Operator and Entanglement Dynamics in Asymmetric Quantum Systems

Charles Nicholas Stahl

November 1, 2018

Advised by Professor David Huse

Second Reader: Professor Shivaji Sondhi

Submitted in partial fulfillment
of the Requirements for the
Degree of Bachelor of Arts

Abstract

Thermalization is an important aspect in quantum physics from condensed matter to black holes. It allows initially local information to be spread and hidden throughout a system. This spreading happens at a finite speed, and can be quantified using the butterfly velocity v_B or the entanglement velocity v_E. These speeds are well-studied, and are independent of each other up to the constraint $v_B > v_E$. Although it is possible to have a direction-dependent v_B, little work has been done to study systems like this. In this thesis we study two systems on spin chains with asymmetric butterfly velocities, which we call $v_{B\pm}$. In the first, a system with a time-independent Hamiltonian, we study v_B through operator spreading. We show that the system is slightly asymmetric, with $v_{B+} > v_{B-}$. The second system is a quantum circuit with random unitary dynamics. Using entanglement dynamics to measure the butterfly velocity, we show that these systems can have v_{B+}/v_{B-} arbitrarily large.

I pledge my honor that this paper represents my own work in accordance with University regulations.

Acknowledgments

Thank you, Professor Huse, for making this project possible. Your availability and willingness to meet made my senior thesis process tractable, interesting, and so much fun. From patiently explaining the necessary background material, even when I had just asked the same questions last week, to suggesting useful new directions to explore and helping me interpret the results, you made my foray into quantum information dynamics painless and engaging.

Professor Sondhi, thank you for all you have done to expand my physics knowledge, from advising my summer and future plans, to agreeing to advise an extra math class this semester, to being my second reader, I will always appreciate your helpfulness throughout my last two years in the department.

Thank you Witherspoon 517+, for simultaneously backing me up and pushing me forward since freshman year. I can't imagine being placed in a better freshman hall. I can never catch a break with you guys, but that constant needling got me here, so I can't really complain.

To the seniors and assorted juniors who have made the physics department a comfortable place for me, thank you for getting me through four years of classes and independent work without losing my mind. I'm lucky to have been in such a cohesive department with so much positive encouragement between students.

Mom, Dad, Maria, thank you for being the smartest, funnest, and most interesting family I've ever had. You have instilled in me a love for science, a love for wit, and a love for knowledge that have made me who I am today. Mom and Dad, I know I would never be here without you. And Maria, I am so lucky to have been able to spend so much time with you over the past two years on campus. You're an inspiration.

And as always, thank you Haley for five wonderful years of love and support. Knowing that, no matter how hard a day has been, you will be ready to listen and understand on the phone gives me a fabulous feeling of peace. I can't wait to see where our next five years of adventures take us. I love you to the moon and back!

Contents

1 Introduction

In quantum mechanics, information is special in that it is always conserved. Conservation of information, which is defined as the ability to recover initial conditions, is ensured by the unitary evolution. For any time evolution operator $U(t)$, there is an inverse operator $U^\dagger(t)$ such that $U^\dagger(t)U(t) = 1$, which brings a system back to its initial conditions.

This is emphatically not the case in classical mechanics. Fluid mechanics and thermodynamics are both examples of dissipative systems. The fluid flow equations with non-zero viscosity are dispersive, so that local features in the flow die away. In thermodynamics, a system with general initial conditions will equilibrate to one which can be described by a small number of parameters, such as temperature, pressure, and chemical potential.

Black holes are an extreme example of systems with information loss, at least when treated classically. The no-hair theorems [1, 2, 3] state that black holes can also be described by only a few degrees of freedom. The connection between these classical systems is thermalization or equilibration, where any initial state approaches one describable by a small number of parameters. This implies that there must be some way to obtain thermalization in quantum systems.

In fact, there are two robust, long-term behaviors that isolated quantum systems can demonstrate, thermalization and localization [4]. In a localized system, initially local perturbations can be robust to the dynamics of the system and not diffuse away [5]. Thermalizing systems, on the other hand, equilibrate to a thermal equilibrium state. This equilibration occurs in the thermodynamic limit of large size and long time.

But then where does the information encoded in the initial conditions go? In thermodynamics, we couple systems to baths, so that the degrees of freedom in the bath can absorb the information. Connecting the system to a bath seems to solve this problem, but quantum mechanically we have only extended the problem to the larger Hilbert space that contains the system and bath.

The resolution is that the initial information in a subsystem spreads throughout the system in a process called scrambling. The full information about the initial conditions is present within the entire system at later times, but the number of degrees of freedom associated with the system is too large, and they may be too far apart, to make recovering this information feasible. Black holes, with their extremal information-hiding, are the fastest scramblers [6, 7].

For thermalization to work on systems in any state, it must work on eigenstates of the

many-body Hamiltonian. Since eigenstates don't evolve in time, subsystems must locally look thermal when the whole system is in an eigenstate, even before the long-time limit. This argument is referred to as the Eigenstate Thermalization Hypothesis (ETH) [8, 9, 10, 4]. Of course, localized systems do not obey the ETH because they do not thermalize. This thesis will study thermalizing systems.

Since information initially localized in a subsystem spreads to other degrees of freedom outside the subsystem, the system acts as a bath for its subsystems. These degrees of freedom become entangled with the subsystem so that the state of the subsystem cannot be described without knowledge of the whole, making it look thermal. Ref. [4] suggests the fundamental characteristic of a bath is its ability to entangle its degrees of freedom with those of the subsystem.

After realizing that thermalization leads to information spreading and local information loss, the next point of study becomes its dynamics. Scrambling in black holes can be described using holography [6, 11]. Scrambling has been studied in conformal field theories [12] and spin chains with integrable and non-integrable Hamiltonian evolution [13, 14, 15, 16]. Random unitary dynamics [17, 18, 19, 20, 21] provide another setting to study spreading.

We will present two velocity scales of the information dynamics, the butterfly velocity and the entanglement velocity. Both of these quantify the scrambling of a system, but do not have a set relationship. In fact, the butterfly velocity can be made arbitrarily large with respect to the entanglement velocity [20]. The result of this thesis is that the butterfly velocity itself need not be the same in both directions on a 1-d system, and that one butterfly velocity can be made arbitrarily large with respect to the other.

We discuss these dynamics in two different types of systems: those with time-independent Hamiltonians (Secs. 2 and 3) and quantum circuits with random unitary dynamics (Secs. 4 and 5). In the Hamiltonian systems we use operator spreading to quantify thermalization, where an initially local operator becomes non-local when evolved in time. We show that a particular Hamiltonian that has slightly asymmetric butterfly velocities.

In the quantum circuits we study entanglement dynamics, quantified by the entanglement entropy, instead of operator dynamics. We show that the butterfly velocity as defined by operator spreading can also be calculated from the entanglement dynamics in our systems. We show that the quantum circuits can have arbitrarily asymmetric butterfly velocities.

2 Operator Spreading in Time-Independent Hamiltonian Systems

We will start our discussion of time-independent Hamiltonians with a brief review of unitary time evolution of states and operators. We then define measures of how far an initially local operator has spread over a system. From these we define the butterfly speed v_B, which measures how fast operators spread.

Referring back to the introduction, we want a measure of the information spreading in a system. We can define information through commutation, because if two measurements commute, then they do not affect each other. This is equivalent to saying that performing one measurement does not send information to the other measurement. For example, in quantum field theory a correlation function for space-like separated points may be non-zero, but commutation between operators at these points must always vanish, which is to say that information can not travel faster than the speed of light. The velocities we describe will measure how fast the commutators grow.

2.1 Background: Evolution in Time

All systems considered in this thesis will exist on spin chains, one-dimensional collections of local quantum degrees of freedom. Initially we consider systems with $q = 2$ degrees of freedom at each site, such as a chain of spin-$\frac{1}{2}$ particles. Later, we will consider sites with more degrees of freedom.

Under a time-independent Hamiltonian H, states of the system evolve in the Schrödinger picture as

$$|\psi(t)\rangle = U(t)\,|\psi(0)\rangle\,, \quad U(t) = e^{-iHt}. \tag{1}$$

Instead of evolving states, it is possible to evolve operators in the Heisenberg picture. In order to preserve the time dependence of expectation values, the operators must evolve as

$$A(t) = U^\dagger(t)\,A(0)\,U(t) = e^{iHt}A(0)\,e^{-iHt}. \tag{2}$$

There is a slight collision of terminology here. Both A and U are operators, but they play different roles. In this thesis the time evolution operator will always be called a "unitary operator," or later a "gate," so that the term "operator" without qualification always refers

to the operator that evolves in time.

One remark worth making about time evolution in different pictures is that operators are not the only the only important Hermitian matrix in the system. The density matrix ρ describes the state of the system, and in pure states is $\rho = |\psi\rangle \langle\psi|$. Density matrices are more general than kets, though, because they can represent mixed states. In the Schrödinger picture ρ evolves the way its construction would imply

$$\rho(t) = e^{-iHt}\rho(0)e^{iHt} \tag{3}$$

while it does not evolve in the Heisenberg picture. This is the opposite of observables, so when discussing the evolution of matrices it is necessary to specify whether they are observables or density matrices.

2.2 Pauli Strings and Pauli Weight

This subsection is based on [17]. Since it is possible to add Hermitian operators and multiply them by constants, they live in a vector space that can be described using some basis. When discussing operator spreading it is convenient to decompose operators that may act on very high dimensional Hilbert spaces into the Pauli basis. The basis operators are tensor products of Pauli matrices. Eventually we will decompose operators that are initially local, but any operator can be decomposed in this manner.

For single sites with Hilbert spaces of complex dimension q, the space of Hermitian operators is q^2-dimensional. For $q = 2$ the basis operators are X, Y, Z, I. In general, there will be two Hermitian operators X and Z that obey $ZX = \exp(2\pi i/q)XZ$ and $Z^q = X^q = I$. Then an arbitrary basis operator will be of the form

$$\sigma^\mu = \phi X^{\mu_1} Z^{\mu_2}, \tag{4}$$

where $\mu_1, \mu_2 \in \{0, 1, \ldots, q-1\}$ and ϕ is a phase to preserve Hermiticity. We can extend this description to multiple sites by taking tensor products of L basis operators with subscript ν representing L μ indices. Under the matrix norm $||M|| = \mathrm{tr}(M^\dagger M)/q^L$, this basis is orthonormal:

$$\begin{aligned}\frac{1}{q^L}\mathrm{tr}(\sigma^{\mu\dagger}\sigma^\nu) &= \frac{1}{q^L}\mathrm{tr}(Z^{\mu_2\dagger}X^{\mu_1\dagger}X^{\nu_1}Z^{\nu_2}) \\ &= \delta_{\mu\nu}.\end{aligned} \tag{5}$$

A general operator $A = \sum_\nu c_\nu(0)\sigma^\nu$ evolves into

$$A(t) = U^\dagger(t) A\, U(t) = \sum_\nu c_\nu(t)\sigma^\nu. \tag{6}$$

Due to the orthonormality, the coefficients are

$$c_\nu(t) = \frac{1}{q^L}\mathrm{tr}(\sigma^{\nu\dagger} A(t)) \tag{7}$$

and obey

$$\frac{d}{dt}\left(\sum_\nu |c_\nu|^2\right) = 0$$

due to unitarity. We will consider normalized operators with $\sum_\nu |c_\nu(t)|^2 = 1$.

With $c_\nu(t)$ in hand, we can define the Pauli weight $W(i,t)$ as how many of the Pauli strings in the decomposition end on site i, weighted by their coefficients in Eq. 6,

$$W(i,t) = \sum_\nu |c_\nu(t)|^2\, \delta(\mathrm{end}(\nu) = i). \tag{8}$$

The delta function constrains the sum to be only over ν such that σ^ν has a non-identity at site i and identities at all sites right of i. This gives a measure of how far the operator has spread. Reference [17] refers to this quantity as ρ to emphasize its conservation and hydrodynamic evolution. It is possible to define an analogous quantity with the sum over strings that begin on site i and have identities on all sites left of i. In that case the quantity in equation 8 can be called $W_R(i,t)$ while the weight of sites that start on i is $W_L(i,t)$.

It is helpful to go through an example to explain the Pauli weights. Write

$$A = \frac{1}{\sqrt{2}} X_0 \otimes Y_1 \otimes Y_2 \otimes I_3 + \frac{1}{\sqrt{2}} I_0 \otimes I_1 \otimes Z_2 \otimes Z_3,$$

where the subscript designates the site on which each operator acts. This can be shortened to

$$A = \frac{1}{\sqrt{2}} XYYI + \frac{1}{\sqrt{2}} IIZZ. \tag{9}$$

Then the right (ending) Pauli weights are $W_R(i=0) = W_R(i=1) = 0$, $W_R(i=2) = W_R(i=3) = \frac{1}{2}$ because one component of the operator ends on site 2 and one ends on site 4. The left (starting) weights are $W_L(i=0) = W(i=2) = \frac{1}{2}$, $W_L(i=1) = W_L(i=3) = 0$.

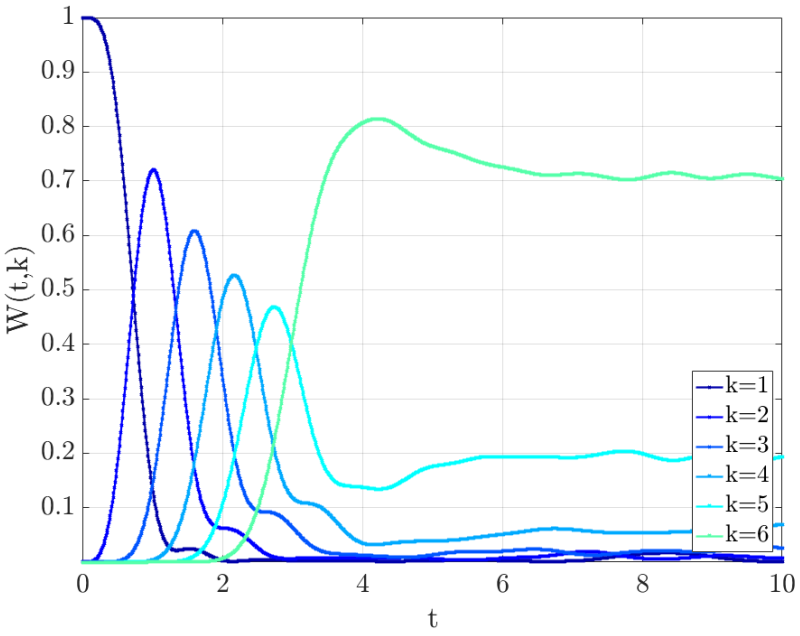

Figure 1: **Operator spreading as shown by the Pauli weight** in an Ising model with transverse field, from [22]. The end weight starts on site 1 and moves sequentially across the system to site 6. Note that the weights at any time sum to 1.

This decomposition is particularly useful when the initial operator is local at site j. This means that initially all strings in the Pauli decomposition contain non-identity operators only at site j and $W(i,0) = \delta_{ij}$. Then the end weights describe how far the operator has spread throughout the system due to the unitary dynamics, which is an essential feature that allows the systems to thermalize. Figure 1 shows the Pauli end weight evolution for an initially local operator in an Ising spin chain with longitudinal field in a strongly chaotic phase [22]. The weight moves from site to site and eventually reaches the last site.

The Pauli decomposition is related to the Out-of-Time-Ordered Commutator (OTOC). Consider an initial operator, say, $A_0 = Z_0 = ZII\ldots$, where the subscript now indicates the initial location of the operator. This will commute with probe operators that are local at other sites, B_i. However, in general the time-evolved $A_0(t)$ will include Pauli strings that have non-identity operators at site i. This can be seen through the Baker-Campbell-Hausdorff expansion [23] of the time evolution

$$\begin{aligned} A_0(t) &= e^{iHt} A_0 e^{-iHt} \\ &= \sum_k \frac{(it)^k}{k!} [H, [H, \ldots [H, A_0] \ldots]], \end{aligned} \tag{10}$$

where the dots represent that there are k total commutators taken with H. In general H will contain matrix elements that connect the initially local operator to the non-local strings. Then, $A_0(t)$ can fail to commute with B_i, which is still local at i and has not been evolved in time.

The extent to which these two fail to commute can be measured by the OTOC on operators A_0 and B_i normalized to $A_0^2 = B_i^2 = 1$,

$$C(i,t) = \frac{1}{2} \operatorname{Tr} \rho \left| [A_0(t), B_i] \right|^2 .$$

For a pure state this is equivalent to [17, 21]

$$\begin{aligned} C(i,t) &= \frac{1}{2} \langle \psi | \left| [A_0(t), B_i] \right|^2 | \psi \rangle \\ &= 1 - \operatorname{Re} \langle \psi | A_0(t) B_i A_0^\dagger(t) B_i^\dagger | \psi \rangle. \end{aligned} \tag{11}$$

Some sources define the OTOC using $[A,B]^2$ instead of $|[A,B]|^2$ [**Jonay**, 23, 18]. Others use the acronym to denote the out-of-time-order correlator, $F(i,t) = \left\langle A_0(t) B_i A_0^\dagger(t) B_i^\dagger \right\rangle$ [**Who**], which is related to $C(i,t)$ through Eq. 11. If ρ is taken to be a thermal state at infinite temperature, the density matrix is proportional to the identity and the OTOC becomes

$$C(i,t) = \frac{1}{2} \operatorname{Tr} \left| [A_0(t), B_i] \right|^2 . \tag{12}$$

This definition can be thought of as independent of the state of the system. We will adopt this as our convention for the OTOC.

To see the relation between this quantity and $W(i,t)$, first consider the case of $q = 2$. If the probe is X_i there will be two classes of Pauli strings that commute with it: those with an identity at site i and those with the operator X at site i. Then $C(i,t)$ will be the sum of the squares of the $c_\mu(t)$ for which the operator at site i is Y or Z. If we average over choice of probe (X, Y, or Z), we arrive at

$$\bar{C}(i,t) = \frac{2}{3} \sum_\nu |c_\nu(t)|^2 \, \delta(\text{condition on } \nu), \tag{13}$$

where the condition is that the operator at site i is not the identity. This is similar to $W(i,t)$ but measures the weight at i. For arbitrary q the prefactor is $\frac{q^2-2}{q^2-1}$.

If the operator $A_0(t)$ is sufficiently random, all q^{2L} coefficients will have expectation values of the same order, with $|c_\nu|^2 \approx q^{-2L}$. There will be $q^{2L} \frac{q^2-1}{q^2}$ operators that meet the

condition, so for these random operators in the large L limit

$$\begin{aligned}\bar{C}(i,t) &= \frac{q^2-2}{q^2-1} q^{2L} \frac{q^2-1}{q^2} \frac{1}{q^{2L}} \\ &= \frac{q^2-2}{q^2}.\end{aligned}$$

For $q = 2$ there is a particularly easy way to calculate the value in expression 13. Since $\mathrm{Tr}(X) = \mathrm{Tr}(Y) = \mathrm{Tr}(Z) = 0$ and $\mathrm{Tr}(I) = 2$, in order to find the non-identity weight at site i start by tracing over the degrees of freedom at that site to obtain $\tilde{A}_0(t) = \frac{1}{2}\,\mathrm{Tr}_i\, A_0(t)$. Only the strings with the identity at i survive this operation so

$$\sum_\nu |c_\nu(t)|^2\, \delta(\text{condition}) = \frac{3}{2}\bar{C}(i,t) = \mathrm{Tr}\, \tilde{A}_0^\dagger(t)\tilde{A}_0(t). \tag{14}$$

We will take this as our definition of the OTOC and relabel it as $C(i,t)$. Using this definition, $\max C(i,t) = 1$, and $\langle C(i,t)\rangle = \frac{3}{4}$ for a random operator with unit norm.

2.3 Information Velocities

Since we are studying the spreading of information throughout a system, we want to quantify how fast the information is spreading. In a relativistic system, for example, no information can travel past the speed of light. This is stated precisely by saying that operators at space-like separated points in spacetime must commute. Only within the light cone can two operators have non-zero commutation. Lack of commutation is used as a measure of information transfer because it quantifies how much the two measurements affect each other.

In a spin chain evolving under a Hamiltonian without relativity, there is no requirement that any two operators must commute. This is because of Eqn. 10, which can connect distant operators at very small t. However, the Lieb-Robinson bound [24] limits the largest possible commutator of two normalized observables $A_0(t)$ and B_i to

$$|[A_0(t), B_i]| \leq K_0 e^{-(|i| - v_{LR}t)/\xi_0}, \tag{15}$$

where K_0 and ξ_0 are constants and v_{LR} is the Lieb-Robinson velocity, set by the Hamiltonian. Outside $x = v_{LR}$, commutators must be exponentially small, so this is a reasonable "light cone."

Another velocity that limits the spreading of information in spin chains is the butterfly

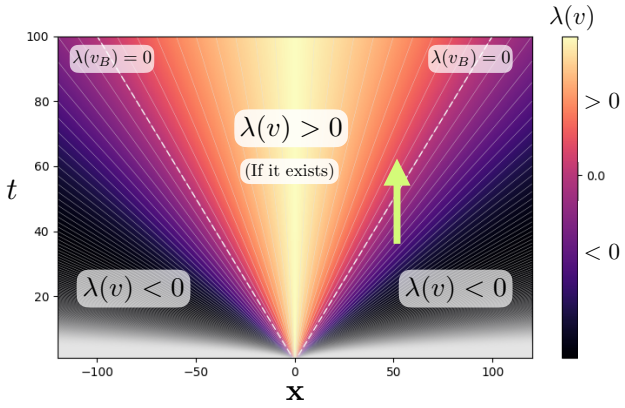

Figure 2: **Velocity dependent Lyapunov exponents** along rays of constant velocity, from [25]. For $v < v_B$, inside the light cone, $\lambda > 0$ or is not defined. Outside the light cone $\lambda < 0$ and perturbations decay exponentially. The speed v_B is where $\lambda = 0$ and perturbations remain of order 1.

velocity, the fastest velocity at which perturbations can propagate without decaying exponentially in time. A useful definition of v_B comes from Ref. [25]. In systems large enough so that the site index i can be replaced with a continuous position x, the OTOC behaves as [25]

$$C(x,t) \sim e^{\lambda(v)t} \quad \text{for} \quad x = vt. \tag{16}$$

The condition $x = vt$ tracks the OTOC along rays of constant velocity, as in the lines emanating from the origin in Fig. 2. The fact that there is a single v_B, or possibly one in each direction, is due to the fact that $\lambda(v)$ must be convex, and can only cross 0 once in each direction.

For small velocities in classical chaotic systems, λ can be positive, in which case it is the Lyapunov exponent. In analogy, we call the $\lambda(v)$ the velocity-dependent Lyapunov exponent. Typical $\lambda(v)$ in classical and quantum systems can be seen in Fig. 3.

Note that, for a system with high temperature such that the density operator $\rho \propto \mathbb{I}$, the definitions of v_{LR} and v_B coincide. In general v_B is a "low-energy analog" of v_{LR} [23]. These two speeds can be different, but are not measuring something fundamentally different about the Hamiltonian.

Information dynamics in quantum circuits does have another velocity scale, defined by the entanglement velocity v_E. We do not study the entanglement velocity in our Hamiltonian system, but we do in quantum circuits in Sec. 5. The entanglement velocity is defined in terms of entanglement entropy, which is defined in Sec. 4. Even without the quantitative

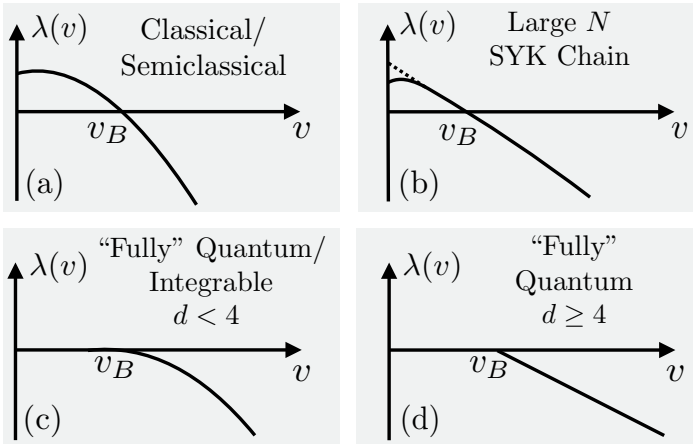

Figure 3: **Velocity-dependent Lyapunov exponents** in classical and quantum systems. The positive λ corresponds to chaos. Figure taken from [25].

definitions, we can still have some understanding of v_E.

Two systems A and B are entangled if their joint state $|\Psi\rangle_{AB}$ can not be written as a product state $|\psi_A\rangle_A |\psi_B\rangle_B$. If two parts of a spin chain are initially unentangled and then coupled, they will gradually become entangled. The spins closer to the boundary will become entangled first. The entanglement reaches farther spins at the speed v_E.

In general, $v_B/v_E > 1$, but the ratio can be made arbitrarily large, as in Ref. [20]. In Sec. 5 we recreate the systems from that reference, but also build systems in which v_B is different in different directions. Before going into more detail about v_E, though, we will present systems with time-independent Hamiltonian in which there are two different values of v_B.

3 Dynamics in an Asymmetric Hamiltonian

The previous section, and much of the current literature, focuses on symmetric Hamiltonian systems, in that the butterfly velocity is the same for perturbations traveling to the left or to the right.

In this section we study the operator dynamics of an asymmetric time-independent Hamiltonian system. First we construct the local 3-site Hamiltonian through its action on the computational basis. After chaining together these local terms to define a multi-site Hamiltonian we find the behavior of the Pauli weights and OTOC. Eventually we show that this system has two distinct butterfly velocities, one for operator fronts spreading in each direction.

3.1 3-Site Hamiltonian

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 n small, and then putting this Hamiltonian on each set of n sites,

$$H_{\text{tot}} = \sum_{i=0}^{L-n} H_n^{(i)}, \tag{17}$$

where $H_n^{(i)}$ is a n-site Hamiltonian acting on sites i through $i+n-1$. A common choice is $n=2$ but this will not suffice, because 2-site Hamiltonians are always symmetric with respect to their operator dynamics. Unitarity preserves information so for any weight the 2-site Hamiltonian moves from site i to $i+1$ it must 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 $U(t) = e^{-iHt}$. There are multiple ways to construct H from $U(T)$. One way is to just take the matrix logarithm, for example by using Mathematica. Since this method is not very physical, we can instead look at eigenstates. $U(t)$ and H will have the same eigenstates, with the eigenvalues of H given by $E_j = i\log(\phi_j)$, where ϕ_j is the eigenvalue of $U(t=1)$. The Hamiltonian is completely specified by its eigenstates and eigenvalues. Since scalar logarithms are simpler than matrix logarithms, this method is much more intuitive.

One asymmetric unitary operator is the 3-site cyclic swap S_{123}. S_{123} is a unitary operator

such that

$$S_{123}\left|\alpha\beta\gamma\right\rangle = \left|\gamma\alpha\beta\right\rangle, \tag{18}$$

where $|\alpha\beta\gamma\rangle$ is a product state with state $|\alpha\rangle$ on site 0, $|\beta\rangle$ on site 1, and $|\gamma\rangle$ on site 2. 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 in a Floquet system or quantum circuit, out of 2-site swap gates $S_{123} = S_{12}S_{23}$. Each 2-site swap interchanges two states, so the action is

$$\begin{aligned} S_{12}S_{23}\left|\alpha\beta\gamma\right\rangle &= S_{12}\left|\alpha\gamma\beta\right\rangle = \left|\gamma\alpha\beta\right\rangle \\ &= S_{123}\left|\alpha\beta\gamma\right\rangle. \end{aligned} \tag{19}$$

It is also possible to build S_{123}t out of a time-independent Hamiltonian, so that $U(t=1) = e^{-iH_3} = S_{123}$. H_3 is the 3-site term we are looking for.

For a system with spin-$\frac{1}{2}$ objects on the sites there is an 8-dimensional Hilbert space, which can be decomposed into a spin-$\frac{3}{2}$ subspace with 4 states and 2 spin-$\frac{1}{2}$ subspaces with 2 states each. Since the eigenstates of S_{123} only pick up a phase under unitary evolution, they will be states in which individual particle states differ only by phases, so that the phase from the dynamics effectively permutes the states.

There are four states that are symmetric with respect to the three sites, and therefore do not change in time and do have 0 energy. Before normalization these are

$$\begin{aligned} &\left|\psi_{0,0}\right\rangle = \left|000\right\rangle, \quad \left|\psi_{0,1}\right\rangle = \left|100\right\rangle + \left|010\right\rangle + \left|001\right\rangle, \\ &\left|\psi_{0,3}\right\rangle = \left|111\right\rangle, \quad \left|\psi_{0,2}\right\rangle = \left|011\right\rangle + \left|101\right\rangle + \left|110\right\rangle, \end{aligned} \tag{20}$$

where, for example, $|001\rangle$ is a product state with sites 0 and 1 in state $|0\rangle$ and site 3 in state $|1\rangle$. $|1\rangle$ is the spin-up state and $|0\rangle$ is the spin-down state.

Of the four other states, two should have positive energy and two should have negative energy. Since $U(t=3) = 1$, their eigenvalues of $U(t=1)$ must be cube roots of unity. This sets the energies to $E_\pm = \pm\frac{2\pi}{3}$ so they pick up a phase $\phi_\pm = e^{-iE_\pm} = e^{\mp i\frac{2\pi}{3}}$. Using

condition 18, we can show that the positive energy states are

$$\begin{aligned} |\psi_{+,1}\rangle &= |100\rangle + \phi_- |010\rangle + \phi_+ |001\rangle , \\ |\psi_{+,2}\rangle &= |011\rangle + \phi_- |101\rangle + \phi_+ |110\rangle , \end{aligned} \tag{21}$$

while the negative energy states are

$$\begin{aligned} |\psi_{-,1}\rangle &= |100\rangle + \phi_+ |010\rangle + \phi_- |001\rangle , \\ |\psi_{-,2}\rangle &= |011\rangle + \phi_+ |101\rangle + \phi_- |110\rangle . \end{aligned} \tag{22}$$

For example. the evolution of $|\psi_{+,1}\rangle$ is

$$\begin{aligned} U(1) |\psi_{+,1}\rangle &= \phi_+ \left(|100\rangle + \phi_- |010\rangle + \phi_+ |001\rangle\right) \\ &= \phi_+ |100\rangle + |010\rangle + \phi_- |001\rangle \\ &= S_{123} |\psi_{+,1}\rangle , \end{aligned} \tag{23}$$

with similar results for the other 3 non-zero energy states.

To write the Hamiltonian as a matrix we have to choose a basis. In the eigenbasis, of course, the Hamiltonian is diagonal. A more useful basis is the computational basis, which has states $|000\rangle$, $|001\rangle$, $|010\rangle$, $|011\rangle$, $|100\rangle$, $|101\rangle$, $|110\rangle$, $|111\rangle$. The strings inside each ket can be interpreted as binary number, so that the states can be written as $|0\rangle$, $|1\rangle$, $|2\rangle$, etc.

Then the Hamiltonian is

$$H_3 = T \, \mathrm{diag}(0, 0, 0, 0, E_+, E_+, E_-, E_-) \, T^\dagger, \tag{24}$$

where $\mathrm{diag}(\dots)$ is the Hamiltonian in its eigenbasis and T is the transformation matrix between the two bases, which can be found from the form of the 3 eigenstates in Eqs. 20,

21, and 22, Altogether, the Hamiltonian is

$$H_3 = \frac{2\pi i}{3\sqrt{3}} \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{25}$$

This Hamiltonian is antisymmetric and purely imaginary. One effect of this is that if the system starts in a state with real coefficients all coefficients stay real because

$$\dot{\psi} = -iH\psi \tag{26}$$

is purely real. Before moving on we note that this commutes with the total spin-Z operator

$$S_Z = \text{diag}\left(-\frac{3}{2}, -\frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, -\frac{1}{2}, \frac{1}{2}, \frac{1}{2}, \frac{3}{2}\right) \tag{27}$$

so total spin-Z is conserved. It also commutes with the other components of spin and therefore also total spin S^2.

There are a few checks we can perform on this Hamiltonian, First, it should be symmetric under a simultaneous rotation of all three spins in real space, so that it can be written as $H_3(\boldsymbol{S}_1, \boldsymbol{S}_2, \boldsymbol{S}_3)$ as a function invariant to simultaneous rotations of the three spin vectors. Furthermore it should be antisymmetric under the interchange of any two spins (equivalent to reversing the direction of propagation). The only function of three vectors that has this property is the triple product $H_3 = \boldsymbol{S}_1 \cdot (\boldsymbol{S}_2 \times \boldsymbol{S}_3)$, where multiplication of components is interpreted as tensor products.

The representation of the Hamiltonian as a triple product provides another reason for the spectrum of the Hamiltonian. As previously mentioned the Hilbert space of the 3 spins decomposes into one spin-$\frac{3}{2}$ space and two spin-$\frac{1}{2}$ spaces. The state $|111\rangle$ is part of the spin-$\frac{3}{2}$ space. Since the three spins point in the same direction, their triple product vanished. Furthermore, since the other three states in the spin-$\frac{3}{2}$ subspace are related to $|111\rangle$ by rotation and the Hamiltonian is symmetric in this respect they also must have energy 0. Of the two spin-$\frac{1}{2}$ pairs, one pair has positive energy and one has negative energy.

Writing the jth component of spin on site i as $S_{i,j}$, this triple product is

$$\begin{aligned} H &= \boldsymbol{S}_1 \cdot (\boldsymbol{S}_2 \times \boldsymbol{S}_3) \\ &= S_{1,1} \otimes S_{2,2} \otimes S_{3,3} - S_{1,1} \otimes S_{2,3} \otimes S_{3,2} + S_{1,2} \otimes S_{2,3} \otimes S_{3,1} - \cdots \\ &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & -i & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & i \\ 0 & 0 & 0 & 0 & -i & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & i & 0 & 0 \\ 0 & 0 & -i & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & i & 0 & 0 & 0 & 0 \\ -i & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & i & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} + \dots \end{aligned} \tag{28}$$

which the reader can check is indeed proportional to Eq. 25.

Exponentiating the Hamiltonian is another check. This gives the time evolution operator for one time step

$$U(t=1) = e^{-iH_3} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{29}$$

This has the properties of of condition 18. Furthermore, application three times gives $[U(t=1)]^3 = U(t=3) = 1$.

We can learn more about this Hamiltonian by watching the evolution of a single state. If the system starts in the state $|100\rangle$, the coefficients for the other states with equal total S_Z will both change from 0 while the other coefficients will stay 0 due to conservation of S_Z. Furthermore we know that the coefficients should remain real, so we can plot the actual coefficients instead of their magnitudes. From the definition of the unitary operator we know the state will become $|010\rangle$ at time 1. Fig. 4 shows this evolution. An important point to realize is that the coefficient of $|001\rangle$ also becomes non-zero at early time, as it

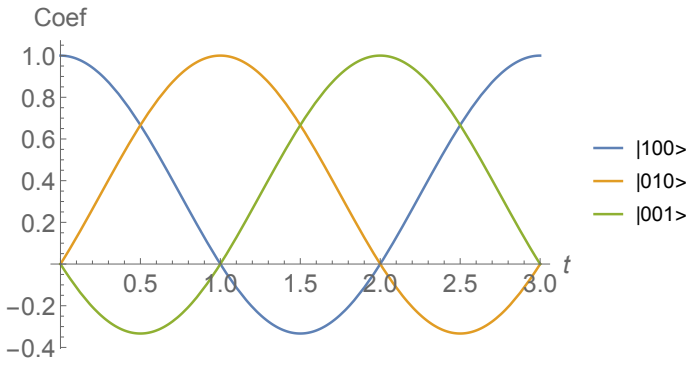

Figure 4: **Evolution of coefficients in 3-site system** if the system starts in state $|100\rangle$. Note that the system initially moves to states $|010\rangle$ and $|001\rangle$ with equal coefficient squared, although at time $t = 1$ the state is entirely $|010\rangle$.

must for evolution under a time-independent Hamiltonian if it is going to become non-zero in the future. In fact, at early time the coefficients for $|010\rangle$ and $|001\rangle$ increase with the same magnitude.

3.2 Multi-Site Hamiltonian

The 3-site system is periodic, and furthermore is not large enough to effectively study operator spreading. We need a way to extend this Hamiltonian to more sites. One way would be to apply H_3 repeatedly for one time step to each individual triplet of spins. This would equivalently apply S_{123} to each triplet and shuffle the states. However this is not a time-independent Hamiltonian, but rather a Floquet system. To preserve time independence we can apply the 3-site Hamiltonian to all triplets simultaneously, as in Eqn. 17. Note that this means that $U(t)$ will no longer simply shuffle the states as in Eqn. 18. This subsection will explore the dynamics of this multi-site Hamiltonian.

Since the perturbations on states move from site $0 \to 1 \to 2 \to 0$, after chaining multiple Hamiltonians together the perturbations to states should move faster from high i to low i, which we will call the backward direction. Heuristically, this is because weight can "jump" directly from site 2 to site 0 when moving backward, but has to move through site 1 when it is moving forward. Once we get to the butterfly velocity, which is defined in terms of perturbations to operators, we expect the forward velocity to be faster, since operators evolve in the opposite direction from states.

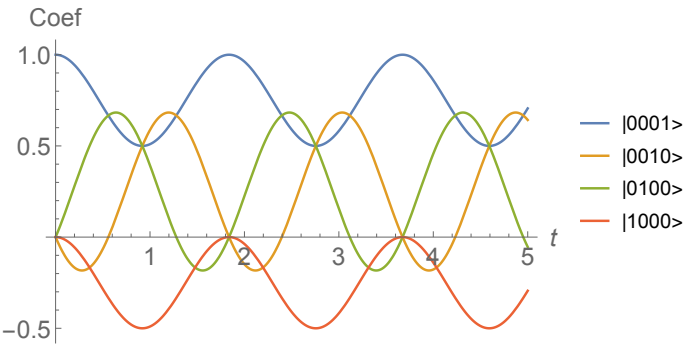

Figure 5: **Evolution of the 4-site system** when the initial state is $|0001\rangle$. Each curve is the coefficient of one of the basis states. The system is periodic with period $\tau = 3\sqrt{\frac{3}{5}}$, and is never fully in one of the other basis states.

3.2.1 Construction

Explicitly, the extension of this Hamiltonian to 4 sites is, using the notation in Eq. 17,

$$H_4 = H_3^{(0)} \otimes \mathbb{I}_1^{(3)} + \mathbb{I}_1^{(0)} \otimes H_3^{(1)}, \tag{30}$$

so that the 3-site Hamiltonian acts on each contiguous triplet of spins. This Hamiltonian still preserves each component of total spin. Although we are not simply swapping states, the behavior is still periodic. Fig. 5 shows that when the system starts in state $|0001\rangle$ it returns to that state with period $\tau = 3\sqrt{\frac{3}{5}}$ but never fully reaches any other basis state.

With 5 sites, adding the third triplet destroys the simple periodic behavior of the system, although it is still quasiperiodic. The Hamiltonian is now

$$H_5 = H_3^{(0)} \otimes \mathbb{I}_2^{(3)} + \mathbb{I}_1^{(0)} \otimes H_3^{(1)} \otimes \mathbb{I}_1^{(4)} + \mathbb{I}_2^{(0)} \otimes H_3^{(2)}, \tag{31}$$

so that there is again one 3-site Hamiltonian on each contiguous triplet. Starting in $|00001\rangle$, the coefficients follow the pattern of figure 9. At first the evolution is similar to the $n = 1$ case, with $|10000\rangle$ and then $|01000\rangle$ reaching near maximal. This suggests that the system retains some of its swap gate-like behavior. However the system is never fully in any basis state after $t = 0$, due to the lack of periodicity.

By directly diagonalizing the Hamiltonian we can study the evolution of a single component of the state. The coefficient of $|00001\rangle$ is

$$c_1(t) = \frac{1}{10}\left(3\cos\left(\frac{2}{3}\sqrt{\frac{5}{3}}\pi t\right) + 5\cos\left(\frac{2\pi t}{3\sqrt{3}}\right) + 2\right) \tag{32}$$

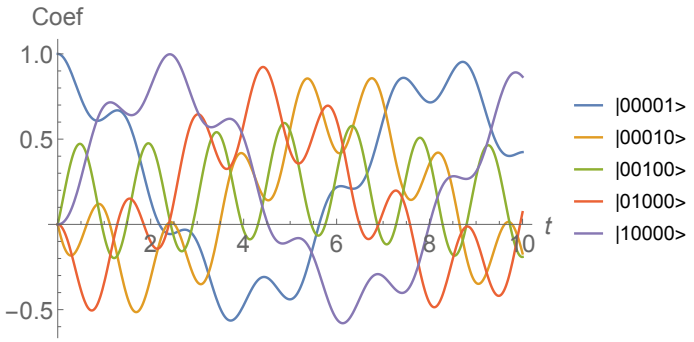

Figure 6: **Evolution of the 5 site system** when the system starts in state $|00001\rangle$. Periodicity is ruined by the third triplet.

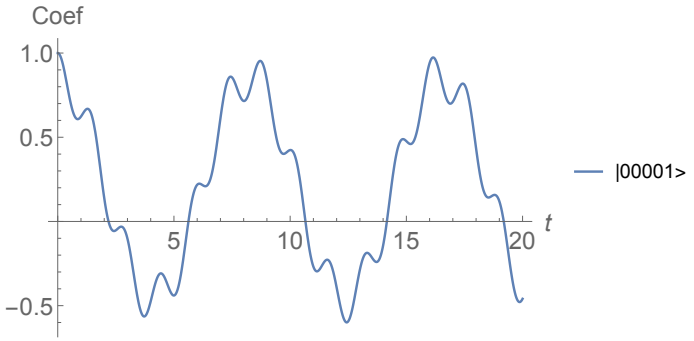

Figure 7: **Simplified view of Fig. 6**, showing only the coefficient for $|00001\rangle$ when that is the starting state. Even with the longer range of time, the coefficient never returns to 1.

which is shown in figure 7. Although it appears to be quasi-periodic, it cannot ever reach 1 for $t \neq 0$ or be truly periodic because its the periods of the two cosine functions are not rationally related.

Instead of putting H_3 on every triplet, we can only chain together the odd triplets. Fig. 8 schematically shows these two configurations for $L = 5$. We will call these sparse Hamiltonians, and they are only possible for odd L. For example the sparse Hamiltonian for $L = 5$ is

$$H'_5 = H_3^{(0)} \otimes \mathbb{I}_2^{(3)} + \mathbb{I}_2^{(0)} \otimes H_3^{(2)}. \tag{33}$$

Figure 8: **Dense and sparse Hamiltonians**. The ellipses show which states are connected by 3-site Hamiltonians.

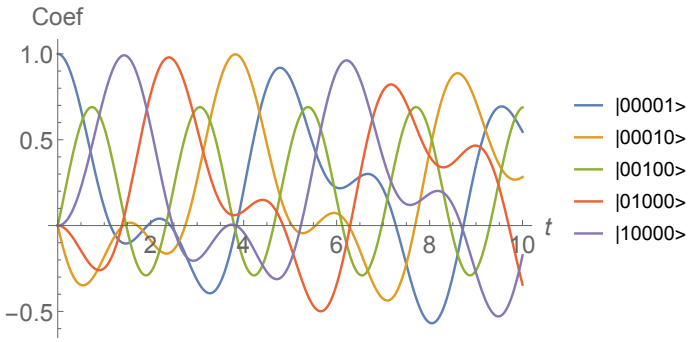

Figure 9: **Evolution of the sparse 5-site system**. Although the coefficients for most states are not periodic, the coefficient for $|00100\rangle$ is.

Like the dense $L = 5$ Hamiltonian, this system is not periodic. A plot of the coefficients over time the the starting state $|00001\rangle$ is shown in Fig. 9. An interesting detail is that the coefficient on state $|00100\rangle$ is actually periodic. This is the last vestige of periodicity left over from the small systems, and provides a small argument for using the dense Hamiltonian, since the aperiodic systems spread more fully.

Note that the previous analysis has focused only on individual states, not operator spreading. Furthermore, for larger systems, looking at all coefficients becomes unwieldy. We can solve these problems simultaneously by using the tools developed in the previous section for quantifying operator spreading.

Before doing this, though, we will show that this Hamiltonian is quantum chaotic through an analysis of its spectrum. One method of diagnosing quantum chaos that is robust to finite size effects is the two-gap level statistics [26]. From the energy eigenvalues E_n define the gaps $\delta_n = E_{n+1} - E_n > 0$. When there is a conserved quantity that commutes with the Hamiltonian, only use the eigenenergies from states with the same quantum numbers, since different subspaces do not interact. The level statistics are

$$0 \leq r_n = \frac{\min\{\delta_n, \delta_{n-1}\}}{\max\{\delta_n, \delta_{n-1}\}} \leq 1. \tag{34}$$

In a random matrix the eigenvalues repel each other, so r_n are not close to 0.

To calculate out level statistics we find r_n and keep all but those for highest and lowest n. We also include a small term

$$H' \propto \sum_i S_i \cdot S_{i+1} \tag{35}$$

to break the rotational symmetry of the system, giving us more states to work with. This

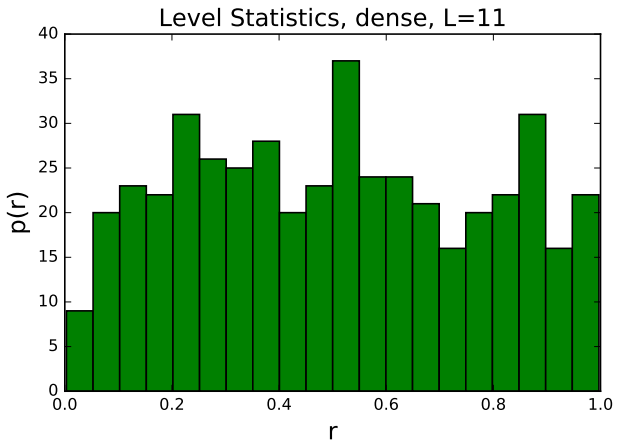

Figure 10: **Level statistics for dense Hamiltonian.** The suppression of ratios close to 0 implies that this Hamiltonian is chaotic.

small term does not noticeably affect any results in this section. The distribution of r_n is shown in Fig. 10. The distribution does appear to be suppressed at small r, so we conclude that this Hamiltonian is quantum chaotic.

3.2.2 Pauli String Weight

We can study the operator dynamics of this system by extending to large L and evolving operators that are identities on all sites except one end. The operators in this section are interpreted as observables, so they will evolve in the Heisenberg picture and cycle in the opposite direction than the states. Since the Hamiltonian is SO(3) symmetric, it does not matter if the perturbation is X, Y, or Z. For convenience, we will use Z. We start by discussing the Pauli weight, and delay discussing the OTOC and butterfly velocity to Sec. 3.2.3.

The first Hamiltonian we will discuss is the $L = 11$ sparse Hamiltonian. Although we will eventually use the dense one to study v_B, we can extract interesting behavior from the sparse one. For the right-propagating wave, with $A(t = 0) = Z_0 = Z \otimes \mathbb{I} \otimes \mathbb{I} \cdots$, the weight all starts on site 0. As it evolves, it reaches peaks for even sites, but does not rise above 1/10 for the odd sites. The successive peaks fall off in size but dominate the Pauli strings until the last site takes over. This evolution can be seen in Fig. 11).

To understand this behavior, realize that $W(i, t)$ only measures the weight of operators that end on site i, not those that reach past it. So, if all components of the operator that are non-identity on site i are also non-identity on site $i+1$, $W(i, t)$ will be very small. Since our Hamiltonian connects site 0 with site 1 only through terms that also connect site 0 to

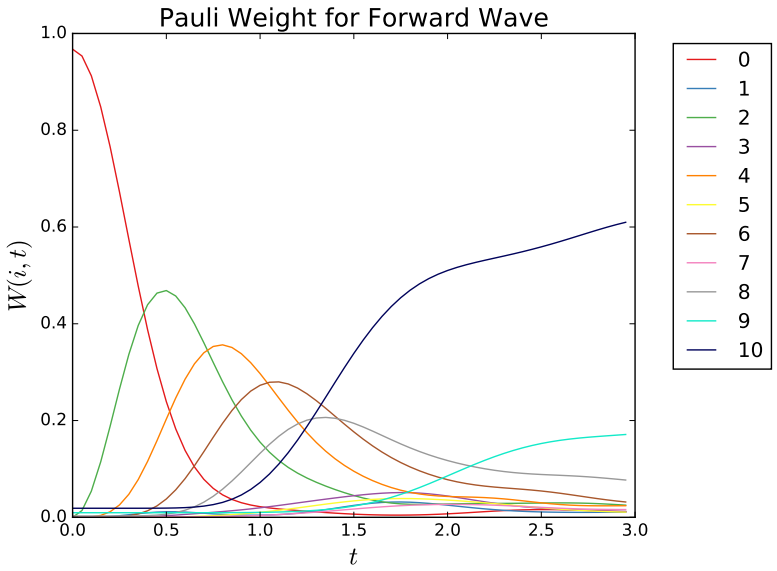

Figure 11: **Weight of operators that end on site** i in the sparse Hamiltonian. The weight starts on site and ends up mostly on site 10, with the large peaks occurring at even i.

site 2, this even-site hopping behavior makes sense.

For the left-propagating waves, with $A(t = 0) = Z_{L-1}$, the initial decaying signal is more difficult to make out but still present. Like in the right wave, it touches only even sites. The signal travels faster and decays faster. It is dominated by a large weight of sites that start on site 9 around $t = 1$. By $t = 3$ the first site has started to dominate the Pauli weights. These weights appear in Fig. 12.

Both the faster speed of the signal and the early dominance of the weight on site 9, which would be analogous to weight on site 1 in the right-moving case, can be understood through the asymmetric 3-site Hamiltonian and Fig. 4. First recall that evolution of operators proceeds in the opposite direction as evolution of states. Then, starting on site 0, a small amount of weight would initially move to site 1 before the entire weight moves to site 2. The fact that the entire weight moves to site 2 allows the evolution in Fig. 11 to be so smooth. However, if the weight starts on site 10, a small amount moves to site 8 before the entire weight moves to site 9. This explains both main differences between the left and right propagating signals.

Although the long-time behavior is interesting, it is strongly affected by finite size effects and is not useful in extracting operator dynamics. For this we use the early-time behavior, in the range in which Pauli weights are still growing polynomially. For right propagation

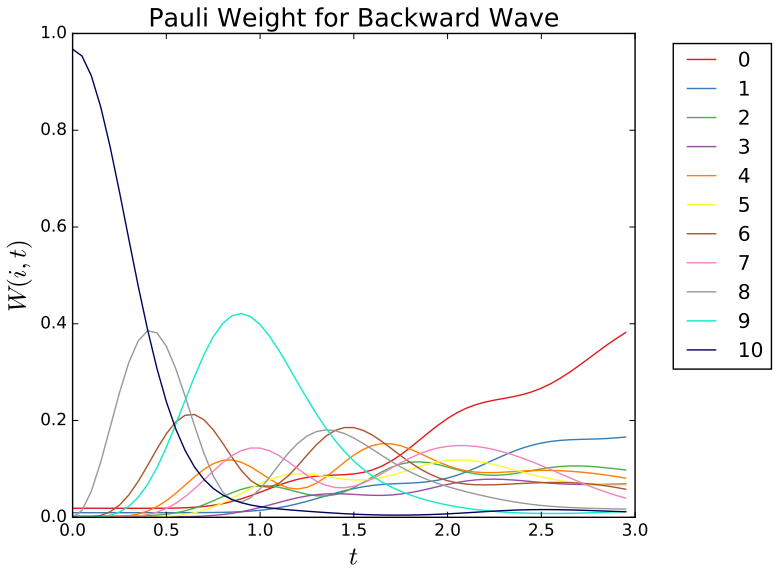

Figure 12: **Weight of operators that begin on site** i. The weight now starts on site 10 and ends mostly on site 0, but takes longer to get to site 0.

the initial behavior is $W(t;i) = \exp(a + b\log(t)) = e^a b^t$ with a, b given by

i	b	a
0	-0.22806104	-0.64537868
1	3.506984	-0.7706456
2	1.81723584	1.24814523
3	5.61941276	-0.85107562
4	3.84629108	1.71062193
5	7.69020229	-1.63644758
6	5.8674097	1.33960991
7	9.7051822	-2.99655938
8	7.8830393	0.37761653
9	11.6271597	-4.90046795
10	9.88377179	-1.04727877

The interesting data here is that the exponents are close to even integers. Even sites start with $b = 0$ at site 0 and increase by 2 for each site, while odd sites start with $b = 4$ at site 3. Figure 13 shows this behavior, while figure 14 shows the analogous behavior for the left propagating wave. The exponents there follow a similar pattern.

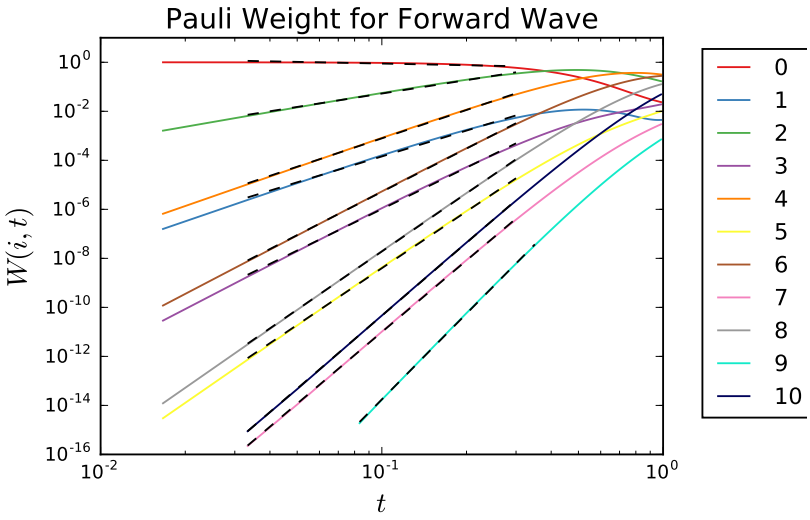

Figure 13: **Early-time leading edge weights for right propagating wave.** All weights initially grow or decay polynomially. Odd sites i are paired with site $i-3$, as explained below.

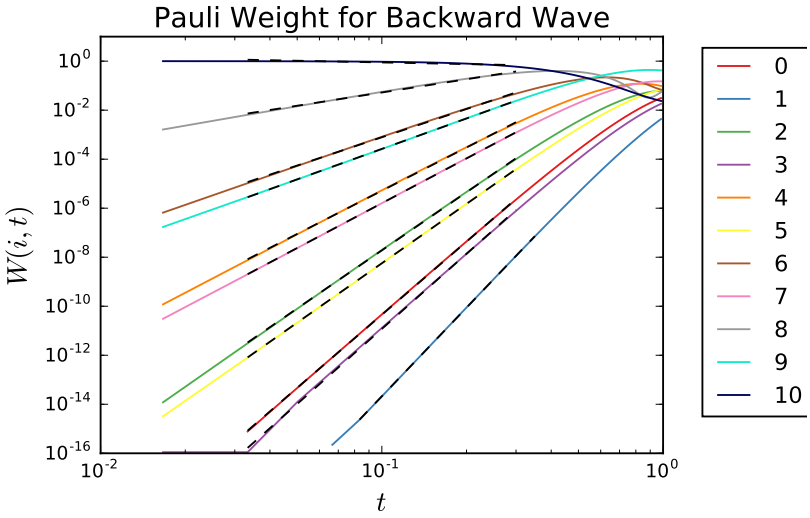

Figure 14: **Early-time leading edge weights for left propagating wave.** Odd sites are again paired with an even site 3 sites away.

To understand these exponents, treat t in the time evolution operator as a perturbation,

$$\begin{aligned} U(t) &= e^{iHt} \\ &= 1 - iHt - \frac{1}{2}H^2t^2 + \frac{i}{3!}H^3 + \frac{1}{4!}H^4t^4 + \dots \end{aligned} \tag{36}$$

The term linear in t contains a single H, which has matrix elements that connect site 0 to sites 1 and 2. The term quadratic in t contains a second H, which connects site 0 to sites 3 and 4 through site 2. The Pauli weight is quadratic in $A(t)$, which explains why the exponents for the OTOC are even.

Although this explains why the exponents for even sites are $0, 2, 4 \dots$, it suggests that the exponents for odd sites should be $2, 4, 6 \dots$. The reason they are not is the same reason that the weight on these sites is suppressed in the full evolution: they are hidden to first available order by the sites in front of them. To see the t^2 dependence on site 1, for example, we need the weight of all operators that are non-zero at that site, not just the ones that end there. We can find this information in $C(i, t)$, the normalized OTOC.

3.2.3 OTOC and Butterfly Velocity

After discussing the behavior of the Pauli weights in the sparse and dense Hamiltonians, we will discuss the OTOC, which encodes the amount of operator weight on a site, and from this explore the butterfly velocity. We do not find an explicit v_B but do show that it is different in the two directions.

First we show the full time evolution of the OTOC for the sparse and dense Hamiltonians, for left and right propagating waves. These can be found in Figs. 15 through 18. The most most obvious difference from the Pauli weight is that these weights asymptote to a non-zero value for all sites. As discussed in Eq. 13 the OTOC measures the weight of non-identity operators on a site, and asymptotes to 1/2, or 3/4 in our normalization, in random matrices. The operators in this Hamiltonian do not asymptote to 3/4 because of their conservation laws [22, 21].

Other interesting features include the persistence of the fast-moving signal in the left-moving wave of the sparse Hamiltonian and the clear asymmetry in early behavior between left- and right-moving OTOCs. The OTOCs for perturbations starting on site 0 consistently rise faster than the signals propagating from site 10. The dense Hamiltonian also transports weight faster than the sparse Hamiltonian.

Like for the Pauli weights we can study the early behavior. Since the forward and

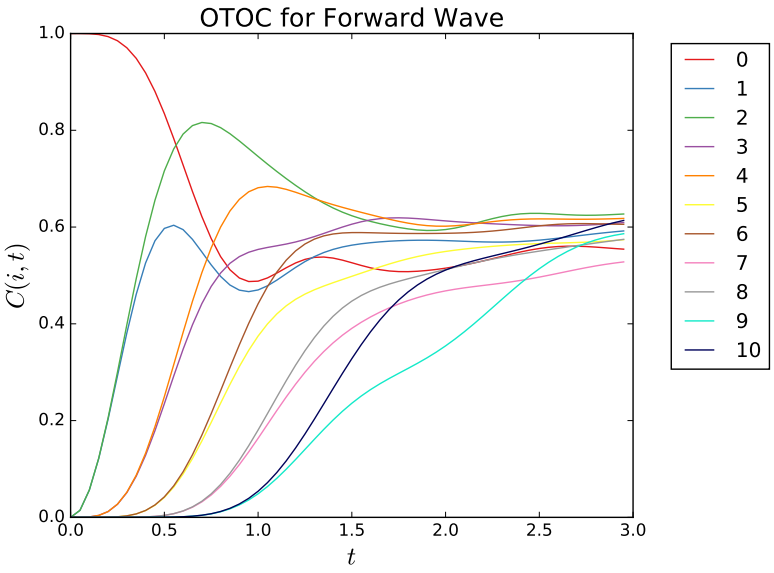

Figure 15: **OTOC for sparse Hamiltonian,** right moving perturbation. Note that odd sites are paired with neighbors now, because we are measuring the total weight on a site.

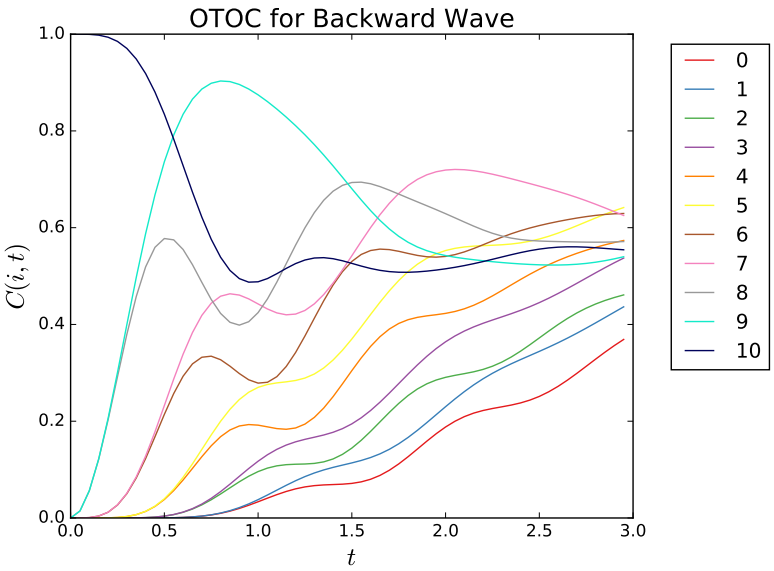

Figure 16: **OTOC for sparse Hamiltonian,** left moving perturbation. Once again odd sites are paired with their neighbors, and again the backwards perturbation moves slower.

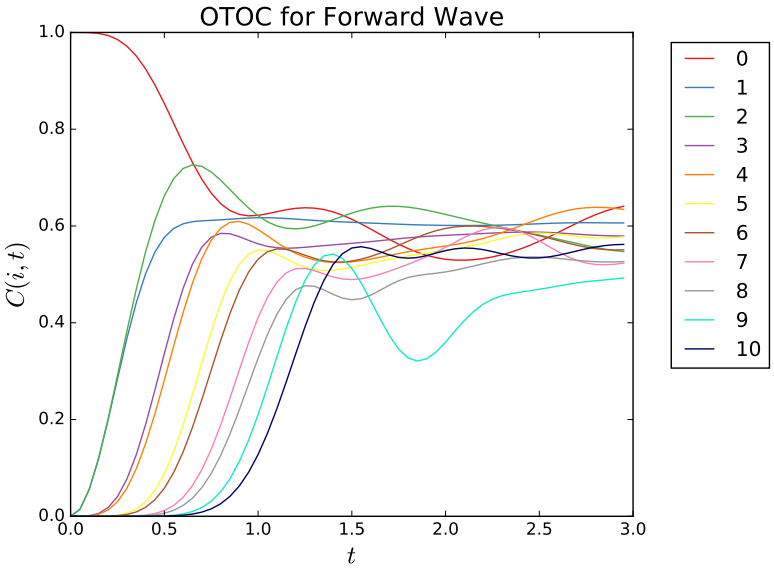

Figure 17: **OTOC for dense Hamiltonian,** right moving perturbation. Although odd sites are still paired with their neighbors at early time, each pair diverges more clearly due to having more 3-site Hamiltonians to move weight. Also, unlike the sparse analog, the pairs peak in the correct order in the forward direction.

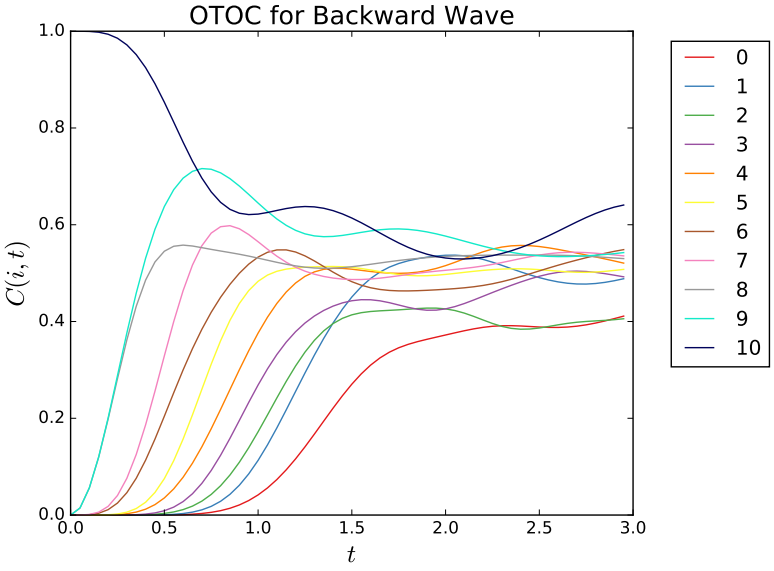

Figure 18: **OTOC for dense Hamiltonian,** left moving perturbation. Once again the pairs diverge, and once again the backward perturbation is slower.

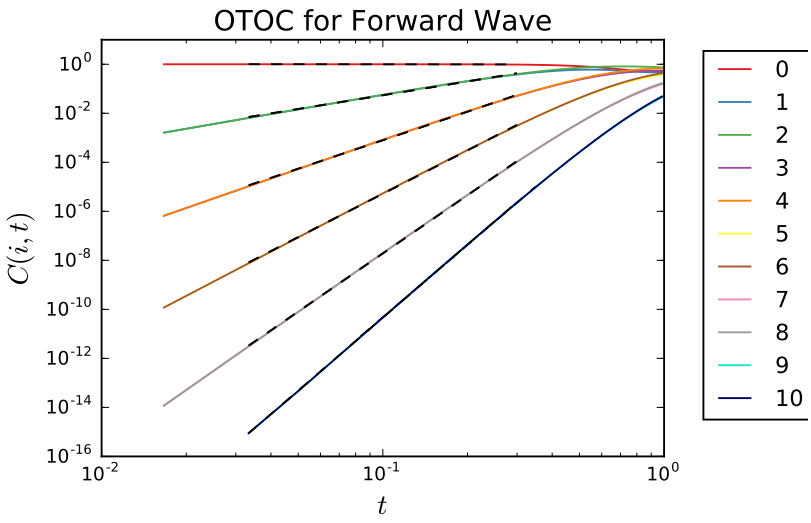

Figure 19: **Early-time OTOC for sparse Hamiltonian.** Sites 1 and 2 are on top of each other. The same is true of 3 and 4, etc.

backward behavior is so similar, we will only show the forward waves. For the sparse Hamiltonian the pairs of sites are so close they are indistinguishable, in Fig. 19. The exponents for the polynomial fits are

i	b	a
0	-0.0110169	-0.02903279
1	1.85885984	1.36784057
2	1.87645224	1.41529219
3	3.86681506	1.76926911
4	3.87270585	1.78515861
5	5.87947281	1.37400747
6	5.88219513	1.38134381
7	7.8910584	0.40044216
8	7.89252802	0.40439999
9	9.80995422	-1.17920962
10	9.88377179	-1.04727877

Although these are not integers, they are much more regular than the exponents for the Pauli weight were. They are close to the exponents suggested by the BCH decomposition in Eq. 10, which is the same as the number of 3-site terms needed to connect site i to the origin, as in Fig. 8.

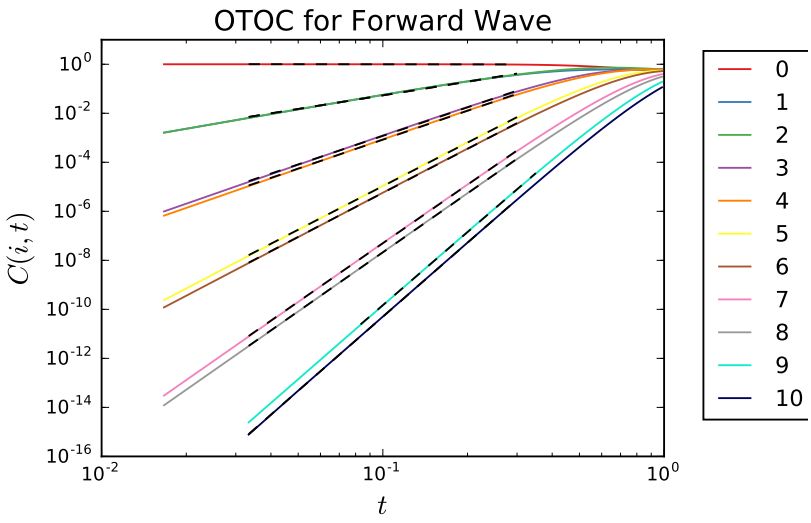

Figure 20: **Early-time OTOC for dense Hamiltonian.** Each pair is slightly separated now.

The early-time OTOCs for the dense Hamiltonian are in Fig. 20. Since the added 3-site Hamiltonians move some weight to the odd sites, the coefficients are not as close as there were in the sparse case.

To find the butterfly velocity, we have to first find the velocity-dependent Lyapunov exponents $\lambda(v)$. To do this we plot $C(vt,t)$, where $vt = i$, for different values of v, as in Eq. 16. For a sufficiently large system there should be a $v = v_B$ such that $C(vt,t)$ does not decay with t. Our system suffers from finite size effect that prevent us from finding v_B exactly, but we can still explore the v dependence of $C(vt,t)$.

We will focus on the dense Hamiltonian, as it has more asymmetry. Unsurprisingly, $C(vt,t)$ suffers from the same odd/even effects we have been seeing all along. Fig. 21 shows the evolution of the weights, plotted on a semilog plot with respect to the site index, so that the slope is $\lambda(v)/v$. The problem is that weight too quickly reaches the even sites and does not reach the odd sites fast enough.

We could still use these plots to calculate $\lambda(v)$ using only even or only odd sites, but there are few enough sites that the calculation will have high uncertainty. Alternatively we can boost the odd sites by connecting site 0 to site 1 with a term not connected to site 2. This will change the overall dynamics but hopefully only by bringing neighboring sites closer to each other, so that it does not change the dynamics averaged over even and odd sites.

The simplest rotationally symmetric term with the desired property is $H' = \boldsymbol{S}_i \cdot \boldsymbol{S}_j$. The

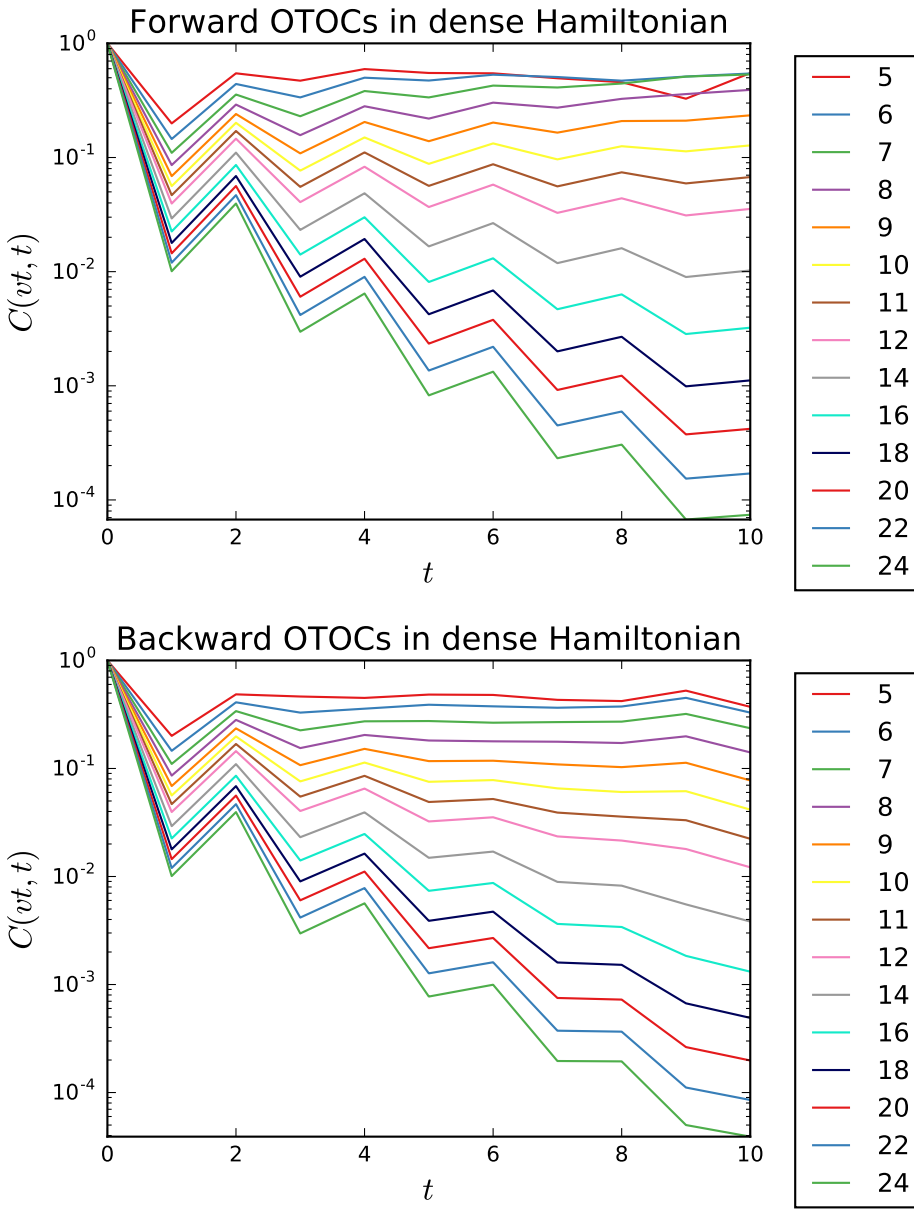

Figure 21: **Velocity-dependent OTOC decay** in dense Hamiltonian. The jaggedness comes from odd/even effects. Each line corresponds to a velocity, given in the legend. The overall behavior is exponential decay as expected.

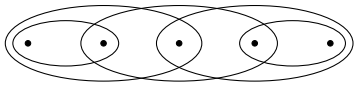

Figure 22: **Dense Hamiltonian with 2-site terms**. This diagrammatically shows how the 2-site terms connect the end sites to their nearest neighbors.

coefficient of H' should be chosen to maximize the smoothness of the OTOC. Empirically, a very effective coefficient is 4. Smaller constants left too much of the jaggedness, and larger constants introduced jaggedness in the other direction. In fact, the lack of an ideal coefficient can be seen on the first site in both plots, where the first site receives too much weight at low velocity and not enough at high velocity. We also add a 2-site term to spins 9 and 10, to have the same effect on the backwards wave. Together these terms replace the picture in Fig. 8 with one that looks like Fig. 22.

The total Hamiltonian is

$$H = \frac{2\pi}{3\sqrt{3}} \sum_{i=0}^{L-2} \boldsymbol{S}_i \cdot (\boldsymbol{S}_{i+1} \times \boldsymbol{S}_{i+2}) + 4\,(\boldsymbol{S}_0 \cdot \boldsymbol{S}_1 + \boldsymbol{S}_{L-2} \cdot \boldsymbol{S}_{L-1})\,. \tag{37}$$

This perturbation is large enough to noticeably change the plots of the Pauli weight and the OTOC, by shuffling weight back and forth between the initial two sites early in the evolution. However the early time behavior, averaged over even and odd sites, is similar enough to not change $\lambda(v)$ dramatically.

Using the new Hamiltonian, the OTOC decay becomes much more linear, although still not perfectly so, as shown in Fig. 23. Fig. 24 shows the v dependence of $\lambda(v)$. In a large enough system, $\lambda(v)$ should approach 0 for finite v. However, due to the small size of our system we were not able to effectively probe these slow speeds. However, it is clear that $\lambda_+ > \lambda_-$ for all v. Therefore, for these time-independent Hamiltonian systems, $v_{B+} > v_{B-}$.

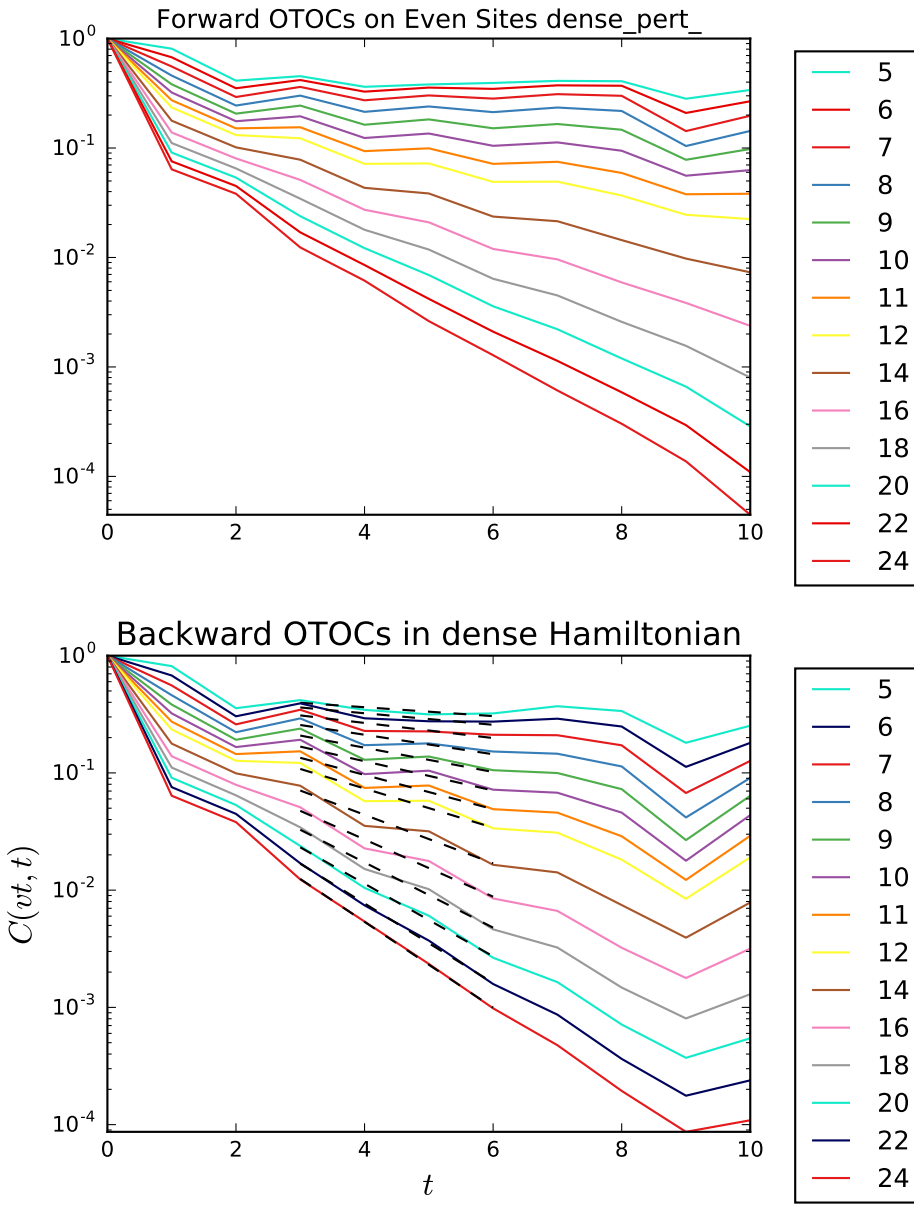

Figure 23: **Velocity-dependent OTOC decay** in dense Hamiltonian with initial and final dot product perturbations. Although there is still jaggedness on the edges, the linear behavior in the bulk allows us to extract $\lambda(v)$.

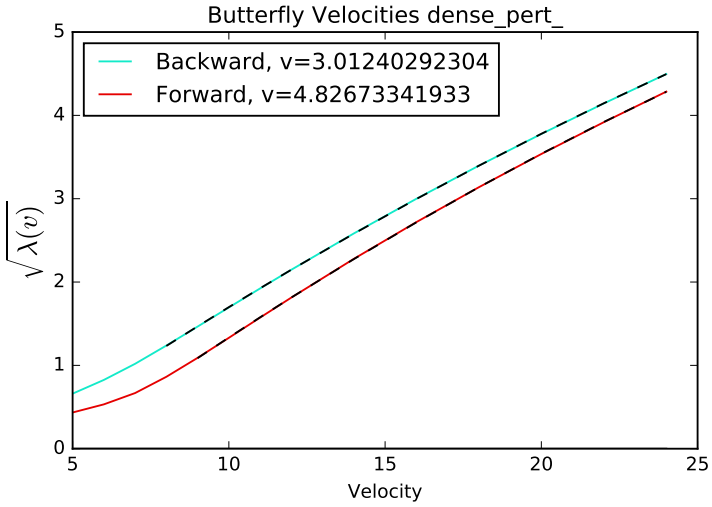

Figure 24: **Velocity dependent Lyapunov exponents** for the dense Hamiltonian with end perturbations. The butterfly velocity should be where these curves cross the origin, such $\lambda(v_B) = 0$. This does not happen, due to finite size effects, but we can see that $\lambda_+ > \lambda_-$ for all v, which implies that $v_{B+} > v_{B-}$.

4 Entanglement and Operator Spreading in Quantum Circuits

One drawback to studying systems with time-independent Hamiltonians is that their dynamics are constrained by their conservation laws [21]. Systems with no conservation laws would not have these constraints. Quantum circuits have no explicit Hamiltonian, but rather evolve unitarily at discrete time steps, making them ideal for studying unconstrained dynamics. Furthermore, while numerical simulation of the Hamiltonian system is limited to small systems with large finite size effects, quantum circuits can have solvable dynamics that can be efficiently simulated.

The circuits provide solvable systems in certain limits, after averaging over circuit architecture. In this limit, which will be discussed later in this section, it is easier to discuss information dynamics in terms of entanglement entropy rather than operator spreading. This is not to say that operator spreading does not apply in circuits, or that entanglement dynamics do not apply to Hamiltonian systems. Rather, this is an easier tool to use for these systems.

In this section we first introduce bitartite entanglement entropy and the constraints it must satisfy. We then discuss the construction of quantum circuits, and the behavior of the entanglement entropy in these circuits. We describe a solvable limit from Ref. [18]. We show that it is possible to calculate v_B entirely from the entanglement entropy, and that this velocity is the same as v_B calculated from operator spreading.

4.1 Entanglement Entropy

Quantum entanglement has applications to branches of physics from high energy and quantum information theory to experimental studies of cold atomic gases. Entanglement entropy can be used to study topological properties in systems with mass gaps [27]. In conformal field theories entropy informs the scale of the renormalization group flow [28, 29]. Although entanglement is so widely studied, its dynamics are less well understood. The dynamics of the entanglement are closely related to the speed at which information travels or spreads.

Let us briefly review entanglement in composite systems. First assume the full system AB can be decomposed into subsystems A and B, and AB is in the pure state $|\Psi\rangle_{AB}$. In the spin chains considered here, the chain is often divided such that A is all spins on one side of a cut and B is all spins on the other side. Another common decomposition is having

A be all spins between two cuts, and B being all other spins. If the states of A and B are entangled it is impossible to write $|\Psi\rangle_{AB} = |\psi_A\rangle_A |\psi_B\rangle_B$. If this is possible then $|\Psi_{AB}\rangle$ is not entangled and is called a product state. Subscripts on kets represent the space that the ket lives in, but this is clear from the name of the state (uppercase for the full state, subscripts otherwise) so we will drop the outside subscripts.

We can assess the amount of entanglement by looking at the density matrix $\rho_{AB} = |\Psi\rangle \langle\Psi|$. All density matrices satisfy $\operatorname{Tr} \rho = 1$. Generic pure state density matrices further satisfy

$$\operatorname{Tr} \rho^2 = \operatorname{Tr} |\psi\rangle \langle\psi| \, |\psi\rangle \langle\psi| = \operatorname{Tr} |\psi\rangle \langle\psi| = 1. \tag{38}$$

We can further construct the reduced density matrices ρ_A and ρ_B as the full density matrix ρ_{AB} traced over subsystem B and A, respectively. If $|\Psi\rangle$ is a product state, the reduced density matrices are $\rho_A = |\psi_A\rangle \langle\psi_A|$, etc. However, if the states are entangled the reduced density matrices do not have such a nice form. They will still have trace 1, but $\operatorname{Tr} \rho^2 < 1$ for a mixed state.

A maximally entangled state will be of the form

$$|\Psi\rangle = \sum_i^N \frac{1}{\sqrt{N}} |\psi_{A,i}\rangle |\psi_{B,i}\rangle , \tag{39}$$

where N is the dimension of the larger of the two Hilbert spaces and i labels orthogonal states. $\operatorname{Tr} \rho_{AB}^2 = 1$ because the full state is pure. The reduced density matrices are

$$\rho_A = \sum_i^N \frac{1}{N} |\psi_{A,i}\rangle \langle\psi_{A,i}| , \tag{40}$$

with $\operatorname{Tr} \rho_A^2 = \sum_i \frac{1}{N^2} = \frac{1}{N}$, with similar results for B. This provides the constraints $\frac{1}{N} \leq \operatorname{Tr} \rho^2 \leq 1$.

Bipartite entanglement entropy provides a more general way to quantify the entanglement and is defined as the quantum entropy of one of the reduced density matrices. As long as the full state is pure, this is equivalent for either subsystem. There are multiple quantum entropies. The nth Renyi entropy of density matrix ρ is

$$S_n = \frac{1}{1-n} \log \left(\operatorname{Tr} \rho^n \right) . \tag{41}$$

In the limit $n \to 1$ this becomes the von Neumann entropy

$$\begin{aligned} S_{vN} &= \frac{1}{1-(1+\epsilon)} \log\left(\operatorname{Tr} \rho^{1+\epsilon}\right) \\ &= -\frac{1}{\epsilon} \log \operatorname{Tr}\left[\rho \rho^{\epsilon}\right] \\ &= -\frac{1}{\epsilon} \log \operatorname{Tr}\left[\rho\left(1+\epsilon \log \rho\right)\right] \\ &= -\frac{1}{\epsilon} \log\left(1+\epsilon \operatorname{Tr} \rho \log \rho\right) \\ &= -\operatorname{Tr} \rho \log \rho, \end{aligned} \tag{42}$$

the analogue of the classical Shannon entropy. The calculation uses the small x expansions of a^x and $\log(1+x)$. All Renyi entropies are maximized by maximally mixed states, with entropy $N \log q$ for N-site systems with q-dimensional Hilbert spaces at each site. They are also minimized by pure states, which have entropy 0.

For the spin chain, we can define many different subsystems, each with a corresponding entanglement entropy. The following description is largely taken from [18]. Consider a spin chain of N sites with dimension q. Sites are labeled by $i = 1, \ldots N$, while the bonds between sites are labeled by $x = 1, \ldots N-1$. After cutting the system at bond x, define the entropy across this cut as the bipartite entanglement entropy of all sites to the right of x. If the whole chain is in a pure state, this is equal to the bipartite entanglement entropy of all sites to the left of x.

We can define a function $S(x) = -\operatorname{Tr} \rho_x \log \rho_x$ where ρ_x is the density matrix of the system with all sites left of x traced out. As long as the full system is in a pure state, this is the von Neumann entanglement entropy of the two subsystems divided by the bond at x. For convenience logarithms are taken base q because $\log_q q = 1$. $S(x)$ will observe constraints not already made clear due to the fact that subsystems defined by adjacent x have heavy overlap.

Classically, for an arbitrary system decomposable into subsystems A and B, the entropies satisfy $\max(S(A), S(B)) \leq S(AB) \leq S(A) + S(B)$. In quantum mechanics, this is replaced by the subadditivity of the von Neumann entropy

$$|S(A) - S(B)| \leq S(AB) \leq S(A) + S(B). \tag{43}$$

If we take subsystem A to be the single site between cuts x and $x+1$ and subsystem B to

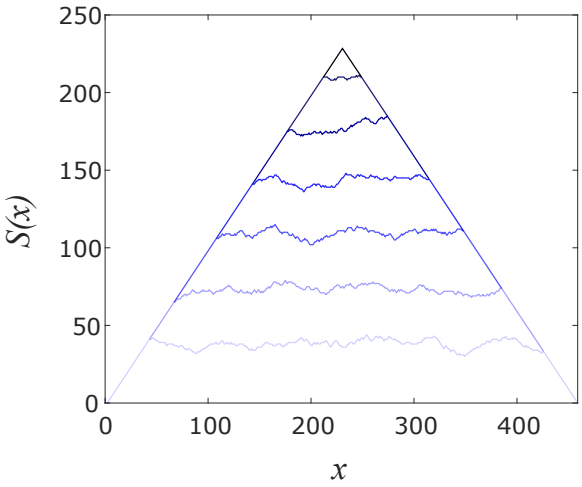

Figure 25: **Bitartite entanglement entropy at different times,** (t = 340, 690, 1024, 1365, 1707, 2048 and 4096). This data comes from a circuit model, but the behavior is believed to be general. Figure from [18].

be all sites right of $x+1$, this becomes

$$|S_1 - S(x+1)| \leq S(x) \leq S_1 + S(x+1), \tag{44}$$

where S_1 denotes the entropy of the single site between cuts x and $x+1$. After some rearranging this can be written $|S(x+1) - S(x)| \leq S_1$. However, since the single site is q dimensional, $S_1 \leq \log_q q = 1$. The preceding arguments taken together give the constraint

$$|S(x+1) - S(x)| \leq 1. \tag{45}$$

A finite system has $S(x)$ pinned at its endpoint, because the entanglement entropy of a single site is bounded by 1. Then the maximally entangled state has $S(x) = \min\{x, L-x\}$. As the entanglement approaches this value it takes forms as in Fig. 25. The flat section has a constant growth rate $\frac{\partial S(x,t)}{\partial t}$ so the border position y moves at a constant speed defined by the growth rate and the slope. Since the maximal slope is $\frac{\partial S(x,t)}{\partial x} = 1$, this velocity is

$$v_E = \frac{\partial y}{\partial t} = \frac{\partial S}{\partial t}\left(\frac{\partial S}{\partial x}\right)^{-1} = \frac{\partial S}{\partial t}. \tag{46}$$

This entanglement speed is equivalent to the entanglement growth rate, under certain assumptions not mentioned here. This calculation is repeated more generally and in more detail in Sec. 4.4.

Refs. [17, 22, 21, 18] discuss the speed of entanglement in brickwork models using the

OTOC and operator density. Reference [30] quantifies the scrambling using the operator entanglement entropy opEE of the time evolution operator.

4.2 Circuit Architectures

Like the earlier systems, quantum circuits consist of chains of Hilbert spaces. Instead of evolving under a time-independent Hamiltonian, they evolve with unitary operators, called gates, at discrete times. If a gate acts on a site at time t it is said to 'fall' on that site. Simple circuits only contain 2-site gates which act on the product of Hilbert spaces at adjacent sites. Once again to clarify the term "operator", the "unitary operator," "unitary gate," or just "gate" will be the time evolution operator of the circuit, while "operator" without qualification will refer to the observable that is evolving in time.

We could study the dynamics of a single circuit, in analogy with the single Hamiltonian studied in the previous section. However, it is also instructive to look at ensembles of circuits, with some type of randomness. This section presents two sources of randomness, the choice of gates and their locations in the circuit.

One source is the unitary operator chosen at each site at each time. Different gates result in different architectures. If the Hilbert space at each site is q dimensional, the space of 2-site unitary matrices is $q^2 \times q^2$ dimensional. To find representative behavior, each circuit is made by drawing gates from some distribution. The circuits can then be averaged over these choices. A common choice is the Haar distribution, which is invariant to rotations in the space of 2-site operators. Since these act on a q^2-dimensional Hilbert space, there are q^4 independent operators, including one trivial operator, the identity.

Another source of randomness is the placement of the gates. The "brickwork" model [17] uses a deterministic architecture, in which at odd times there is a gate to the right of every odd site, while for even times there is a gate to the right of every even site. Figure 26 shows an example. The "random" architecture [18] chooses a random site for a single gate at each time step $t = n/(L\gamma)$, where γ is the gate rate per site and n is an integer. The placement is then averaged over.

The random architecture is equivalent to placing gates at each site in continuous time with Poisson-distributed time steps, as the length L is taken to infinity with the gate rate density γ held constant. [18]. This thesis focuses on a generalization of this model, described in Sec. 5.

To study operator spreading, we might want to know how the Pauli end-weight $W(t, i)$ evolves in time in these circuits. Recall that, given an observable $\mathcal{O}(t)$ evolving in time, this

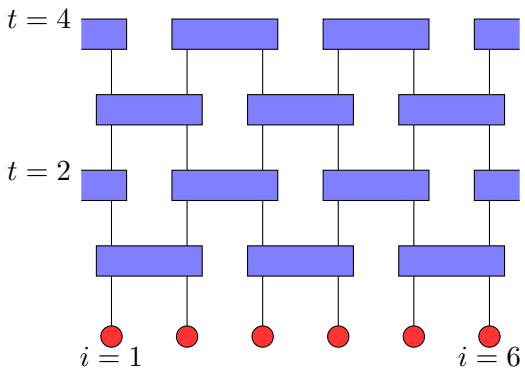

Figure 26: **Brickwork circuit architecture**. The red circles represent the initial states of the system, while the blue rectangles are unitary gates that evolve the system in time. The Hilbert space at each site is q-dimensional, meaning the gates are $q^2 \times q^2$ unitary matrices, chosen from the Haar distribution. Adapted from [17].

measure how far $\mathcal{O}$ has spread. Consider a unitary gate falling on sites i and $i+1$. Some components of the operator $\mathcal{O}(t)$ will be non-identity at site i but identities after i, and will have weight $W(t,i)$ by definition. There are q^2-1 operators like this. A randomly chosen unitary gate will evolve the operator on sites i and $i+1$ to a random operator that is not the identity on both sites. There are q^4-1 of these, q^2-1 of which are identities on the second site.

If some component of the observable ends on site $i+1$, the unitary can move the end back, if it evolves the observable at $i+1$ to the identity. Again, q^2-1 of the random final operators are identities on the second site, so that the probability of moving weight from site $i+1$ to i is the same as not moving the weight from i to $i+1$. The probability that a random gate advances $W(t,i)$ is $p = \frac{q^2}{q^2+1}$, so that $1-p = \frac{1}{q^2+1}$. After averaging over the possible circuits, this leads to the dynamics of $W(t,i)$, from [17]

$$\begin{aligned} W(t+1,i) &= (1-p)\left(W(t,i)+W(t,i+1)\right) \\ W(t+1,i+1), &= p\left(W(t,i)+W(t,i+1)\right). \end{aligned} \tag{47}$$

4.3 Deterministic Limit

Note the the dynamics in Eq. 47 are degenerate in the large q limit to deterministic movement of end sites with the application of each gate. This implies that this limit may be interesting, and indeed it is. This is also seen by studying the entropy growth.

Recall the entropy bound from Eq. 45. This means that when a gate falls at bond x,

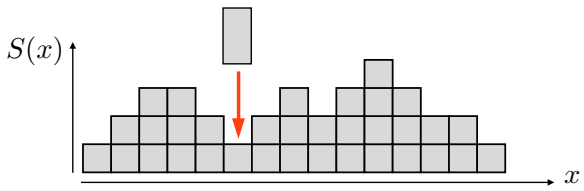

Figure 27: **Tetris-like model for large-q chain.** The gate at cut x adds enough entropy so that $S(x)$ is one greater than either of its neighbors. Figure from [18].

the maximum entanglement possible across cut x is

$$S(x) \leq \min\{S(x-1), S(x+1)\} + 1,$$

while the gate does not affect $S(x-1)$ or $S(x+1)$. It would be helpful to find a solvable limit of quantum circuits under which this inequality is saturated because it would completely specify the evolution of the entanglement.

Such a limit is found by taking $q \to \infty$. Ref. [18] includes a detailed proof, but the key is that all unitary gates that do not saturate the bound obey some polynomial equation, so these operators define a measure zero subset on the Haar distribution. We know then that, with probability 1, the random two-site gate will maximize bipartite entanglement entropy.

Taken together, these facts mean that after a gate across cut x, the new entanglement entropy is

$$S(x, t+1) = \min\{S(x-1, t), S(x+1, t)\} + 1. \tag{48}$$

Then if at any time $S(x,t)$ becomes integer valued at all x, it remains so for the rest of the evolution. There are a few pictures that make this integer-valued evolution more intuitive.

4.3.1 Surface Growth Picture

One model for the entropy growth described above is the Tetris-like surface growth picture. Here, the entropy is represented by a piecewise-constant function with the height given by the entropy across each cut, as in Fig. 27, taken from [18]. In general, it is possible for the entropy across two adjacent cuts to be equal. However, Fig. 28, also from [18], shows that flat sections can be destroyed but not created.

If the system has been evolving for a long time, then, it makes sense then to only consider states that have no flat sections, so that $S(x) = S(x-1) \pm 1$ for all x. In this case, instead of a picture like Fig. 27 it is possible to represent the entropy at each bond as a point, with

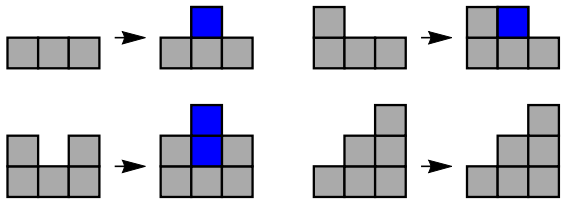

Figure 28: **Several possibilities for local changes** in the large q model. If three adjacent cuts all have equal entropy, the two flat sections annihilate each other. There is no way to generate flat sections. Figure from [18].

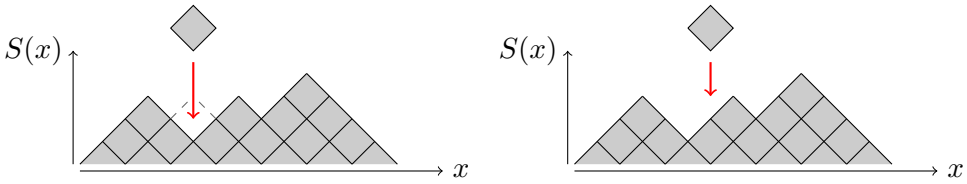

Figure 29: **Another picture of the configuration in Fig. 27**. Here the cut positions are top or bottom vertices of the squares. On the left, the gate at time t results in entropy growth at x because at time $t-1$, $S(x-1) > S(x) < S(x+1)$. On the right the gate has no effect. Although $S(x) < S(x+1)$, the case $S(x) < S(x-1)$ is not satisfied. The bottom row of the picture is not made of squares because the state has been evolving for a long time.

a diagonal slope connecting bonds. The slope between cuts (at a site) will always be ± 1. Instead of 2×1 rectangles, the gates are then squares coming down point-first, as in Fig. 29. If the gate falls on a local minimum, it lifts the entropy at that bond. If the gate falls on a local maximum or a non-extremum it has no effect.

In this picture, we can describe an entropy curve as a series of up and down steps, with each step defined at a site (between two cuts), For example, the initial state in Fig. 29 would be *uuudduuduudddd* and the final state would be *uuudududuudddd*. To make this look like a classical spin system (as in the Ising model), we can a state s, where s_i is the local slope at site i, with $u \to 1$ and $d \to -1$. This encoding provides simple expressions for the coarse-grained slope and correlation, which we will use in Sec. 4.4.

4.3.2 Minimal Cut Picture

Using the fact that a productive gate raises the entropy at a cut by 2, we can use another picture of the dynamics based on Fig. 26. Assume the initial state is a product state so the entropy at each cut is initially 0 and evolve with a brickwork circuit. Then after one of the first gates (for example between $i = 3$ and $i = 4$) the entropy at that cut will become 1. After any of the subsequent gates the entropy at the affected cut increases by 2.

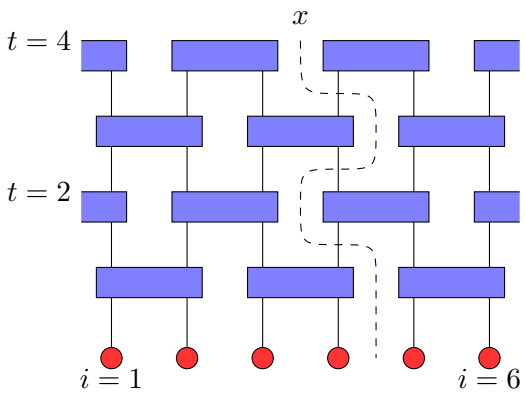

Figure 30: **Minimal cut picture.** Assuming an initial product state, $S(x)$ is given by the minimal number of legs that must be cut to draw a line from the bottom of the circuit to x. Cutting through gates is not allowed (or equivalently costs two cuts) and the initial point need not be x. Figure adapted from [18].

The minimal cut picture [18] reproduces this behavior by considering cuts through the circuit, as in Fig. 30. To find the entanglement across x, start in an unentangled state. If the initial state is entangled, expand the circuit backward by including gates for $t < 0$ that evolve a product state into the initial state at time $t = 0$. Then, starting at the top of the circuit at x, find the path through the circuit that cuts through no gates and through the fewest of the lines defined by sites, called legs. The path need not end at x. The entanglement at x is the number of legs cut by this path.

For the brickwork circuit, this reproduces the entropy calculation quoted above, with first gates generating 1 unit of entanglement and subsequent gates generating 2. However, in circuits where not every gate is productive, the min-cut picture still works, and agrees with the surface growth picture.

Consider the circuit in Fig. 31. After the initial gate acts between sites 3 and 4, $S(x)$ becomes a local maximum, with so that all gates after it become unproductive. In the surface growth picture this is represented by multiple gates falling on a local maximum. This is to show that the min-cut picture agrees with the surface growth picture, and provides another set of intuition for the dynamics of this class of circuits.

Since cuts through gates can considered as having cost 2, there is one more equivalent picture for the min-cut procedure: at the location of each gate, cross, or "tangle," the two relevant strands. Fig. 32 shows this picture for the same circuit as in Fig. 31.

Instead of asking about the entanglement from x to the bottom of the circuit, we can find the entanglement between x and a specific bond position y at the bottom of the circuit, where x and y are separated by time t. This is the entanglement of the time evolution

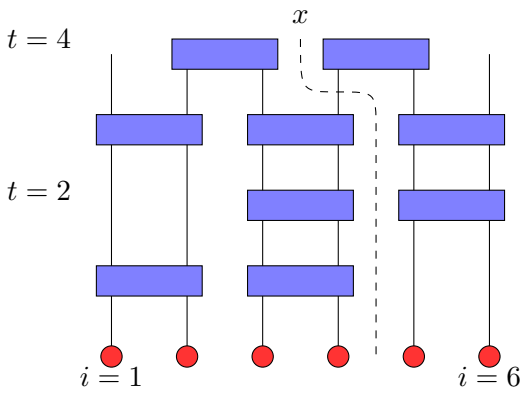

Figure 31: **Minimal cut with degenerate gates.** All gates at cut x after the first gate there produce no entanglement because $S(x)$ is a local maximum.

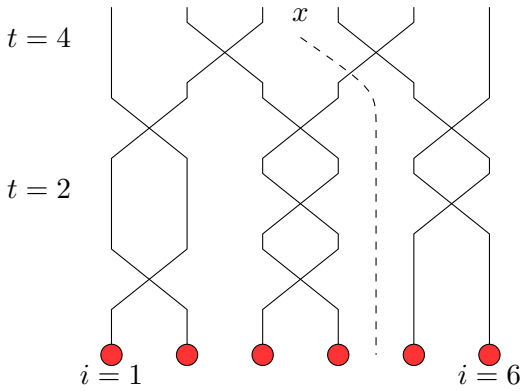

Figure 32: Another representation of the min-cut picture, with gates replaced by tangled strands. The advantage of this view is that gates clearly look like objects whose crossing has twice the cost of crossing a strand.

operator $U(t)$, and will be denoted $S_U(y, x, t)$ [21]. Then the entanglement is just the cost of the minimal cut from x to y. There are two reasons this is a useful quantity. Once is in finding the entanglement of operators, as the space of operators behaves like a doubled space of state [22, 21]. Alternatively, if we want to find $S(x)$ for an arbitrary initial entanglement, we can compute $S(x, y)$, and then $S(x) = \min_y\{S(x, y) + S_0(y)\}$.

4.4 Coarse Graining and Long Wavelength Dynamics

It is possible to abstract out even more of the detailed dynamics to consider the large scale behavior. To do this, interpret $S(x)$ as a continuous real valued function. The entanglement growth rate can be calculated under the assumption that the large scale slope $\frac{\partial S}{\partial x}$ is constant. We will go through this procedure for the random circuit architecture, where the results are exact in the large L limit. In Sec. 5 we calculate this function for different architectures, but we only calculate an approximation to this function because we will work under an inexact assumption, that individual sites are uncorrelated. However, that assumption is exact in the random circuit.

For an entropy surface with constant slope $m \equiv \frac{\partial S}{\partial x}$ and no correlations, each step from one site to the next has probability $\frac{1+m}{2}$ of being up and $\frac{1-m}{2}$ of being down. Consider a gate operating on cut x at time t. For the gate to increase the entropy $S(x)$, it must be the case that $S(x) < S(x-1), S(x+1)$. The probability of this is $\frac{1+m}{2}\frac{1-m}{2} = \frac{1-m^2}{4}$. In this case we have $S(x, t+1) = S(x, t) + 2$, because the gate increases the entropy to be great than that of its neighbors. Then if the gates arrive at each cut with a rate γ, the entanglement growth rate is

$$\frac{\partial S}{\partial t} = \gamma\frac{1-m^2}{2} \equiv \Gamma(m). \tag{49}$$

Useful checks of this formula are that the entropy does not increase at maximal or minimal slope $m = 1, -1$, and that $\Gamma(0) = \gamma/2$, $1/4$ the brickwork value. The rate $\Gamma(0)$ makes sense because in the case of the brickwork circuit all gates are guaranteed to raise the entropy, while here only $1/4$ will have an effect.

4.4.1 Entanglement Velocity

$\Gamma(m)$ contains encodes a significant amount of information about the entanglement growth of the system [21], and is useful in calculating both the entanglement velocity and the butterfly velocity. We will discuss this function in more detail and show how these

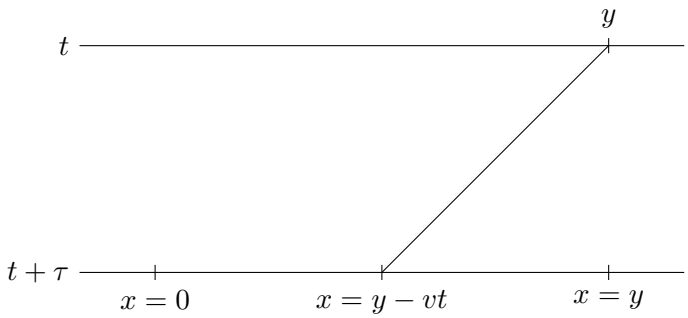

Figure 33: **Graphical interpretation of the Legendre transform.** Since the architecture is spacially homogeneous, S_U depends only on the length and slope of the cut and is $S_U(y, x, \tau) \equiv G(v)\tau$, where $v = (x - y)/\tau$.

speeds can be extracted.

Consider a linear section of the entanglement curve $S(x, t)$, and its evolution through time τ. We know that $S(y, t + \tau) = \min_x\{S_U(y, x, \tau) + S(x, t)\}$. If we assume the gate architecture is translationally invariant (which is true in the large-scale limit) then $S_U(y, x, \tau)$ depends only on the slope $v = \frac{y-x}{\tau}$, with the overall value scaled by the length of the cut from x to y. Graphically, this corresponds to Fig. 33. Since we assumed the initial state was linear, we can write it as $S(x, t) = mx = mv\tau + my$. From the linearity of S_U, and the fact that its only other dependence is on v, we can define $S_U(y, x, \tau) \equiv G(v)\tau$.

The final entanglement $S(y, t + \tau) = \min_x\{G(v)\tau + mv\tau\}$ can be differentiated with respect to τ to obtain.

$$\Gamma(m) = \min_v\{G(v) + mv\}. \tag{50}$$

Ref. [21] contains this equation with an extra factor of m_{eq}, the equilibrium slope. In the large-q random circuits we are considering, this slope is 1 due to the saturation of Eq. 45.

One useful fact about $\Gamma(m)$ is that it is convex. If it were not, then there would be some m^* such that $\Gamma(m^* + \epsilon), \Gamma(m^* - \epsilon) < \Gamma(m^*)$. However, if this is the case, we can always draw a jagged path from x to y alternating between slopes $m^* + \epsilon$ and $m^* - \epsilon$. This path with have average slope m but have a higher growth rate than $\Gamma(m)$. Therefore we have reached a contradiction and $\Gamma(m)$ is convex.

We can use the coarse-grained dynamics to reproduce the calculation of the entanglement velocity in Eq. 46. The entanglement velocity is the speed at which the kinks between the flat section $m = 0$ and the sloped sections $m = \pm 1$ move in the coarse grained version of Fig. 25. The sloped section does not move up, while the flat section moves up at $\Gamma(0)$. To

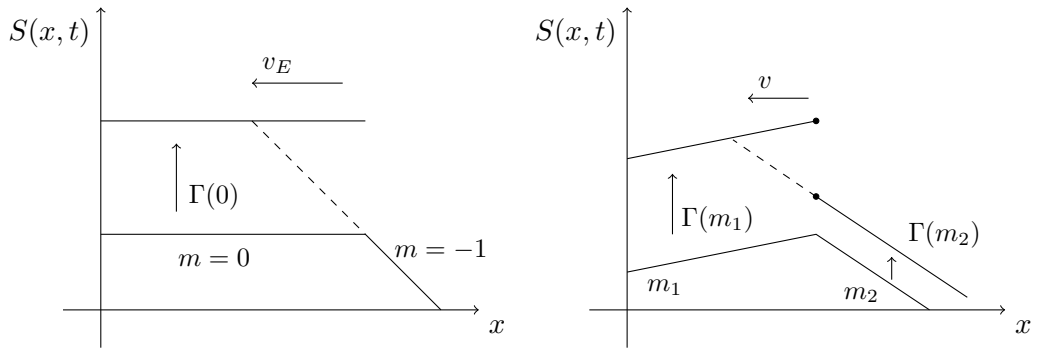

Figure 34: **Entanglement velocity in the surface growth picture.** The left figure shows a flat slope rising at $\Gamma(0)$. The $m=-1$ section does not grow. The flat slope would rise to the shown line, but the right end is held down by the entanglement bounds. The dashed line shows the corrected $S(x)$. On the right both sections grow at the growth rates defined by their slopes. To maintain continuity the kink moves at $v=-\frac{\Gamma(m_2)-\Gamma(m_1)}{m_2-m_1}$.

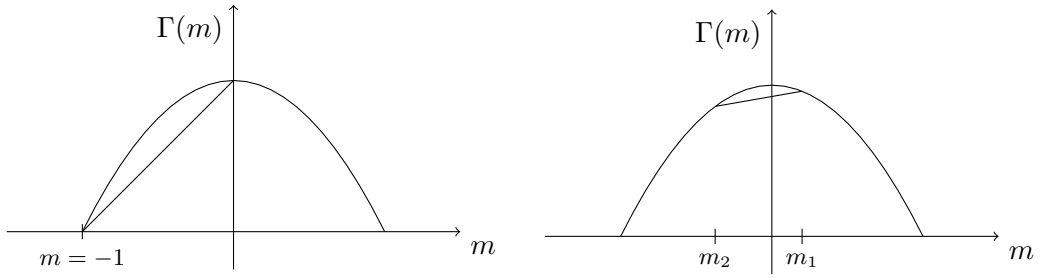

Figure 35: **Calculation of kink speed from chord,** using the growth rate for random circuits, Eq. 49.Note that the chord slopes are positive, while the corresponding velocities (Fig. 34) are negative.

preserve the continuity of $S(x)$, the kinks have to move at $v_E=\Gamma(0)$. Fig. 34 graphically shows this calculation of v_E.

The same figure can be used to calculate the velocity of a kink between arbitrary slopes, as long as the kink is a local maximum. In this case the vertical distance between the two solid line endpoints is $(\Gamma(m_1)-\Gamma(m_2))t$. This must be equal to the total vertical rise along the two slopes, $(m_1-m_2)\left(\frac{-v}{t}\right)$, where $\frac{-v}{t}$ is the horizontal distance between the initial and final positions of the feature. Altogether, $v=-\frac{\Gamma(m_2)-\Gamma(m_1)}{m_2-m_1}$. For $m_1=0, m_2=-1$ this reduces to $v_E=\Gamma(0)$.

This expression has a nice interpretation in terms of the growth rate $\Gamma(m)$. It is the slope a chord drawn on the growth rate function between points corresponding to the two slopes, as in Fig. 35.

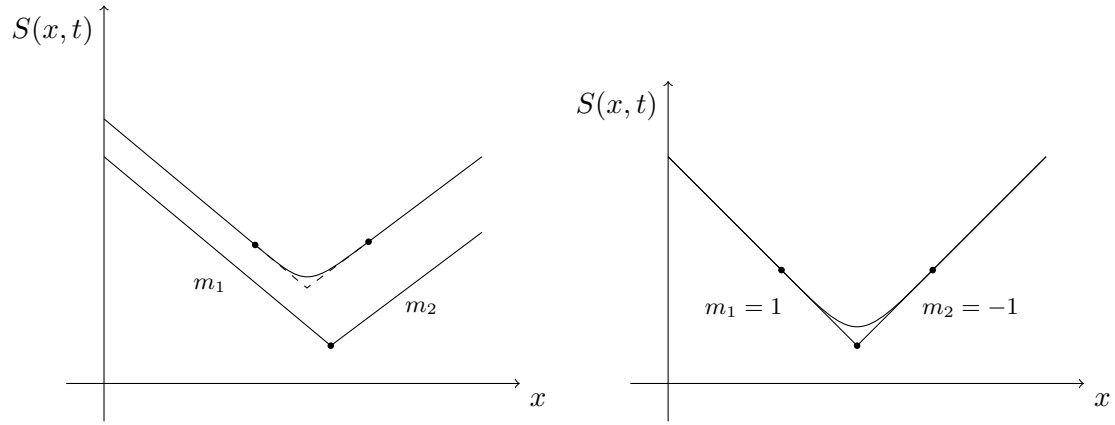

Figure 36: **Velocity of tangent points.** On the left, the left and right points move at $v = -\Gamma'(m_1)$ and $v = -\Gamma'(m_2)$ respectively. On the right these two speeds are the butterfly speeds, $-\Gamma'(\pm 1) = \mp v_B$. Since both slopes are extremal, the linear sections do not grow.

4.4.2 Butterfly Velocity

The same calculation for the velocity of a kink does not work if the kink is a local minimum, as these features do not remain sharp. This is because a perturbation to the kink will travel up faster than either side of the kink, filling in the kink with a smooth curve. There is a well defined point where this curve hits the linear section with slope m at a tangent. It is possible to calculate the speed at which this point moves (see Fig. 36). This tangent point is the limiting case of a cusp where the two slopes m_1 and m_2 approach each other. This turns the calculation of the chord slope into a calculation of the tangent to the entanglement growth curve at m, $\Gamma'(m)$.

Although it is not obvious, the butterfly is given by the speed of tangent points on maximal slope sections, $v_B = -\Gamma'(m_{\max})$. This connection holds in general, but is clearer in the random circuits considered here. Since $\Gamma(m)$ is convex, this is the largest slope attainable by a chord or tangent line, so v_B is the fastest possible speed for a feature on the entanglement function.

When the slope is near its largest possible value, so that the entropy increases at nearly every site, it is possible to isolate the behavior of the few down steps. Figure 37 shows such a configuration. Since gates can only act between a down step and an up step, these "particles" follow a deterministic behavior for a given architecture and control the entropy growth in the circuit.

Consider a series of consecutive gates with its first gate acting between sites i and $i+1$ and its last gate between j and $j+1$. Series of gates like this (called staircases) are discussed

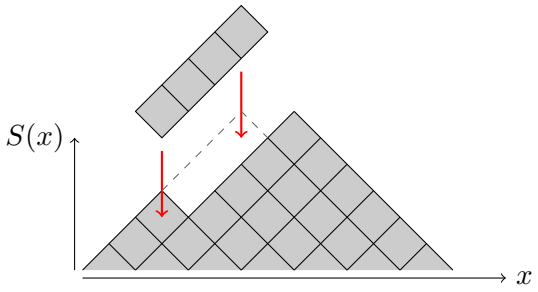

Figure 37: **Near-maximal slope**. The single down step acts like a particle in the system. The 4 consecutive gates have the effect of moving the particle 3 sites to the right.

in Sec. 5. If there are no down steps in this region, the gate has no effect. If there is a single down step somewhere between sites i and j, inclusive, that step gets moved to site $j+1$. More down steps will interact with each other but at the present we are only considering configurations with isolated down steps.

Now consider an operator with the last non-identity contribution at a site between i and j, inclusive. With probability 1, the series of gates in the last paragraph move the end of the operator to site $j+1$. This demonstrates that operator ends and down step "particles" have the same dynamics. The speed of the end of the operator is v_B. Since the particle has the same dynamics as the operator end, it also moves at velocity v_B.

All that remains is to show how the speed of the particle is connected to $\Gamma(m)$. If there are N down steps in a section of length L, then the slope is $m = 1 - \frac{2N}{L}$. Furthermore, in time t each down slope moves $v_B t$, so the entropy growth is $\Gamma(m) = \frac{2Nv_B}{L}$. Plugging in for m, this gives

$$\Gamma(m) = v_B \left(\frac{2N}{L} \right) = v_B(1-m), \tag{51}$$

$$\left. \frac{\partial \Gamma}{\partial m} \right|_{m=0} = -v_B. \tag{52}$$

The identification of particle dynamics with operator end dynamics assumed that the particles are far enough from each other that they do not interfere with each other's dynamics. This is valid in the large m limit as long as the particles are not attracted to each other. They could only affect each other if they are closer than a single series of gates. Then the early gates move the first particle until it is touching the second. At this point there are 2 down steps in a row and the next gate has no effect. The following gate lands on a local minimum and is able to move the second down slope to the end of the series of gates.

An equivalent description of these dynamics is that when a series of gates falls over two particles, the first one moves to just behind the second one, and then the second one moves to the end of the series of gates. Therefore the second one moves the distance it would without the first, while the first moves less than it would have if the second had not gotten in its way. This means that the particles are uncorrelated when far apart, and repel when they get too close. Therefore the uncorrelated assumption is correct.

The interpretation of v_B as the maximal slope of $\Gamma(m)$ does not imply any upper bound on v_B, but it does imply $v_B \geq v_E$. Recall that $\Gamma(m)$ is convex. Then the minimal slope possible at $m = -1$ is obtained by a straight line with slope

$$v_{B,\min} = \frac{\Gamma(0) - \Gamma(-1)}{0 - (-1)} = \Gamma(0), \tag{53}$$

which is just the entanglement velocity.

We have assumed so far that $\Gamma(m)$ is symmetric. There is no constraint that this must be true, however. When it is asymmetric, the two butterfly velocities are $v_{B+} = \Gamma'(m_{\max})$ and $v_{B-} = \Gamma'(m_{\min})$, where in our systems $m_{\max} = 1$, $m_{\min} = -1$. However since v_E is defined as $\Gamma(0)$, there is no way to define an asymmetric v_E. We discuss asymmetric systems in the next section.

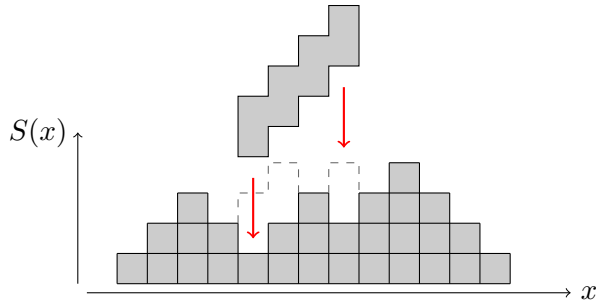

Figure 38: **Staircase circuit architecture,** in which the gate at site x is always followed by ones at sites $x+1, x+2$, and $x+3$ making this a 4-stair. Note that not all gates are productive, only the ones that fall on sites that are local minima when they fall.

5 Dynamics in Asymmetric Staircase Circuits

We can use the tools from the previous section to analyze the asymmetry of information speeds in various quantum circuits. First, though, we need to build an asymmetric circuit. There are a few possibilities. The circuits could have 3-site unitary operators, with the gates chosen so that their dynamics are asymmetric. An example gate would be the 3-site swap discussed earlier. Alternatively, we could change the probabilities of gates falling on sites so that if a gate falls across bond x, the next gate is more likely to fall across bond $x+r$ for small r. This would lead to correlations in gate locations.

The architecture considered in this section will be a limiting case of correlated gates, suggested in [21]. We will again use the solvable large-q limit and gates will fall in sets of n. Each n gates will fall consecutively across bonds $x, x+1 \ldots x+n-1$. The first site of each staircase is chosen randomly. These sets of gates are called staircases, and when the bond location is increasing it is called a right staircase. If there are n gates in a staircase it is called an n-stair. They are called staircases because using the surface growth picture in Fig. 27 they look like steps, as in Fig. 38. In the picture in which the entropy has slope up or down at each site, as in Fig. 29, n-stair look like $n \times 1$ rectangles tilted 45°, as in Fig. 39. Since each staircase consists of multiple gates, there is an ambiguity in whether the rate γ defines the number of gates per bond per time step or the number of staircases. We will use the convention that it is the number of gates per bond per unit time, so that γ/n is the rate of staircases.

Ref. [20] uses configurations like this, but in equal proportion with left staircases. The combination of left and right staircases allows for arbitrarily small values of v_E/v_B. In this section we show that including only right (or left) staircases leads to two distinct butterfly

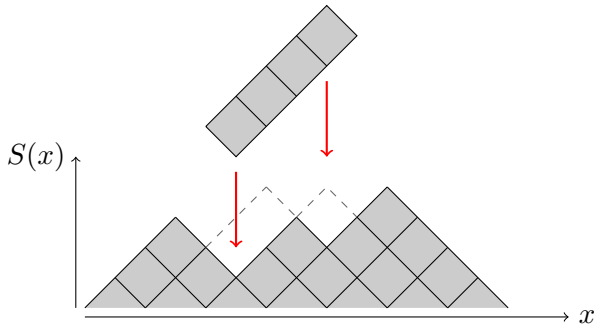

Figure 39: **Another picture of the staircase architecture.** Note that this picture results in the same final state as in Fig. 38.

velocities. One can be made arbitrarily large compared to v_E, while the other approaches v_E.

5.1 Behavior Ignoring Correlations

The dynamics of the staircase circuits become solvable after ignoring correlations in up and down steps in $S(x)$ and ignoring second and higher order derivatives in $S(x,t)$. Combining these two assumptions, we arrive at uncorrelated entropy environments, which may be described only by their slope, $\langle s_i \rangle = m$. From vanishing correlation we have $\langle s_i s_{i+1} \rangle = \langle s_i \rangle \langle s_{i+1} \rangle = m^2$. The correlation of the true steady state is small, and may be treated as a perturbation. This is done in Sec. 5.2.

5.1.1 Small Stairs

The smallest stairs are 1-stairs, which are just individual gates. As in the derivation of the random growth rate in Eq. 49, the probability of a gate raising the entanglement at a cut is $\frac{1-m^2}{4}$, and each gate raises the entanglement by 2, so

$$\Gamma_1(m) = \gamma \frac{1-m^2}{2}, \tag{54}$$

where the subscript represents the size of the staircases. The random architecture does not generate any correlation, so this growth rate is exact for a long enough system with constant slope.

2-stairs consist of one gate acting at cut x and one at cut $x+1$. The entropy production of these gates is affected by the slope between the two cuts and the slopes on either side. There are 8 possible configurations of those three slopes, but only 4 result in entropy growth, as

Initial and Final Configuration	Probability	Productivity
$d\,u\,d \to u\,d\,d$	$\frac{1-m}{2}\frac{1+m}{2}\frac{1-m}{2}$	2
$d\,u\,u \to u\,u\,d$	$\frac{1-m}{2}\frac{1+m}{2}\frac{1+m}{2}$	4
$d\,d\,u \to d\,u\,d$	$\frac{1-m}{2}\frac{1-m}{2}\frac{1+m}{2}$	2
$u\,d\,u \to u\,u\,d$	$\frac{1+m}{2}\frac{1-m}{2}\frac{1+m}{2}$	2

Table 1: The four configurations that result in entropy growth for 2-stairs, the relative proportions of the initial states assuming an uncorrelated entropy distribution, and the growth in entropy generated by a 2-stair falling on that configuration. The four configurations that do not result in entropy growth are $u\,u\,u, d\,d\,d, u\,d\,d$, and $u\,u\,d$.

shown in table 1. Weighting each configuration by its probability and the entropy generated, and then multiplying by the staircase rate $\frac{\gamma}{2}$, the growth rate is approximately

$$\begin{aligned}\Gamma_2(m) &= \frac{\gamma}{2}4\frac{1+m^2}{4}\frac{1+m}{2} + \frac{\gamma}{2}2\frac{1-m^2}{4}\left(\frac{1-m}{2}+\frac{1-m}{2}+\frac{1+m}{2}\right)\\ &= \frac{\gamma}{2}\frac{1-m^2}{2}\left(1+m+\frac{3-m}{2}\right)\\ &= \gamma\frac{1-m^2}{2}\frac{5+m}{4},\end{aligned} \tag{55}$$

where γ is the rate of gates, so the rate of 2-stairs is $\frac{\gamma}{2}$. Although there will be some correlation created by the 2-stair architecture, it will be small so we ignore it for now.

5.1.2 Larger Stairs

Now that we have calculated $\Gamma_1(m)$ and $\Gamma_2(m)$, direct calculation of $\Gamma_n(m)$ for larger n will become increasingly unwieldy. Even for $n = 3$ there will be 4 slopes to consider and 16 possible initial configurations, with 11 generating some entanglement. The probabilities will be quadratic in m. Luckily, there is a better way.

We can determine the growth rate for arbitrary length stairs through a recursive relationship. Consider a staircase made of n gates. Like in the $n = 2$ case, its growth rate will be proportional to the stair rate $\frac{\gamma}{n}$ multiplied by the entanglement generated by each staircase, so we can write

$$\Gamma_n(m) = \frac{\gamma}{n}R_n(m), \tag{56}$$

where $R_n(m)$ is the average entropy production of an n-stair. To find an equation for $R_n(m)$, note that the first $n-1$ gates have the same entanglement production as the $(n-1)$-stair. All final states end in a down slope, so the nth gate will produce another 2 units of entropy

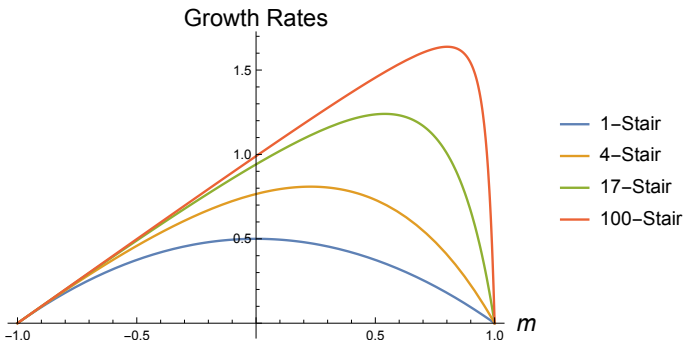

Figure 40: **Approximate growth rates** for 1-, 4-, 17-, and 100-stairs as a function of slope m. As stair length increases, the growth rate asymptotes to the function $\frac{\partial S}{\partial t} = m+1$.

if the last step is up. However, if all $n+1$ initial steps are up no entanglement is generated.

This behavior is captured by the recursive formula

$$R_n(m) = R_{n-1}(m) + 2\frac{1+m}{2} - 2\left(\frac{1+m}{2}\right)^{n+1}, \tag{57}$$

along with the base case $R_0(m) = 0$. The solution is

$$R_n(m) = \frac{1+m}{1-m}\left((1+m)\left[\left(\frac{1+m^2}{2}\right)^n - 1\right] + n(1-m)\right). \tag{58}$$

We can check that this reproduces the familiar cases

$$R_1(m) = \frac{1-m^2}{2}, \tag{59}$$

$$R_2(m) = \frac{1-m^2}{2}\frac{5+m}{2}. \tag{60}$$

Figure 40 contains a graph of some growth rates with γ set to unity, so that $\Gamma(m) = \frac{1}{n}R_n(m)$, as a function of m.

5.1.3 Infinite Stairs

Note that with increasing stair length the last term takes over and the growth rate asymptotes to the function $\Gamma_\infty = \gamma(m+1)$. This is the most asymmetric behavior, with the ratio v_{B+}/v_{B-} increasing without limit. Furthermore v_{B-}/v_E approaches 1 from above, and v_{B+}/v_E increases without limit.

There are two ways to reason about this behavior. The first, which matches the order of limits of the current reasoning, is to start with infinite spatial support, which we have

been assuming, and take the stair length $n \to \infty$. Alternatively, start with a finite spin chain with length L and periodic boundary conditions, and set the stair length to ∞. Now a single gate acts on site $1, 2, 3 \ldots L$, before wrapping around to act at site 1 again.

In both cases the rates $\Gamma_\infty(0) = \gamma$ and $\Gamma_\infty(1) = 2\gamma$ can be understood easily. 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. At near-maximal slope nearly all slopes are u, 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.

Interestingly, this argument applies regardless of correlations, so that while the growth rates in this section are only approximate for finite $n > 1$, the approximation actually becomes better at very large n. The only argument might be that if a long stair falls across multiple down steps, it will not generate the maximal entanglement and the sharp peak in $\Gamma_\infty = \gamma(m+1)$ might be rounded out. However, the earlier argument that these particles are uncorrelated means that the in the large m limit the particles do not interact.

5.1.4 Entanglement and Butterfly Velocities

The entanglement velocity can be immediately read from the entanglement growth rate. Again ignoring the correlations, it is just

$$v_E = \Gamma_n(0) = \frac{\gamma}{n}\left(n + \frac{1}{2^n} - 1\right), \tag{61}$$

which starts at .5 at $m = 0$ and asymptotes to 1, as it should. Since there is no way to define asymmetric entanglement velocities this is as far as we will take the analysis of v_E.

In these models we can extract asymmetric butterfly velocities by looking at the slopes of $\Gamma_n(m)$ at $m = \pm 1$. The left-moving butterfly velocity, v_{B-} stays constant with respect to n, while the right moving v_{B+} increases without bound. The interpretation of these behaviors comes from Sec. 4.4, using the speed of operator edges. If a long stair falls across an operator's left edge, it can only move the edge one site, since by the time it moves the edge the next gate already acts farther right. However whenever a staircase hits a right edge it moves the edge to the end of the staircase, which in the large n limit is arbitrarily far.

Since the operator edge dynamics are equivalent to down step "particle" dynamics, we

can discuss the latter, which is more directly related to $\Gamma(m)$. For $n \to \infty$, in a system of length L, the maximal growth rate occurs when the slope is near-maximal, with only a single down step. This corresponds to a slope of $\frac{L-2}{L}$. Then with only a single local minimum in the entropy function, the local minimum moves one site to the right with each gate. Every gate that acts raises the entropy, resulting in an entropy gain of 2 per gate. Of course for a slope of 1, there is still no entropy generation.

If there are 2 down steps, for a slope of $\frac{L-4}{L}$, the entropy generation is almost the same. One local minimum still moves to the right with the leading edge of the staircase. However, once per staircase (once every L gates), one down step is next to the other down step. This results in no entropy growth for that step, for an average entropy gain of $2\frac{L-1}{L}$.

This pattern continues as the slope decreases. With ρL down steps, the slope is $\frac{L-2\rho L}{L}$ and the average growth rate is $2\frac{L-\rho L}{L} = m+1$. At $m = -1$ there is no growth, as expected. At $m = 0$, the growth rate is 1. This matches the discussion in Sec. 5.1.2.

When we combine left and right stairs to recover symmetry, in the limit of large n, the entanglement growth rate becomes $\Gamma(m) \to 1$, with $\Gamma(1) = \Gamma(-1) = 0$. Since the slopes at $m = \pm 1$ become infinite, these slopes have arbitrarily high v_B/v_E. The lack of bound on v_B/v_E was shown using these symmetric staircase circuits in Ref. [20].

5.2 Steady States with Correlation

The first correction to make to the previous section is to incorporate a non-zero correlation in the steady state. For example, the system could be slightly ferromagnetic or antiferromagnetic, meaning that after an up slope the next slope is slightly more or less likely to be up, respectively. These terms are used in analogy with the Ising model; there is no actual magnetism in the circuit. An antiferromagnetic state has more local minima, so more gates will generate entanglement leading to a higher growth rate.

The correlation affects the growth rate for 2-stairs through the probabilities in table 1. In a system with non-zero correlation they depend on both the slope and correlation, as shown below. We then show how to find the steady state in Secs. 5.2.2 and 5.2.3.

When looking at a single site, the probabilities of u or d are simple: $p(u) = \frac{1+m}{2}, p(d) = \frac{1-m}{2}$. Previously we used the naïve assumption that, when looking at 2 sites, $p(uu) = p(u)p(u)$. However this is only true without correlations. The nearest neighbor correlation $C_{i,i+1}$ completely captures this dependence. Since our system is translationally invariant, the correlation cannot depend on position so $C_{i,i+1} = C_1$.

Ignoring finite-size effects, we have

$$\begin{aligned}
\langle s_i s_{i+1}\rangle &= p(uu) + p(dd) - p(ud) - p(du),\\
m &= p(uu) - p(dd)\\
C_1 &= \langle s_i s_{i+1}\rangle - \langle s_i\rangle \langle s_{i+1}\rangle = \langle s_i s_{i+1}\rangle - m^2,\\
&= p(uu) + p(dd) - p(ud) - p(du) - (p(uu))^2 - (p(dd))^2 + 2p(uu)p(dd). \qquad (62)
\end{aligned}$$

Along with the constraint $p(ud) = p(du)$ due to the fact that their must be an even number of domain walls, this gives 4 constraints on 4 probabilities, so the 2-site probabilities are completely constrained. After more manipulation we arrive at

$$p(ud) = p(du) = \frac{1 - m^2 - C_1}{4}, \qquad (63)$$

$$p(uu) = \frac{(1+m)^2 + C_1}{4}, \qquad (64)$$

$$p(dd) = \frac{(1-m)^2 + C_1}{4}. \qquad (65)$$

These reduce to our old expressions for $C_1 \to 0$.

If there were correlation in the random circuit then these new probabilities would affect the growth rate. As we will show in Sec. 5.2, there is no correlation in that architecture. However, we can use C_1 to calculate the first order correction to the 2-stair growth rate.

Referring back to table 1, the only quantities that have changed are the probabilities. Although we have not explicitly calculated the 3-site probabilities, we can get the C_1 correction using Bayes' rule. For example,

$$\begin{aligned}
p(ddu) &= p(dd) * p(u|d) = p(dd)\frac{p(du)}{p(d)}\\
&= \frac{(1-m)^2 + C_1}{4}\frac{1 - m^2 - C_1}{4}\frac{2}{1-m}\\
&= \frac{\left((1-m)^2 + C_1\right)\left(1 - m^2 - C_1\right)}{8(1-m)}, \qquad (66)
\end{aligned}$$

where $p(u|d)$ is the probability that $s_i = u$ given $s_{i-1} = d$. This value is the same as it would be if we had started from the right side and calculated $p(ddu) = p(du) * p(d;d)$, where $p(d;d)$ is the probability that $s_i = d$ given $s_{i+1} = d$.

After calculating the other relevant probabilities and going through the same process

used to generate Eq. 55, we find

$$\Gamma_2^{(1)}(m) = \frac{\gamma}{2}\frac{\left(1-m^2-C_1\right)\left((1+m)(5+m)\right)+C_1}{4(1+m)}, \tag{67}$$

which reduces to Eq. 55. The superscript notes that this is the C_1 correction to the growth rate. The n-stair growth rate will have corrections from correlations up to C_n. The analysis of these corrections will proceed similar to the analysis for C_1.

5.2.1 Ergodicity and Stability

To describe the steady state and extract a correlation we use the language of Markov processes. A Markov process is one in which the future state depends only on the current state, not the past. Label the states s_i and define $p_{i,t}$ as the probability that the system is in state s_i at time t. When there is a constant probability S_{ij} of transitioning from state j to state i it is possible to write the transition matrix S such that $p_{i,t+1} = S_{ij}p_{j,t}$. Since the product of the transition matrix and a probability vector gives the probabilities at the next time step, the transition matrix for t time steps is just S^t. Under certain conditions the multi-step transition matrix approaches a constant matrix with all columns equal to the same vector p^*,

$$\lim_{t\to\infty} S^t = S^* = \begin{bmatrix} \vdots & \vdots & & \vdots \\ p^* & p^* & \cdots & p^* \\ \vdots & \vdots & & \vdots \end{bmatrix}. \tag{68}$$

Then the probability after a long time is the vector p^* for any initial state.

The transition matrix can also be written as $S = 1 + T$. The columns of T must sum to 0 to preserve probabilities. From $S^*p^* = p^*$, it must be true that $Tp^* = 0$. This definition provides an easier route to finding p^*.

5.2.2 Stable State in Random Architecture

This analysis can be used to find the correlations present in the steady states of staircase architectures. Consider the 1-stair circuit, and enumerate 2-site (3-cut) states by the slope at the 2 sites: $s_1 = d\,d,\ s_2 = d\,u$, etc. Since at every time step there is an equal probability of a gate falling at any site, the transition matrix is the matrix product of single-cut transition

matrices $S_N = \prod_N S_1 \otimes$, where the 2-site transition matrix is

$$S_1 = \begin{bmatrix} 1 - \frac{1+m}{2}\Gamma & 0 & \frac{1-m}{2}\Gamma & 0 \\ \frac{1+m}{2}\Gamma & 1-\Gamma & 0 & \frac{1-m}{2}\Gamma \\ 0 & \Gamma & 1-\Gamma & 0 \\ 0 & 0 & \frac{1+m}{2}\Gamma & 1 - \frac{1-m}{2}\Gamma \end{bmatrix}. \tag{69}$$

The m dependence comes from the possibility of a gate acting on the left or right cut, which depends on the probability of the next slope being up or down. The equilibrium state is

$$v^* = \begin{pmatrix} \frac{(1-m)^2}{4} \\ \frac{1+m}{2}\frac{1-m}{2} \\ \frac{1-m}{2}\frac{1+m}{2} \\ \frac{(1+m)^2}{4} \end{pmatrix} = \begin{pmatrix} \frac{1-m}{2} \\ \frac{1+m}{2} \end{pmatrix} \otimes \begin{pmatrix} \frac{1-m}{2} \\ \frac{1+m}{2} \end{pmatrix}, \tag{70}$$

which is uncorrelated, showing that the assumption of lack of correlation (used in equation 69) is consistent. Markov's theorem states that if all states are reachable from all other states then the system is ergodic. An ergodic system contains only one equilibrium state, so the uncorrelated state is the unique equilibrium state.

5.2.3 Stable States in Larger Staircases

Since the previous analysis relied on the assumption of lack of correlation and then showed its consistence, that will not work for larger stairs, which in fact do have a non-zero correlation. Instead of trying to calculate the correlation on an infinite system, which is complicated by having to know the previous and subsequent slopes, which rely on the correlation, we can find the steady state in small finite systems and extrapolate any measured quantities from these to large systems.

We do this by enumerating all possible states for our system of size L and slope m and building a transition S matrix between states. Since gates have an equal probability of landing on each set of n sites, the transition matrix element S_{ij} will be $1/L$ times the number of ways state i can be obtained after state j using a single n-stair. Then the eigenvector of this matrix with eigenvalue 1 will be the steady state.

The hard part is constructing the transition matrices. For small systems it is possible to explicitly enumerate all states and calculate the probabilities of transitioning to other states from these. Luckily it is possible to calculate transition matrices for larger systems

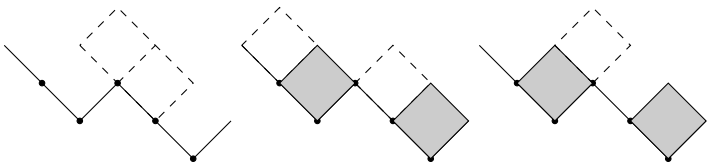

Figure 41: **Three possible transitions** from the state *ddu*. Two different gate configurations result in the state *dud*. Recall that the system is periodic.

recursively from the small case. This step is made easier by calculating the matrices for all slopes, not just a specific one. we can then take only the subspace of slope m. We will go through this procedure for 2-stairs.

The case for $L = 2$ is almost trivial. There are 4 states: *dd*, *du*, *ud*, *uu*. Recall that the systems are periodic, so for example the first state would look like *dddddd*... if viewed over multiple periods. The first and last state have no local minima, so gates cannot change the states and these are stationary. For the middle two states, label the cut before the listed sites cut 0 and the cut between the listed sites cut 1. Then for the state *du* cut 0 is a local maximum and cut 1 is a local minimum. If the first gate in the two stair lands on cut 1, the state will become *ud*, and will finish in state *du* after the second gate. Otherwise the first gate, on cut 0, will have no effect and the final state will be *ud*, after the second gate falls on the local minimum. State *ud* has a similar analysis. The result is that from either $m = 0$ state, the probability of transition to the either $m = 0$ state is 1/2.

The 2-site transition matrix for 2-stairs is then

$$T_{2,2} = \frac{1}{2}\begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 2 \end{bmatrix} \tag{71}$$

The steady state for $m = 1$ is *uu*, the steady state for $m = -1$ is *dd*, and the steady state for $m = 0$ is $\frac{1}{2}(ud + du)$. The result that steady states for extremal slopes are all up or all down is general.

For the 3-site transition matrix we will show the calculation of 2 columns and then display the full matrix. Once again the state *ddd* cannot transition, so the first column is $(1, 0, 0 \ldots)$. The second state in our basis is *ddu*, with $m = -1/3$. If the two gates fall on cuts 0 and 1, there is no change. If they land on cuts 1 and 2 or on sites 2 and 3 the final state is *dud*. These three possibilities are shown in Fig. 41. Both final states have

$m = -1/3$, since all final states have to have the same slope as the initial state. This is what allows us to decompose the full transition matrix into subspaces.

The full 3-site, 2-stair transition matrix is

$$T_{2,3} = \frac{1}{3}\begin{bmatrix} 3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 2 & 0 & 0 & 0 \\ 0 & 2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 2 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 3 \end{bmatrix} \tag{72}$$

There are two keys to building the larger transition matrices. First, instead of the transition matrix for gates at all 3 cuts, we can look at only transitions due to gates on cuts 1 and 2,

$$T_{2,3}^{(1)} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \tag{73}$$

The other 2 components of the transition matrix will be this matrix, with states permuted by 1 or two spaces. This permutation can be achieved using the cyclic swap matrix S_3 from Sec. 3, so that $T_{2,3}^{(2)} = S_3\, T_{2,3}^{(1)}\, S_3^{-1}$, etc.

The usefulness of this decomposition comes when looking at the analogous matrix for 4 sites. The matrix is large and its calculation is tedious, but the second key is realizing that it is just $T_{2,3}^{(1)} \otimes I_2$. This is because gates on the cuts 1 and 2 only affect the first 3 slopes. Since the computational basis is arranged such that *ddud* is immediately followed by *dduu*, each 2×2 block of a larger matrix is the identity times the corresponding element of the smaller matrix.

We can now build the full $T_{2,4}$ from $T_{2,4}^{(1)}$ and its permutations. From subspaces of

constant slope we can find the steady states, for example the $m = 0$ steady state

$$s_{m=0} = \frac{1}{5}dduu + \frac{1}{10}dudu + \frac{1}{5}duud + \frac{1}{5}uddu + \frac{1}{10}udud + \frac{1}{5}uudd. \tag{74}$$

Given the steady state it is possible to calculate the steady state correlation by averaging the correlations in each component of the steady state, weighted by their coefficients. The only hitch here is that the obvious formula for correlation, $C_1 = \langle s_i s_{i+1} \rangle - \langle s_i \rangle^2$ does not work for finite systems if the expectations are interpreted as averages. To see why, consider a 6-site state with slope 0. If the first 3 slopes are up, the last 3 have to be down to maintain the state. This introduces a negative bias into the overall correlation, equal to $1/(L-1)$, where L is the length of the system.

It is possible to correct for this bias by adding back the bias, so the true correlation of, for example, the state $uududu$ is

$$\begin{aligned} C_1 &= \langle s_i s_{i+1} \rangle - \langle s \rangle^2 + \frac{1}{L-1} \\ &= \frac{1-1-1-1-1+1}{6} - \frac{2^2}{6^2} + \frac{1}{6-1} \\ &= -\frac{1}{3} - \frac{1}{9} + \frac{1}{5} \\ &= -\frac{11}{45}. \end{aligned} \tag{75}$$

After repeating this calculation for other states we can find the steady states in various size systems. For $m = 0$, these are

$$\begin{aligned} L = 4, \qquad & C_1 = 0.1333 \\ L = 6, \qquad & C_1 = 0.1103 \\ L = 8, \qquad & C_1 = 0.1171 \\ L = 10, \qquad & C_1 = 0.1163. \end{aligned} \tag{76}$$

To actually show that the correlation asymptotes to a value we would have to take these calculations out to very large L. However, we can see now that the correlation will be positive and ≈ 0.11 or 0.12.

For nonzero m we have fewer choices of L. For example, for $m = \pm\frac{1}{2}$, L must be a

multiple of 4. The steady state correlations for $m = \frac{1}{2}$ are

$$\begin{aligned} L = 4, \qquad & C_1 = 0.0833 \\ L = 8, \qquad & C_1 = 0.0760 \\ L = 12, \qquad & C_1 = 0.0703, \end{aligned} \tag{77}$$

while for $m = -\frac{1}{2}$,

$$\begin{aligned} L = 4, \qquad & C_1 = 0.0833 \\ L = 8, \qquad & C_1 = 0.1151 \\ L = 12, \qquad & C_1 = 0.1017. \end{aligned} \tag{78}$$

The relevant information here is that both correlations are smaller than at $m = 0$, but the correlation is significantly larger in the negative slope system. We can directly simulate the circuit dynamics to test these results.

5.3 Numerical Simulation

To simulate the entropy growth in infinite spin chains with non-zero slope, we use finite chains with periodic boundary conditions, where the entropy at one end of the chain is equal to the entropy at the other with an offset. The spin chains are initialized with an uncorrelated entanglement function with each slope chosen to be up or down with $p(u) = (1 + m)/2$. This results in a sample slope that is near the given slope. Then, for each time step, a random gate location is chosen. The entanglement across that cut, and the next $n - 1$, is calculated using the update rule from Eq. 48.

Since larger gates will generate non-zero correlation, the initial $S(x)$ is not a valid steady state. However the dynamics should bring the system into a steady state, so growth rates are calculated for only the latter part of the simulation. Fig. 45 shows that there is a time after which the correlations reach this steady state.

5.3.1 Measuring Growth Rates

To calculate the growth rates we start in the initial uncorrelated state and evolve with 400 n-stairs, then average the growth rate over the next 1600 stairs. This procedure is repeated 100 times to calculate statistics on the mean. Growth rates for stairs of various lengths are shown in Fig. 42. As length increases, the growth rate follows the same pattern

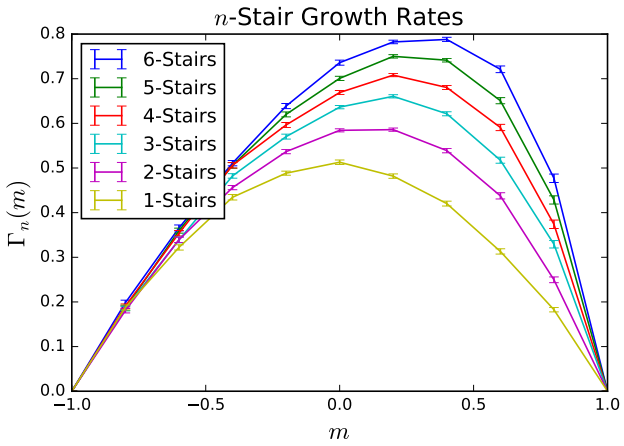

Figure 42: **Growth rates for different length stairs.** All growth rates were calculated using a 100-site spin chain with offset periodic boundary conditions. Rates were calculated from the application of 2,000 gates total, or 20 per site, averaged over the last 80% of the gates in order to build up correlations, then averaged over 100 trials.

as predicted, increasing with the maximum moving right.

For 1-stairs (the random architecture), the measured growth rate is not significantly different than predicted, seen in Fig. 43. This implies that the lack of correlation is exact, and that the finite size effects are small enough at $L = 100$ to not significantly affect $\Gamma_1(m)$.

Stairs of length greater than 1 do however generate correlations. Figure 44 shows the growth rates for 2- and 6-stairs. The notable difference in growth rate here is due to correlations created by the staircases. Although the effect is larger for larger stairs, this pattern must reverse eventually since the correlations have no effect on infinite stairs.

5.3.2 Measuring Correlations

In Sec. 5.2 we found the correlation in the steady state is larger at $m = 0$ than at $m = \pm.5$, and larger at $m = -.5$ than at $m = .5$. We can check this numerically, by initializing a state with 0 correlation, evolving it, and measuring the correlation at later times. In general the correlation grows to the steady state value and then levels off. Fig. 45 shows the evolution of the correlation in 1- 2- 3- and 10-stair circuits, averaged over 40 trials. As expected, the random circuit produces no significant correlation. Larger stairs do, and this should be taken into account when calculating growth rates.

The correlation also depends on the slope of the entanglement curve. This must be true because $C_1 = 0$ for $m = \pm 1$, and the correlation should vary smoothly with slope. Figure 46 shows this dependence. All curves were created by first evolving by 4 gates

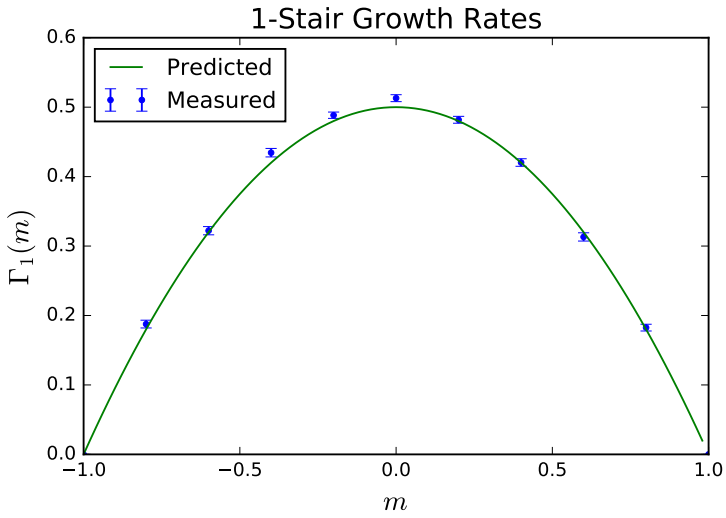

Figure 43: **Measured and analytic growth rates** for the 1-stair architecture. Note that the measured rate is not different than predicted.

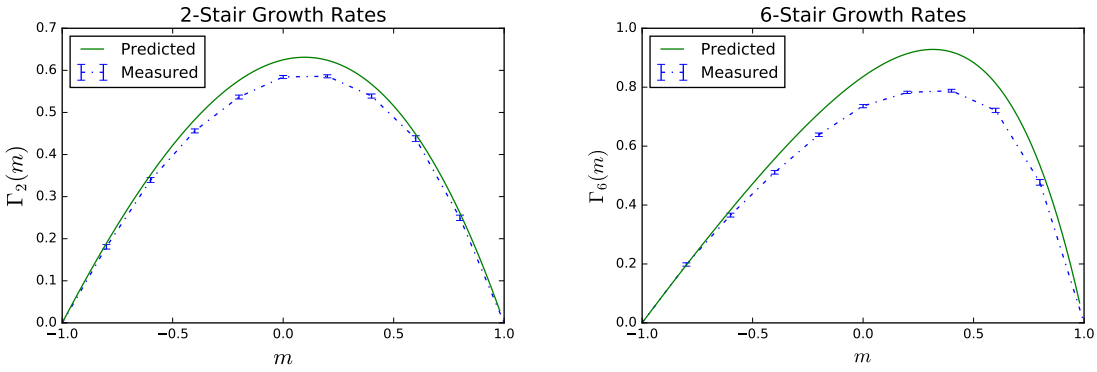

Figure 44: **Measured and predicted growth rates** for the 2- and 6-stair architecture. The measured rate is now significantly lower than predicted, implying that the correlations built up by the stairs act to lower their entropy production.

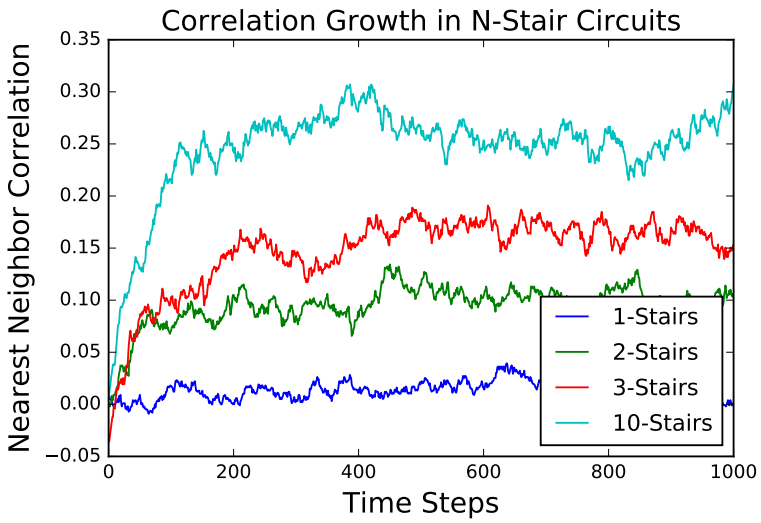

Figure 45: **Correlation Growth for 0 slope.** Each curve corresponds to an initially 0 entanglement curve evolved with the given circuit architecture. Time in this plot is normalized so that one time step corresponds to one N-stair. This way, all curves reach steady-state entanglement by about 400 steps.

per site to remove the initial conditions, and then sampling over 1600 time steps, for 100 separate trials.

The 3-stair circuit indeed generates more correlation than the 1-stair circuit. The solid curve is a simple quadratic function of the slope, which shows that the correlation function is asymmetric, and that negative slope states have slightly more correlation in the steady state.

We can insert this empirical correlation into the growth rate correlation to get a closer estimate of the true growth rate. Fig. 47 compares this procedure to the earlier calculation. The systematic error appears to now be in the other direction, overestimating the growth rate of these circuits.

The nearest neighbor correlation is not the only correlation in the system. A given step can also be correlated with a step ℓ sites away from it. Again because the system is translationally invariant this correlation is dependent on distance, and can be called C_ℓ. Since the n-stair architecture is itself highly correlated to a scale of n, it would be unsurprising if C_ℓ was significantly for $\ell \geq n$.

Fig. 48 shows C_ℓ for various circuit architectures. C_ℓ is positive for $\ell \geq n$, dips negative for the next few correlation lengths, and approaches 0 for long distance correlation. The negative dip may be due to finite size effects which constrain the sum of all correlations. This suggests that it might be possible to have a circuit architecture with a mix of staircases

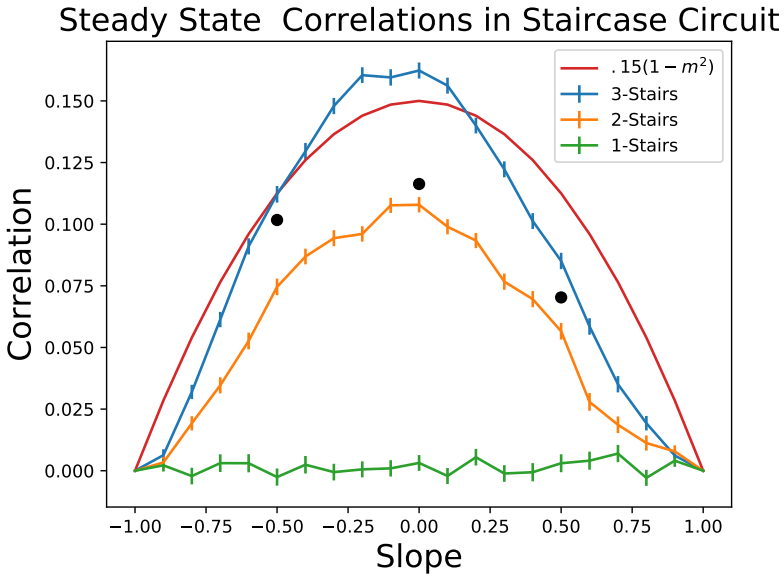

Figure 46: **Correlations created by 1-, 2-, and 3-stair circuits.** The correlation in 1-stair circuits is consistent with 0. Larger stairs generate significant correlation. The analytic function, $.15(1-m^2)$, acts as a guide to show that the correlations are asymmetric with respect to m. The black dots are the correlations calculated in Sec. 5.2.3. Although they are not within the error, they still have the right shape.

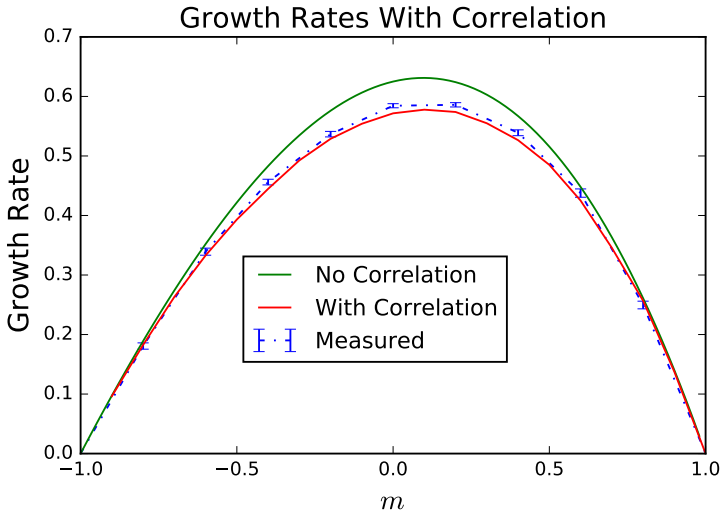

Figure 47: **Growth rates with and without correcting for non-zero correlation** in the steady state. Including the correlation removes a significant portion of the error.

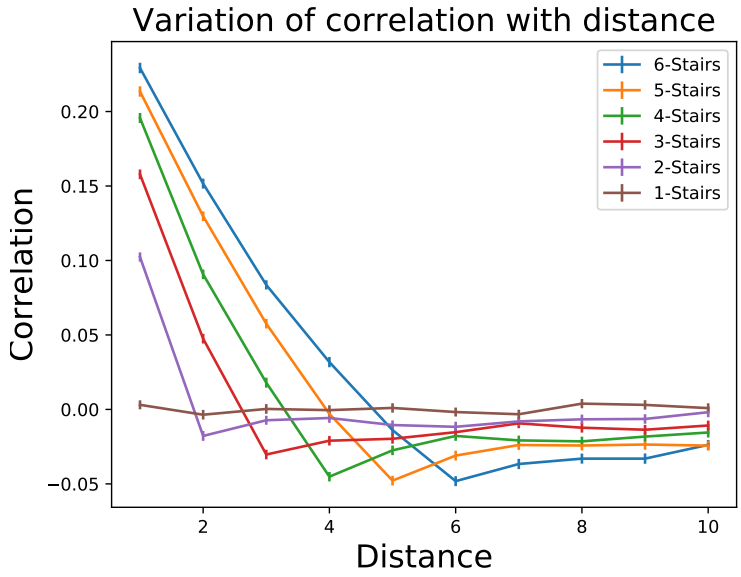

Figure 48: **Variable-distance correlation for various architectures.** Note that the correlation is generally positive for $\ell < n$, until $n = 6$.

of different lengths, such that correlations introduced by different length stairs cancel each other out.

Instead of looking at the variation of correlation with respect to distance, we can look at variation with respect to stair length, as in Fig. 49. This does not include new information, but presents the information from Fig. 48 differently. In particular, it appears that the correlation begins to level off for large stairs. Once again, the negative correlation for $n \approx \ell$ is assumed to be due to finite-size effects.

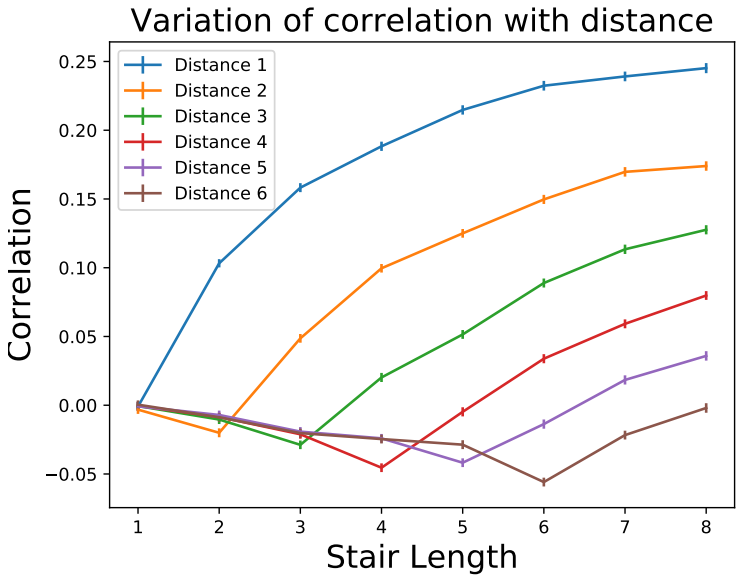

Figure 49: **Dependence of correlation on stair length** for different ℓ. Note that the correlation appears to level off for long stairs.

6 Conclusion

6.1 Summary

In this thesis, we analyzed the operator and entanglement dynamics in two asymmetric systems, a system with a time-independent Hamiltonian and a quantum circuit. In the Hamiltonian system we used the OTOC to diagnose operator spreading and move toward finding a butterfly velocity. Studying operator spreading worked particularly well in the Hamiltonian system because the OTOC has a simple form for a spin chain with spin-$\frac{1}{2}$ sites. In the circuit we used the entanglement dynamics to assess the entanglement and butterfly velocities. Entanglement dynamics were well suited to the circuit because of a simplification in the large q limit, which allowed exact solvability after averaging over circuits. We provided an argument that v_B as calculated using entanglement dynamics or operator spreading is the same in the circuit models.

We constructed the Hamiltonian out of local 3-site terms, with the individual terms asymmetric. We explored the dynamics of the Hamiltonian before explicitly calculating velocity dependent Lyapunov exponents $\lambda(v)$. Since v_B is defined by $\lambda(v_B) = 0$, we would expect $\lambda(v)$ to have a zero. However, due to the finite size of our system, the $\lambda(v)$ we measured did not have a zero. Despite this difficulty, $\lambda(v)$ was greater for right moving signals than for left moving signals for all velocities, implying that $v_{B+} > v_{B-}$ when these velocities are well defined.

The quantum circuits we considered are staircase circuits, where n gates fall sequentially on consecutive sites. We used the large-q limit to calculate entanglement growth rates of these circuits in the approximation that the staircases do not generate correlations in the entanglement $S(x)$. We also moved toward calculating the correlations in the steady state.

We then simulated these circuits to show that this approximation is incorrect, but does capture the gross shape of $S(x)$. We then measured the correlation of the simulated circuits, and showed that inserting this empirical correlation removes most of the error in growth rate. We also demonstrated that it is possible to analytically predict the correlation over different distances for different stair lengths, and that only a finite number of correlations affect the growth rate.

6.2 Outlook

For both systems, there are several improvements that can be made, both obvious and subtle. The simplest extension to the Hamiltonian model is to extend the chain to more sites. This is computationally difficult, but might reward the user with a close approximation to v_B.

Other extensions include different Hamiltonians. For example, instead of having a multisite Hamiltonian that is a chain of 3-site Hamiltonians, one could build it out of 4-site Hamiltonians. Then each individual term could be the Hamiltonian whose unitary evolution is the 4-site swap.

In a broader analysis, it could be worthwhile to search the space of possible Hamiltonians for the most asymmetric system. This thesis only studies the Hamiltonian that it does because the 3-site swap is an asymmetric unitary operator. However there could very well be a different 3-site Hamiltonian that, when chained together, gives a higher ratio of v_{B+}/v_{B-}. Searching the space of Hamiltonians would be difficult because the criterion for asymmetry is non-linear.

In the circuit system, the main improvement to be made is to better calculate the correlation, and to possibly come up with a closed form expression for C_ℓ for n-stairs. This thesis indicates that if this were completed the growth rated could be calculated exactly, because the main discrepancies between calculated and measured growth rates were explained by the measured correlation. Furthermore, this thesis did not include detailed discussion of the function $G(v)$, which encodes the shape of $S(x,t)$ as t grows [22].

Another topic discussed in Refs. [18, 21] is entanglement growth for higher dimensional systems. It might be interesting to introduce asymmetric, or direction-dependent, entanglement dynamics into these systems.

References

[1] Werner Israel. "Event Horizons in Static Vacuum Space-Times". In: *Phys. Rev.* 164 (5 1967), pp. 1776–1779. DOI: 10.1103/PhysRev.164.1776.

[2] Werner Israel. "Event horizons in static electrovac space-times". In: *Communications in Mathematical Physics* 8.3 (1968), pp. 245–260. ISSN: 1432-0916. DOI: 10.1007/BF01645859.

[3] B. Carter. "Axisymmetric Black Hole Has Only Two Degrees of Freedom". In: *Phys. Rev. Lett.* 26 (6 1971), pp. 331–333. DOI: 10.1103/PhysRevLett.26.331.

[4] Rahul Nandkishore and David A. Huse. "Many-Body Localization and Thermalization in Quantum Statistical Mechanics". In: *Annual Review of Condensed Matter Physics* 6.1 (2015), pp. 15–38. DOI: 10.1146/annurev-conmatphys-031214-014726.

[5] P. W. Anderson. "Absence of Diffusion in Certain Random Lattices". In: *Phys. Rev.* 109 (5 1958), pp. 1492–1505. DOI: 10.1103/PhysRev.109.1492.

[6] Yasuhiro Sekino and L Susskind. "Fast scramblers". In: *Journal of High Energy Physics* 2008.10 (2008), pp. 065–065. DOI: 10.1088/1126-6708/2008/10/065.

[7] Juan Maldacena, Stephen H. Shenker, and Douglas Stanford. "A bound on chaos". In: *Journal of High Energy Physics* 2016.8 (2016). DOI: 10.1007/jhep08(2016)106.

[8] J. M. Deutsch. "Quantum statistical mechanics in a closed system". In: *Phys. Rev. A* 43 (4 1991), pp. 2046–2049. DOI: 10.1103/PhysRevA.43.2046.

[9] Mark Srednicki. "Chaos and quantum thermalization". In: *Physical Review E* 50.2 (1994), pp. 888–901. DOI: 10.1103/physreve.50.888.

[10] Marcos Rigol, Vanja Dunjko, and Maxim Olshanii. "Thermalization and its mechanism for generic isolated quantum systems". In: *Nature* 452.7189 (2008), pp. 854–858. DOI: 10.1038/nature06838.

[11] Stephen H. Shenker and Douglas Stanford. "Black holes and the butterfly effect". In: *Journal of High Energy Physics* 2014.3 (2014). DOI: 10.1007/jhep03(2014)067.

[12] Pasquale Calabrese and John Cardy. "Evolution of entanglement entropy in one-dimensional systems". In: *Journal of Statistical Mechanics: Theory and Experiment* 2005.04 (2005), P04010. DOI: 10.1088/1742-5468/2005/04/p04010.

[13] Maurizio Fagotti and Pasquale Calabrese. "Evolution of entanglement entropy following a quantum quench: Analytic results for the XY chain in a transverse magnetic field". In: *Physical Review A* 78.1 (2008). DOI: 10.1103/physreva.78.010306.

[14] Andreas M Läuchli and Corinna Kollath. "Spreading of correlations and entanglement after a quench in the one-dimensional Bose–Hubbard model". In: *Journal of Statistical Mechanics: Theory and Experiment* 2008.05 (2008), P05018. DOI: 10.1088/1742-5468/2008/05/p05018.

[15] Hyungwon Kim and David A. Huse. "Ballistic Spreading of Entanglement in a Diffusive Nonintegrable System". In: *Physical Review Letters* 111.12 (2013). DOI: 10.1103/physrevlett.111.127205.

[16] Jens H. Bardarson, Frank Pollmann, and Joel E. Moore. "Unbounded Growth of Entanglement in Models of Many-Body Localization". In: *Phys. Rev. Lett.* 109 (1 2012), p. 017202. DOI: 10.1103/PhysRevLett.109.017202.

[17] C. W. von Keyserlingk, Tibor Rakovszky, Frank Pollmann, and S. L. Sondhi. "Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws". In: *Physical Review X* 8.2 (2018). DOI: 10.1103/physrevx.8.021013.

[18] Adam Nahum, Jonathan Ruhman, Sagar Vijay, and Jeongwan Haah. "Quantum Entanglement Growth under Random Unitary Dynamics". In: *Physical Review X* 7.3 (2017). DOI: 10.1103/physrevx.7.031016.

[19] David A. Huse Adam Nahum Jonathan Ruhman. "Dynamics of entanglement and transport in 1D systems with quenched randomness". In: (2017). eprint: https://arxiv.org/abs/1705.10364.

[20] Adam Nahum, Sagar Vijay, and Jeongwan Haah. "Operator Spreading in Random Unitary Circuits". In: *Physical Review X* 8.2 (2018). DOI: 10.1103/physrevx.8.021014.

[21] Cheryne Jonay, David A. Huse, and Adam Nahum. "Coarse-grained dynamics of operator and state entanglement". In: (2018). eprint: https://arxiv.org/abs/1803.00089.

[22] Cheryne Jonay. *Operator Spreading and Entanglement in a Quantum Chaotic Spin Chain*. Senior Thesis, Princeton University. 2017.

[23] Daniel A. Roberts and Brian Swingle. "Lieb-Robinson Bound and the Butterfly Effect in Quantum Field Theories". In: *Physical Review Letters* 117.9 (2016). DOI: 10.1103/physrevlett.117.091602.

[24] Elliott H. Lieb and Derek W. Robinson. "The finite group velocity of quantum spin systems". In: *Communications in Mathematical Physics* 28.3 (1972), pp. 251–257. DOI: 10.1007/bf01645779.

[25] Vedika Khemani, David A. Huse, and Adam Nahum. "Velocity-dependent Lyapunov exponents in many-body quantum, semi-classical, and classical chaos". In: (2018). eprint: https://arxiv.org/abs/1803.05902.

[26] Vadim Oganesyan and David A. Huse. "Localization of interacting fermions at high temperature". In: *Physical Review B* 75.15 (2007). DOI: 10.1103/physrevb.75.155111.

[27] Alexei Kitaev and John Preskill. "Topological Entanglement Entropy". In: *Physical Review Letters* 96.11 (2006). DOI: 10.1103/physrevlett.96.110404.

[28] Horacio Casini, Marina Huerta, and Robert C. Myers. "Towards a derivation of holographic entanglement entropy". In: *Journal of High Energy Physics* 2011.5 (2011). DOI: 10.1007/jhep05(2011)036.

[29] Robert C. Myers and Aninda Sinha. "Holographic c-theorems in arbitrary dimensions". In: *Journal of High Energy Physics* 2011.1 (2011). DOI: 10.1007/jhep01(2011)125.

[30] Tianci Zhou and David J. Luitz. "Operator entanglement entropy of the time evolution operator in chaotic systems". In: *Physical Review B* 95.9 (2017). DOI: 10.1103/physrevb.95.094206.