time dependence is encoded in the coefficients $c(\mathcal{S}, t)$. The normalization is $\mathrm{tr}[W(t)^{\dagger}W(t)] = 2^{NL}$, so $\sum_{\mathcal{S}}|c(\mathcal{S}, t)|^2 = 1$. Under Heisenberg time evolution in a chaotic quantum system, an initially localized operator grows and eventually equilibrates, as far as local probes are concerned. The initial configuration cannot be recovered from local data, a manifestation of scrambling. The OTOC is designed precisely to quantify this process since the deviation of the correlator from its initial value for certain positions indicates that the Heisenberg operator has developed a nontrivial component at that site. To understand the relationship between the OTOC and the operator string picture, we consider the following averaged OTOC,

$$F(r, t) = \frac{1}{4N}\sum_{a,\alpha}\frac{1}{2^{NL}}\mathrm{Tr}(W(t)^{\dagger}\sigma_{r,a}^{\alpha}W(t)\sigma_{r,a}^{\alpha}), \qquad (5)$$

where a runs from 1 to N and α from 0 to 3, representing the identity and the three Pauli matrices.

Using the decoupling channel identity, we obtain that

$$\phi(r, t) = 1 - F(r, t) = \sum |c(\mathcal{S})|^2 \frac{1}{N}\left(\sum_a w(\mathcal{S}_{r,a})\right), \qquad (6)$$

where $\mathcal{S}_{r,a}$ is the r, a Pauli operator in the string $\mathcal{S}$ and the weight is $w(\sigma^{\alpha}) = 1 - \delta_{\alpha 0}$. Therefore, $\phi(r, t)$ measures the

average number of nontrivial local operators within a single spin cluster located at r, growing from 0 and saturating to $3/4$ in a thermalizing system. Similarly, the average squared commutator $C(r, t)$ equals $\frac{8}{3}\phi(r, t)$. There are two implications. First, $C(r, t)$ does not depend on the phase information of the coefficients $c(\mathcal{S})$ [which are, in fact, real for Hermitian $W(t)$] since it is completely determined by the probability distribution $|c(\mathcal{S})|^2$. Second, $C(r, t)$ is not concerned with the specific operator configuration $\mathcal{S}$ since only the number of nontrivial operators matters. Both features significantly simplify the calculations.

In general, dealing with the Schrödinger equation is overcomplicated for the purposes of studying operator dynamics due to the unimportant phase information and associated high operator entanglement entropy. However, deriving a closed set of dynamical equations for the operator probability $|c(\mathcal{S})|^2$ from the Schrödinger equation is difficult. Random circuit models are useful to overcome this difficulty by introducing disorder as a dephasing mechanism, making the probability distribution dynamics tractable both numerically and analytically, for example, in recent studies of the Haar random brickwork circuit. This paper introduces another class of random circuit models, the BCC, describing the dynamics of a system of coupled spin clusters with interaction that is random in both space and time. BCC can be regarded as a smoother version of the random brickwork circuit, in which the interactions are only between pairs of spins (or, more generally, few-body interactions) even in the large-N limit; thus, it is naturally more tied to holographic and SYK models. While in the random brickwork circuit, the dimension of the on-site Hilbert space does not qualitatively affect the operator dynamics, in the BCC, we expect a smooth crossover from the random brickwork circuit result for small N to holographic and SYK physics in the large-N limit.

III. BROWNIAN COUPLED CLUSTER MODEL

Figure 1 shows a schematic of the BCC model. It is best described using discrete time steps dt, with the limit $dt \to 0$ taken later. For small dt, the whole time evolution unitary breaks into pieces,

$$U(t) = \prod_{m=1}^{t/dt} \exp\left(-i\sum_r H_r^{(m)} - i\sum_{rr'} H_{rr'}^{(m)}\right), \quad (7)$$

with m a discrete time index. The on-site terms and the bond terms read

$$H_r^{(m)} = \sum_{a,b,\alpha,\beta} J_{m,r,a,b}^{\alpha\beta}\sigma_{r,a}^{\alpha}\sigma_{r,b}^{\beta},$$

$$H_{rr'}^{(m)} = g\sum_{a,b,\alpha,\beta} \tilde{J}_{m,r,r',a,b}^{\alpha\beta}\sigma_{r,a}^{\alpha}\sigma_{r',b}^{\beta}, \quad (8)$$
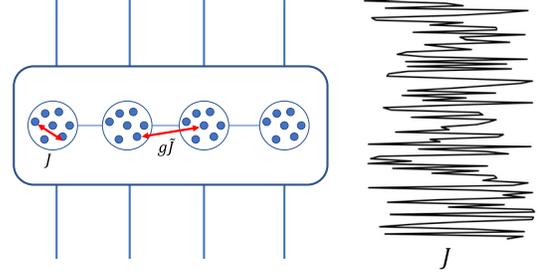

FIG. 1. The Brownian coupled cluster model. Spins within the same cluster interact with each other and also interact with the spins in the neighboring clusters. The intracluster coupling J and the intercluster coupling $\tilde{J}$ are random in both space and time.

where α and β are the Pauli matrix indices running from 0 to 3 (including the identity, for convenience), a and b from 1 to N label the spins in the cluster, r and r' label clusters (sometimes called sites), and $\langle rr' \rangle$ stands for nearest neighbors. At each time step, the models contains two sets of uncorrelated random variables J and $\tilde{J}$ with mean zero and variance $\{1/[8(N-1)]\}dt$ and $[1/(16N)]dt$, respectively.

With the help of the random couplings, one can derive a master equation for the averaged probability distribution $h = \overline{|c(\mathcal{S})|^2}$. To simplify the calculation, we assume that h only depends on the operator weight $w_r = \sum_a w(\mathcal{S}_{r,a})$ of each cluster instead of on the details of the operator configuration $\mathcal{S}$. This approximation is valid after a short relaxation time even though W starts as a specific operator. To proceed further, introduce the operator weight probability

$$\tilde{h}(\mathbf{w}) = h(\mathbf{w})D(\mathbf{w}), \quad (9)$$

with D the number of operators with weight configuration $\mathbf{w}$,

$$D(\mathbf{w}) = \prod_r \binom{N}{w}_r 3^{w_r}. \quad (10)$$

The operator weight probability is a properly normalized probability distribution over the $(N+1)^L$ possible weight strings.

The derivation of the master equation for $\tilde{h}$ is recorded in the Appendix A, with the result being

$$\partial_t \tilde{h} = \sum_r [-\gamma_r^+(w)\tilde{h}(\mathbf{w}) + \gamma_r^+(w+1)\tilde{h}(\mathbf{w} + \mathbf{e}_r)]$$

$$+ [-\gamma_r^-(w)\tilde{h}(\mathbf{w}) + \gamma_r^-(w-1)\tilde{h}(\mathbf{w} - \mathbf{e}_r)]$$

$$+ \sum_{\langle rr' \rangle} [-\gamma_b^+(w_r, w_{r'})\tilde{h}(\mathbf{w}) + \gamma_b^+(w_r+1, w_{r'})\tilde{h}(\mathbf{w} + \mathbf{e}_r)]$$

$$+ [-\gamma_b^-(w_r, w_{r'})\tilde{h}(\mathbf{w}) + \gamma_b^-(w_r-1, w_{r'})\tilde{h}(\mathbf{w} - \mathbf{e}_r)]$$

$$+ [r \leftrightarrow r']. \quad (11)$$

The evolution equation manifestly conserves the total probability, $\sum_{\{\mathbf{w}\}} \tilde{h} = 1$ for all time, independent of the specific form of the functions γ^+ and γ^-. For this particular problem, these functions are

$$\gamma_r^+(w) = \frac{1}{N-1}w(w-1), \quad \gamma_r^-(w) = \frac{1}{N-1}3(N-w)w,$$

$$\gamma_b^+(w_1, w_2) = \frac{g^2}{2N}w_1 w_2, \quad \gamma_b^-(w_1, w_2) = \frac{g^2}{2N}3(N-w_1)w_2.$$

(12)

In the following subsections, we analyze this master equation in the infinite-N limit, study its large-N expansion, and compare the result with small-N results. Complementing these analytical results are numerical simulations of the master equation for 200 spin clusters using tensor network methods.

A. Infinite-N limit

In the infinite-N limit, the master equation (11) can be approximated by a Fokker-Planck equation,

$$\partial_t \tilde{h} = \sum_r \partial_{\phi_r}[\alpha(\phi_r)\tilde{h}(\phi)] + \partial_{\phi_r}^2[\beta(\phi_r)\tilde{h}(\phi)] + \mathcal{O}\left(\frac{1}{N^2}\right)$$

$$\alpha(\phi_r) = (4\phi_r - 3)\left(\phi_r + \frac{g^2}{2}(\phi_{r-1} + \phi_{r+1})\right) + \frac{1}{N}4(\phi_r - 1)\phi_r$$

$$\beta(\phi_r) = \frac{1}{4N}(3 - 2\phi_r)[2\phi_r + g^2(\phi_{r-1} + \phi_{r+1})],$$

(13)

where $\phi_r = w_r/N$ is the scaled operator weight.

Using the Ito stochastic calculus, the Fokker-Planck equation can be mapped to a Langevin equation,

$$\partial_t \phi_r = -\alpha(\phi_r) + \sqrt{2\beta(\phi_r)}\eta_r(t),$$

(14)

with $\langle \eta_r(t)\eta_{r'}(t')\rangle = \delta_{rr'}\delta(t - t')$. This mapping explicitly demonstrates that the noise η arises from the deterministic master equation for the operator weight probability as a $1/N$ effect. It is important that this noise η is conceptually different from the randomness of the Brownian circuit introduced to obtain the master equation; it originates purely from the quantum fluctuation in the BCC. Later, we show that the noise, although suppressed at large N, has a drastic effect on the operator dynamics.

First, we study the infinite-N limit in which the noise is set to zero and the stochastic Langevin equation becomes deterministic. After taking the continuum limit of the Langevin equation, in which $\phi(r, t)$ is assumed to vary slowly with respect to r, we obtain a FKPP-type equation,

$$\partial_t \phi(r, t) = [3 - 4\phi(r, t)]\left(\frac{g^2}{2}\partial_r^2\phi(r, t) + (1 + g^2)\phi(r, t)\right),$$

(15)

describing a growth-diffusion-saturation process. For simplicity of presentation, we hereafter refer to Eq. (15) as a FKPP equation. There are two fixed points of the dynamics, an unstable solution $\phi(r, t) = 0$ and a stable solution $\phi(r, t) = \frac{3}{4}$. The stable solution describes the equilibrium state where every operator string is equally probable. An initially localized operator configuration translates to an initial condition for the FKPP-type equation, which is the unstable solution everywhere away from the initial local operator.

Similar to the FKPP equation, Eq. (15) admits a traveling-wave solution $\phi(r, t) = f(r - vt)$ when the velocity v is larger than $v_c = \sqrt{18g^2(1 + g^2)}$. Ahead of the wave front, $r \gg vt$, the traveling wave decays exponentially with r. For an initial operator profile that is sufficiently localized, the wave front travels with the minimal velocity v_c and approaches the traveling-wave solution at late times. A detailed analysis can be found in Appendix C. Ahead of the wave front, the traveling wave decays as $\exp(6(1 + g^2)(t - r/v_c))$, consistent with a sharp wave front. Therefore, the infinite-N limit of the Brownian coupled cluster model exhibits a well-defined Lyapunov exponent. The butterfly velocity and the Lyapunov exponent are

$$v_B = \sqrt{18g^2(1 + g^2)},$$

$$\lambda_L = 6(1 + g^2).$$

(16)

Within the framework described by Eq. (3), the infinite-N limit has a broadening exponent $p = 0$.

The existence of a Lyapunov exponent in the infinite-N limit is in sharp contrast with the random brickwork circuit model result, where the diffusive-spreading nature of the wave front is independent of the dimension of the on-site Hilbert space. The reason for this difference is that the brickwork model has no notion of few-body interactions within an on-site cluster due to the use of Haar random unitary matrices in the circuit. In Fig. 6(b), we explicitly verify the sharp wave front by numerically solving Eq. (15).

B. Large-N expansion

Having established the purely exponential growth of the squared commutator in the infinite-N limit, we now investigate the behavior away from this limit. Comparing with the infinite-N limit, the large-N expansion affects Eq. (15) in two significant ways. First, $\phi(r, t)$, in principle, only takes discrete values $0, \frac{1}{N}, \frac{2}{N} \ldots$. Therefore, in Eq. (15), $\phi(r, t)$ is set to zero when it is below $1/N$. This hard cutoff allows the traveling wave to propagate with velocity

smaller than v_c, and as such, the cutoff is important for obtaining the correction to the butterfly velocity. Second, the noise term in the Langevin equation (14) becomes important. The deterministic differential equation (15) is augmented by a multiplicative noise term

f_{noise}

$$= \sqrt{\frac{1}{N}(3 - 2\phi(r,t))\left(\frac{g^2}{2}\partial_r^2\phi(r,t) + (1+g^2)\phi(r,t)\right)}\eta(r,t).$$

(17)

Because of its multiplicative nature, the noise term only affects the physics when $\phi(r,t)$ is nonzero, and therefore, the noise does not violate the causal structure of the noiseless FKPP equation. The effect of the noise is most prominent near the forward edge of the wave, with its most important effect being to make the position of the wave front a random variable described by a biased random walk. The resulting noise-averaged front spreads diffusively with the diffusion constant D in addition to the drift $v_B(N)t$. Following an analysis of the original noisy FKPP equation [52,53], which we review in Appendix C, we are able to obtain the scaling of the $v_B(N)$ and D in the large-N limit,

$$\delta v_B = v_B(N) - v_B(\infty) \sim -\frac{1}{(\log N)^2},$$

$$D \sim \frac{1}{(\log N)^3}.$$

(18)

This result is remarkable, indicating that the system approaches the infinite-N limit very slowly. It also shows that at large but finite N, the broadening exponent becomes $p = 1$ instead of 0.

In each realization of the noise, $\phi(r,t)$ still grows exponentially. But as $\sqrt{2Dt}$ grows larger than the width of the traveling wave, the exponential growth of $\phi(r,t)$ is smoothed out by the diffusive movement of the wave front's position, leading to a diffusive broadening of the noise-averaged wave front. To quantitatively understand the effect of the noise induced by finite N on the wave front, we approximate the traveling-wave solution in a single realization of the noise by the following phenomenological model,

$$\phi(r,t) = \begin{cases} \frac{3}{4} & \text{if } r < v_Bt + r_0 + \frac{v_B}{\lambda}\log\frac{4}{3} \\ 0 & \text{if } r > v_Bt + r_0 + \frac{v_B}{\lambda}\log(N) \\ e^{\lambda_L(t-(r-r_0)/v_B)} & \text{otherwise,} \end{cases}$$

(19)

which accounts for the saturation behind the wave front and the growth ahead of the wave front.

Using this simple model, the noise-averaged squared commutator, which is proportional to the noise-averaged $\phi(r,t)$, is the convolution of $\phi(r,t)$, with a Gaussian distribution describing the diffusive motion,

$$C(r,t) = \frac{1}{\sqrt{4\pi Dt}}\int d\Delta x \frac{8}{3}\phi(r+\Delta x,t)e^{-[(\Delta x^2)/(4Dt)]}. \quad (20)$$

To simplify the notation, we introduce the dimensionless units $\tau = \lambda_L t$, $u = (\lambda_L/v_B)r$ and $\xi = [(D\lambda_L)/v_B^2]$, with ξ describing the strength of the noise. The result of the convolution is

$$C(u,\tau) = \text{erf}\left(\frac{\log(4/3) - z}{\sqrt{4\xi\tau}}\right) + 1$$
$$- \frac{4}{3}e^{\xi\tau - z}\text{erf}\left(\frac{\log(4/3) + 2\xi\tau - z}{\sqrt{4\xi\tau}}\right)$$
$$+ \frac{4}{3}e^{\xi\tau - z}\text{erf}\left(\frac{\log(N) + 2\xi\tau - z}{\sqrt{4\xi\tau}}\right), \quad (21)$$

where $\text{erf}(x)$ is the error function $(2/\sqrt{\pi})\int_0^x e^{-t^2}dt$, and $z = u - u_0 - \tau$ is the position in the traveling frame with some unimportant offset u_0 determined by the initial condition.

The next step is to analyze the behavior of Eq. (21) in space-time. It exhibits a light-cone structure with a butterfly velocity independent of ξ since the butterfly velocity is entirely set by the cutoff approximation and does not depend on the diffusion constant explicitly. This can be seen from the fact that $C(\tau, \tau)$ asymptotically approaches $\frac{1}{2}$. The space-time of the t-r plane can be approximately divided into three regions based on the behavior of Eq. (21), as illustrated in Fig. 2. The region near the wave front is the diffusive region, where the last two terms of Eq. (21) roughly cancel each other and $C(u,\tau)$ is dominated by the single error function,

$$\lim_{\tau\to\infty} C(\tau + z, \tau) = 1 + \text{erf}\left(-\frac{z}{\sqrt{4\xi\tau}}\right). \quad (22)$$

In the limit that $\sqrt{4\xi\tau} \ll z \ll \tau$, $C(r,t) \sim \exp\{-[(r - r_0 - v_bt)^2]/4Dt\}$, consistent with the universal form with the broadening exponent $p = 1$. This result clearly demonstrates that the wave front spreads diffusively. The growth behavior near the wave front is dominated by the noise, and the original Lyapunov exponents do not enter.

This case should be contrasted with the chaotic region where the first two error functions are far from saturation, but the last error function is already saturated. In this region, the squared commutator is $C(u,\tau) \sim \frac{4}{3}e^{\xi\tau - z}$, a pure growth form with a modified Lyapunov exponent $\lambda_L(1 + D\lambda_L/v_B^2)$. The size of this region scales as $\log N$, and the value of C in this region can be arbitrarily small in the long-time limit since it is enclosed by two lines with a
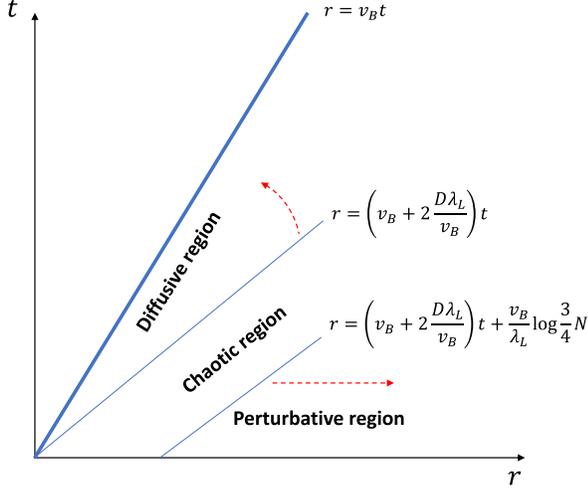
FIG. 2. The scrambling "phase diagram" in space-time. An overall unimportant spatial offset is omitted. The bold line marks the wave front. The space-time is roughly divided into three regions. In the diffusive region, the squared commutator exhibits a diffusively broadening wave front, quantitatively different from the exponential growth in the large-N limit. Away from the wave front, there is a region where the squared commutator grows exponentially with a modified Lyapunov exponent. Further ahead of the wave front, the chaotic region gives way to the perturbative region. As N increases, the chaotic region expands as indicated by dashed arrows, and it eventually dominates the wave-front behavior in the infinite-N limit.

bigger velocity $v = v_B(1 + 2\xi)$ than the butterfly velocity. Therefore, it is difficult to extract this region from numerical data of finite-N spin chains.

There is also a third region, which we denoted as the perturbative region, where z is the largest scale in the system. In this case, the squared commutator is infinitesimally small and behaves as $[8/(3N)]\sqrt{[(\xi\tau)/(\pi z^2)]}\exp(-z^2/4\xi\tau)$.

C. Small-N limit

The large-N analysis presented in the last section cannot be naively generalized to the case of small N. In the small-N limit, the master equation (11) is still valid, but the approach of approximating the master equation with the Fokker-Planck equation (13) to derive the Langevin equation (14) is not.

Instead, we take a rather different approach by considering the probability of the operator string ending on a specific site at a given time, similar to what was studied in the random brickwork model. This probability $\rho(r, t)$ is defined as

$$\rho(r, t) = \sum_{\{w\}} \left(\tilde{h}(\mathbf{w}, t)(1 - \delta_{w_r, 0}) \prod_{s > r} \delta_{w_s, 0} \right). \quad (23)$$

Note that the sum of $\rho(r, t)$ is conserved.

From Eq. (11), one can derive the rate equation for $\rho(r, t)$ as

$$\partial_t \rho(r, t) = -\xi \rho(r, t) + \sum_l \gamma_b^-(0, l)\rho_l(r - 1, t)$$
$$+ \sum_l \gamma_b^+(1, l)\rho_{l1}(r + 1, t), \quad (24)$$

where a subindex on ρ indicates a restriction: The operator string must end with that particular configuration. For example, $\rho_{l1}(r + 1, t)$ is the probability of the operator string with $w = l$ on site r, $w = 1$ on site $r + 1$, and $w = 0$ for all sites beyond $r + 1$.

Now, we use the approximation of local equilibrium, which is crucially different from the large-N case, to relate $\rho_l(r, t)$ and $\rho_{l'l}(r, t)$ to $\rho(r)$. The approximation of local equilibrium states that all the local Pauli strings instantly have the same probability to appear once the operator front reaches there, as a result of small on-site d.o.f. This idea suggests that $\rho_l(r, t) = [3^l/(4^N - 1)]\binom{N}{l}\rho(r, t)$ and $\rho_{l1}(r, t) = (3^l/4^N)\binom{N}{l}[3N/(4^N - 1)]\rho(r, t)$. With this approximation, we obtain a closed equation for $\rho(r, t)$,

$$\partial_t \rho(r, t) = -\xi \rho(r, t) + \frac{9N}{8} \frac{4^N}{4^N - 1} g^2 \rho(r - 1, t)$$
$$+ \frac{9N}{8} \frac{1}{4^N - 1} g^2 \rho(r + 1, t). \quad (25)$$

The conservation law of $\rho(r, t)$ determines that $\xi = [(9N)/8][(4^N + 1)/(4^N - 1)]$. In the continuum limit, the equation reads

$$\partial_t \rho(r, t) = -\frac{9N}{8} g^2 \partial_r \rho(r, t) + \frac{9N}{16} \frac{4^N + 1}{4^N - 1} g^2 \partial_r^2 \rho(r, t). \quad (26)$$

This result leads to

$$\rho(r, t) = \frac{1}{2\sqrt{\pi D t}} \exp\left(-\frac{(x - v_B t)^2}{4 D t}\right), \quad (27)$$

with

$$v_B = \frac{9N}{8} g^2,$$

$$D = \frac{9N}{16} \frac{4^N + 1}{4^N - 1} g^2. \quad (28)$$

The average squared commutator is related to $\rho(r, t)$ as

$$C(r, t) = 2 \int_r^\infty \rho(s, t) ds = 1 + \text{erf}\left(\frac{v_B t - x}{\sqrt{4 D t}}\right). \quad (29)$$

The final result is consistent with the Haar random circuit result and has a broadening exponent $p = 1$.

The above analysis relies on the local equilibrium approximation, which holds in the small-N limit. However, as N increases, the relaxation time $t_{\rm loc}$ increases as well, scaling as $\log N$. Consequently, Eq. (29) needs to be modified to incorporate the delay in the relaxation. For finite $t_{\rm loc}$, the operator on the rightmost site never reaches full local equilibrium, and the conditional probability $\rho_l(r, t)/\rho(r, t)$ is always biased towards small l compared with the equilibrium value $[3^l/(4^N-1)]\binom{N}{l}$, with the extreme case $\rho_l(r, t)/\rho(r, t) = \delta_{l,1}$. Therefore, to incorporate the effect of finite $t_{\rm loc}$, we write $\rho_l(r, t)$ as

$$\frac{\rho_l(r,t)}{\rho(r,t)} = \left[\left(\frac{3^l}{4^N-1}\binom{N}{l} - \delta_{l,1}\right)(1 - e^{-\Delta t/t_{\rm loc}}) + \delta_{l,1}\right], \quad (30)$$

which approaches the equilibrium distribution as $t_{\rm loc} \to 0$ and approaches the extreme case as $t_{\rm loc} \to \infty$. Here, Δt can be interpreted as $1/v_B$, the timescale for operator expansion.

Based on Eq. (30), the summation of ρ_l in Eq. (24) can be written as

$$\sum_l \gamma_b^-(0, l)\rho_l(r-1, t)$$
$$= \left[\left(\frac{9N}{8}\frac{4^N}{4^N-1} - \frac{3}{2}\right)g^2(1 - e^{-\Delta t/t_{\rm loc}}) + \frac{3}{2}g^2\right]\rho(r-1, t). \quad (31)$$

Applying a similar argument to ρ_{l1}, we obtain

$$\sum_l \gamma_b^+(1, l)\rho_{l1}(r+1, t)$$
$$= \left(\frac{9N}{8}\frac{1}{4^N-1}g^2(1 - e^{-\Delta t/t_{\rm loc}}) + \frac{g^2}{2N}\right)\rho(r+1, t), \quad (32)$$

leading to modified v_B and D,

$$\tilde{v}_B = v_B(1 - e^{-\Delta t/t_{\rm loc}}) + \left(\frac{3}{2} - \frac{1}{2N}\right)g^2 e^{-\Delta t/t_{\rm loc}},$$
$$\tilde{D} = D(1 - e^{-\Delta t/t_{\rm loc}}) + \left(\frac{3}{4} + \frac{1}{4N}\right)g^2 e^{-\Delta t/t_{\rm loc}}. \quad (33)$$

The above analysis suggests that Eq. (28) overestimates both v_B and D for N larger than 1, and the deviation increases as N. We also note that Eq. (29) relies on the assumption of local equilibrium and is, in general, not valid at finite N. To study the operator dynamics and scrambling at finite N, we directly simulate the master equation using the matrix-product form, as discussed in the next section. We see that, indeed, the results deviate from the prediction in Eq. (28) as N increases and gradually crosses over to the large-N results discussed in the last section.

D. Finite-N results from tensor network simulation

The large-N analysis, together with the small-N analysis, convincingly demonstrates that, away from infinite N, the behavior of the squared commutator is dominated by the error function near the wave front, leading to a broadening exponent $p = 1$. The effect of N is mostly encoded in the butterfly velocity and the diffusion constant. But the N dependence of v_B and of D obtained from the two limits are not consistent. The small-N analysis suggests that both v_B and d increase with N, while the large-N analysis indicates that v_B saturates to a certain value and D decreases with N. Therefore, it is interesting to study how the two quantities interpolate between the two limits, especially the diffusion constant, which is expected to be nonmonotonic.

For this purpose, we directly simulate the master equation (11) by representing the probability distribution as a MPS with physical dimension $N + 1$. Comparing with the usual TEBD method [54,55], simulating a stochastic process with MPS (S-MPS) [56] is quite different. In Appendix D, we discuss the difference and introduce several useful techniques, including the canonical form for S-MPS along with the truncation schemes that are helpful for preserving the 1-norm of the S-MPS instead of the 2-norm. Notably, the truncation scheme developed by White et al. [57] can be directly applied here to exactly preserve the 1-norm for all time.

Generally, S-MPS requires higher bond dimension to capture local observables accurately compared with unitary MPS with the same entanglement because the 1-norm normalization appropriate to S-MPS tends to amplify errors. Nevertheless, the entanglement entropy is still a good measure for determining whether the probability distribution can be represented efficiently as a MPS. Initially, the probability distribution is a product state with operator weight 1 in the center cluster of the system and operator weight 0 elsewhere. In the early growth region ahead of the light cone, the probability is not affected by the stochastic evolution, and it continues to enjoy low entanglement. Inside the light cone, the probability distribution reaches the steady state where every operator string is equally probable and also admits a simple product state representation. Therefore, the entanglement entropy only accumulates around the wave front. In practice, we find that the entanglement entropy never exceeds one bit, allowing us to obtain the whole scrambling curve up to $N = 10$. The resulting scrambling curve can be fit with the error function with almost perfect quality to extract the butterfly velocity and the diffusion constant. The fitting result is shown in Appendix D.

The result of the butterfly velocity is shown in Fig. 3(a) together with that obtained from small-N and large-N analysis for $g = 1$. We find that at $N = 1$ (single spin in the spin cluster), v_B from S-MPS agrees with the small-N analysis perfectly. As N increases, v_B deviates from the linear growth predicted at small N and smoothly connects
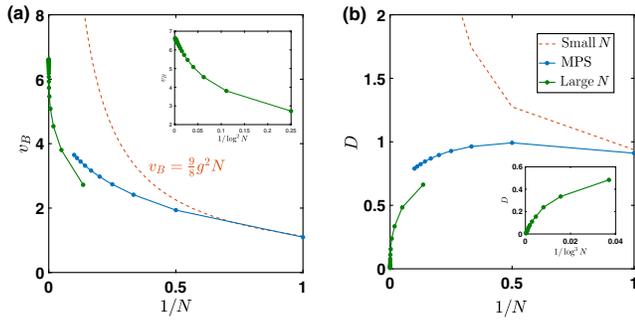
FIG. 3. Comparison between the butterfly velocity v_B (a) and the wave-front diffusion constant D (b) obtained from MPS simulation of the probability distribution to the small-N and large-N analysis. When N is small, v_B and D agree with the small-N result, relying on the local equilibrium assumption. As N increases, the approximation breaks down, and both v_B and D cross over to the large-N result.

to results from the large-N analysis. Figure 3(b) summarizes the N dependence of the diffusion constant from the different methods of analysis. The result from the S-MPS indeed exhibits nonmonotonic behavior. At $N = 1$, it agrees with the small-N analysis. It peaks at $N = 3$ and drops as N increases, approaching the result from the large-N analysis. Therefore, the numerical results agree with the analyses above from both limits. In principle, for large enough N, one should be able to identify the chaotic region.

E. Comparing with Haar random brickwork circuit models

It is instructive to compare the Brownian circuit model studied here with the previously studied random brickwork circuit models. By designation, in the brickwork circuit, the Haar random unitary matrices equilibrate the operator string on the two sites it connects to immediately if there is nontrivial weight. In this case, the analysis in Sec. III C becomes exact; the entire region ahead of the wave front is governed by the error function. On the other hand, in the BCC model studied here, the operator string takes a finite time to reach equilibrium even locally, the timescale being about $\log N$. The direct consequence is that, although near the wave front the behavior of the squared commutator is dominated by the error function, there is still a region in space-time ahead of the wave front where $C(r,t)$ grows exponentially, as illustrated in Fig. 2. As N increases, this chaotic region expands and finally dominates the wave front in the infinite-N limit.

IV. IMPLICATIONS FOR LOCAL HAMILTONIAN SYSTEMS

The analysis so far has given two results. First, there is a random circuit model with a parameter N such that, at infinite N, the model exhibits exponential growth of the squared commutator with $p = 0$. Second, for any noninfinite N, the

dynamics of the model inevitably crosses over to a diffusively broadened wave front with $p = 1$. It is quite plausible that any sufficiently generic random circuit model with finite on-site Hilbert space will also exhibit a diffusively broadened wave front with $p = 1$ (with this already being established for the BCC and the random brickwork circuit). The key question is, what aspects of this analysis hold when the couplings are not random in time? We now argue that $p = 1$ is generic for chaotic quantum many-body systems in $d = 1$ with finite local Hilbert space dimensions.

The argument has two thrusts. First, we directly numerically simulate a small-N Ising spin chain with conserved energy. Combining large-scale numerical simulations with a new analysis technique, our previous result is improved to show that the system asymptotically approaches $p = 1$. Second, based on previous work in energy-conserving systems showing the existence of noiseless FKPP-type equations governing the spreading of chaos at the large-N and/or weak coupling model, we argue that quantum fluctuations inevitably introduce multiplicative noise into these equations. The physics of the noisy FKPP equation then naturally leads to $p = 1$. For the latter argument, recall that we were careful to distinguish the noise in the $1/N$ corrected FKPP equation, which was a manifestation of quantum fluctuations, from the space-time random couplings in the microscopic Hamiltonian.

Although we focus on energy-conserving systems here, we conjecture that our analysis also applies to Floquet models where the couplings are not random in time, but energy is not conserved because the Hamiltonian is time dependent. In the case of conserved energy, it also makes sense to talk about noninfinite-temperature states. We briefly discuss how the story might be modified in this case, with a focus on the physics of the chaos bound.

A. Wave-front broadening for small-N energy-conserving systems

According to the analysis in Sec. III B, for finite on-site Hilbert space dimensions, each contour of the squared commutator intersects with the chaotic region for a limited amount of time and eventually merges into the diffusive region. The chaotic regime would suppress wave-front broadening, and the diffusive broadening is only clearly visible when the diffusive regime is much larger than the chaotic regime. This result suggests a strong finite-size effect that will hinder extraction of the broadening exponent p from fitting the squared commutator with the universal growth form. This new insight obtained from analyzing the BCC model is consistent with what we observed in our earlier result [11] from MPO calculation of the squared commutator, where the value of p drifts upwards along the contour and the asymptotic value is not addressed.

To unambiguously analyze the broadening of the wave front, we improve our numerical result by pushing both the system size and the simulation time so that the wave front

travels through about 200 sites. We measure the spatial difference δx between two contours of the squared commutator as a function of time. For the two contours we choose, we check that a MPO with bond dimension $\chi = 8$ and one with a bond dimension $\chi = 16$ give identical results. Then, assuming that the general form in Eq. (3) applies, the asymptotic value of the slope is related to the broadening exponent

$$\frac{d \log \delta x}{d \log t} = \frac{p}{p+1}. \tag{34}$$

We perform this analysis on the mixed-field Ising chaotic spin chain described by the Hamiltonian,

$$H = -\frac{1}{E_0}\left(J\sum_{r=1}^{L}Z_rZ_{r+1} + h_x\sum_{r=1}^{L}X_r + h_z\sum_{r=1}^{L}Z_r\right), \tag{35}$$

where X_r and Z_r are local Pauli operators. The parameters are set to $J = 1$, $h_x = 1.05$, and $h_z = 0.5$. The overall normalization factor $E_0 = \sqrt{4J^2 + 2h_x^2 + 2h_z^2}$. The result is shown in Fig. 4. The inset of Fig. 4(b) plots δx with t on a log-log plot. The slope gradually increases and approaches $1/2$ in the large space-time limit, indicating that the wave front broadens diffusively, $p = 1$, at the largest sizes and times. The initial deviation may be due to the early-time microscopic physics where $C(r,t)$ behaves as $t^x/x!$. However, the fact that the deviation persists to an intermediate scale suggests that there may exist a chaotic region in the space-time causing a strong finite-size effect. Extracting this region for local Hamiltonian systems is an interesting future research direction.
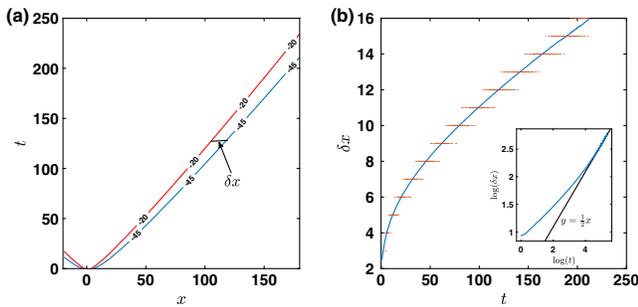

FIG. 4. The diffusive broadening wave front in the mixed-field Ising chain. (a) The contour of $\log C(r,t) = -20$ and $\log C(r,t) = -45$. The system size and the time limit allow the front to travel through about 200 sites. The contours obtained from MPO simulation with bond dimensions 8 and 16 are identical, showing the excellent convergence of our method. (b) The spatial distance δx between the two contours increases with time, showing definitive broadening of the wave front. In the inset, we plot δx with t on a log-log plot. The slope of the curve approaches $1/2$, strong evidence of diffusive broadening, and the broadening exponent $p = 1$.
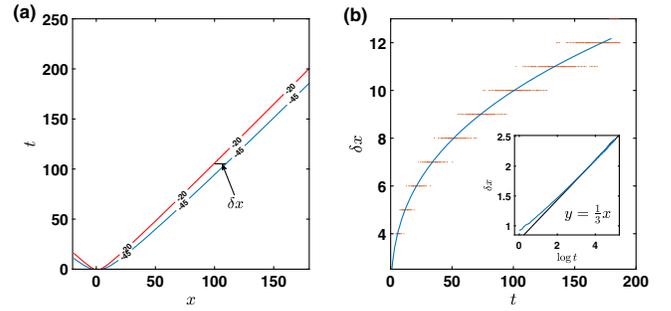

FIG. 5. The same figure as Fig. 4 but for a transverse-field Ising model describing noninteracting fermions. The slope of the curve on a log-log plot approaches $1/3$, agreeing with exact results for this model.

To further validate this method, the same analysis is performed for the transverse-field Ising model (setting h_z to zero), which describes noninteracting fermions and has a broadening exponent $p = 1/2$. The result is shown in Fig. 5 where one can indeed see that the slope of the curve on a log-log plot approaches $1/3$.

B. Conjectured wave-front broadening for large-N energy-conserving systems

Given that one generic energy-conserving model exhibits $p = 1$, it is plausible that this is a universal behavior among local chaotic Hamiltonians in one dimension. To bolster this conjecture and to give a physical picture for it in one limit, it is useful to return to the noiseless FKPP equation. Indeed, a number of different models have been shown to have operator growth described by a noiseless FKPP-type equation, in some cases linearized and in some cases fully nonlinear, at large N or weak coupling. What we argue is that such equations should inevitably be augmented by a noise term describing quantum fluctuations, which has the form considered in this work, implying that this broad class of large-N models also has $p = 1$ at the largest sizes and times. Assuming this is true, $p = 1$ then occurs at large and small N and weak and strong coupling, so it is reasonable to conjecture that it is a universal property of one-dimensional chaotic systems.

The argument proceeds in three steps of increasing specificity. The background assumptions are that one has a closed dynamical equation governing $\phi \propto C$ and some parameter N that measures the local d.o.f. First, quantum fluctuations are expected to add a noise term to any approximate set of closed equations governing the dynamics of the OTOCs of simple operators. This is because operator growth is not deterministic in a quantum system since a single Heisenberg operator is a superposition of many different complex operator strings. Second, the specific form of the noise term must be multiplicative for local Hamiltonians, meaning proportional to some power of ϕ, because operator growth arises from the failed

cancelation between U and $U^\dagger$ due to the insertion of the perturbation W. Operator growth never spontaneously occurs far away from the current support of $W(t)$ precisely because U and $U^\dagger$ cancel in far away regions. Third, the form of the multiplicative noise term should be $\sqrt{\phi/N}$. This form is necessary to ensure that the noise is most important when ϕ is small, which should be true since larger values of ϕ are more self-averaging. This term also guarantees that the associated Fokker-Planck equation has a sensible $1/N$ expansion, i.e., with no unusual powers of $1/N$ appearing. Then, assuming the dynamical equation for C includes saturation effects, we have all the components necessary for the noisy FKPP analysis to apply.

One caveat here is that the analysis is framed in terms of large-N models. While FKPP-type equations have also been derived in weak coupling approximations, it is less clear how to identify the precise role of the N parameter in that case. One simple intuition is that N should arise because the dynamics is effectively coarse grained over a long length scale corresponding to the inelastic mean free path. Identifying $N \sim \ell$, with ℓ some kind of inelastic mean free path, would then predict butterfly velocity corrections and a diffusion constant going like $\delta v_B \sim 1/\log^2 \ell$ and $D \sim 1/\log^3 \ell$. It would be interesting to better understand the situation at weak coupling.

C. Finite temperature

Here, we make some comments about the dynamics of operator growth at noninfinite T. The key point is that the FKPP equation and its noisy counterpart make no particular reference to temperature, so it is reasonable to suppose that they could hold at noninfinite T. The noiseless FKPP equation has already been derived at finite T for a variety of models; temperature only enters insofar as the parameters in the FKPP equation are temperature dependent. Thus, the results on the BCC from small to large N provide some hints on the interplay between quantum fluctuation and local scrambling.

The manifestation of this interplay is the distinction between the diffusive region and the chaotic region in the space-time structure of the OTOC. In the infinite-N limit, quantum fluctuations are completely suppressed, and the local scrambling time of about $\log N$ is infinite. In this case, the chaotic region occupies the entire space-time. For large but finite N, as soon as quantum fluctuations are present, the wave front broadens diffusively, while the chaotic region occurs ahead of the front. In the small-N limit, with the local equilibrium assumption implying that the local scrambling time is short, the diffusive region extends to the entire space-time. This result is also consistent with the analytical results on the random brickwork circuit.

Now, consider a local Hamiltonian system, say, a spin-1/2 system. Assuming the mixed-field Ising behavior is generic, we show that, at infinite temperature, the wave front also broadens diffusively. But it is difficult to tell whether there exists a chaotic region in addition to the diffusive region ahead of the wave front due to the microscopic details affecting the behavior of OTOC at early times. On the other hand, at finite temperature and assuming a separation of timescales, there is a thermally regulated version of the OTOC whose growth rate is bounded by the temperature [58]. If we assume the same bound applies to the nonthermally regulated object [59], we would have

$$\frac{d \log C}{dt} \leq 2\pi T. \qquad (36)$$

Given this chaos bound, the diffusive broadening form

$$C(r, t) = \exp\left(-\frac{(r - v_B t)^2}{4Dt}\right) \qquad (37)$$

can only be valid up to a finite distance ahead of the front because its growth rate diverges in the large r limit. Imposing the chaos bound, we can estimate the maximum distance from the wave front for which the diffusive behavior is valid. The growth rate of the diffusive growth from is

$$\gamma(r, t) = \frac{v_B}{2D}\left(\frac{r}{t} - v_B\right) + \frac{1}{4D}\left(\frac{r}{t} - v_B\right)^2. \qquad (38)$$

Demanding that $\gamma(r, t) \leq 2\pi T$ predicts that the diffusive region can at most persist up to the space-time line given by

$$\frac{r}{v_B t} = \sqrt{1 + \frac{4D(2\pi T)}{v_B^2}}. \qquad (39)$$

One might also imagine that even at infinite temperature, the logarithmic derivative cannot be too large; for example, it should be bounded by the size of the microscopic couplings. In the BCC, this was indeed true at large N where the crossover from the diffusive to chaotic region was roughly where the rate of growth in the diffusive region was approaching the Lyapunov rate (which itself was of order the microscopic scale). A similar argument might also apply when the system has a Lyapunov exponent less than $2\pi T$. In any event, at finite T, there may generically be a region ahead of the wave front, at least at large but finite N or weak coupling, where the Lyapunov exponent is still visible. One issue is that we could run into the perturbative region before any exponent can be extracted. More generally, one should carefully study the thermally regulated commutators to see the precise consequences of the chaos bound, something we will report on in forthcoming work.

V. CONCLUSION AND OUTLOOK

In this work, we studied a random time-dependent Hamiltonian model with a large-N limit in which we could

study in detail the way the infinite-N Lyapunov exponent gave way to a diffusively broadened scrambling wave front. Based on this model and an analysis of large-scale MPO simulations in a time-independent Hamiltonian model, we conjectured that the local operator growth wave front broadens diffusively in generic local chaotic Hamiltonians with finite local Hilbert space dimensions. We also showed how a modified stochastic MPS formalism could be used to simulate the operator dynamics for all times after averaging over different Hamiltonian realizations in the random model. A unifying element was the emergence of a noiseless FKPP equation at infinite N and a corresponding noisy FKPP equation at finite N. The noise was an effect of quantum fluctuations, and it ensured that both large N and small N exhibited $p = 1$ dynamics.

It is straightforward to extend the BCC model to any dimension or, indeed, to any graph. In higher dimensions, there will still be a Lyapunov exponent and sharp wave front at infinite N. Finite-N corrections will then introduce noise into the FKPP-type equation. The analog of the cutoff on ϕ is an extended cutoff front where $\phi = 1/N$. This front will then experience some random dynamics with a constant drift (the butterfly velocity) and noisy local dynamics. Although the general long-distance structure may be complex (e.g., in high dimensions the noise may be relevant or irrelevant), one expects KPZ-like dynamics in low dimensions. It will be interesting to analyze the higher-dimensional case in more detail and possibly also study the model on more general graphs.

In terms of future directions, we have several works in progress. One is to consider the effect of a conserved $U(1)$ symmetry on the operator spreading. This effect has already been studied in the random brickwork circuit [38,39], but we anticipate new interesting physics associated with the interplay with the large-N effects. Another direction is a study of the entanglement dynamics in the model as a function of N, as well as a study of the noise physics in a large-N Hamiltonian model.

APPENDIX A: DERIVATION OF THE MASTER EQUATION OF THE BROWNIAN COUPLED CLUSTER

In this Appendix, the derivation of the master equation governing the operator dynamics of the Brownian coupled cluster model is presented. Consider a chain of L clusters

with open boundary conditions where each cluster contains N qubits. At every time step, all the qubits within a cluster interact, and all the qubits between neighboring clusters interact. The time evolution operator is stochastic and obeys

$$
\begin{aligned}
U(t + dt) &- U(t) \\
&= -\frac{N(L + g^2(L - 1))}{2} U(t)dt \\
&- iA \sum_{r=1}^{L} \sum_{b>a=1}^{N} \sum_{\alpha,\beta} \sigma_{r,a}^{\alpha} \sigma_{r,b}^{\beta} U(t) dB_{r,a,b,\alpha,\beta} \\
&- ig\sqrt{\frac{N-1}{2N}} A \sum_{r=1}^{L-1} \sum_{a,b=1}^{N} \sum_{\alpha,\beta} \sigma_{r,a}^{\alpha} \sigma_{r+1,b}^{\beta} U(t) d\tilde{B}_{r,a,b,\alpha,\beta}.
\end{aligned}
$$

Here, r labels the cluster, a, b are labels within a cluster, and α, β label Pauli matrices. The coupling constant g can be used to dial the relative strength of the within-cluster and between-cluster interactions. Some of the other factors are chosen for convenience, and the sum over a, b is unconstrained in the between-cluster term. The coefficient A is determined by demanding that $UU^\dagger = 1$ on average, leading to

$$
A = \sqrt{\frac{1}{8(N-1)}}. \tag{A1}
$$

Given an operator W, the Heisenberg operator is $W(t) = UWU^\dagger$. We may expand $W(t)$ in a complete basis of operators,

$$
W(t) = \sum_{\mathcal{S}} c(\mathcal{S})\mathcal{S}, \tag{A2}
$$

where $\mathcal{S}$ is a product of cluster operators $\mathcal{S}_r$, with each $\mathcal{S}_r$ a product of Pauli operators within the cluster. The coefficients $c(\mathcal{S})$ can be determined from

$$
c(\mathcal{S}) = \frac{1}{2^{NL}} \text{tr}\big(\mathcal{S}W(t)\big). \tag{A3}
$$

We study the average operator probabilities,

$$
h(\mathcal{S}) = \overline{c(\mathcal{S})^2}. \tag{A4}
$$

To determine the equation of motion of $h(\mathcal{S})$, we must compute $dh(\mathcal{S})$. There are two kinds of terms, depending on whether dB or $d\tilde{B}$ appears in the same trace or not. When they appear in the same trace, we find

$$dh_1 = -2N(L+g^2(L-1))hdt + 2A^2\left(\sum_{r,a<b,\alpha,\beta} q_{r,a,b}^{\alpha,\beta}(\mathcal{S})\right)hdt$$
$$+ 2g^2\frac{N-1}{2N}A^2\left(\sum_{r,a<b,\alpha,\beta} \tilde{q}_{r,a,b}^{\alpha,\beta}(\mathcal{S})\right)hdt,$$

$$dh_2 = -\frac{2}{8(N-1)}\sum_{r=1}^{L}\sum_{a<b}\sum_{\alpha,\beta} q_{r,a,b}^{\alpha,\beta}(1-q_{r,a,b}^{\alpha,\beta})h(\sigma_{r,a}^\alpha \sigma_{r,b}^\beta \mathcal{S})dt$$
$$-\frac{2g^2}{16N}\sum_{r=1}^{L-1}\sum_{a,b}\sum_{\alpha,\beta} \tilde{q}_{r,a,b}^{\alpha,\beta}(1-\tilde{q}_{r,a,b}^{\alpha,\beta})h(\sigma_{r,a}^\alpha \sigma_{r+1,b}^\beta \mathcal{S})dt.$$

$$(A7)$$

where $q_{r,a,b}^{\alpha,\beta} = \pm 1$ depending on whether $\sigma_{r,a}^\alpha \sigma_{r,b}^\beta$ commutes or anticommutes with $\mathcal{S}$ and $\tilde{q}_{r,a,b}^{\alpha,\beta} = \pm 1$ depending on whether $\sigma_{r,a}^\alpha \sigma_{r+1,b}^\beta$ commutes or anticommutes with $\mathcal{S}$.

Let w_r denote the total weight of $\mathcal{S}$ on site r, $w_r = 0, \ldots, N$. The sums above can be written in terms of the w_r. The first sum is

$$\sum_{r,a<b,\alpha,\beta} q_{r,a,b}^{\alpha,\beta} = 16\sum_{r=1}^{L}\frac{(N-w_r)(N-w_r-1)}{2}. \quad (A5)$$

If $\mathcal{S}$ has a nonzero Pauli on spin r, a or r, b, then $\sum_{\alpha,\beta} q_{r,a,b}^{\alpha,\beta} = 0$; otherwise it is 16. Thus, we need to count the number of pairs a, b, which are both the identity operator. This number is $\{[(N-w_r)(N-w_r-1)]/2\}$. Similarly, the second sum is

$$\sum_{r,a<b,\alpha,\beta} \tilde{q}_{r,a,b}^{\alpha,\beta} = 16\sum_{r=1}^{L-1}(N-w_r)(N-w_{r+1}). \quad (A6)$$

Putting everything together gives

$$dh_1 = -2N(L+g^2(L-1))hdt$$
$$+\frac{2}{N-1}\sum_{r}(N-w_r)(N-w_r-1)hdt$$
$$+\frac{2g^2}{N}\sum_{r}(N-w_r)(N-w_{r+1})hdt.$$

When the dB or $d\tilde{B}$ factors appear in different traces, then we get terms connecting $h(\mathcal{S})$ to $h(\sigma\sigma\mathcal{S})$. These terms are

To proceed further, let us assume that h depends only on the total weight w_r on site r and not on the particular operator $\mathcal{S}$. The dh_1 terms already manifestly depend just on the total weight. The dh_2 terms connect probabilities for different weights. The dh_1 term can be written

$$dh_1 = \sum_{r=1}^{L}\frac{-2[(2N-1)w_r - w_r^2]}{N-1}hdt$$
$$-\frac{g^2}{N}\sum_{\langle rr'\rangle}(2Nw_r - w_r w_{r'})hdt + r \leftrightarrow r'. \quad (A8)$$

To compute the dh_2 terms, we can consider a particular operator of the desired weight. Consider, first, the on-site terms. Suppose $\mathcal{S}_r = \sigma_{r,1}^x \ldots \sigma_{r,w_r}^x I_{r,w_r+1} \ldots I_{r,N}$. Now, consider all $\sigma_{r,a}^\alpha \sigma_{r,b}^\beta$. If $\sigma\sigma$ commutes with $\mathcal{S}$, then the dh_2 contribution vanishes. If $\sigma\sigma$ anticommutes with $\mathcal{S}$, then the $q(1-q)$ factor is -2. If $\mathcal{S}_r$ is the identity on one of a or b, say b, then the anticommuting terms are $\sigma^y\sigma^\beta$ and $\sigma^z\sigma^\beta$. Of these eight, two keep the weight the same and six increase the weight by 1. There are $w_r(N-w_r)$ choices of a, b in this class. The contribution is thus

$$-2w_r(N-w_r)[2h(\mathbf{w})dt + 6h(\mathbf{w}+\mathbf{e}_r)dt]. \quad (A9)$$

Here, $\mathbf{e}_r = (0, \ldots, 0, 1_r, 0, \ldots, 0)$ is a unit vector that adds 1 to weight w_r.

If $\mathcal{S}_r$ is a nonidentity on both a and b, then the anticommuting terms are $(y$ or $z)$ $(I$ or $x)$ and $(I$ or $x)$ $(y$ or $z)$. Of these eight, four keep the weight the same and four decrease the weight by 1. There are $w_r(w_r-1)/2$ such choices of a, b. The total contribution from these terms is thus

$$-2\frac{w_r(w_r-1)}{2}[4h(\mathbf{w})dt + 4h(\mathbf{w}-\mathbf{e}_r)dt]. \quad (A10)$$

The total on-site contributions are thus

$$dh_{2,\text{on-site}} = \frac{1}{2(N-1)}\sum_{r=1}^{L}w_r(N-w_r)[2h(\mathbf{w})dt + 6h(\mathbf{w}+\mathbf{e}_r)dt] + \frac{1}{2(N-1)}\sum_{r=1}^{L}\frac{w_r(w_r-1)}{2}[4h(\mathbf{w})dt + 4h(\mathbf{w}-\mathbf{e}_r)dt].$$

For the non–on-site contributions, we choose $\mathcal{S}$ to be an operator with w_r σ^xs on cluster r and w_{r+1} σ^xs on cluster $r+1$. If $\mathcal{S}$ is the identity on r, a, then the anticommuting operators are $\sigma^\alpha_{r,a}\sigma^y_{r+1,b}$ and $\sigma^\alpha_{r,a}\sigma^z_{r+1,b}$. Of these eight, two keep w_r the same and two increase w_r by 1. There are $(N - w_r)w_{r+1}$ such a, b. The contribution is

$$-2(N - w_r)w_{r+1}[2h(\mathbf{w})dt + 6h(\mathbf{w} + \mathbf{e}_r)dt]. \quad (A11)$$

Similarly, if $\mathcal{S}$ is the identity on $r + 1, b$, then by the same logic, we find a contribution of

$$-2w_r(N - w_{r+1})[2h(\mathbf{w})dt + 6h(\mathbf{w} + \mathbf{e}_{r+1})dt]. \quad (A12)$$

Finally, suppose $\mathcal{S}$ is not the identity on both r, a and $r + 1, b$. Then, the anticommuting operators are $(y$ or $z)$ $(I$ or $x)$ and $(I$ or $x)$ $(y$ or $z)$. Of these eight, four leave both weights unchanged, two decrease weight w_r by 1, and two decrease weight w_{r+1} by 1. There are $w_r w_{r+1}$ such a, b. The contribution is

$$-2w_r w_{r+1}[4h(\mathbf{w})dt + 2h(\mathbf{w} - \mathbf{e}_{r+1})dt + 2h(\mathbf{w} - \mathbf{e}_r)dt]. \quad (A13)$$

The total non–on-site contribution is thus

$$dh_{2,\text{non-on-site}} = \frac{g^2}{4N}\sum_{\langle r,r'\rangle} Nw_{r'}(2h(\mathbf{w}) + 6h(\mathbf{w} + \mathbf{e}_r))$$
$$- w_r w_{r'}(6h(\mathbf{w} + \mathbf{e}_r) - 2h(\mathbf{w} - \mathbf{e}_r)) + r \leftrightarrow r'. \quad (A14)$$

By combining Eqs. (A8), (A11), and (A14), we obtain the complete equation of motion for $h(\mathbf{w})$ in the case where $h(\mathcal{S})$ depends only on the weights w_r of P at the different clusters. It is, however, convenient to take one more step and include degeneracy factors.

The number of operators with weights w_r is

$$D(\mathbf{w}) = \prod_r \binom{N}{w_r}3^{w_r}. \quad (A15)$$

While $h(\mathbf{w})$ is the probability of a single operator with weights w_r, the object $\tilde{h}(\mathbf{w})$, defined by

$$\tilde{h}(\mathbf{w}) = D(\mathbf{w})h(\mathbf{w}), \quad (A16)$$

is the total probability of all operators with weight w_r, i.e., the probability of weights w_r.

The equation of motion for $\tilde{h}$ can be obtained from that of h. The dh_1 terms immediately translate to $d\tilde{h}_1$ terms since the weights are the same on both sides of the equation. However, the dh_2 must by modified. We replace each $h(\mathbf{w})$ with $\tilde{h}(\mathbf{w})/D(\mathbf{w})$. Then, the various rates are modified by ratios of $D(\mathbf{w})$ to $D(\mathbf{w} \pm \mathbf{e}_r)$. These ratios are

$$\frac{D(\mathbf{w})}{D(\mathbf{w} + \mathbf{e}_r)} = \frac{w_r + 1}{3(N - w_r)} \quad (A17)$$

and

$$\frac{D(\mathbf{w})}{D(\mathbf{w} - \mathbf{e}_r)} = \frac{3(N - w_r + 1)}{w_r}. \quad (A18)$$

Thus, we have

$$\frac{d\tilde{h}_1}{dt} = \sum_{r=1}^{L} \frac{-2[(2N - 1)w_r - w_r^2]}{N - 1}\tilde{h}$$
$$- \frac{g^2}{N}\sum_{\langle r,r'\rangle}(2Nw_{r'} - w_r w_{r'})\tilde{h} + r \leftrightarrow r',$$

$$\frac{d\tilde{h}_{2,\text{on-site}}}{dt} = \sum_{r=1}^{L} \frac{w_r}{(N - 1)}[\tilde{h} + \frac{w_r + 1}{N - w_r}\tilde{h}(\mathbf{w} + \mathbf{e}_r)]$$
$$+ \sum_{r=1}^{L} \frac{w_r(w_r - 1)}{(N - 1)}[\tilde{h} + \frac{3(N - w_r + 1)}{w_r}\tilde{h}(\mathbf{w} - \mathbf{e}_r)], \quad (A19)$$

and

$$\frac{d\tilde{h}_{2,\text{non-on-site}}}{dt} = \frac{g^2}{2N}\sum_{\langle rr'\rangle}3(N - w_r + 1)w_{r'}\tilde{h}(\mathbf{w} - \mathbf{e}_r)$$
$$+ Nw_{r'}\tilde{h} + w_{r'}(w_r + 1)\tilde{h}(\mathbf{w} + \mathbf{e}_r) + r \leftrightarrow r'. \quad (A20)$$

Combining everything together, we obtain the master equation of $\tilde{h}$ presented in Eq. (11) in the main text.

APPENDIX B: MORE DETAILS ON THE BROWNIAN COUPLED CLUSTER IN THE INFINITE-N LIMIT

In this Appendix, some additional details on the infinite-N noiseless limit of the BCC model are presented. The continuum limit of Eq. (14) resembles a FKPP-type partial differential equation, Eq. (15).

To justify the continuum approximation, here we directly study the original discrete ordinary differential equation on the lattice. Starting with

$$\partial_t\phi(r, t) = 3\left(\phi(r, t) + \frac{g^2}{2}(\phi(r - 1, t) + \phi(r + 1, t))\right), \quad (B1)$$

we look for the traveling-wave solution $\exp(\lambda(t - r/v))$ in the small ϕ limit, so the nonlinearity can be safely ignored. The relation between the velocity v and the growth rate is
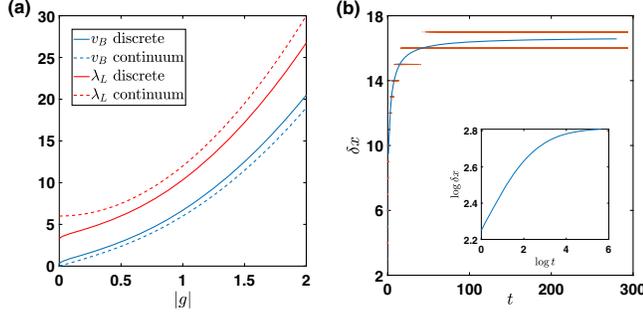
FIG. 6.   (a) The difference of v_B and λ_L between the discrete infinite-N BCC and the continuum approximation. (b) Direct simulation of the discrete model by numerically solving the differential equation (B1). The spatial distance between two contours of $\log C(r, t)$ saturates in the late time, showing that the front is sharp. On a log-log plot, the slope of the curve decreases to zero. This result is in sharp contrast to the local Hamiltonian systems, where the slope increases to $1/2$ and $1/3$ for the chaotic case and noninteracting case, respectively, as shown in Figs. 4 and 5.

$$\lambda = 3 + 3g^2 \cosh \frac{\lambda}{v}. \tag{B2}$$

The minimum velocity for the positive growth rate is the butterfly velocity v_B, and the corresponding growth rate is the Lyapunov exponent. In Fig. 6(a), we plot v_B and λ_L as a function of g and compare them with the analytical result from the continuum limit. Overall, the result from the continuum approximation tracks that from the discrete model. The main difference occurs in the limit that $g \ll 1$: Analyzing Eq. (B2) shows that

$$v_B \sim -\frac{3}{2 \log |g| + 1},$$
$$\lambda_L \sim 3 - \frac{3}{2 \log |g| + 1}, \tag{B3}$$

while the continuum approximation predicts that $v_B \sim 3\sqrt{2}|g|$ and $\lambda_L \sim 6(1 + g^2)$.

Furthermore, we directly simulate the discrete nonlinear ordinary differential equation. The result is shown in Fig. 6(b). As time increases, the spatial difference between two contours of $\log C(r, t)$ saturates. This result explicitly verifies the sharp wave front and the exponential growth of the squared commutator.

APPENDIX C: BUTTERFLY VELOCITY AND DIFFUSION CONSTANT FROM THE NOISY FKPP EQUATION

In this Appendix, we discuss how to obtain Eq. (18) for large but finite-N BCC by analyzing Eq. (15) with the hard-cutoff approximation and the noise term Eq. (17). This material is essentially a review of the analysis of Brunet

et al. [52,53]. To simplify the notation, we rescale ϕ, r, and t as follows:

$$\phi \to \frac{4}{3}\phi,$$
$$r \to \sqrt{2(1 + g^2)/g^2},$$
$$t \to 3(1 + g^2)t. \tag{C1}$$

After the rescaling, Eq. (15) becomes

$$\partial_t \phi = (1 - \phi)(\partial_r^2 \phi + \phi), \tag{C2}$$

and the noise term becomes

$$f_{\text{noise}} = \frac{(2 + 2g^2)^{1/4}}{(3g)^{1/2}} \sqrt{\frac{1}{2N}(2 - \phi)(\partial_r^2 \phi + \phi)} \eta(r, t). \tag{C3}$$

In this section, we mainly focus on Eqs. (C2) and (C3).

1. Noiseless case

We first discuss the case without noise, corresponding to the infinite-N limit of the BCC model. Equation (C2) is similar to the Fisher-KPP equation,

$$\partial_t \phi = \partial_r^2 \phi + \phi(1 - \phi), \tag{C4}$$

describing a growth-saturation process occurring in a wide class of systems, including population dynamics, combustion, and reaction-diffusion systems. One of the interesting features of the Fisher-KPP equation is that it admits traveling-wave solutions $\phi(r, t) = w(r - vt)$ that the initial configurations converge to [60]. We expect that Eq. (C2) obtained from unitary dynamics also falls into the FKPP universality class because the linearized version of Eq. (C2) is the linearized Fisher-KPP equation and because Eq. (C2) also has saturation physics so that $\phi = 1$ is a stable solution.

For the Fisher-KPP equation, given an initial condition $\phi(r, 0)$ that is sufficiently well localized, it asymptotically approaches the traveling-wave solution,

$$\lim_{t \to \infty} \phi(z + m(t), t) = w_v(z),$$
$$\lim_{t \to \infty} \frac{m(t)}{t} = v, \tag{C5}$$

where $m(t)$ is the position of the wave front defined by the equation $\phi(m(t), t) = $ constant.

Traveling waves.—The first question to answer is, what is the shape of the traveling-wave solution $w_v(z)$ for different velocities? It can be determined from the following equation,

$$-v\partial_z w(z) = (\partial_z^2 w(z) + w(z))(1 - w(z)), \tag{C6}$$

with the boundary conditions $w(\infty) \to 0$ and $w(-\infty) \to 1$. In the region that $z \gg 0$ and $w(z) \ll 1$, the shape can be understood from the linearized equation. For each velocity, there are two modes $e^{-\gamma z}$ and $e^{-z/\gamma}$, where $\gamma + 1/\gamma = v$. At the critical velocity $v_c = 2$, $\gamma = 1/\gamma = 1$, and the two modes are ze^{-z} and ze^{-z}. On the level of the linear equation, for a given velocity, any combination of the two modes is valid. The effect of the nonlinearity can be regarded as setting a boundary condition for the linearized equation, say, at $z = 0$,

$$w(0) = \alpha(v), \qquad w'(0) = \beta(v), \qquad (\text{C7})$$

where $\alpha(v)$ and $\beta(v)$ are tied to each other based on the solution to the full nonlinear equation. The above boundary condition forces both decay modes to appear, with the slower decaying mode dominating the behavior of $w(z)$ in the large-z limit. In the case where $v < 2$, γ becomes complex, and both modes can appear as long as the combination is real. Therefore, we have

$$w(z) \sim \begin{cases} e^{-\gamma z}, \gamma < 1 & \text{if } v > 2 \\ ze^{-\gamma z} & \text{if } v = 2 \\ ae^{-\gamma z} + a^* e^{-\gamma^* z} & \text{if } v < 2. \end{cases} \qquad (\text{C8})$$

In the current context, ϕ is interpreted as the operator weight and is always positive. Therefore, the last case where $w(z)$ oscillates is physically irrelevant, setting the minimal physical velocity to $v_c = 2$. However, the last case becomes important for the noisy case discussed below.

Approaching the traveling waves.—Great efforts have been made to understand the relationship between the initial configuration and the final traveling wave it asymptotes to. It is found that for an initial perturbation which is sufficiently localized, $\phi(r, t)$ approaches the critical traveling wave in the long time limit. This result can be understood from the linearized equation,

$$\partial_t \phi = \partial_r^2 \phi + \phi. \qquad (\text{C9})$$

Using Green's function, we can write down the general solution as

$$\phi(r, t) = \int dr' \frac{e^{t - \frac{(r'-r)^2}{4t}}}{2\sqrt{\pi}\sqrt{t}} \phi(r', 0), \qquad (\text{C10})$$

where $\phi(r', 0)$ is given by the initial condition. Starting with $\phi(r, 0) = e^{-\lambda|r|}$, we obtain

$$\lim_{t \to \infty} \phi(vt + z, t) \sim \begin{cases} \frac{1}{\sqrt{t}} e^{t(1 - \frac{v^2}{4}) - \frac{vz}{2}} & \text{if } v < 2\lambda \\ e^{(1 - \lambda v + \lambda^2)t - \lambda z} & \text{if } v \geq 2\lambda. \end{cases} \qquad (\text{C11})$$

The velocity of the wave front is determined by choosing v to cancel the t dependence in the exponent so that ϕ approaches a constant in the traveling frame. As a result,

$$v_B = \begin{cases} 2, & \text{if } \lambda \geq 1 \\ \lambda + \frac{1}{\lambda} & \text{if } \lambda < 1. \end{cases} \qquad (\text{C12})$$

This result demonstrates that an initial, sufficiently localized configuration travels with the minimal velocity $v_c = 2$; i.e., the leading term of the wave-front position $m(t)$ is $2t$. The asymptotic form is

$$\phi(v_c t + z, t) \sim \frac{1}{\sqrt{t}} e^{-z}. \qquad (\text{C13})$$

From this result, one can also obtain the subleading term of $m(t)$ by canceling the $1/\sqrt{t}$ prefactor, which gives rise to $m(t) \sim 2t - \frac{1}{2}\log(t)$. Then, $\phi(m(t) + z, t) \to e^{-z}$.

The linearized equation gives the right velocity. However, the subleading term in $m(t)$ is not correct. The nonlinearity forces the waveform at the critical velocity to decay like ze^{-z}, more slowly than the e^{-z} form obtained above. To take into account the nonlinear effects, we note that $\partial_r \phi(r, t)$ is also a solution to the linearized equation, and we can combine $\partial_r \phi(r, t)$ and $\phi(r, t)$ to obtain a new solution $\tilde{\phi}$ that minimizes the effect of the nonlinear term $\phi(r, t)^2$. By expanding Eq. (C13) to the next order, one finds that

$$\tilde{\phi}(v_c t + z, t) = (\phi + \partial_r \phi) \to \frac{z}{t^{3/2}} e^{-z}, \qquad (\text{C14})$$

which indeed has the corrected asymptotic behavior as a function of z. We can again obtain the subleading term in $m(t)$ by canceling the time dependence,

$$m(t) \sim 2t - \frac{3}{2}\log(t). \qquad (\text{C15})$$

The second term was found by Bramson [60] and turns out to be independent of the specific shape of the initial condition as long as it is sufficiently localized.

To incorporate a simple saturation mechanism into the linearized equations and enforce the asymptotic shape of the traveling-wave solution, Berestycki *et al.* [61] recently solved the linearized equation with the following moving boundary condition,

$$\phi(m(t), t) = \alpha, \qquad \partial_r \phi(m(t), t) = \beta, \qquad (\text{C16})$$

in order to obtain the vanishing correction of $m(t)$. They found that

$$m(t) \sim 2t - \frac{3}{2}\log(t) - \frac{3\sqrt{\pi}}{\sqrt{t}} + \cdots. \qquad (\text{C17})$$

The same correction has been identified in other models and is therefore expected to be universal, applying to the original Fisher-KPP equation and Eq. (C2) as well.

2. Noisy case

The above picture applies to the infinite-N limit of the BCC model. As stated in the main text, $1/N$ expansion away from the limit has two main effects. First, we need the cutoff $\phi(r, t)$ to 0 whenever it is below $1/N$; second, we need to include the noise term Eq. (C3) in Eq. (C2).

Cutoff velocity.—We first discuss the effect of the cutoff without considering the noise. As discussed above, due to the positivity of ϕ, the velocity of the traveling wave can never go below $v_c = 2$. However, with the cutoff, the part of ϕ below $1/N$ is set to zero by hand, and a smaller velocity becomes possible. The scaling behavior of the velocity as a function N can be obtained following Ref. [52]. When $v < 2$, the tail of the traveling wave acquires an oscillating part with a long wavelength in addition to the decay,

$$w(z) \sim \sin(\gamma_I z) e^{-\gamma_R z}, \tag{C18}$$

with γ_I and γ_R the real and imaginary parts of γ, respectively. Even with the cutoff, $w(z)$ should still remain positive until it decays to $1/N$. This case imposes a constraint on the wavelength of the oscillation part, which is determined by γ_I, requiring $\gamma_I < [\pi/(\log N)]$. In consequence, $v > v_c - [\pi^2/(\log^2 N)]$. Therefore, the velocity correction scales as $1/\log^2 N$, which is consistent with the numerical result presented in the inset of Fig. 3(a).

To quantify the above picture, and especially to understand how different initial conditions approach the asymptotic traveling wave with the cutoff, Brunet *et al.* [52] introduced a third boundary condition $\phi(m(t) + L, t) = 0$, with $L \sim \log N$, to the linearized equation in addition to Eq. (C16). This new boundary condition is to account for the hard cutoff occurring at the leading edge of the traveling wave. They also set $\alpha = 0$ for simplicity, which seems unnatural but does not affect the shape of the traveling wave in the large-z limit. In the following, we repeat their analysis in some detail. With this setup, it is convenient to go to the traveling frame by performing the substitution $\phi(r, t) = w(r - m(t), t)$. Then, the boundary conditions are

$$\partial_t w = \dot{m}(t)\partial_z w + \partial_z^2 w + w,$$
$$w(0, t) = 0, \qquad w(L, t) = 0,$$
$$\partial_z w(0, t) = 1. \tag{C19}$$

We first approximate $\dot{m}(t)$ as its asymptotic value v, which is to be determined. By setting the time derivative of $\dot{m}$ to zero, the boundary conditions uniquely determine the form of the asymptotic traveling wave,

$$w(z) = \frac{L}{\pi}\sin\frac{n\pi}{L}z \exp\left(-\frac{v}{2}z\right), \tag{C20}$$

with the velocity $v = 2\sqrt{1 - (\pi^2/L^2)}$.

Understanding the asymptotic form, we restore $\dot{m}(t)$ in the equation and study the full dynamics of $w(z, t)$. To the leading order of L, $w(z, t)$ can be written as a superposition of eigenmodes,

$$w(z, t) = \sum a_n \frac{L}{\pi}\sin\frac{n\pi}{L}z$$
$$\times \exp\left(-z + \frac{\pi^2}{L^2}(1 - n^2)t + vt - m(t)\right), \tag{C21}$$

where a_n is obtained from Fourier expanding the initial condition, and $m(t)$ is tuned so that $\partial_z w(0, t) = 0$ for all time. All modes with $n > 1$ decay exponentially. In the long-time limit, since $m(t) \to vt$, we obtain

$$w(t \to \infty, z) = a_1\frac{L}{\pi}\sin\frac{\pi z}{L}\exp(-z + vt - m(t)),$$
$$a_1 = \frac{2\pi}{L^2}\int_0^L w(0, z)\exp(z)\sin\frac{\pi z}{L}dz. \tag{C22}$$

Therefore,

$$m(t) - vt = \log a_1, \tag{C23}$$

in order to match the boundary condition. In other words, the relaxation to the asymptotic traveling wave causes an additional shift to the wave-front position.

Noise-induced diffusive motion.—Now, let us analyze the role of the noise term on top of the hard cutoff. Based on Eq. (C3), the noise scales as $\sqrt{\phi/N}$ and therefore is most prominent at the leading edge of the traveling wave where $\phi \sim (1/N)$. We argue that the wave front obeys

$$m(t) \sim vt + \delta vt + X, \tag{C24}$$

where X is a random diffusive process with $\langle X \rangle = 0$ and $\langle X^2 \rangle = 2Dt$.

Different from the deterministic model studied above, in the actual noisy equation, it is possible that $\phi(r, t)$ is on the order of $1/N$ even when $z > L$. Let $L + \delta$ denote the maximal distance that $w(r, t) \neq 0$, where δ is a random variable. Then, in the region that $L < z < L + \delta$, the noise term scales as e^δ/N, significantly larger than $w(z, t)$ itself, which scales as $1/N$. The wave-front shift caused by the noise in a unit time is

$$\Delta z(\delta) \sim \log\left(1 + \frac{2\pi}{L^2}\int_{L-\delta}^{L} e^{\delta+z-L}\sin\frac{\pi z}{L}dz\right)$$

$$\sim \log\left(1 + \frac{e^\delta}{L^3}\right). \qquad (C25)$$

Based on a phenomenological reaction-diffusion model, Brunet *et al.* [53] obtained the probability for a large δ to appear in a unit time as

$$p(\delta) \sim e^{-\delta}, \qquad (C26)$$

decaying with the natural decay constant in the system. Then, the additional velocity correction and the diffusion constant can be calculated straightforwardly as

$$\delta v \sim \int \Delta z(\delta)p(\delta)d\delta \sim \frac{1}{L^3},$$

$$D \sim \int (\Delta z(\delta))^2 p(\delta)d\delta \sim \frac{1}{L^3}. \qquad (C27)$$

This suggests that the system approaches its infinite-N limit extremely slowly.

APPENDIX D: S-MPS SIMULATION OF STOCHASTIC PROCESSES

In this Appendix, we present more details on using matrix product state techniques to simulate the master equation, Eq. (11).

1. Matrix product state for simulating a stochastic process

The key idea of S-MPS is to represent a probability distribution in a matrix product form and update the MPS based on the master equation for the probability distribution. A probability distribution ρ (viewed as a diagonal density matrix) and a quantum state ψ have similar structures, both containing L indices with dimension d, where L is the number of sites and d is the number of physical states per site. But the normalization is different. The normalization of a quantum state is $\langle\psi|\psi\rangle = 1$, a 2-norm condition, while the normalization of a probability distribution is $\text{Tr}(\rho) = 1$, a 1-norm condition. Furthermore, each element of ρ must be positive. Most conventional MPS techniques, such as the canonical form, are built around the 2-norm structure. They can still be applied to a probability distribution ρ for small system sizes, but numerical stability issues are encountered in larger systems. One reason is that the 1-norm is much larger than 1 if the 2-norm is kept to 1. Therefore, conventional MPS techniques need modification for simulating large-scale stochastic processes.

Decomposition.—The first goal is to decompose a probability into a matrix product form that facilitates the calculation of local observables, similar to the infamous canonical form of matrix product states. To realize this goal, consider a probability distribution $\rho^{\alpha\beta}$, where α represents the state in the first site and β represents the rest. The system is assumed to consist of L sites with open boundary conditions. In the following, we use superscripts for physical d.o.f. and subscripts for auxiliary indices.

Step 1: Break ρ into two pieces, $\rho^\alpha_{(l),m}$ and $\rho^\beta_{(r),m}$, so that $\sum_m \rho^\alpha_{(l),m}\rho^\beta_{(r),m} = \rho^{\alpha\beta}$. Here, (l) and (r) stand for left and right, as we have in mind a sweeping procedure. This can be achieved using the usual Schmidt decomposition or a LQ decomposition.

Step 2: Do a local basis change on the auxiliary dimension so that $\sum_\alpha \rho^\alpha_{(l),m} = \delta_{m,1}$. This change can be achieved by performing a LQ decomposition on $\sum_\alpha \rho^\alpha_{(l),m}$,

$$\sum_\alpha \rho^\alpha_{(l),m} = \sum_{m'}\lambda\delta_{m',1}Q_{m'm}, \qquad (D1)$$

where Q is unitary and λ is a number. Then, $\rho^\alpha_{(l),m}$ and $\rho^\beta_{(r),m}$ are updated as follows,

$$\rho^\alpha_{(l),m} \rightarrow \frac{1}{\lambda}\sum_{m'}\rho^\alpha_{(l),m'}Q^{-1}_{m'm},$$

$$\rho^\beta_{(r),m} \rightarrow \lambda\sum_{m'}Q_{mm'}\rho^\beta_{(r),m'}. \qquad (D2)$$

Now, we factor out the first tensor $\rho(1)^\alpha_m = \rho^\alpha_{(l),m}$ to yield the first matrix in the matrix product form.

Step 3: The goal is to factor out the next tensor from $\rho^\beta_{(r),m}$. First, rewrite it as $\rho^{\alpha\beta'}_{m_1}$ by dropping the (r) label and breaking β into α and β', with α representing the state on site 2 and β' the states on sites 3 to L. The label m_1 indicates that this auxiliary index is associated with the first link in the MPS. Because of the previous steps, $\rho^{\alpha\beta'}_{m_1=1}$ is the reduced probability distribution for sites 2 to L. We again decompose it into $\sum_m \rho^\alpha_{(l),m_1,m}\rho^\beta_{(r),m}$ by, say, a SVD. Keep in mind that the auxiliary bond m now has dimension d^2 since the combined index αm_1 has dimension d^2.

Step 4 (optional): Perform a local basis change on the auxiliary space m so that $\rho^\alpha_{(l),m_1=1,m} = 0$ for $m > d$. This change can be achieved by performing a LQ on $\rho^\alpha_{(l),m_1=1,m}$ and using the unitary matrix Q to update the tensors as follows,

$$\rho^\alpha_{(l),m_1=1,m} = \sum_{m'}\rho^\alpha_{(l),m_1=1,m'}Q_{m'm},$$

$$\rho^\alpha_{(l),m_1>1,m} \rightarrow \sum_{m'}\rho^\alpha_{(l),m_1>1,m'}Q^{-1}_{m'm},$$

$$\rho^{\beta'}_{(r),m1,m} \rightarrow \sum_{m'}Q_{mm'}\rho^{\beta'}_{(r),m'}. \qquad (D3)$$

This step is not necessary to bring the probability distribution into the S-MPS canonical form, but it is important for the purpose of truncation as discussed later.

Step 5: Similar to step 2, we perform another local basis change to make sure that $\sum_\alpha \rho^\alpha_{(l),m_1=1,m} = \delta_{m,1}$ by doing a LQ decomposition on $\sum_\alpha \rho^\alpha_{(l),m_1=1,m}$ and updating the tensors $\rho^\alpha_{(l),m_1,m}$ and $\rho^\alpha_{(r),m}$ as in Eq. (D2).

Step 6: Iterate step 3 through step 5 until site L is reached. The whole procedure is called a right sweep, and it produces the 1-norm right canonical form of S-MPS. One could equally well start with site L and perform a left sweep by changing the LQ decomposition to a QR decomposition to obtain the left canonical form.

Local observables: To measure a local observable at site i, one can first right sweep from site 1 to $i - 1$ and then left sweep from site L to site $i + 1$. Then, the probability distribution is in the so-called mixed canonical form,

$$\rho^{\alpha\gamma\beta} = \sum_{mm'} \rho^\alpha_m \rho^\gamma_{mm'} \rho^\beta_{m'}, \tag{D4}$$

where γ is the index for the states at site i, α is for the states to the left of i, and β is for the states to the right of i. By construction, $\rho^\gamma_{1,1}$ is the reduced probability distribution for site i, $\sum_{\alpha\beta} \rho^{\alpha\gamma\beta}$, from which the calculation of local observables is straightforward. Thus, we have realized the goal sketched in Fig. 7, where the calculation of local observables is reduced to manipulating local data.

Simulating the master equation and truncation.—To simulate a local master equation such as Eq. (11), we view the generator of the stochastic evolution as a non-Hermitian Hamiltonian, which explicitly conserves 1-norm. Then, we can roughly follow the usual time-evolving block decimation (TEBD) steps to update the S-MPS after a short time step but with the sweeping procedures replaced by those described in the last section.

Similar to TEBD, the bond dimension of the S-MPS typically grows rapidly with time, and truncation is always necessary. Consider the situation after updating the tensors at site i and site $i + 1$. The probability distribution is in the following form,
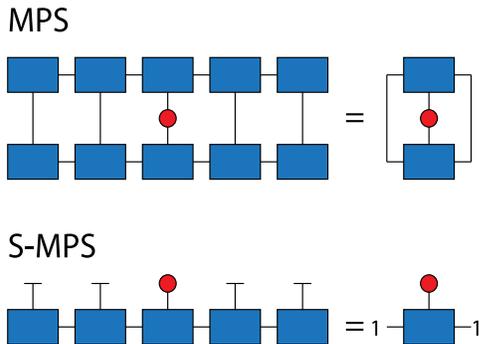
MPS

S-MPS

FIG. 7.   Procedures for measuring local observables in a MPS in canonical form (top) and a S-MPS in canonical form (bottom).

$$\rho^{\alpha\gamma_1\gamma_2\beta} = \sum_{m_l,m_r} \rho^\alpha_{m_l} \rho^{\gamma_1\gamma_2}_{m_l m_r} \rho^\beta_{m_r}, \tag{D5}$$

where γ_1 and γ_2 represent the states on site i and $i + 1$, respectively, and, by construction, $\rho^{\gamma_1\gamma_2}_{11}$ is the reduced probability distribution of site i and site $i + 1$. Assuming that the stochastic simulation starts with a S-MPS with bond dimension χ, the dimension of the middle matrix $\rho^{\gamma_1\gamma_2}_{m_l m_r}$ is χd; i.e., the bond dimension at the bond linking sites i and $i + 1$ is χd. The goal is to reduce it back to χ by breaking $\rho^{\gamma_1\gamma_2}_{m_l m_r}$ into two pieces, $\rho^{\gamma_1}_{m_l,m}$ and $\rho^{\gamma_2}_{m,m_r}$, where the dimension of m is χ. Then, one can continue to right sweep or left sweep to update the next bond.

There are two comparable methods to achieve this. The most straightforward method is to perform a SVD on $\rho^{\gamma_1\gamma_2}_{m_l m_r}$ by regarding it as a matrix with dimension χd. After keeping only the leading χ singular values, the bond dimension is reduced back to χ.

The second method is one developed by White *et al.* in a recent paper studying density matrix dynamics. They first break $\rho^{\gamma_1\gamma_2}_{m_l m_r}$ into two pieces and perform a single right- or left-sweeping step (step 3 to step 5) on the left or right piece, so $\rho^{\gamma_1\gamma_2}_{m_l m_r}$ is expressed as

$$\rho^{\gamma_1\gamma_2}_{m_l m_r} = \sum_{mm'} \rho^{\gamma_1}_{m_l,m} Q_{mm'} \rho^{\gamma_2}_{m',m_r}, \tag{D6}$$

with the properties that $\rho^{\gamma_1}_{1,m>d} = 0$ (due to step 4 in the last section), $\sum_{\gamma_1} \rho^{\gamma_1}_{1,m} = \delta_{m,1}$, and similarly for $\rho^{\gamma_2}_{m',1}$. In general, the rank of the matrix Q is χd.

As they pointed out, given Eqs. (D5) and (D6), the right lower $(\chi - 1)d \times (\chi - 1)d$ section of Q does not affect the reduced probability distribution of site 1 to site i, and site $i + 1$ to site L, precisely due to step 4 in the sweeping procedure. This result gives some freedom to manipulate that section of the matrix in order to reduce the rank of Q without affecting local observables (they were concerned with developing a truncation scheme that respects local observables). For example, if one performs a SVD on that section and keeps the leading $\chi - 2d$ singular values, then the resulting Q matrix would have rank χ, as required. One could also follow White *et al.* and reduce the rank while minimizing the error occurring in the correlation functions between the left part $(1 \sim i)$ and the right part $(i + 1 \sim L)$ of the system.

Comparing the two methods, the second one is more appealing since a truncation step at a particular bond does not change the reduced probability distribution of the left and right parts of the system. Furthermore, the truncation scheme automatically preserves conserved quantities, like the 1-norm or the total amount of some conserved charge, for all time regardless of the bond dimension. However, the accuracy of the time dependence of local quantities still depends on the bond dimension. After one sweeps through the system and performs the truncation on each bond, only
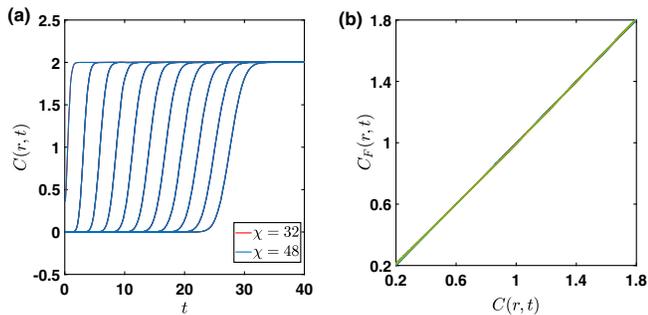
FIG. 8. (a) Up to $N = 10$ spins in each cluster and $L = 200$ clusters, the numerical data for the squared commutator converge excellently with the bond dimension of the S-MPS. Bond dimension $\chi = 32$ and bond dimension $\chi = 48$ give rise to almost identical results for all timescales. This result allows us to access the entire scrambling behavior from early growth to late-time saturation. (b) Near the wave front, $C(r, t)$ perfectly agrees with the fitting function $C_{\text{fit}}(r, t) = 1 + \text{erf}((v_B t - r - r_0)/\sqrt{4Dt})$ for all N.

local observables with support on up to two neighboring sites remain unaffected. Furthermore, as time increases in the simulation, the errors made in observables with larger support could feed back to the dynamics of the two-site observables.

In practice, for our goal of calculating the squared commutator from the master equation, Eq. (11), we find that the results obtained from these two truncation schemes are similar, and the first scheme is 2 to 3 times faster since it requires fewer steps of SVD.

2. Applying S-MPS to the master equation of the Brownian couple cluster

We apply the technique described above to simulate the master equation, Eq. (11). We first check the convergence of the resulting squared commutator with the bond dimension of the S-MPS. As shown in Fig. 8, a S-MPS with bond dimension $\chi = 32$ already produces converged results for all timescales for $N = 10$ spins within a single cluster and a total of $L = 200$ clusters. This result allows us to access both early growth and late-time saturation of scrambling in the BCC. We then fit $C(r, t)$ in the region near the wave front with an error function $1 + \text{erf}[(r - v_B t - r_0)/\sqrt{4Dt}]$ to extract the butterfly velocity v_B and the diffusion constant D. The fitting quality is shown in Fig. 8(b).

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

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

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

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

[5] J. M. Deutsch, *Quantum Statistical Mechanics in a Closed System*, Phys. Rev. A **43**, 2046 (1991).

[6] M. Srednicki, *Chaos and Quantum Thermalization*, Phys. Rev. E **50**, 888 (1994).

[7] H. Tasaki, *From Quantum Dynamics to the Canonical Distribution: General Picture and a Rigorous Example*, Phys. Rev. Lett. **80**, 1373 (1998).

[8] M. Rigol, V. Dunjko, and M. Olshanii, *Thermalization and Its Mechanism for Generic Isolated Quantum Systems*, Nature (London) **452**, 854 (2008).

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

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

[11] S. Xu and B. Swingle, *Accessing Scrambling Using Matrix Product Operators*, arXiv:1802.00801.

[12] D. A. Roberts, D. Stanford, and A. Streicher, *Operator Growth in the SYK Model*, J. High Energy Phys. 06 (2018) 122.

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

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

[15] Y.-Z. You and Y. Gu, *Entanglement Features of Random Hamiltonian Dynamics*, Phys. Rev. B **98**, 014309 (2018).

[16] X. Chen and T. Zhou, *Operator Scrambling and Quantum Chaos*, arXiv:1804.08655.

[17] B. Swingle, G. Bentsen, M. Schleier-Smith, and P. Hayden, *Measuring the Scrambling of Quantum Information*, Phys. Rev. A **94**, 040302(R) (2016).

[18] G. Zhu, M. Hafezi, and T. Grover, *Measurement of Many-Body Chaos Using a Quantum Clock*, Phys. Rev. A **94**, 062329 (2016).

[19] 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.

[20] N. Y. Halpern, *Jarzynski-like Equality for the Out-of-Time-Ordered Correlator*, Phys. Rev. A **95**, 012120 (2017).

[21] N. Y. Halpern, B. Swingle, and J. Dressel, *Quasiprobability Behind the Out-of-Time-Ordered Correlator*, Phys. Rev. A **97**, 042105 (2018).

[22] M. Campisi and J. Goold, *Thermodynamics of Quantum Information Scrambling*, Phys. Rev. E **95**, 062127 (2017).

[23] B. Yoshida and A. Kitaev, *Efficient Decoding for the Hayden-Preskill Protocol*, arXiv:1710.03363.

[24] M. Gärttner, 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*, Nat. Phys. **13**, 781 (2017).

[25] K. X. Wei, C. Ramanathan, and P. Cappellaro, *Exploring Localization in Nuclear Spin Chains*, Phys. Rev. Lett. **120**, 070501 (2018).

[26] 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).

[27] E. J. Meier, J. Ang'ong'a, F. A. An, and B. Gadway, *Exploring Quantum Signatures of Chaos on a Floquet Synthetic Lattice*, Phys. Rev. A **100,** 013623 (2019).

[28] D. Chowdhury and B. Swingle, *Onset of Many-Body Chaos in the $O(N)$ Model*, Phys. Rev. D **96,** 065005 (2017).

[29] A. A. Patel, D. Chowdhury, S. Sachdev, and B. Swingle, *Quantum Butterfly Effect in Weakly Interacting Diffusive Metals*, Phys. Rev. X **7,** 031047 (2017).

[30] S. Sachdev and J. Ye, *Gapless Spin-Fluid Ground State in a Random Quantum Heisenberg Magnet*, Phys. Rev. Lett. **70,** 3339 (1993).

[31] A. Kitaev, *A Simple Model of Quantum Holography*, in *KITP Progr. Entanglement Strongly-Correlated Quantum Matter* (unpublished).

[32] 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.

[33] D. J. Luitz and Y. Bar Lev, *Information Propagation in Isolated Quantum Systems*, Phys. Rev. B **96,** 020406(R) (2017).

[34] 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 (2016).

[35] M. Heyl, F. Pollmann, and B. Dóra, *Detecting Equilibrium and Dynamical Quantum Phase Transitions in Ising Chains via Out-of-Time-Ordered Correlators*, Phys. Rev. Lett. **121,** 016801 (2018).

[36] C.-J. Lin and O. I. Motrunich, *Out-of-Time-Ordered Correlators in a Quantum Ising Chain*, Phys. Rev. B **97,** 144304 (2018).

[37] A. Nahum, J. Ruhman, and D. A. Huse, *Dynamics of entanglement and transport in one-dimensional systems with quenched randomness*, Phys. Rev. B **98,** 035118 (2018).

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

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

[40] B. Swingle and D. Chowdhury, *Slow Scrambling in Disordered Quantum Systems*, Phys. Rev. B **95,** 060201(R) (2017).

[41] Y. Chen, *Universal Logarithmic Scrambling in Many Body Localization*, arXiv:1608.02765.

[42] R. Fan, P. Zhang, H. Shen, and H. Zhai, *Out-of-Time-Order Correlation for Many-Body Localization*, Sci. Bull. **62,** 707 (2017).

[43] Y. Huang, Y.-L. Zhang, and X. Chen, *Out-of-Time-Ordered Correlators in Many-Body Localized Systems*, Ann. Phys. (Amsterdam) **529,** 1600318 (2017).

[44] V. Khemani, D. A. Huse, and A. Nahum, *Velocity-Dependent Lyapunov Exponents in Many-Body Quantum,*

Semiclassical, and Classical Chaos, Phys. Rev. B **98,** 144304 (2018).

[45] E. Leviatan, F. Pollmann, J. H. Bardarson, D. A. Huse, and E. Altman, *Quantum Thermalization Dynamics with Matrix-Product States*, arXiv:1702.08894.

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

[47] S. H. Shenker and D. Stanford, *Stringy Effects in Scrambling*, J. High Energy Phys. 05 (2015) 132.

[48] R. A. Fisher, *The Wave of Advance of Advantageous Genes*, Ann. Eugen. **7,** 355 (1937).

[49] A. Kolmogorov, I. Petrovsky, and N. Piscounov, *A Study of the Diffusion Equation with Increase in the Amount of Substance*, Bull. Univ. État Moscou **1,** 1 (1937).

[50] 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).

[51] S. Grozdanov, K. Schalm, and V. Scopelliti, *Kinetic Theory for Classical and Quantum Many-Body Chaos*, Phys. Rev. E **99,** 012206 (2019).

[52] E. Brunet and B. Derrida, *Shift in the Velocity of a Front Due to a Cutoff*, Phys. Rev. E **56,** 2597 (1997).

[53] E. Brunet, B. Derrida, A. H. Mueller, and S. Munier, *Phenomenological Theory Giving the Full Statistics of the Position of Fluctuating Pulled Fronts*, Phys. Rev. E **73,** 056126 (2006).

[54] G. Vidal, *Efficient Classical Simulation of Slightly Entangled Quantum Computations*, Phys. Rev. Lett. **91,** 147902 (2003).

[55] G. Vidal, *Efficient Simulation of One-Dimensional Quantum Many-Body Systems*, Phys. Rev. Lett. **93,** 040502 (2004).

[56] T. H. Johnson, S. R. Clark, and D. Jaksch, *Dynamical Simulations of Classical Stochastic Systems Using Matrix Product States*, Phys. Rev. E **82,** 036702 (2010).

[57] C. D. White, M. Zaletel, R. S. K. Mong, and G. Refael, *Quantum Dynamics of Thermalizing Systems*, Phys. Rev. B **97,** 035127 (2018).

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

[59] We know this cannot be true, in general, or at least the conditions for the bound to apply are different. The expectation for the thermally regulated object is that the bound applies after a time independent of distance x, but this cannot be true for the nonthermally regulated object. This is because the nonthermally regulated object always has a perturbative regime where $C \sim (t^x/x!)$, which has an unbounded derivative $(d \log C)/(dt)$ at large x.

[60] M. Bramson, *Convergence of Solutions of the Kolmogorov Equation to Travelling Waves*, Mem. Am. Math. Soc. **44,** 0 (1983).

[61] J. Berestycki, É. Brunet, S. C. Harris, and M. Roberts, *Vanishing Corrections for the Position in a Linear Model of FKPP Fronts*, Commun. Math. Phys. **349,** 857 (2017).