Asymmetric butterfly velocities in local time-independent Hamiltonians

The butterfly velocity v_B is the velocity at which operators spread, the effective speed limit in thermalizing systems. In many 1-D systems it is independent of the direction of spreading, but this need not be the case. In fact, with arbitrarily nonlocal Hamiltonians, or arbitrarily deep circuit models, the ratio of left- and right-going butterfly velocities may be made arbitrarily large. We present two models with asymmetric v_B. The first is a local Hamiltonian which is tunable and includes 3-site interactions. The butterfly velocity can be measured from either the nonlocal right- and left-weights or from the OTOC. In both methods, we show the asymmetry is large enough to be significant. The other model is a class of circuits, which we call n-staircases, where n can be interpreted as the depth of the circuit (Can it?). $n = 1$ is a random architecture with symmetric spreading while $n = \infty$ is completely chiral, so this class of circuits allows us to interpolate between the two extremes.

I. INTRODUCTION

Thermalization and operator spreading have recently been important subjects of study. In systems with time-independent Hamiltonians this spreading is often assumed to be symmetric, but there exists no such general constraint. Asymmetric transport is already seen in "staircase" and "glider" circuits. The existence of these circuits suggests that some asymmetric Hamiltonians exists, but they might not be general. One goal of this paper is then to describe tunable local asymmetric Hamiltonians.

The case of asymmetric circuits remains interesting. On the edge of a 2-D system, spreading can be chiral even with a finite circuit depth [?]. To be completely chiral with only 1 dimension, however, the circuit will have to be of infinite depth. After identifying the depth of the circuit with its locality (how?) we once again have a trade-off between locality and asymmetry. How asymmetric can a circuit of given depth be?

In this paper we will start by discussing a local Hamiltonian with asymmetric spreading. We show that it is a general Hamiltonian, and provide multiple methods for measuring v_B for left and right spreading. A representative plot can be seen in Fig 1. We then discuss staircase circuits in the small- and large-staircase limit and show that in the latter limit the circuit is completely chiral.

Throughout this paper we will use different methods to measure v_B. For the Hamiltonian system, we use two methods, both directly related to the spreading of operators. The first is non-local, measuring the velocity of the peak in the right- and left-weights. The other defines velocity-dependent Lyapunov exponents from the early-time OTOCs. For the circuit we extract v_B from the growth rate of the entanglement entropy.

II. LOCAL HAMILTONIANS

We want a multi-site Hamiltonian in which sites interact only through local interactions. We can accomplish this by defining an n-site Hamiltonian for small n, and
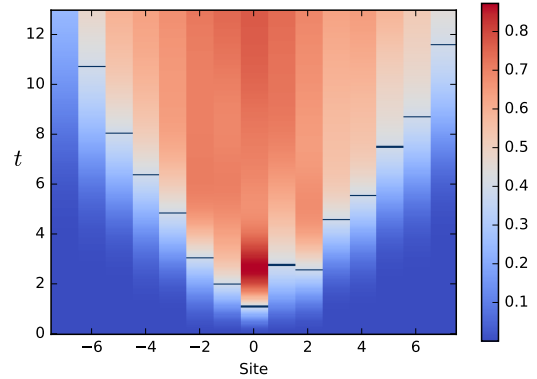

FIG. 1. Evolution of the initially local operator. Data is obtained at $L = 15, h = 0.35$ with the initial operator at the central site, averaged over 5 disorder realizations. The bars indicate the time at which the OTOC passes 0.4, to emphasize the asymmetry.

then putting this Hamiltonian on each set of n sites,

$$H_{\text{tot}} = \sum_{i=1}^{L-n+1} H_n^{(i)}, \tag{1}$$

where $H_n^{(i)}$ is a n-site Hamiltonian acting on sites i through $i + n - 1$. $n = 2$ will not suffice, because 2-site Hamiltonians are always symmetric with respect to their operator dynamics. To see this, consider the amount of information flowing from site to site. Unitarity preserves the total amount of information so if the 2-site Hamiltonian moves some weight from site i to $i + 1$ it must also move an equal amount from site $i + 1$ to i. 3-site Hamiltonian do not have this constraint, though, and can have asymmetric dynamics.

Instead of looking directly for an asymmetric Hamiltonian, we can find a unitary operator $U(t)$ with the dynamics we want. From that operator we can construct a Hamiltonian that gives $U(t) = e^{-iHt}$. One asymmetric unitary operator is the 3-site cyclic swap S_{123}. S_{123} is a

unitary operator such that

$$S_{123} |\alpha\beta\gamma\rangle = |\gamma\alpha\beta\rangle, \qquad (2)$$

where $|\alpha\beta\gamma\rangle$ is a product state with state $|\alpha\rangle$ on site 1, etc. The idea of using this operator is that it can transport a state from site 3 to site 1 in 1 step, but takes two applications to move a state from site 1 to site 3.

One way to build the three site swap gate is out of 2-site swap gates $S_{123} = S_{12}S_{23}$. Each 2-site swap interchanges two states, so the action is

$$\begin{aligned} S_{12}S_{23} |\alpha\beta\gamma\rangle &= S_{12} |\alpha\gamma\beta\rangle = |\gamma\alpha\beta\rangle \\ &= S_{123} |\alpha\beta\gamma\rangle. \end{aligned} \qquad (3)$$

Note that this construction can be extended to n sites to create an n-site swap gate. The gate would be a series of overlapping 2-site swap gates. We will come back to this construction to build our asymmetric circuits. The 2-site gates will not be swaps, but the geometry will be the same. The problem is that here we apply S_{12} after S_{23} while the staircases seem to be the other way around. I think it might have to do with evolving the state vs the operator.

It is also possible to build S_{123} out of a time-independent Hamiltonian, so that $U(t)|_{t=1} = e^{-iH_3} = S_{123}$. H_3 is the 3-site term we are looking for. There are many ways to construct this Hamiltonian, from directly taking the matrix logarithm to analyzing eigenstates. But the simplest way is to note that exchanging site indices gives S_{123}^{-1} while overall SU(2) rotations leave the gate unchanged. This means H_3 should be antisymmetric with respect to site indices and symmetric with respect to SU(2). Therefore H_3 is the triple product of the spin at each site.

The Hamiltonian on the full chain is then

$$H = \sum_{i=1}^{L-2} \mathbf{S}_i \cdot (\mathbf{S}_{i+1} \times \mathbf{S}_{i+2}),$$

The use of 3-site terms has some further consequences. For example, first order perturbation theory will connect site 1 to sites 2 and 3, while second order perturbations connect site 1 to sites 4 and 5. At early time sites 2 will behave the same as site 3, etc., leading to jaggedness in the spreading. We will correct for this by only looking at odd (or even) sites for each analysis.

A. Degeneracy and Generality

As is, the model is not general, with one symptom being a large degeneracy at $E = 0$. This is an effect of various antisymmetries in the model, which we will call R_i, such that $\{H, R_i\} = 0$. Each of these operators maps between states of opposite energy, $R_i|E\rangle = |-E\rangle$. Ref. [?] shows that each Tr R_i provides a lower bound on the

$E = 0$ degeneracy, which we will call N_0. Writing the $E = 0$ states as $|\alpha\rangle$,

$$\text{Tr } R_i = \sum_{\alpha=1}^{N_0} \langle\alpha|R|\alpha\rangle, \qquad (4)$$

But from $R_i^2 = 1$ we have

$$\langle\alpha|R_i|\alpha\rangle = \pm 1. \qquad (5)$$

Thus Tr $R_i < N_0$.

From the design of the model, one such R_i is the inversion operator, I. This can be composed with parity operators such as $P_X = \prod_i X_i$ that commute with both H and I to form new R_i. Neither operator saturates the degeneracy bound for all L, but Tr $I = N_0$ for odd L. If we break the SU(2) symmetry to U(1) but leave the inversion symmetry intact then Tr I is exact. If we add a uniform field in the Z direction, then IP_X is an antisymmetry but I is not. In this case Tr(IP_X) is exact. Taken together, these facts imply that the extra degeneracies for odd L come from interplay between the SU(2) symmetry and inversion antisymmetry.

We can fully break this degeneracy within each U(1) block by introducing a random field in the Z direction, so the total Hamiltonians is

$$H = \sum_{i=1}^{L-2} \mathbf{S}_i \cdot (\mathbf{S}_{i+1} \times \mathbf{S}_{i+2}) + \sum_{i=1}^{L} h_i S_i^z, \qquad (6)$$

where each h_i has a uniform probability distribution on $[-h, h]$. This field breaks the SU(2) symmetry but leaves the U(1) subgroup intact.

As we continue to increase h the model moves through the thermalizing phase and becomes localized. In the large-L limit the transition from ergodic to localized is a phase transition, described in [?]. The transition for the present model can be seen in Fig. 2, showing the ratio of adjacent energy gaps. Note that at smaller L the model also drifts away from GUE statistics at very small h, when the field is no longer large enough to sufficiently lift the $E = 0$ degeneracy. In order to avoid both of these regimes, we will perform all calculations with $h = 0.35$.

B. Right-weight peaks

To measure spreading we evolve initially local operators in the Heisenberg picture. For right-spreading the initial operator will be at site 1, and for left-spreading at site L. We will quantify the asymmetry using two metrics. The first is the weight of all operators with right (left) endpoint on site i, which we will call the right (left) weight. The other is the OTOC, which will be discussed below.

We use the definition of the right weight from [?]. An arbitrary operator $\mathcal{O}$ can be decomposed into Pauli strings $\mathcal{O} = \sum_\nu c_\nu \sigma^\nu$ where each string contains one of
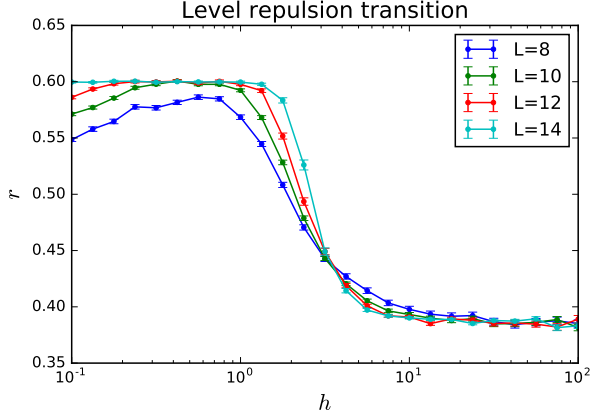
FIG. 2. Phase transition for the model, with level repulsion parameter plotted against field strength. Note that in the thermalizing phase the ratio is 0.6 instead of 0.53 because the statistics are GUE instead of GOE. I think I remember Vedika saying this but I can't find where.

$\{I, X, Y, Z\}$ acting on each site. As the operator evolves in time, so do the c_ν. The right weight is then

$$\rho_r(i,t) = \sum_\nu |c_\nu(t)|^2 \, \delta(\text{RHS}(\nu) = i), \qquad (7)$$

where the delta function ensures that we only count Pauli strings that have their right-most non-identity operator on site i. The left weight $\rho_l(i,t)$ is defined analogously. If $\mathcal{O}$ is initially local on site j then $\rho_r(i,0) = \rho_l(i,0) = \delta_{ij}$. As the operator spreads, the support of ρ_r moves right at $v_{B,r}$ and ρ_L moves left at $v_{B,l}$. Operator broadening manifests itself in the support of both weights increasing in size. At late times both weights should vanish near j.

In the thermalizing phase, the right weights peak as the information front passes. Because of the three-site nature of each term in the Hamiltonian, the right weight and OTOC exhibit an "odd-even" effect where site 3 peaks before 2, etc. It is possible to account for these by averaging judiciously, or by only looking at even (or odd) sites. Once this has been done the peaks do travel balistically. At $L = 13$, there are enough even sites that the asymmetry can be seen. We define $L = 2N + 1$ and number our sites from $-N$ to N. This makes comparing left- and right-weights easier. Since the peaks broaden at late times, we use the time that $\rho_r(i,t)$ reaches half-maximum to measure its velocity. For a picture of the rights weights with their successive peaks, see Fig. 3.

Fig. 4 shows the peaks traveling on even sites. The peaks reach equivalent sites at later times for the left-moving wave, implying $v_{B,l} < v_{B,r}$. We can extract $v_{B,l}$ and $v_{B,r}$ from these curves by fitting linear functions to the peak timings. We find $v_{B,l} = 1.011 \pm 0.072$ and $v_{B,r} = 0.791 \pm 0.062$. If we instead look at odd sites we find $v_{B,l} = 1.106 \pm 0.085$ and $v_{B,r} = 0.958 \pm 0.078$. In both cases the difference is larger than the error, although the difference is less exaggerated for odd sites.
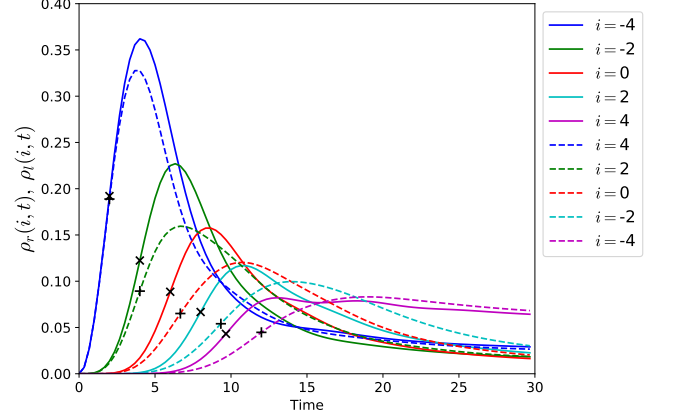

FIG. 3. Right weight at even sites for $L = 13$, $h = 0.35$, averaged over 100 disorder realizations. For $\rho_r(i,t)$, the initial operator is Z_1 while for $\rho_l(i,t)$ the initial operator is Z_{L-1}. The peak travels ballistically. Later peaks are smaller Is this due to broadening?
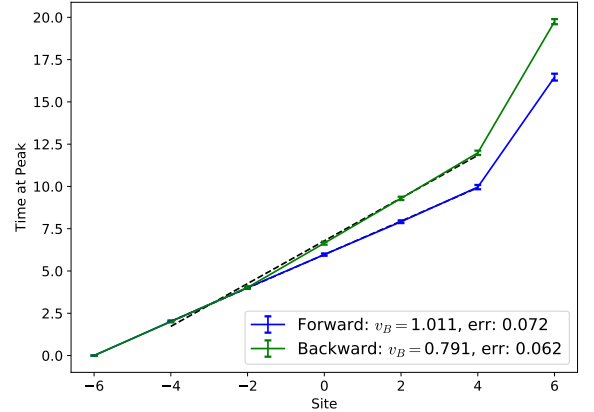

FIG. 4. Time of peak vs. site. The parameters are the same as in Fig. 3. Since this is plot of time as a function of distance, the larger slope in the left weight means that v_B is larger for propagation to the right. For the left-weight we plot against $(-\text{site})$ in order to compare left and right.

C. Velocity-dependent Lyapunov exponents

It is also possible to extract butterfly velocities from the the velocity-dependent Lyapunov exponents, which in turn rely on the OTOC. We define the OTOC as

$$C(i,t) = \frac{1}{2}\langle |[Z_j(t), Z_i(0)]|^2 \rangle_{\beta=0}$$
$$= 1 - \frac{1}{2^L}\text{Re Tr}\,[Z_j(t)Z_i(0)Z_j(t)Z_i(0)] \qquad (8)$$

where j is the site of the initial operator and the expectation value in the top row is with respect to a thermal

ensemble at infinite temperature. As in the right-weight case, we set $j = 1$ to measure $v_{B,r}$ and $j = L - 1$ to measure $v_{B,l}$. The OTOC should be order-1 inside the lightcone and exponentially small outside the lightcone defined by v_B.

Move this paragraph to an appendix? From conservation of $S_z^{\rm tot}$, the Hamiltonian and all relevant operators are block-diagonal, with the size of the $i^{\rm th}$ block being $\binom{L}{i}$. For smaller blocks we can compute the trace directly, but for larger blocks this becomes computationally difficult. We then rely on quantum typicality to approximate the trace in the large blocks. For each disorder realization we replace the trace with an average over expectation values in pure states [?]. The pure states are chosen Haar-randomly, and we find that using 5 vectors gives relative errors around 0.05 for blocks larger than 500×500. For smaller blocks we use exact diagonalization.

The VDLEs quantify how fast signals decay along constant-velocity trajectories outside the lightcone. In particular, if the OTOC is measured along the ray defined by each site i at time $t_i = i/v$ for some v, then it should decay exponentially,

$$C(i, t) \sim e^{\lambda(v)t} \quad \text{for} \quad i = vt. \tag{9}$$

Ref. [?] gives a thorough exposition and explanation of VDLEs. The name comes from the fact that the Lyapunov exponent defines how fast a signal grows inside a lightcone in a classically chaotic system. Fig. 1 shows the OTOC for an operator initially local at the center of the chain.

In the current system, the OTOCs are influenced by the previously-mentioned odd-even effects. We can once again look only at even sites for sufficiently large L to calculate $\lambda(v)$. Then v_B is the point at which $\lambda(v)$ smoothly goes to 0.

Fig. 5 illustrates the process of finding $\lambda(v)$ and shows the VDLEs for the right-going and left-going OTOCs, estimated from the even sites. Note that in the top plot the independent variable is distance, so the slope of the best fit line is $\lambda(v)/v$. Finite-size effects slightly perturb $\lambda(v)$ around v_B, but we can see that $v_{B,l} \sim 0.5$ and $v_{B,r} \sim 0.9$. Analysis of the odd sites is less clean, but suggests $v_{B,l} \sim 0.5$ and $v_{B,r} \sim 0.7$. These are not very close to the velocities estimated from the right weight. Since v_B is well defined only in the thermodynamic limit, we expect these methods to agree as $L \to \infty$.

III. CIRCUIT MODELS

We will now move away from Hamiltonians to quantum circuits, which give us more direct control over the asymmetry. Before discussing asymmetric circuits we will explain how v_B can be extracted from the growth of entanglement. We will then show that this method is particularly tractable in the large-q limit before applying this method to staircase circuits.
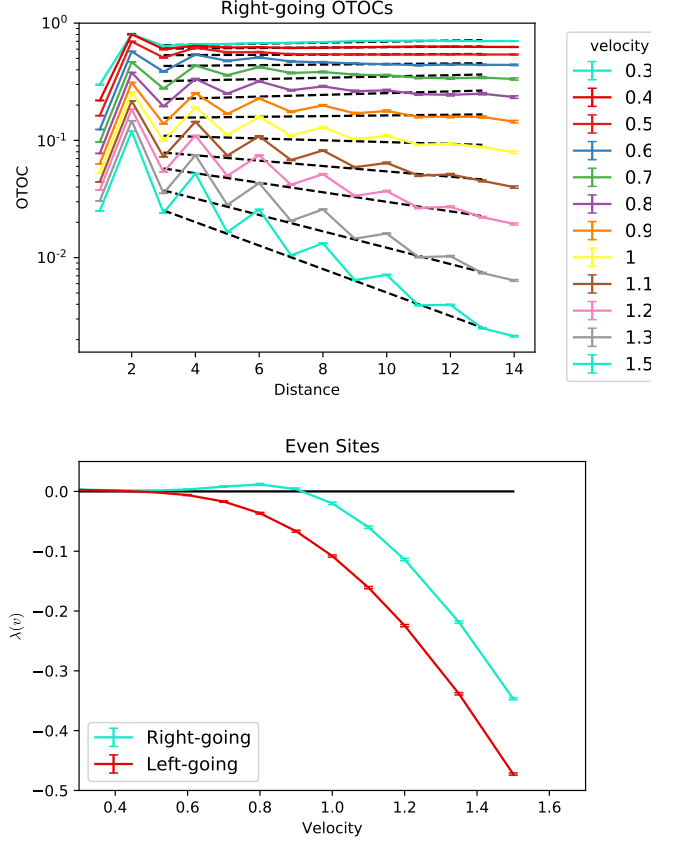

FIG. 5. Velocity-dependent Lyapunov exponents extracted from the OTOC on odd sites. The parameters are $L = 15, h = 0.35$, while the OTOC is as defined in the text with initial operators at sites 1 and $L - 1$ respectively. Statistics are obtained from 100 disorder realizations. The top figure shows the early-time right-moving OTOCs on a semilog plot. VDLEs are obtained from the best fit line through odd sites. The lower figure shows $\lambda(v)$. Since $\lambda_r(v) > \lambda_l(v)$, we know $v_{B,r} > v_{B,l}$.

Consider a spin chain of L sites, each with dimension q. Sites are labeled by $i = 1, \ldots, L$, while the bonds between sites are labeled by $x = 1, \ldots, L - 1$. Define the entropy function $S(x)$ as the bipartite entanglement entropy across bond x.

After course-graining, the entanglement becomes a continuous function $S(x, t)$. Given a circuit architecture, the entanglement growth rate is to first order only a function of the slope, so we can write [?]

$$\frac{\partial S}{\partial t} = \Gamma \left(\frac{\partial S}{\partial x} \right). \tag{10}$$

It is useful to define the entropy density $s = \partial S/\partial x$, which is so-called because the equilibrium entropy is $S(x, t) = s_{\rm eq} \min\{x, L - x\}$. In the next subsection we will show that the maximal slope is 1. In general the equilibrium slope can be smaller than 1, but since our systems are noisy, $s_{\rm eq} = s_{\rm max} = 1$.
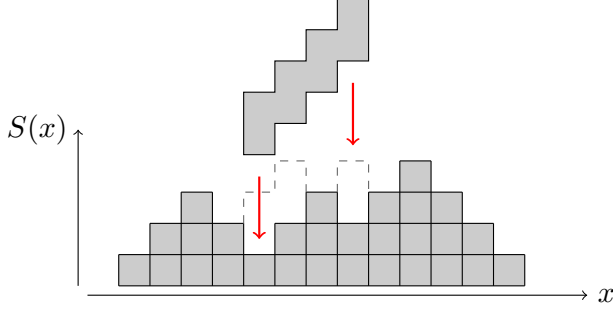
FIG. 6. A 4-staircase falling on an example entropy function. Note that each gate raises $S(x)$ by 2 iff $S(x)$ is a local minimum when that gate falls. So the second gates does in fact hit a local minimum because it acts after the first gate.

This function encodes the butterfly velocity as the derivative $\Gamma'(s)|_{s_{\text{ext}}}$, where s_{ext} is one of the extremal entropy densities, 1 or -1 in this case. A brief explanation of why this is the case is given in the appendix, while a stronger argument can be found in Ref. [?] It follows that any $\Gamma(s)$ with asymmetry at the endpoints will have asymmetric butterfly velocities.

A. Staircase circuits

We will consider entropy functions $S(x,t)$ defined on a periodic system of length L. The boundary conditions are $S(x + L,t) = S(x,t) + sL$, allowing for an overall entanglement density. Subadditivity tells us $|S(x + 1) - S(x)| \leq S_1$, where S_1 is the entropy at a single site. If we take our logarithms with base q, then $S_1 \leq 1$.

If a gate acts on bond x, it can increase the bipartite entanglement entropy $S(x)$, up to the constraint $|S(x + 1) - S(x)| \leq 1$. In the large-q limit, a Haar-randomly chosen gate will, with probability 1, maximally increase the entanglement across the bond it acts on [?]. Given the previous constraint, this means that if a gate acts at bond x at time t, then $S(x,t + 1) = \min\{S(x - 1,t) + 1, S(x + 1,t) + 1\}$. Should we explain why? Since individual circuits have deterministic behavior, we average over circuit architecture.

It suffices to consider integer-valued $S(x)$ with $|S(x) - S(x - 1)| = 1$ for all x. A state of this form can be described as a series of up and down steps at each site. If a gate falls on bond x, it adds two units of entropy to $S(x)$ iff the step before is down and the step after is up.

Staircase circuits of length n are defined by always having strings of n gates act on sites x through $x + n - 1$ in succession. For $n = 1$ this is just a random architecture, but large n results in more asymmetric circuits. Fig. 6 shows an example. It is already possible to see the origin of the asymmetry. The 4-staircase can only be perfectly effective if it hits the microstate (down, up, up, up, up). Since this microstate has positive slope, it is more likely
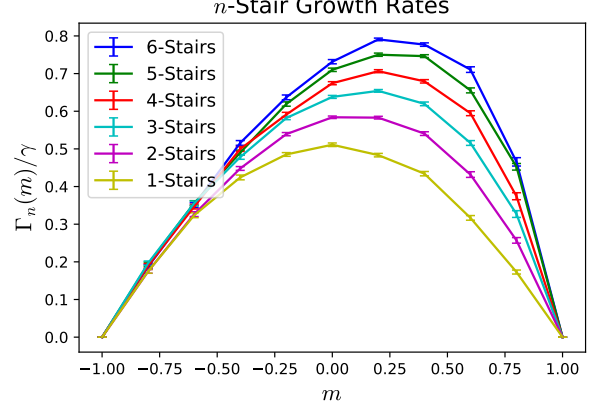

FIG. 7. Empirical growth rate as a function of slope for n-stair circuits. The right/forward and left/backward butterfly velocities are the slopes of these curves at their endpoint, indicating that as the left v_B stays constant, the right v_B increases. The appendix includes an argument that the right v_B is unbounded in the large-n limit. All growth rates were calculated using a 100-site spin chain with offset periodic boundary conditions. Rates were calculated from the application of 2,000 gates total, or 20 per site, averaged over the last 80% of the gates in order to build up correlations, then averaged over 100 trials.

to be found when the total slope is larger.

The gate rate γ is defined as the number of gates per unit time, not staircases.

B. Asymmetric v_B

For small n, we can simulate the circuit directly. For the growth rate curves of n-stair circuits for $n \leq 6$ see Fig. 7. The asymmetry is evident for all $n > 1$, and the asymmetry continues to increase as n increases. However, simulating larger circuits becomes computationally unwieldy.

It is also possible to approximate these growth rates by assuming the up and down steps are uncorrelated. Then the probability that a randomly placed gate will result in growth is $(1 - s^2)/2$. Any asymmetry comes from the fact that for longer staircases, a gate acting on site x is more likely to have been preceded by a gate on site $x - 1$ so the step from $x - 1$ to x is more likely to be a down step.

Under this assumption the growth rate is

$$\Gamma_n(s) = \frac{\gamma}{n}\frac{1 + s}{1 - s}\left((1 + s)\left[\left(\frac{1 + s}{2}\right)^n - 1\right]\right.$$
$$\left. + n(1 - s)\right). \qquad (11)$$

Then $v_{B,l} = \gamma$, while $v_{B,r} = \frac{1}{2}\gamma(n + 1)$. This produces successively more asymmetric growth rates as n increases. However, the correlation in the true steady state,

and therefore the error of this approximation, also increases. It is possible to correct for the correlation term-by-term in correlation length, but this quickly becomes tedious. For 2-stairs, including the nearest-neighbor correlation removes most of the error in $\Gamma(s)$.

Luckily, as n becomes very large or approaches the size of the system, the correlations again become unimportant. To see this we can consider the growth rate at $s = -1, 0$, and 1 for $n = L$ stairs. The approximate growth rate is $\Gamma_\infty(s) = \gamma(s+1)$, so we want to show the true growth rate approaches the approximation. Near $s = -1$, the entropy profile consists mostly of down steps, with isolated up steps. Then the circuit generates entanglement every time a staircase falls. This rate is γ/L, which approaches 0 as L becomes large.

In the $m = 0$ case, after a gate falls between sites i and $i+1$, s_{i+1} will be a down slope regardless of whether the gate generated entanglement. Then the next gate falls across sites $i+1$ and $i+2$. At site $i+2$ $s_{i+2} = u$ with probability $\frac{1}{2}$, so on average $\frac{1}{2}$ of the gates produce 2 units of entanglement and $\Gamma_\infty(0) = \gamma$. At near-maximal slope nearly all slopes are up, except at the site to the right of the most recent gate. Then the next gate falls at a local minimum with probability 1, and all gates produce 2 units of entanglement, so $\Gamma_\infty(1) = 2\gamma$. These growth rates are all consistent with $\Gamma_\infty(s) = \gamma(s+1)$. Using that function, we obtain $v_{B,l} = \gamma$ and $v_{B,r} = \infty$.

IV. CONCLUSION

Asymmetry in v_B is possible, but is limited by locality. In time-independent Hamiltonian systems locality is measured by the range of interactions, while in circuits it is related to the depth. This paper studied both types of systems, showing that local Hamiltonians can support asymmetric spreading and probing spreading in nearly-local circuits.

The advantages of the Hamiltonian system are that it is a general model. Each site is only a 2-level system. The random field allows the model to be tuned within the thermalizing phase, which could be useful in watching the asymmetry dissipate as the system approaches the phase transition.

The system studied in this paper is not the only possible 2-nearest-neighbor system. Another interesting direction of research would be how to find maximally asymmetric Hamiltonians for a given interaction length. With more sites per term it would also be possible to study 2-D systems with anisotropic v_B.

The class of circuit models studied here provide the opportunity to interpolate between symmetric circuits and completely chiral circuits by varying n. One important generalization is to consider finite q, the dimension of the Hilbert space at each site. Ref. [?] suggest that this will lead to a slower v_B as well as $v_E < v_B$. The separate velocity scale v_E leads to operator spreading in individual circuits, not just the ensemble averages this paper discusses.

Appendix A: Butterfly velocity from $\Gamma(s)$

To see why $\Gamma'(s_{\text{ext}})$ gives v_B, consider a region in which the entropy profile is piecewise linear, with $S'(x < x^*) = s_1$ and $S'(x > x^*) = s_2$. Also note $\Gamma(s)$ is always convex. This is certainly true for the functions considered above, and is also true in general. Then if $s_1 < 0 < s_2$, the transition between the region stays sharp. If it did not, a curve would form with intermediate slopes $s_1 < S'(x) < s_2$. But from convexity all of this region would have a faster entanglement growth than $\min\{s_1, s_2\}$ and the peak would reform. Furthermore, this peak location x^* travels at velocity $\dot{x}* = -\frac{\Gamma(s_2)-\Gamma(s_1)}{s_2-s_1}$, the slope of the chord connecting $\Gamma(s_2)$ and $\Gamma(s_1)$.

If instead, $s_2 < 0 < s_1$, the kink does not remain sharp. The sharp point x^* becomes a smooth curve, running tangent to the two linear sections at x_L^* and x_R^*. By a similar argument to the above, these features travel at $-\Gamma'(s_{1,2})$. Then the convexity shows that the fastest velocities in the system are $-\Gamma'(s_{\text{ext}})$. It remains to be shown that v_B cannot be slower than this, but this is just an appendix.

Should something go here?