Entanglement Dynamics for Different Measures of Entanglement

Sam Cree
Department of Physics, Stanford University, Stanford, CA 94305
(Dated: June 17, 2020)

Submitted as coursework for PH470, Stanford University, Spring 2020

In quantum many-body systems, evolution in time has a tendency to increase the entanglement between any two regions. The precise dynamics of this entanglement growth can be a valuable lens into understanding a system's dynamical properties. In this work, we review some results in this field, particularly focusing on the roles played by different measures of entanglement. We pedagogically introduce a family of measures of entanglement known as the *Rényi* entropies. We review some qualitative features of entanglement growth in different systems, and particularly study a case studied recently where different Rényi entropies behave very differently. In certain systems, the presence of a conserved charge that spreads diffusively in time results in a slower growth of some Rényi entropies than might otherwise be expected, namely growing as $\sim \sqrt{t}$ rather than the more typical $\sim t$.

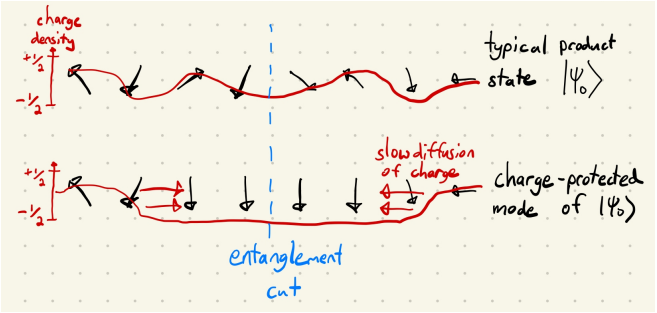

FIG. 1: This shows pictorially the argument for why Rényi entropies with $\alpha > 1$ will grow slowly in certain systems with conserved quantities, as discussed in the text. In this case, the system is a spin-1/2 chain with total spin conserved. Above is shown a typical product state of a spin chain in black, and below is a charge-protected mode of this state. The charge density shown in red is the charge associated with the conserved spin. Because charge is conserved, a region with no spin can only have trivial dynamics unless there is an in flux of charge from some other region. Thus entanglement cannot be generated across the cut for the charge-protected mode due to its void of charge around the cut, until such time has passed that the charge has had time to diffuse into the middle of the chain.

I. INTRODUCTION

In every branch of physics, major developments often arise after the introduction of new theoretical or experimental tools that allow us to ask new questions about physical systems. A number of recent developments has allowed us to ask many more theoretical questions about quantum many-body dynamics, which can be said to fall broadly into two categories – tools that allow the study of a greater variety of systems (such as tensor network methods, and analytical tools for studying random circuits), and tools that allow us to ask new questions of familiar systems (such as probing quantum chaos via out-of-time-ordered correlators). Here we investigate a recent development of the latter category, namely the study of *entanglement dynamics* of many-body quantum systems.

Entanglement is a distinct type of correlation present in quantum systems, and is known to play a key role in many parts of physics – including quantum foundations, black hole information theory, thermodynamics, quantum computing and many more. There are many more ways for a system to be entangled than to be unentangled, which gives measures of entanglement an entropic-like tendency to grow in time for generic systems. The details of how this entanglement growth occurs can reveal distinct characteristics of the system in question, such as indicating whether a system exhibits localization, distinguishing integrable systems from thermalizing systems, and even signalling the existence of a local conserved charge.

One thing that sets entanglement apart from most other tools for studying quantum many-body dynamics is that it is not quantified via correlation functions of observables – rather, it is an inherently nonlinear quantity. To see why entanglement must be nonlinear, just consider two distinct unentangled states $|\phi\rangle$ and $|\psi\rangle$. Their superposition can be strongly entangled despite both individual states being completely unentangled, so for any measure of entanglement S, we have $S(|\phi\rangle) + S(|\psi\rangle) = 0 \neq S(|\phi\rangle + |\psi\rangle)$.

How do we construct a measure of entanglement for a many-body system? First, we make a "cut" through the system to partition it into two regions. If one has access to only one of these regions, then the state of the sites in that region are not described by a pure quantum state but by a *density matrix* – which is essentially a proba-

bility distribution over a set of pure states, with the randomness arising from the ignorance about the state of the other region. This probability distribution is known as the *entanglement spectrum*, and completely characterizes the strength of the correlations between the two regions in the state $|\psi\rangle$. Loosely, the more uncertain this probability distribution, the more entanglement between these regions. There are many different classical ways of characterizing the "uncertainty" (or "entropy") of a probability distribution, each of which corresponds to a measure of entanglement when applied to the entanglement spectrum of a state.

A particularly well-studied family of measures is known as the Rényi entropies[1], which are a family parametrized by a nonnegative real number α. For $\alpha > 1$, these entropies are primarily indicators of how small the largest probability in the entanglement spectrum is. At $\alpha = 1$, it reduces to the well-known von Neumann entropy (or "entanglement entropy"), which measures the "typical order of magnitude" for a probability taken from this distribution. Finally, for smaller $\alpha < 1$, the Rényi entropies are more dominated by the smaller probabilities in the spectrum, reflecting more the number of probabilities in the distribution rather than their magnitudes.

A typical state (i.e. one chosen at random according to a uniform distribution) will almost always have the maximal amount of entanglement according to all of these Rényi entropies[2] – much like the entropy of a thermodynamic system[16]. That means that for generic time evolution of a quantum system, we should expect all of the Rényi entropies to increase. The most straightforward way to study the dynamics of entanglement is to begin with a state that is fine-tuned to have minimal initial entanglement – such as product states – and see how it dynamically approaches the equilibrium value.

In this report, we will summarize some of the scenarios where entropies with different α play different roles in characterizing the dynamical behaviour of many-body systems. In particular, we will discuss a recent finding[3–5] that in certain systems with a conserved quantity that spreads out diffusively in time, the Rényi entropies at $\alpha > 1$ grow much more slowly than the the von Neumann entropy (i.e. the $\alpha = 1$ case).

Here, we give more technical details to support the discussion above. In section II, we introduce in mathematical detail the Rényi entropies. We then overview some features of the dynamics of these entropies in section III. Finally, we discuss a particular case in which higher Rényi entropies behave qualitatively differently to the usual von Neumann entropy in section IV.

II. ENTANGLEMENT

In the context of lattice systems, we will study entanglement by partitioning the lattice sites into two connected regions A and B. In other words, we will try to quantify how much the degrees of freedom on either side of a boundary in the lattice are dependent on and interconnected with one another. We will also assume that the overall state of the system is a pure state – this is a convenient choice, because the theory of entanglement for pure states is much simpler than that of mixed states[17]. We will see that in this case, the entanglement between the degrees of freedom in regions A and B will be completely characterized by a set of real numbers $\{\lambda_k\}$ that we will refer to as the *entanglement spectrum*, also known as the *Schmidt values*.

The Hilbert space for the lattice system is simply a tensor product of the local Hilbert spaces in each region, $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$. A state $|\psi\rangle \in \mathcal{H}$ is called a *product state* with respect to this decomposition if it can be written as

$$|\psi\rangle = \left|\psi^A\right\rangle \otimes \left|\psi^B\right\rangle . \tag{II.1}$$

Physically, this means that the state is uncorrelated across the two regions; any measurement of a local observable $\mathcal{O}_A$ in $\mathcal{H}_A$ will depend only on $\left|\psi^A\right\rangle$,

$$\left(\left\langle\psi^A\right| \otimes \left\langle\psi^B\right|\right)\left(\mathcal{O}_A \otimes \mathbb{1}_B\right)\left(\left|\psi^A\right\rangle \otimes \left|\psi^B\right\rangle\right) = \left\langle\psi^A\right| \mathcal{O}_A \left|\psi^A\right\rangle , \tag{II.2}$$

and similarly for local observables in $\mathcal{H}_B$. Such a state is *unentangled*, as the state in $\mathcal{H}_A$ can be completely described without any reference to $\mathcal{H}_B$.

More generally, any state in $\mathcal{H}$ can be written as a linear combination of product states. In particular, we can show that any state has a *Schmidt decomposition* defined by

$$|\psi\rangle = \sum_k^{\min(d_A,d_B)} \lambda_k \left|\psi_k^A\right\rangle \otimes \left|\psi_k^B\right\rangle , \tag{II.3}$$

where $\lambda_k \geq 0$ satisfy $\sum_k \lambda_k^2 = 1$, $\left|\psi_k^A\right\rangle$ ($\left|\psi_k^B\right\rangle$) is an orthonormal set of states in $\mathcal{H}_A$ ($\mathcal{H}_B$), and d_A (d_B) is the dimension of $\mathcal{H}_A$ ($\mathcal{H}_B$).

The derivation is as follows. Suppose we have an orthonormal basis $\left\{\left|\phi_i^A\right\rangle\right\}_{i=1}^{d_A}$ for $\mathcal{H}_A$ and a basis $\left\{\left|\phi_j^B\right\rangle\right\}_{j=1}^{d_B}$ for $\mathcal{H}_B$, with $\dim(\mathcal{H}_A) = d_A$ and $\dim(\mathcal{H}_B) = d_B$. Then $\left\{\left|\phi_i^A\right\rangle \otimes \left|\phi_j^B\right\rangle\right\}_{i,j}$ is a basis for the Hilbert space $\mathcal{H}$, and we can write any $|\psi\rangle$ as

$$|\psi\rangle = \sum_{i,j} c_{i,j} \left|\phi_i^A\right\rangle \otimes \left|\phi_j^B\right\rangle . \tag{II.4}$$

Now we can perform a singular value decomposition on the matrix c_{ij}. A singular value decomposition for any $d_A \times d_B$ matrix M_{ij} (including non-square matrices!) is given by

$$M_{ij} = \sum_k^{\min(d_A,d_B)} U_{ik} \lambda_k V_{kj}^\dagger , \tag{II.5}$$

where λ_k are nonnegative real numbers, and $U^\dagger U = V^\dagger V = \mathbb{1}$, i.e. if M is square then U and V are unitary[18].

Doing this for c_{ij} in eq. (II.4) gives

$$|\psi\rangle = \sum_{i,j,k} U_{i,k}\lambda_k V^\dagger_{kj} \left|\phi^A_i\right\rangle \otimes \left|\phi^B_j\right\rangle . \tag{II.6}$$

Now, we can define a new (partial) basis for each space given by $\left|\psi^A_i\right\rangle = \sum_i U_{ik}\left|\phi^A_i\right\rangle$ and $\left|\psi^B_k\right\rangle = \sum_j V^\dagger_{kj}\left|\phi^B_j\right\rangle$, and we see that eq. (II.3) is satisfied.

By performing a partial trace on B, one finds that

$$\rho_A = \mathrm{Tr}_B(|\psi\rangle\langle\psi| \tag{II.7}$$

$$= \sum_k \lambda_k^2 \left|\psi^A_k\right\rangle\!\left\langle\psi^A_k\right| \tag{II.8}$$

i.e. the Schmidt values also play the role of the eigenvalues of the reduced density operator ρ_A (and likewise for $A \leftrightarrow B$).

If only one Schmidt value λ_k is nonzero, then it must equal one, and $|\psi\rangle$ is a product state of the form eq. (II.1). Otherwise, the state is entangled, because local observables will not factorize like in eq. (II.2). The strength of this entanglement is completely characterized by the entanglement spectrum $\{\lambda_k\}$, and one can construct specific measures of entanglement strength via functions of this spectrum. The most common measure is the von Neumann entropy, defined as

$$S = -\sum_k \lambda_k^2 \log \lambda_k^2, \tag{II.9}$$

where all logarithms in this work are defined in base 2 for simplicity when working with qubits and spin-$\frac{1}{2}$ chains. A more general family of entropies parametrized by some real $\alpha \in (0, \infty)$ is given by[19]

$$S_\alpha = \frac{1}{1-\alpha}\log \sum_k \lambda_k^{2\alpha}, \tag{II.10}$$

known as the *Rényi entropies*. One can verify using l'Hôpital's rule that

$$S = \lim_{\alpha\to 1} S_\alpha. \tag{II.11}$$

What kind of information does each S_α contain about the entanglement spectrum? To answer this, let us consider the limiting cases, $S_0 := \lim_{\alpha\to 0} S_\alpha$ and $S_\infty := \lim_{\alpha\to\infty} S_\alpha$. S_0 is known as the *Schmidt rank* or the *Hartley entropy*, and one can verify that it is simply

$$S_0 = \log\left(\text{number of nonzero } \lambda_k\right). \tag{II.12}$$

The Rényi entropies are also monotonically non-increasing with α, as can be seen by taking a derivative $\frac{\mathrm{d}S_\alpha}{\mathrm{d}\alpha} \le 0$. This also means that all of the Rényi entropies are upper bounded by the maximum possible value of S_0, which is $\log d$ where $d := \min(d_A, d_B)$. In fact, this maximum is obtained for all α when the entanglement spectrum is flat, i.e. when all Schmidt values are equal,

$$\forall k, \lambda_k = \frac{1}{d} \quad \Longrightarrow \quad S_\alpha = \log d. \tag{II.13}$$

A state with a flat entanglement spectrum is said to be *maximally entangled*, and the reduced density matrix on the smaller subsystem is the maximally mixed state $\rho = \frac{1}{d}\mathbb{1}$.

For S_∞, we can use l'Hôpital's rule again to find that it depends only on the largest Schmidt value $\lambda_{\max}$, as[20]

$$S_\infty = -\log \lambda^2_{\max} \tag{II.14}$$

Note that $\sum_k \lambda_k^{2\alpha} \ge \lambda^{2\alpha}_{\max}$, which gives a upper bound on S_α for $\alpha > 1$, as

$$S_\alpha \le \frac{\alpha}{1-\alpha}\log\lambda^2_{\max} = \frac{\alpha}{\alpha-1}S_\infty, \quad \alpha > 1 \tag{II.15}$$

Combining this with the monotonicity property $\frac{\mathrm{d}S_\alpha}{\mathrm{d}\alpha} \le 0$ from before, we see that for $\alpha > 1$, S_α is "sandwiched" by S_∞,

$$S_\infty \le S_\alpha \le \frac{\alpha}{\alpha-1}S_\infty, \quad \alpha > 1 \tag{II.16}$$

$$-\log\lambda^2_{\max} \le S_\alpha \le \frac{\alpha}{1-\alpha}\log\lambda^2_{\max}, \quad \alpha > 1$$

Note that this bound does not apply to the von Neumann entropy, S; in fact it can be arbitrarily large for any fixed $\lambda_{\max}$ so long as $\lambda_{\max} < 1$[21].

To summarize this discussion, as α increases, the Rényi entropy becomes more and more dominated by the very largest Schmidt value, to the point that for $\alpha > 1$ it is bounded from above and below by a function of $\lambda_{\max}$. In contrast, for smaller α the entanglement is more "egalitarian" and counts all Schmidt values more evenly.

III. OVERVIEW OF ENTANGLEMENT DYNAMICS

Suppose that our state has some dependence on time, $|\psi\rangle = |\psi(t)\rangle$. Then our Schmidt decomposition will also depend on time, as

$$|\psi\rangle = \sum_k^{\min(d_A,d_B)} \lambda_k(t)\left|\psi^A_k(t)\right\rangle \otimes \left|\psi^B_k(t)\right\rangle . \tag{III.1}$$

Although much of the important information about the evolution of $|\psi(t)\rangle$ is contained in the time-dependence of the basis vectors, we wish to focus on the dynamics of the *entanglement*, which as discussed above is fully characterized by the spectrum $\lambda_k(t)$. Each of the Rényi entropies discussed above are now time-dependent functions that measure the evolution of the entanglement over time, as

$$S_\alpha(t) = \frac{1}{1-\alpha}\log\sum_k \lambda_k(t)^{2\alpha}. \tag{III.2}$$

We will now discuss a few examples to give a sense for how the different Rényi entropies represent different information about the entanglement dynamics of a system. It will generally be assumed that the initial state

$|\psi(0)\rangle$ is a product state, meaning that its entanglement spectrum is just $\{1, 0, 0, 0 \dots\}$ and the Rényi entropies are trivial: $\forall\alpha,\ S_\alpha(0) = 0$.

Let's start with an extreme example, namely the dynamics of the Schmidt rank, $S_0(t)$, in a local Hamiltonian system. We can heuristically argue that it will behave discontinuously as,

$$S_0(t) = \begin{cases} 0, & \text{when } t = 0 \\ S_0^f, & \text{when } t \neq 0 \end{cases}. \tag{III.3}$$

To see this, note that the form of the state for any time t is,

$$|\psi(t)\rangle = e^{-iHt}|\psi(0)\rangle = \sum_{n=0}^{\infty} \frac{(-iHt)^n}{n!}|\psi(0)\rangle. \tag{III.4}$$

This time-evolved state will always remain in the subspace spanned by states of the form $H^n|\psi(0)\rangle$. Because H has local interactions, each application of H does not generate much entanglement between different regions, and the first few terms will only be weakly entangled. Terms with higher order applications of H^n will be more strongly entangled, but at early times will be suppressed by t^n. But the Schmidt rank S_0 cares only about whether terms are nonzero, not about how large they are. These highly entangled states in the sum will require a large number of terms in the Schmidt decomposition, and unless there are finely-tuned cancellations with other terms (which will only happen at a measure-zero set of t values), we should expect these terms to also contribute to the Schmidt decomposition of $|\psi(t)\rangle$. The late-time value S_0^f will depend on details of H, in particular the size of the subspace spanned by $H^n|\psi\rangle$.

Of course, at small t, the amplitude of these terms will remain very small, which will mean that $S_\alpha(t)$ will remain small for $\alpha > 0$; the discontinuous behaviour above is unique to $S_0(t)$[22]. For small t the $n = 0$ term in eq. (III.4) will dominate, which ensures that the $\alpha > 1$ will remain close to zero due to eq. (II.16). $S_\alpha(t)$ for $0 < \alpha \leq 1$ will also remain small, because of the exponential tail in higher-order terms in the expansion.

The evolution of S_α for nonzero α is more subtle, and depends on the details of H itself. The most straightforward to interpret is the von Neumann entropy, S, because of some of its nice mathematical properties. Unlike the other entropies, it exhibits *subadditivity*, which guarantees that shifting the cut by a single site will not alter S by more than one unit. This property provides the von Neumann entropy a nice physical interpretation: it counts the number of effective qubits that are separated across the entanglement cut, i.e. if you take N EPR pairs and put one particle of each on either side of the cut, the von Neumann entropy will be N[23].

This qubit picture is quite literally correct for integrable systems! Such systems can be completely described by an effective quasiparticle description, in which entanglement is generated across the cut when a quasiparticle passes across it[6]. This means that as quasiparticles "free stream" throughout the system with constant velocity, the entanglement spreads linearly in time. These quasiparticles carry both information and energy, meaning the behaviour of these two quantities is very similar in integrable systems.

It turns out that this is not always the case, however. For chaotic systems, energy typically spreads diffusively, i.e. it covers a distance proportional to $\sqrt{t}$ like for a random walk. Nonetheless, the entanglement continues to spread ballistically, as first pointed out by Kim and Huse[7]. In chaotic systems it is believed generally that the following holds[8],

$$\frac{\mathrm{d}}{\mathrm{d}t}S = v_E s_{\mathrm{th}} A, \tag{III.5}$$

where A is the surface area of the cut (which is one in the 1D systems that we consider here), s_{th} is the equilibrium entropy density of the system, and v_E is a constant known as the entanglement velocity. The form of the equation suggests a picture of an inward "flux" of entanglement with density s_{th} that flows across the cut at a speed v_E. This picture (known as the *entanglement tsunami* picture) is useful in interpreting the above equation, but it is not a perfect analogy; entanglement is not conserved, so it does not so much "flow" as it is generated over time[9].

Another class of systems with interesting entanglement dynamics is that of many-body localized systems. In such systems, large disorder disrupts the spreading of energy, as well as operator spreading, and similarly, it provides an obstacle to the generation of entanglement. It is believed generally that in such systems, entanglement grows logarithmically in time[10]. Furthermore, there is an interesting crossover between such highly-disordered localized systems and more weakly disorded thermalizing systems. In systems with isolated regions that locally resemble the localized phase, the entanglement entropy grows not linearly but as a power law $S(t) \sim t^\beta$, with $0 < \beta < 1$[11,12].

All this discussion has just been for the von Neumann entropy, which is the entanglement measure that generally receives the most attention in the literature. What about other Rényi entropies?

In some systems, the distinction is trivial! A unitary time evolution operator that maps tensor products of Pauli operators into other tensor products of Pauli operators is said to be an element of the *Clifford group*. This kind of dynamics is known to be efficiently simulated classically[13], which makes them numerically useful in many circumstances. This result is a consequence of the *stabilizer formalism*, which also has an interesting consequence for our purposes – namely, that the entanglement spectrum of a stabilizer state is always flat, i.e. all of the Rényi entropies are equal to one another[14]!

However, in the example that we are about to explore in some depth, the Rényi entropies have markedly different behaviour depending on whether $\alpha > 1$ or not.

IV. RÉNYI ENTROPIES WITH $\alpha > 1$ AS A PROBE OF TRANSPORT

While chaotic systems generally have linear $S(t)$, the behaviour of $S_\alpha(t)$ depends on the behaviour of conserved quantities. In particular, certain systems with conserved quantities that spread diffusively will have $S_\alpha(t)$ grow as $\sqrt{t}$ rather than linearly at late times for $\alpha > 1$. This was originally suggested by Rakovszky et al[3], with evidence from charge-conserving random unitary circuits and the tilted field Ising model (at least, under typical initial conditions and at late times). It was more rigorously shown in a later work by Huang[4] that this can be shown in general for sufficiently typical states of any charge-conserving (not necessarily random!) unitary circuit with diffusive energy transport. This was then expanded upon in a work by Zhou and Ludwig[5] following very similar logic but in a Heisenberg picture framework. We will outline Huang's argument here.

The key requirement in this work, aside from the existence of a local conserved charge with diffusive dynamics, is that there is a *local charge sector with trivial dynamics*. Let us clarify what we mean by this.

Consider a spin-$\frac{1}{2}$ lattice with overall conservation of spin. If the system begins with all spins in the $|\downarrow\rangle$ state, then it must be completely unaffected by the dynamics, because the subspace of states in that spin sector is only one-dimensional. Now suppose a certain region is in the $|\downarrow\rangle$ state, but the rest of the system is not. Then provided the time evolution is *local*, that region will still be mostly unaffected by the time evolution, as the spin must be locally conserved. Gradually, positive spin can leak in from other parts of the system, and eventually the dynamics will become nontrivial again. But so long as a region is in this sector of minimum spin, the dynamics remain trivial, and thus *no entanglement can be generated across a cut in this region*.

Not all systems with conservation laws are like this! Consider a similar system but with each site now having *two* independent qubits on it, one of type a and one of type b, i.e. each site of the lattice has a four-dimensional local Hilbert space $\mathcal{H} = \mathcal{H}_a \otimes \mathcal{H}_b$. Then if only the spins of type a are conserved by the dynamics, the above property no longer applies! Entanglement can still be generated even in a region of minimum a spin, because of the degrees of freedom afforded by the b qubits.

Our argument for slow growth of higher Rényi entropies relies on the fact that entanglement cannot be generated in a region with appropriate charge, so to proceed we must assume that there is a *local charge sector with trivial dynamics*. Without this assumption, a similar argument would follow showing that the Rényi entropies still have a term contributing of the form $\sim \sqrt{t}$, but also a more dominant linear term that overshadows it[3].

Under this assumption, then, let us outline the derivation of the Rényi entropy growth. For any time t, the slow diffusion of charge guarantees that a typical initial state will contain a mode that is protected against entanglement; which guarantees the existence of a term in the Schmidt decomposition $\lambda_{\max}$ with comparatively large amplitude. The relationship between an unentangled mode of a state $|\psi\rangle$ and an upper bound on its Rényi entropy arises from the Eckart-Young theorem[15]:

Theorem IV.1. *For any state $|\psi\rangle$ whose largest Schmidt value is λ_{max}, and any product state $|\phi\rangle = |\phi_A\rangle \otimes |\phi_B\rangle$, the following inequality holds.*

$$|\langle\phi|\psi\rangle| \leq \lambda_{max} \tag{IV.1}$$

This in turn gives an upper bound to the Rényi entropies according to eq. (II.16).

For a given time t, we can apply this theorem to the state $|\psi(t)\rangle$ to obtain a bound on the Rényi entropies at that time. We should choose a product state $|\phi(t)\rangle$ to have as large overlap as possible with $|\psi(t)\rangle$, to give the strongest possible lower bound on $\lambda_{\max}$ (and thus the strongest possible upper bound on the growth of the Rényi entropies). Note that $|\langle\phi(0)|\psi(0)\rangle| = |\langle\phi(t)|\psi(t)\rangle|$, so as long as the two states start at $t = 0$ with significant overlap, this will remain true at late times. The key trick of Huang's is to make a clever choice of product state $|\phi(0)\rangle$. We pick one as shown in fig. 1, with a region of zero charge in the middle near the cut between regions, and make it look as much like $|\psi(0)\rangle$ as possible everywhere else. This state is completely unentangled to begin with, and the fact that this remains true as time goes on is thanks to our assumption of a local charge sector with trivial dynamics; no entanglement can generate if there's no charge near the cut! It will continue to be unentangled up to a time of order $\sqrt{t}$ when the charge has had time to "leak" into the void.

Most of the technical difficulties in the paper arise from the small but nonzero chance that charge leaks in towards the cut much faster than is typical. The resolution to this challenge is to consider a slightly modified version of the time evolution operator U, which we call $\tilde{U}$, such that $\tilde{U}$ does not produce any entanglement across the cut (i.e. simply remove any interactions between sites on either side). By evolving with $\tilde{U}$, an initial product state will remain a product state, so $\left|\tilde{\phi}(t)\right\rangle := \tilde{U}\,|\phi(0)\rangle$ can be used for the Eckart-Young theorem eq. (IV.1). So long as only a small amount of charge leaks into the void all the way to the cut, $\left|\tilde{\phi}(t)\right\rangle$ will be very similar to $|\phi(t)\rangle$, which in turn will be not *too* different to $|\psi(t)\rangle$.

How large should the void be? We have two conflicting factors here. A large void reduces the chance of charge leaking all the way to the centre, which means that $|\phi(t)\rangle$ is more similar to $\left|\tilde{\phi}(t)\right\rangle$. However, it also reduces the overlap $|\langle\phi(0)|\psi(0)\rangle|$, because less of the state resembles $|\psi\rangle$! This in turn reduces the overlap $\left|\left\langle\tilde{\phi}(t)\middle|\psi(t)\right\rangle\right|$ and leads to a weaker bound via eq. (IV.1). The optimal choice of state $|\phi(0)\rangle$ (i.e. the one that leads to the tightest bound) turns out to be one with a void of width $\sim \sqrt{t}$.

One final subtlety to mention is that this argument only holds for typical initial states[24], but can fail for specific choices of $|\psi(0)\rangle$. To understand this, suppose that $|\psi(0)\rangle$ is a tensor product of alternating Pauli z eigenstates. Then any state with a central void will have *zero* overlap with $|\psi(0)\rangle$, which completely eliminates the usefulness of the Eckart-Young theorem. Such initial states can lead to linear growth of the Rényi entropies[3].

[1] A. Rényi, in *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, Volume 1: Contributions to the Theory of Statistics* (University of California Press, Berkeley, Calif., 1961), pp. 547–561, URL https://projecteuclid.org/euclid.bsmsp/1200512181.

[2] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993), URL https://link.aps.org/doi/10.1103/PhysRevLett.71.1291.

[3] T. Rakovszky, F. Pollmann, and C. von Keyserlingk, Physical Review Letters **122** (2019), ISSN 1079-7114, URL http://dx.doi.org/10.1103/PhysRevLett.122.250602.

[4] Y. Huang, *Dynamics of Rényi entanglement entropy in local quantum circuits with charge conservation* (2019), arXiv:1902.00977, URL http://arxiv.org/abs/1902.00977.

[5] T. Zhou and A. W. W. Ludwig, *On the diffusive scaling of Rényi entanglement entropy* (2019), arXiv:1911.12384, URL http://arxiv.org/abs/1911.12384.

[6] P. Calabrese and J. Cardy, Journal of Statistical Mechanics: Theory and Experiment **2005**, P04010 (2005), ISSN 1742-5468, URL http://dx.doi.org/10.1088/1742-5468/2005/04/P04010.

[7] H. Kim and D. A. Huse, Physical Review Letters **111** (2013), ISSN 1079-7114, URL http://dx.doi.org/10.1103/PhysRevLett.111.127205.

[8] M. Mezei and D. Stanford, Journal of High Energy Physics **2017** (2017), ISSN 1029-8479, URL http://dx.doi.org/10.1007/JHEP05(2017)065.

[9] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, Journal of High Energy Physics **2016** (2016), ISSN 1029-8479, URL http://dx.doi.org/10.1007/JHEP02(2016)004.

[10] G. D. Chiara, S. Montangero, P. Calabrese, and R. Fazio, Journal of Statistical Mechanics: Theory and Experiment **2006**, P03001 (2006), URL https://doi.org/10.1088%2F1742-5468%2F2006%2F03%2Fp03001.

[11] R. Vosk, D. A. Huse, and E. Altman, Phys. Rev. X **5**, 031032 (2015), URL https://link.aps.org/doi/10.1103/PhysRevX.5.031032.

[12] A. Nahum, J. Ruhman, and D. A. Huse, Physical Review B **98** (2018), ISSN 2469-9969, URL http://dx.doi.org/10.1103/PhysRevB.98.035118.

[13] D. Gottesman, *Stabilizer codes and quantum error correction* (1997), arXiv:quant-ph/9705052, URL https://arxiv.org/abs/quant-ph/9705052.

[14] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, Physical Review X **7** (2017), ISSN 2160-3308, URL http://dx.doi.org/10.1103/PhysRevX.7.031016.

[15] C. Eckart and G. Young, Psychometrika **1**, 211 (1936).

[16] In a system with a time-independent Hamiltonian, energy will be conserved, which complicates this statement. The entanglement will not grow to the maximum amount, but rather the maximum amount subject to the constraint that energy is conserved from the initial state, which may be at some finite temperature. In fact at low temperatures, the late-time state of the system will look locally similar to the ground state, which has weak entanglement for a non-critical local Hamiltonian – so to see a significant range of entanglement growth we will usually be at high energies/temperatures.

[17] For example, even just determining whether a bipartite mixed state is entangled or not is generally an NP-hard problem.

[18] This decomposition is closely related to the polar decomposition, $M = HR$ with H Hermitian and $R^\dagger R = \mathbb{1}$; simply set $H = U\lambda U^\dagger$ and $R = UV^\dagger$. When M is normal, i.e. unitarily diagonalizable, it can also be compared with the diagonalization $M = ODO^\dagger$ with D diagonal and O unitary; specifically note that λ is more restricted in the singular value decomposition as it must be real and non-negative, in exchange for the additional freedom that U can be distinct from V.

[19] Both this and the von Neumann entropy are more commonly defined in terms of the reduced density matrix ρ_A, but we write them here in terms of the entanglement spectrum to emphasize that the latter completely characterizes the entanglement properties of the bipartite decomposition.

[20] This is also true even if the largest Schmidt value is not unique.

[21] Let there be n Schmidt values equal to $\frac{1-\lambda_{\max}^2}{n}$, so that $1 = \sum_k \lambda_k^2$. Then as $n \to \infty$, $S \to \infty$ with $\lambda_{\max}$ fixed.

[22] It is also unique to Hamiltonian systems, and in discrete circuits $S_0(t)$ can actually be quite a useful non-trivial probe of entanglement, e.g. in a work by Nahum et al[14].

[23] This is quite well-defined in a resource theoretic sense, as for a pure state $|\psi\rangle$, the von Neumann entropy is equal to the number of EPR pairs one could distill from the state $|\psi\rangle$ via local operations on either side of the cut and classical communication, or the number of EPR pairs one would require to produce the state $|\psi\rangle$ under the same restrictions.

[24] In Huang's work, it was specified that initial states were tensor products of $|\pm\rangle$ states, but the logic extends to typical states.