

MIT Open Access Articles

Void formation in operator growth, entanglement, and unitarity

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: Journal of High Energy Physics. 2021 Mar 16;2021(3):159

Published Version: [https://doi.org/10.1007/JHEP03\(2021\)159](https://doi.org/10.1007/JHEP03(2021)159)

Publisher: Springer Berlin Heidelberg

Permanent Link: <https://hdl.handle.net/1721.1/132072>

Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of use: <https://creativecommons.org/licenses/by/4.0/>



Void formation in operator growth, entanglement, and unitarity

Hong Liu and Shreya Vardhan

*Center for Theoretical Physics, Massachusetts Institute of Technology,
Cambridge, MA 02139, U.S.A.*

E-mail: hong_liu@mit.edu, vardhan@mit.edu

ABSTRACT: The structure of the Heisenberg evolution of operators plays a key role in explaining diverse processes in quantum many-body systems. In this paper, we discuss a new universal feature of operator evolution: an operator can develop a void during its evolution, where its nontrivial parts become separated by a region of identity operators. Such processes are present in both integrable and chaotic systems, and are required by unitarity. We show that void formation has important implications for unitarity of entanglement growth and generation of mutual information and multipartite entanglement. We study explicitly the probability distributions of void formation in a number of unitary circuit models, and conjecture that in a quantum chaotic system the distribution is given by the one we find in random unitary circuits, which we refer to as the random void distribution. We also show that random unitary circuits lead to the same pattern of entanglement growth for multiple intervals as in $(1 + 1)$ -dimensional holographic CFTs after a global quench, which can be used to argue that the random void distribution leads to maximal entanglement growth.

KEYWORDS: Quantum Dissipative Systems, Random Systems

ARXIV EPRINT: [1912.08918](https://arxiv.org/abs/1912.08918)

Contents

1	Introduction	1
2	Void formation and implications	4
2.1	Setup	4
2.2	Upper bound on average probability for void formation	7
2.3	Unitarity of entanglement growth for one interval	7
2.4	Mutual information and multi-partite entanglement	9
3	Random void distribution and entanglement growth	12
3.1	Random unitary circuits	12
3.2	Random void distribution	14
3.3	Random void distribution and maximal entanglement growth	16
4	Void formation in two Clifford circuit models	19
4.1	Free propagation model	20
4.2	Perfect tensor model	22
5	Conclusions and discussion	25
A	Derivations in the random unitary circuits	27
A.1	Mapping to a classical Ising partition function	27
A.2	Derivation of the random void distribution	29
A.3	Derivation of the time-evolution of S_2 in random unitary circuits with large q	32
B	Entanglement entropy from the random void distribution	34

1 Introduction

The Heisenberg evolution of operators in a quantum many-body system is in general extremely complicated. But during the last decade, remarkable universalities have been found, such as ballistic growth of operators [1] and growth of operator entanglement [2–5]. Such universal properties have played an important role in diverse problems like scrambling of quantum information, quantum many-body chaos, and entanglement growth during thermalization (see e.g. [6, 7]).

In this paper, we discuss another universal feature of operator evolution: an operator can develop a void during its evolution. More explicitly, considering a spatial region A within the “lightcone” of an initial operator \mathcal{O} , we can decompose its time evolution $\mathcal{O}(t)$ as

$$\mathcal{O}(t) = \mathcal{O}_1(t) + \mathcal{O}_2(t), \quad \mathcal{O}_1(t) = \tilde{\mathcal{O}}_{\bar{A}} \otimes \mathbf{1}_A \tag{1.1}$$

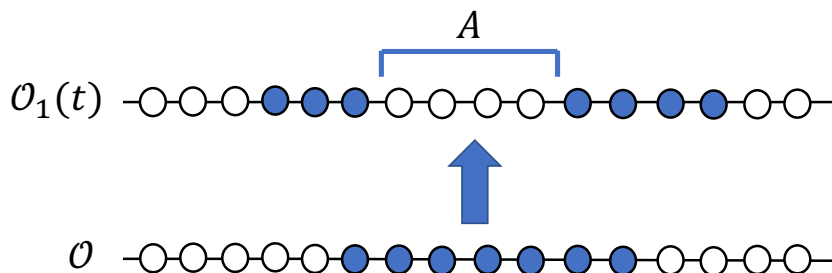


Figure 1. An operator \mathcal{O} in a many-body system can develop a “void” under time evolution, where nontrivial parts of $\mathcal{O}_1(t)$ (introduced in (1.1)) are separated by a region of identity operators. In the example shown above in one spatial dimension, the operator has nontrivial single-site operators at blue sites, and the identity at white sites.

where $\mathbf{1}_A$ denotes the identity operator in A , $\tilde{\mathcal{O}}_{\bar{A}}$ is some operator in \bar{A} (the complement of A),¹ and $\mathcal{O}_2(t)$ is an operator whose projection onto A is orthogonal to $\mathbf{1}_A$. Here we assume that the system has a finite-dimensional Hilbert space and tensor product structure associated with spatial regions. See figure 1 for an illustration. Given the space of all operators is a Hilbert space, we can also associate a weight or “probability” for $\mathcal{O}(t)$ to develop a void in region A

$$P_{\mathcal{O}}^{(A)}(t) = \frac{\text{Tr}(\mathcal{O}_1^\dagger(t)\mathcal{O}_1(t))}{\text{Tr}(\mathcal{O}^\dagger(t)\mathcal{O}(t))}. \tag{1.2}$$

Below we will refer to the presence of $\mathcal{O}_1(t)$ in $\mathcal{O}(t)$ as void formation. Void formation is present in both integrable and chaotic systems, and is required by unitarity. We will show that it has important implications for unitarity of entanglement growth. For example, evolving from an initial product state, it is the contribution from $\mathcal{O}_1(t)$ that ensures $S_A(t) = S_{\bar{A}}(t)$, where S_A is the entanglement entropy for region A . Further, we will show that void formation is responsible for generating mutual information and multipartite entanglement among disjoint regions during the evolution of a system, as we illustrate using a cartoon picture in figure 2. We will also derive a number of general constraints on probability distributions of void formation from unitarity, which are applicable to both chaotic and integrable systems.

To develop intuition for probability distributions of void formation, we study three types of unitary circuit models in one spatial dimension: (i) the random unitary model of [8–10], which can be considered a minimal model for quantum chaotic systems; (ii) a “free propagating” model [11] in which entanglement can only be spread, but not created, which may thus be considered a proxy for free theories; (iii) a circuit built from perfect tensors [12], which may be considered a model for non-chaotic systems. In the free-propagating and perfect tensor models, patterns of void formation depend sensitively on the initial operator. In particular, in the perfect tensor circuit model, void formation exhibits a fractal structure.

¹In the more detailed definition of a void that we give in section 2, we will only include the part of $\tilde{\mathcal{O}}_{\bar{A}}$ which is orthogonal to the identity operator when projected onto any disconnected region of \bar{A} .

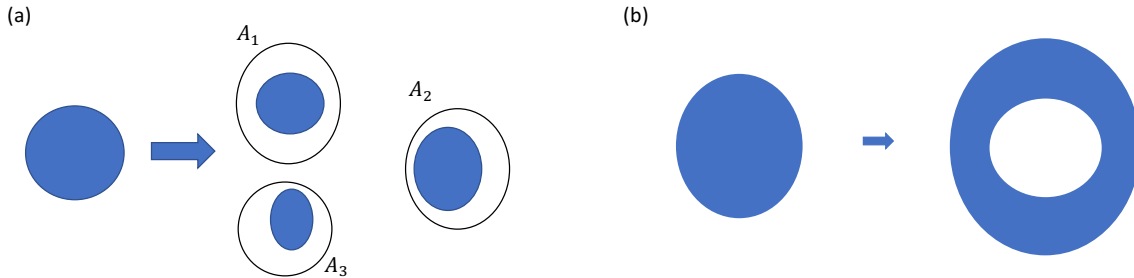


Figure 2. In more than one spatial dimension, void formation can either break up a single operator into disconnected parts, as in (a), or form a hole in an operator, as in (b). In the figure, shaded regions indicate where the operator has nontrivial support. A process like (a) contributes to multipartite entanglement between the three disjoint regions A_1 , A_2 , and A_3 .

In the random unitary circuit, we find at sufficiently late times

$$P_{\mathcal{O}}^{(A)} = \frac{1}{d_A^2} = e^{-2S_{\text{eq}}^{(A)}} \quad (1.3)$$

where d_A is the dimension of local Hilbert space in region A . After the second equality, we have written the expression in a form which is generalizable to continuum systems, with $S_{\text{eq}}^{(A)}$ the equilibrium entropy of A . It is natural to conjecture that (1.3), to which we will refer as the random void distribution from now on, holds for generic operators in general chaotic systems at sufficiently late times. While (1.3) is very small for a macroscopic region A , for certain processes the number of contributing operators can be exponentially large, leading to significant physical effects. Note that $1/d_A^2$ is simply the fraction of all basis operators in the system that are equal to the identity in the region A .

As an illustration, we show that together with the assumption of sharp light-cone growth of operators, the random void distribution (1.3) fully determines the second Renyi entropy of an arbitrary number of disjoint intervals in random unitary circuits in the limit of large one-site Hilbert space dimension. Furthermore, surprisingly, the resulting expression coincides exactly with the von Neumann entropy after a global quench in $(1+1)$ -dimensional holographic systems. On the one hand, this indicates that the random void distribution (1.3) may underlie operator evolution in holographic systems. On the other hand, in the light of the fact that the holographic expression maximizes the evolution of entanglement entropy in all $(1+1)$ -dimensional systems [11], we are led to conclude that together with sharp light-cone growth, the random void distribution maximizes entanglement growth.

The plan of this paper is as follows. In section 2, after describing in detail our general set-up, we discuss implications of void formation for unitarity of entanglement growth and generation of mutual information and multi-partite entanglement, as well as constraints on void formation from requiring unitarity. In section 3, we discuss the random circuit model, derive the random void distribution (1.3), and discuss its implications. In section 4, we discuss void formation in the free propagation and perfect tensor models. We conclude in section 5 with future directions. We have included a number of appendices for technical details.

2 Void formation and implications

In this section, we first describe our general setup, and then derive some simple constraints from unitarity on void formation during Heisenberg evolution.

2.1 Setup

For convenience, we will consider a one-dimensional lattice system with a finite-dimensional Hilbert space at each site. The discussion generalizes immediately to higher dimensions. We comment on generalizations to systems with an infinite local Hilbert space in the discussion section, section 5.

The Hilbert space at a site i will be denoted as \mathcal{H}_i , and is taken to have dimension q . The full Hilbert space is $\mathcal{H} = \otimes \mathcal{H}_i$, and has dimension q^L , where $L \rightarrow \infty$ is the system size. Operators at a single site form a Hilbert space of dimension q^2 , which will be denoted as \mathcal{G}_i . Operators of the full system form a Hilbert space $\mathcal{G} = \otimes_i \mathcal{G}_i$ of dimension q^{2L} . We will use $O_a^i, a = 0, 1, \dots, q^2 - 1$, to denote an orthonormal basis of \mathcal{G}_i which is normalized as

$$\text{tr}((O_a^i)^\dagger O_b^i) = q\delta_{ab}, \quad O_0^i = \mathbf{1}_i \tag{2.1}$$

where $\mathbf{1}_i$ is the identity operator of \mathcal{H}_i , and there is no summation over i in the above equation. Orthogonality with O_0^i implies $O_a^i, a = 1, \dots, q^2 - 1$, are all traceless. A convenient choice of basis which we will use throughout the paper is

$$O_a = X^{s_1} Z^{s_2}, \quad s_1, s_2 = 0, 1, \dots, q - 1 \tag{2.2}$$

where X and Z are respectively the shift and clock matrices.² An orthonormal basis for \mathcal{G} , which will be denoted as $\mathcal{O}_\alpha, \alpha = 0, 1, \dots, q^{2L} - 1$, can be obtained from tensor products of $\{O_a^i\}$. These basis operators satisfy

$$\text{Tr } \mathcal{O}_\alpha^\dagger \mathcal{O}_\beta = \delta_{\alpha\beta} q^L. \tag{2.3}$$

\mathcal{O}_0 is the identity operator for the full Hilbert space \mathcal{H} , and all other \mathcal{O}_α 's are traceless.

Under time evolution,

$$\mathcal{O}_\alpha(t) = U^\dagger(t) \mathcal{O}_\alpha U(t) = \sum_\beta c_\alpha^\beta(t) \mathcal{O}_\beta \tag{2.4}$$

where $U(t)$ denotes the evolution operator. From unitarity of $U(t)$,

$$\sum_\beta |c_\alpha^\beta(t)|^2 = 1. \tag{2.5}$$

We can interpret $|c_\alpha^\beta(t)|^2$ as the probability of operator \mathcal{O}_α evolving to \mathcal{O}_β . Systems with a local Hamiltonian have an effective light-cone speed v_c for how fast an operator can grow with time [13], so that $|c_\alpha^\beta(t)|^2 \approx 0$ for \mathcal{O}_β not contained within the light-cone of \mathcal{O}_α .

For the purpose of not obscuring the conceptual picture with technicalities, throughout this paper, we will consider a particularly simple form of operator evolution in which the

²Explicitly, $Z = \sum_{k=0}^{q-1} e^{2\pi i k/q} |k\rangle \langle k|$ and $X = \sum_{k=0}^{q-1} |k+1\rangle \langle k|$, where addition is defined mod q .

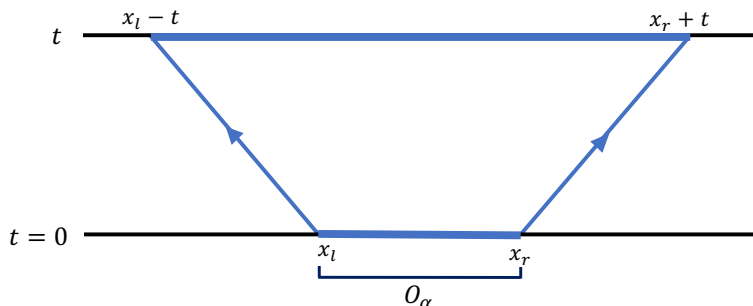


Figure 3. Under time-evolution, the end points of an operator \mathcal{O}_α move at the light-cone speed $v_c = 1$ in opposite directions. The initial operator with endpoints at x_l and x_r can become a superposition of many final operators at time t , each of which have endpoints $x_l - t$ and $x_r + t$.

end points of an operator of interest move at the light-cone speed v_c in opposite directions.³ We will refer to this assumption as sharp light-cone growth. See figure 3. From now on we will set $v_c = 1$. Causality imposed by the light-cone structure will play a key role in the subsequent discussion. Throughout the paper, if not mentioned explicitly, we will always consider q large and ignore subleading $1/q$ corrections. The qualitative picture does not change in more general situations, but the story becomes technically more complex, and will be treated elsewhere.

The statement about light-cone growth does not say anything about the internal structure of an operator under time evolution. A basic question addressed in this paper is: can the operator develop a void in some region between its endpoints, and if so, what are the implications of this process? We say an operator has a void in some region A (which can have disconnected components) if it is a superposition of operators the form $\tilde{\mathcal{O}} \otimes \mathbf{1}_A$, where $\mathbf{1}_A$ denotes the identity operator in A and $\tilde{\mathcal{O}}$ is some operator which has nontrivial support in all disconnected parts of \bar{A} (the complement of A). Then the probability for a basis operator \mathcal{O}_α to develop a void in region A at time t is

$$P_{\mathcal{O}_\alpha}^{(A)}(t) = \sum_{\beta \text{ with void in } A} |c_\alpha^\beta(t)|^2. \tag{2.6}$$

A main goal of this paper is to explore the role of void formation in the entanglement structure of a system. A good observable for this purpose is the evolution of the second Renyi entropy, which can be expressed in terms of operator growth [7–9, 15]. Suppose at $t = 0$, the system is described by a homogeneous pure product state, $\rho_0 = \otimes_i \rho_i$, where i runs over all sites of the system and ρ_i is a pure state which is the same for all sites. It is convenient to choose a basis so that

$$\rho_i = \frac{1}{q} \left(\mathbf{1}_i + Z_i + Z_i^2 + \dots + Z_i^{q-1} \right) \tag{2.7}$$

³Examples include the random unitary circuits in the large q limit [8, 9], Clifford circuit models to be discussed in section 4, and $(1 + 1)$ -dimensional CFTs in the large central charge limit [14]. This was also used as a toy model in [7, 15].

with Z_i the clock matrix at site i , and thus

$$\rho_0 = \frac{1}{q^L} \sum_{\alpha \in I} \mathcal{O}_\alpha \quad (2.8)$$

where I denotes the set of operators which can be built from powers of Z_i 's. Note that the space I is q^L -dimensional, in contrast to the q^{2L} -dimensional full space of operators.

Under time-evolution, the reduced density matrix for some region A is given by⁴

$$\rho_A(t) = \text{Tr}_{\bar{A}} \rho(t) = \frac{1}{q^L} \sum_{\alpha \in I} \text{Tr}_{\bar{A}} \mathcal{O}_\alpha(t) = \frac{1}{q^{|A|}} \sum_{\alpha \in I} \sum_{\beta \in A} c_\alpha^\beta(t) \mathcal{O}_\beta \quad (2.9)$$

where we have used (2.4), and the fact that due to tracelessness of all nontrivial basis operators, only operators of the form $\mathcal{O}_\beta \otimes \mathbf{1}_{\bar{A}}$ with \mathcal{O}_β an operator in region A (denoted by $\beta \in A$) contribute to $\text{Tr}_{\bar{A}} \mathcal{O}_\alpha(t)$. $|A|$ denotes the size of region A . The second Renyi entropy for A can then be written as

$$e^{-S_2^{(A)}(t)} = \text{Tr}_A \rho_A^2(t) = \frac{1}{q^{|A|}} \sum_{\alpha_1, \alpha_2 \in I} \sum_{\beta \in A} c_{\alpha_1}^\beta(t) c_{\alpha_2}^{\beta*}(t). \quad (2.10)$$

We will now make a further simplification by ignoring the off-diagonal terms (i.e. terms with $\alpha_1 \neq \alpha_2$) in (2.10). For chaotic systems, one expects the phases of c_α^β to be random, so that the off-diagonal terms are suppressed by order $O(q^{-|A|})$ compared with diagonal terms [15]. For integrable systems one cannot make this argument. Nevertheless, there are often situations where the off-diagonal terms vanish identically. This is the case for all the explicit examples we discuss in section 3 and section 4. We then find

$$e^{-S_2^{(A)}(t)} = \frac{1}{q^{|A|}} N_A(t), \quad N_A(t) \equiv \sum_{\alpha \in I} \sum_{\beta \in A} |c_\alpha^\beta(t)|^2, \quad S_2^{(A)} = |A| \log q - \log N_A(t). \quad (2.11)$$

$N_A(t)$ has a simple physical interpretation: it is the expected number of operators in the set I contained within region A at time t . Note that $N_A \geq 1$, as $O_\alpha = \mathbf{1}$ always contributes 1 to the above sum. For our later discussion, it is convenient to introduce a function $N(A, B; t)$, defined as the expected number of initial operators in I from some region B that are contained in A at time t , i.e.

$$N(A, B; t) \equiv \sum_{\alpha \in I \cap B} \sum_{\beta \in A} |c_\alpha^\beta(t)|^2, \quad (2.12)$$

and a void formation function $G(A, B; t)$, defined as the expected number of initial operators in I from some region B that develop a void in A at time t , i.e.,

$$G(A, B; t) \equiv \sum_{\alpha \in I \cap B} \sum_{\beta \text{ with void in } A} |c_\alpha^\beta(t)|^2 = \sum_{\alpha \in I \cap B} P_{\mathcal{O}_\alpha}^{(A)}(t). \quad (2.13)$$

$G(A, B; t)$ is closely related to $N(\bar{A}, B; t)$, but in $G(A, B; t)$ the final operators must be supported on all disconnected parts of \bar{A} . Again by definition, $N(A, B; t), G(A, B; t) \geq 1$, as we always have a contribution of 1 from the identity operator.

Throughout the paper, we will denote the union of two regions $A_1 \cup A_2$ simply as $A_1 A_2$.

⁴For notational convenience we will take states of the system to evolve by U^\dagger .

2.2 Upper bound on average probability for void formation

To give some intuition for the motivation behind (1.3), we first discuss the average probability for an operator to become trivial in a given region.

Consider a Hilbert space \mathcal{G} of dimension d and a subspace $\mathcal{G}' \subset \mathcal{G}$ of dimension d' , with P the projector onto \mathcal{G}' . For a vector $|\psi\rangle \in \mathcal{G}$, the probability that it transitions into a state in \mathcal{G}' under a unitary transformation is given by

$$p = \text{Tr} \left(P U |\psi\rangle \langle \psi| U^\dagger \right) . \tag{2.14}$$

The average probability \bar{p} , obtained by averaging $|\psi\rangle$ over all unit vectors in \mathcal{G} with the unitarily invariant Haar measure, is then

$$\bar{p} = \frac{1}{d} \text{Tr} P = \frac{d'}{d} . \tag{2.15}$$

We take \mathcal{G} to be the Hilbert space of all operators with dimension $d = q^{2L}$, and \mathcal{G}' to be the set of operators which are the identity in region A , which is a subspace of dimension $d' = q^{2(L-|A|)}$. We thus conclude that the average probability for an operator to be trivial in region A is

$$\bar{p} = q^{-2|A|} = \frac{1}{d_A^2} . \tag{2.16}$$

Apart from the assumption that U is unitary, (2.16) is independent of U and does not depend on the region A other than through its size. Note that (2.16) gives the probability for an operator to be trivial in A , which is larger than the probability for developing a void in A , as it also includes final operators which are trivial in some disconnected parts of \bar{A} .

Note that if we do not average $|\psi\rangle$ but instead take U to be a random unitary matrix with the Haar distribution, we find the same answer. This is the motivation for the name “random void distribution” for (1.3).

2.3 Unitarity of entanglement growth for one interval

We now present a simple argument which shows that void formation plays a crucial role in ensuring that the entanglement growth of a system after a global quench is compatible with unitarity. The argument can be used to derive a constraint on void formation by requiring unitarity.

To find the second Renyi entropy (2.11) for some finite interval A , we need to find the expected number of operators in I which fall into A as a function of time. The sharp light-cone growth of operators depicted in figure 3 makes the counting very simple [7, 15]. Let us denote the intersection of the past domain of dependence of region A with the $t = 0$ slice as $D(A)$. It then follows from the light-cone structure that all operators in $D(A)$ are contained within region A at time t , while an operator with a nontrivial part outside $D(A)$ at $t = 0$ will evolve into operators with nontrivial parts outside A . See figure 4. Thus only the initial operators in $D(A)$ contribute to $N_A(t)$, and each of them contributes a probability 1. $N_A(t)$ is then given by the number of basis operators in the set $I \cap D(A)$,

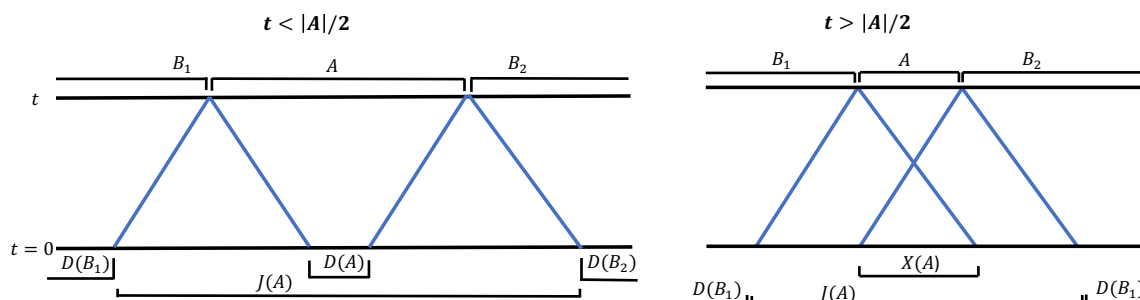


Figure 4. Left: $t < |A|/2$. Right: $t > |A|/2$. $D(A)$ is the intersection of the past domain dependence of A with the $t = 0$ slice, which becomes empty for $t > |A|/2$. $J(A)$ is the region at $t = 0$ which is causally connected with A . At $t < |A|/2$, every operator in $J(A)$ has zero probability of transitioning to a final operator contained in $\bar{A} = B_1 B_2$. At $t \geq |A|/2$, an operator in $J(A)$ can transition to an operator contained in \bar{A} if it develops a void in A .

which leads to

$$N_A(t) = \begin{cases} q^{|D(A)|} = q^{|A|-2t} & t < \frac{|A|}{2} \\ 1 & t > \frac{|A|}{2} \end{cases} \implies S_2^{(A)}(t) = \begin{cases} 2s_{\text{eq}} t & t < \frac{|A|}{2} \\ s_{\text{eq}} |A| & t > \frac{|A|}{2} \end{cases}, \quad s_{\text{eq}} \equiv \log q. \quad (2.17)$$

Physically, under evolution, initial basis operators localized in A move out at the lightcone speed in both directions. When an operator develops a nontrivial part outside A , it ceases to contribute to $N_A(t)$ and increases the entropy of A . At time t , the only operators remaining inside A are those which are initially localized in $D(A)$. When $t > \frac{|A|}{2}$, $D(A)$ is empty, with all nontrivial operators having evolved outside A . The only contribution to $N_A(t)$ is from the identity, and the entropy saturates. One can in fact show that with the sharp light-cone growth, all the Renyi and von Neumann entropies for a single interval A are also given by (2.17), and ρ_A is unitarily equivalent to a reduced density matrix in which the region $A - D(A)$ is maximally entangled with \bar{A} while $D(A)$ remains pure, as in the picture of the “entanglement tsunami” [16].

This discussion provides a simple explanation for the linear growth of entanglement entropy and its saturation, and shows that such behavior has its physical origin in ballistic operator growth, regardless of whether the system is chaotic or integrable.⁵ But at this stage, there is an apparent violation of unitarity. Applying the above discussion to $S_2^{(\bar{A})}(t)$, we obtain the same behavior as (2.17) with $|A|$ replaced by $|\bar{A}| = L - |A| \rightarrow \infty$, which is inconsistent with $S_2^{(A)}(t) = S_2^{(\bar{A})}(t)$ for $t \geq |A|/2$. That is, instead of growing indefinitely, under unitary evolution $S_2^{(\bar{A})}(t)$ must also saturate at $s_{\text{eq}}|A|$ for $t > |A|/2$.

The way out is as follows. Let us denote the region at $t = 0$ which is in causal contact with A as $J(A)$ (see figure 4), and consider an initial operator of the form $\mathcal{O}_{J(A)} \otimes \mathcal{O}_{D(\bar{A})}$ with $\mathcal{O}_{J(A)}$ a nontrivial operator in $J(A)$. In the discussion above, such an operator was assumed to have no contribution to $N_{\bar{A}}(t)$, as naively its time evolution will be nontrivial in A . But this is incorrect; such operators can contribute to $N_{\bar{A}}(t)$ if $\mathcal{O}_{J(A)}$ develops a void in A . Moreover, due to causality, the behavior of $\mathcal{O}_{J(A)}$ in region A under time evolution

⁵We will see some explicit examples for integrable systems in section 4.

— including the probability of developing a void — should be independent of $\mathcal{O}_{D(\bar{A})}$. This means that $N_{\bar{A}}(t)$ can be written in a factorized form

$$N_{\bar{A}}(t) = q^{|\bar{A}|-2t} N(\bar{A}, J(A); t), \quad (2.18)$$

where the factor $q^{|\bar{A}|-2t}$ is the number of basis operators in $I \cap D(\bar{A})$, and the function N was introduced in (2.12). When A is a single interval, it is clear from causality that

$$N(\bar{A}, J(A); t) = G(A, J(A); t) \quad (2.19)$$

where as defined in (2.13), $G(A, J(A); t)$ is defined as the expected number of initial basis operators in $J(A)$ that develop a void in A . For $S_2^{(A)} = S_2^{(\bar{A})}$, we need

$$G(A, J(A); t) = \begin{cases} 1 & t < \frac{|A|}{2} \\ q^{2t-|A|} & t > \frac{|A|}{2} \end{cases}. \quad (2.20)$$

The second line of (2.20) has the simple interpretation that the average probability of the $q^{|J(A)|} = q^{|A|+2t}$ basis operators in $J(A) \cap I$ to develop a void in A is $q^{-2|A|}$. Also note that $q^{2t-|A|} = q^{|X|}$, where the region X is as shown in figure 4. For an operator to develop a void in A , from causality it must be supported in region X . Note that while there are some similarities, equation (2.20) is different in nature from (2.16). Equation (2.16) is a purely kinematic statement, while (2.20) contains the dynamical input of the sharp light-cone growth and refers to an average over a more restricted set of initial operators.

Now consider the following ‘‘Renyi mutual information’’ between regions B_1 and B_2 in figure 4:

$$I_2(B_1, B_2) \equiv S_2^{(B_1)} + S_2^{(B_2)} - S_2^{(\bar{A})}, \quad \bar{A} = B_1 B_2. \quad (2.21)$$

From (2.18) and (2.19), we find that this quantity is fully controlled by the expected number of operators developing a void

$$I_2(B_1, B_2) = \log G(A, J(A); t) = \begin{cases} 0 & t < \frac{|A|}{2} \\ s_{\text{eq}}(2t - |A|) & t > \frac{|A|}{2} \end{cases}. \quad (2.22)$$

This result is intuitively appealing: when an operator develops a void in an interval A , it leads to mutual information between regions separated by A .

We stress that (2.20), and accordingly (2.22), are constrained by unitarity and should apply to any system, integrable or chaotic, which has sharp light-cone growth for the initial set of operators. We will see that (2.20) is indeed satisfied in various exactly solvable unitary circuit models in section 3 and section 4.

2.4 Mutual information and multi-partite entanglement

Here we explore further implications of void formation, as well as general constraints on this process, by examining the Renyi entropy for two and more disjoint intervals. Extending (2.22), we first show that the mutual information between two disjoint finite intervals is determined by the void formation function $G(B, \tilde{B}; t)$ for some appropriate regions B

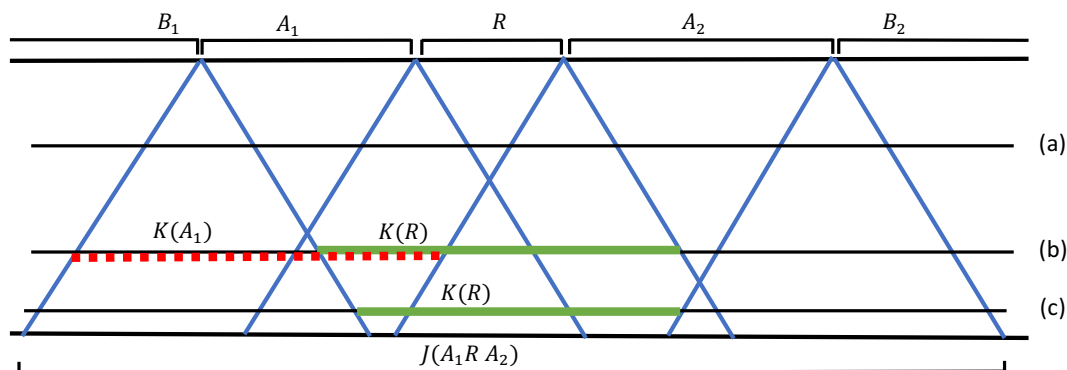


Figure 5. Different situations for two intervals. Different horizontal lines correspond to $t = 0$ slices for (a) $t < |R|/2$; (b) $t > |l_1|/2$; (c) $t > |l_2|/2$. Here we take $|R| < |A_1|$ and thus $l_1 = |A_1|$ and $l_2 = |A_2|$.

and \tilde{B} . In a certain region of the parameter space, the corresponding mutual information is universal, determined by (2.20) from unitarity. We also derive new constraints on void formation from unitarity of entanglement growth for two intervals. For more disjoint intervals, we see that void formation in operator growth leads to multi-partite entanglement, and void formation functions provide new measures for characterizing multi-partite entanglement.

Consider S_2 for a region $A = A_1A_2$ separated by an interval R . Without loss of generality, we can take $|A_1| \leq |A_2|$. For $t < |R|/2$, from causality, $N_A(t)$ of (2.11) factorizes into a product of the functions for A_1 and A_2

$$N_A(t) = N_{A_1}(t)N_{A_2}(t), \quad S_2^{(A)}(t) = S_2^{(A_1)}(t) + S_2^{(A_2)}(t), \quad t < \frac{|R|}{2} \quad (2.23)$$

where $N_{A_1}(t)$ denotes the contribution from initial operators in the region $D(A_1)$ and is given by (2.17). For $t > \frac{|R|}{2}$, initial operators in the region $K(R) = J(R) \cap D(A_1RA_2)$ (see figure 5) can now potentially form a void in region R and thus contribute to $N_A(t)$

$$N_A(t) = N_{A_1}(t)N_{A_2}(t)G(R, K(R); t), \quad (2.24)$$

which gives

$$I_2(A_1, A_2; t) = \log G(R, K(R); t). \quad (2.25)$$

For further discussion, it is convenient to introduce the following notation:

$$l_1 \equiv \max(|A_1|, |R|), \quad l_2 \equiv \max(|A_2|, |R|), \quad K(A_1) \equiv J(A_1) \cap D(B_1A_1R). \quad (2.26)$$

Note that when $t < |R|/2$, we have the factorized form (2.23) in any theory with sharp light-cone growth. For $|A_1|/2 > t > |R|/2$ (which can happen for $|A_1| > |R|$), we have $K(R) = J(R)$, and from (2.20)

$$I_2(A_1, A_2) = s_{\text{eq}}(2t - |R|). \quad (2.27)$$

For $t > \frac{|A_1+|R|}{2}$, $K(R)$ becomes an empty set, and thus $G(R, K(R); t) = 1$. So for $t < l_1/2$ and for $t > \frac{|A_1+|R|}{2}$, $G(R, K(R); t)$ and $I_2(A_1, A_2)$ have a universal form in all systems

with sharp light-cone growth.⁶ For $\frac{|A|+|R|}{2} > t > l_1/2$, $K(R) \subset J(R)$, the behavior of $G(R, K(R); t)$ becomes system-dependent, and so does $I_2(A_1, A_2)$. Now consider the entropy for the region $\bar{A} = B_1RB_2$, for which

$$N_{\bar{A}}(t) = N_{B_1}(t)N_{B_2}(t)N(\bar{A}, J(A_1RA_2); t) . \tag{2.28}$$

Similar to the discussion immediately above, for $t \leq l_1/2$, the constraint (2.20) is enough to ensure $S_2^{(A)} = S_2^{(\bar{A})}$, but for $t > l_1/2$ new constraints arise. We find (see also figure 5)

$$\frac{l_1}{2} < t \leq \frac{l_2}{2} : \quad G(R, K(R); t) = q^{|A_1|-|R|}G(A_1, K(A_1); t) \tag{2.29}$$

$$t > \frac{l_2}{2} : \quad G(R, K(R); t) = q^{|A|-|R|-2t}N(\bar{A}, J(A_1RA_2); t), \tag{2.30}$$

For $t > \frac{|A|+|R|}{2}$, with $G(R, K(R); t) = 1$, we must have

$$N(\bar{A}, J(A_1RA_2); t) = q^{2t+|R|-|A|}, \quad t > \frac{|A| + |R|}{2} . \tag{2.31}$$

If we take A_2 to be the entire semi-infinite region to the right of R , then both regions $K(A_1)$ and $K(R)$ appearing in (2.29) and (2.30) depend only on A_1 and R , and $l_2/2 \rightarrow \infty$, so (2.29) holds for all $t > l_1/2$. We can thus deduce from (2.29) a general constraint on the void formation functions for any two adjacent finite intervals A and B in an infinite system,

$$q^{|A|}G(A, K_l(A, B); t) = q^{|B|}G(B, K_r(A, B); t), \quad t > \max(|A|/2, |B|/2) \tag{2.32}$$

where

$$K_l(A, B) = J(A) \cap D(L_1AB), \quad K_r(A, B) = J(B) \cap D(ABL_2) \tag{2.33}$$

where L_1, L_2 are the semi-infinite regions to the left of A and to the right of B respectively. The different regions appearing in (2.32) are shown in figure 6.

The above discussion can be generalized to express the second Renyi entropy for any number of intervals in terms of appropriate void formation functions. On the one hand, by requiring $S_2^{(A)}(t) = S_2^{(A)}(t)$ for A consisting of an arbitrary number of intervals, one can obtain further constraints on the void formation functions. On the other hand, with the full knowledge of the void formation functions, one should be able to deduce the expression for $S_2^{(A)}$ for any A . We will see an explicit example of this in section 3.3.

Now consider a region consisting of n disjoint intervals $A = A_1 \cdots A_n$ separated by intervals $R_1, \cdots R_{n-1}$, as in figure 7. Then the void formation function $G(A, Q; t)$ gives a contribution to the entropy of the region $\bar{A} = R_0R_1 \cdots R_n$, but does not contribute to the entropy of any region consisting of a proper subset of $\{R_0, \cdots, R_n\}$. Thus void formation in A leads to multi-partite entanglement among all disconnected regions in \bar{A} , which can be captured by the quantity $G(A, Q; t)$.

⁶In the discussion of [17] of Renyi entropies for two-dimensional conformal field theories (CFTs), these are indeed the regimes which are universal for all CFTs.

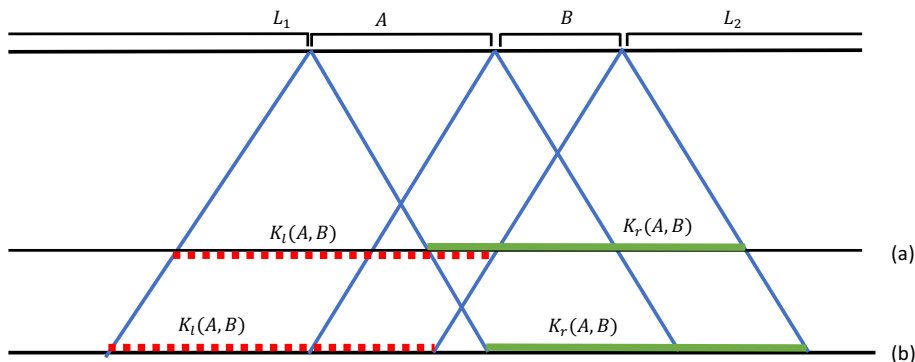


Figure 6. Regions $K_l(A, B)$ and $K_r(A, B)$ of (2.32) at a time $t < (|A| + |B|)/2$ in (a), and $t > (|A| + |B|)/2$ in (b). The condition (2.32) holds at all times $t > \max(|A|/2, |B|/2)$.

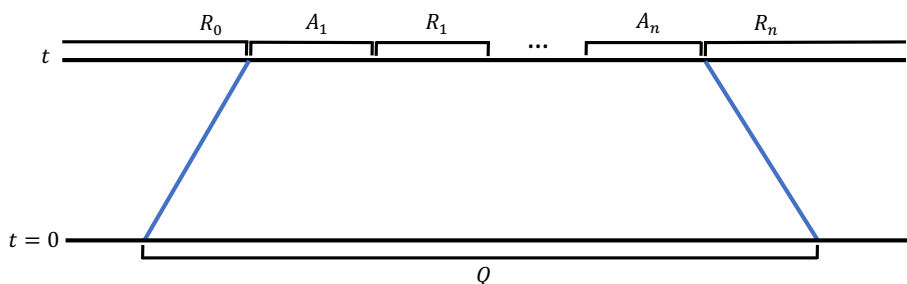


Figure 7. When initial operators in Q develop a void in $A = A_1 \cdots A_n$, this leads to multipartite entanglement among regions R_0, R_1, \dots, R_n . Here Q is $J(A_1 R_1 \cdots R_{n-1} A_n)$.

3 Random void distribution and entanglement growth

In this section, we consider the probability distribution of void formation for a generic operator in the random unitary circuit model in the large q limit. We show that it is given by the random void distribution (1.3). We then show that by assuming the random void distribution for all initial operators, we can correctly obtain the full expression for $S_2^{(A)}$ in the random circuit model for A consisting of an arbitrary number of disjoint intervals. Surprisingly, the resulting expression is found to coincide with the von Neumann entropy for holographic systems. In the next section, we will contrast the random void distribution with the void formation properties of two Clifford circuit models (which may be seen as non-chaotic systems).

3.1 Random unitary circuits

We first describe briefly the setup of the random unitary circuit discussed in [8–10], and its main properties. Consider a time-evolution of the system as in figure 8, where the evolution operator $U(t)$ can be written as

$$U(t) = U_t U_{t-1} \cdots U_0, \tag{3.1}$$

$$U_0 = \cdots \otimes U_0^{0,1} \otimes U_0^{2,3} \otimes \cdots, \quad U_1 = \cdots \otimes U_1^{-1,0} \otimes U_1^{1,2} \otimes \cdots. \tag{3.2}$$

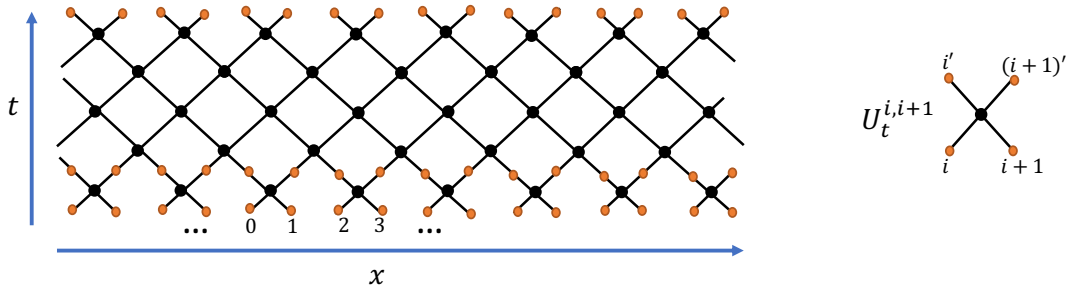


Figure 8. Unitary circuit for the time-evolution operator $U(t)$ (3.1)–(3.2) for $t = 5$. The black circles represent unitary matrices $U_t^{i,i+1}$ acting on two adjacent sites i and $i + 1$. In the random circuit model, $U_t^{i,i+1}$ for each i and t is independently drawn from the Haar ensemble of $q^2 \times q^2$ unitary matrices. Here in the first step of the time-evolution, we couple sites 0 and 1, 2 and 3, and so on, and in the second step we couple -1 and 0, 1 and 2, and so on. We explicitly show lattice sites of the system at $t = 0$ and $t = 1$ with the orange circles.

Here we take discrete time steps, with U_n corresponding to the evolution operator at the n -th step. $U_n^{i,i+1}$ is a $q^2 \times q^2$ unitary matrix that acts on two neighboring sites i and $i + 1$ at n -th step, and each such matrix is averaged over the Haar measure independently. We will consider the limit of large q .

The random unitary circuit can be seen as a discretized model for chaotic quantum systems. It is manifestly unitary and local, but replaces the local interactions between neighboring spins in a realistic Hamiltonian system by random ones. It provides a powerful playground for studying chaotic systems, as many observables such as entanglement entropy, OTOCs, and operator spreading coefficients are analytically calculable, and the resulting behavior has been found to be consistent with numerical results of realistic chaotic spin-chain systems [8, 9, 18–20].

Here are some features of the random circuit model in the large q limit which are relevant for our discussion [8–10]:

1. For a basis operator \mathcal{O}_α (introduced above (2.3)) with right and left endpoints given by x_r and x_l respectively, one finds

$$\sum_{\substack{\beta \text{ with left endpoint} \\ x_l-t, \text{ right endpoint } x_r+t}} \overline{|c_\alpha^\beta(t)|^2} = 1 + O(1/q) \quad (3.3)$$

which implies that the end points of any operator \mathcal{O} move under time evolution in opposite directions with light cone speed 1 as in figure 3. Here and below, an overline denotes an average over the random unitaries.

2. The calculation of the second Renyi entropy $S_2^{(A)}$ for a region A can be reduced to computing the partition function of a classical Ising model on a triangular lattice. For A consisting of a single interval, it has the form (2.17), consistent with the general argument presented earlier.

3. In the large q limit,

$$e^{-\overline{S_2^{(A)}}} = \overline{e^{-S_2^{(A)}}} = \frac{1}{q^{|A|}} \overline{N_A(t)} \quad (3.4)$$

where $N_A(t)$ was defined in (2.11). Here the off-diagonal terms in (2.10) automatically vanish due to random averages. Furthermore, one has

$$\overline{S_2^{(A)}} = \overline{S_n^{(A)}} = \overline{S^{(A)}}, \quad \forall n > 2, \quad n \in \mathbb{N} \quad (3.5)$$

where $S^{(A)}$ denotes the von Neumann entropy. So our discussion below about S_2 can also be understood as being relevant for the von Neumann entropy.

3.2 Random void distribution

We now look at the probability distribution for a generic operator \mathcal{O} to develop a void in some designated region A in the random circuit model at large q .

We can expand \mathcal{O} in terms of basis operators as

$$\mathcal{O} = \sum_{\alpha} a_{\alpha} \mathcal{O}_{\alpha}, \quad \mathcal{O}(t) = \sum_{\alpha} \sum_{\beta} a_{\alpha} c_{\alpha}^{\beta}(t) \mathcal{O}_{\beta}, \quad \sum_{\alpha} |a_{\alpha}|^2 = 1. \quad (3.6)$$

Then under time evolution, the probability for \mathcal{O} to have a void in A is

$$\overline{P_{\mathcal{O}}^{(A)}(t)} = \sum_{\beta \text{ with void in } A} \sum_{\alpha_1, \alpha_2} a_{\alpha_1} a_{\alpha_2}^* \overline{c_{\alpha_1}^{\beta}(t) c_{\alpha_2}^{*\beta}(t)} = \sum_{\alpha} |a_{\alpha}|^2 \overline{P_{\mathcal{O}_{\alpha}}^{(A)}(t)} \quad (3.7)$$

where $P_{\mathcal{O}_{\alpha}}^{(A)}$ was introduced in (2.6), and in the second equality the off-diagonal terms drop out due to the random average. In (3.7) by “ β with void in A ” we mean \mathcal{O}_{β} should be trivial in A and have support in each disconnected part of \bar{A} . For instance, in the case of $A = A_1 A_2 \cdots A_n$ in figure 7, \mathcal{O}_{β} should be the identity in A while being nontrivial in each of R_i 's. From now on, for notational simplicity we will suppress the explicit overline for averages.

$P_{\mathcal{O}_{\alpha}}^{(A)}(t)$ can be expressed as the partition function of a classical Ising model on a triangular lattice with boundary conditions specified by \mathcal{O}_{α} and A . We present the details of the calculation in appendix A.2. For any operator \mathcal{O}_{α} which does not have an initial void, the final result is the random void distribution (RVD) already mentioned in the Introduction section

$$P_{\mathcal{O}_{\alpha}}^{(A)}(t) = \begin{cases} e^{-2s_{\text{eq}}|A|} = \frac{1}{d_A^2} & t \geq \max\left(\frac{|A_i|}{2}, \forall i\right), A \in J^+(\mathcal{O}_{\alpha}) \\ 0 & \text{otherwise} \end{cases} \quad (3.8)$$

where we have taken $A = A_1 A_2 \cdots A_n$ with $|A| = |A_1| + \cdots + |A_n|$, $J^+(\mathcal{O}_{\alpha})$ denotes the region at time t which is causally connected to \mathcal{O}_{α} , and $s_{\text{eq}} = \log q$. See figure 9. That to have a nonzero result we must have $A \in J^+(\mathcal{O}_{\alpha})$ follows simply from causality. If any of the A_i are not in $J^+(\mathcal{O}_{\alpha})$, clearly \mathcal{O}_{α} cannot evolve into an operator which is nontrivial in all disconnected parts of \bar{A} .

Operators with initial voids can in general evolve to final operators that have support in all disconnected parts of \bar{A} by two distinct kinds of processes. We can have processes

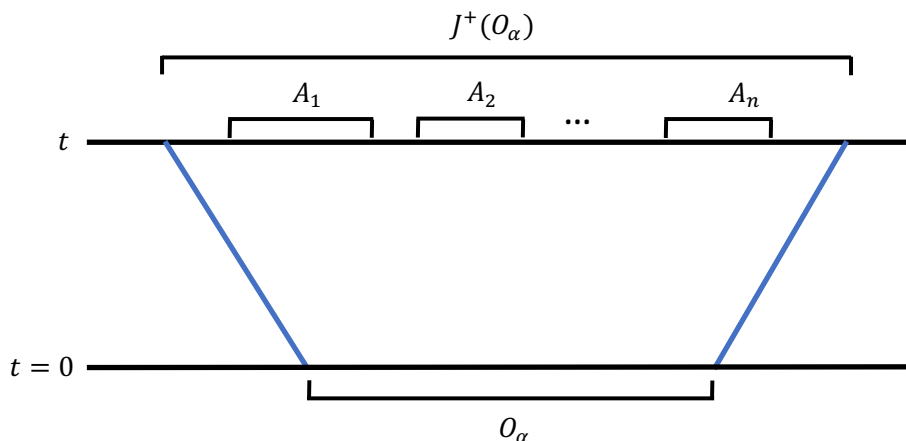


Figure 9. Regions $J^+(\mathcal{O}_\alpha)$ and A_1, A_2, \dots, A_n . In the random void distribution, the probability of forming a void in $A \equiv A_1 A_2 \dots A_n$ depends only on the lengths $|A_1|, \dots, |A_n|$.

where at some intermediate stage, the operator evolves to an operator without a void (if the initial voids are denoted by G_1, \dots, G_m , such processes are allowed by causality only at $t > |G_i|/2$ for all i), which then splits to form a void in A . Another kind of process is one where different disconnected parts of the initial operator evolve to different disconnected parts of \bar{A} without interacting with one another. Depending on the sizes and locations of the G_i , either of these types of processes can give the dominant contribution to $P_{\mathcal{O}_\alpha}^{(A)}(t)$ in the large q limit. When the leading contribution is given by the former process, which can physically be seen as one of genuine “void formation”, we still have (3.8). When one of the latter types of processes is dominant, $P_{\mathcal{O}_\alpha}^{(A)}(t)$ is enhanced compared to (3.8). Nevertheless, for a generic operator (3.6), the part of the operator corresponding to \mathcal{O}_α with an initial void is suppressed due to a much smaller phase space. One can show that the enhancement from disconnected processes never overcomes the suppression. See appendix A.2. Given that (3.8) is independent of \mathcal{O}_α other than through the causality constraint $A \in J^+(\mathcal{O}_\alpha)$, a generic operator \mathcal{O} also satisfies (3.8).

Note that if we assumed that for a given initial operator, the probability of having any of the q^2 basis operators at any site between the endpoints of the operator at time t is the same, then the probability $q^{-2|A|}$ would immediately follow: having a void in A corresponds to fixing the operators at $|A|$ sites in the final operator to be one of q^2 options, while allowing the remaining operators to take any value. A similar ergodicity assumption was used in the operator growth model introduced in [7]. Equation (3.8) also applies to the random circuit model at a finite q in the regime that region A and t are large (so that $q^{|A|}, q^t$ are large), as we will discuss elsewhere.

From (3.8) one can obtain the void formation function (2.13) for any regions A and B . In the simplest case with A consisting of a single interval lying within the future light cone of a region B , we have

$$G_{\text{RVD}}(A, B; t) = \begin{cases} 1 & |B| \leq 2|A| \text{ or } t \leq \frac{|A|}{2} \\ e^{s_{\text{eq}}(|B|-2|A|)} & |B| > 2|A| \text{ and } t > \frac{|A|}{2} \end{cases} . \quad (3.9)$$

We can check that the above expression satisfies the unitarity constraint (2.20) by taking $B = J(A)$, so that $|B| = |A| + 2t$. We can similarly check that it satisfies the constraints (2.29) and (2.30).

For the Renyi entropy of two intervals (2.24)–(2.25), we need $G(R, K(R); t)$ with $K(R) = J(R) \cap D(A_1 R A_2)$, whose behavior depends on the relative sizes of $|A_1|$, $|A_2|$ and R . For example, for $|R| < |A_1|$ we find that

$$G(R, K(R); t) = \begin{cases} 1 & t \leq \frac{|R|}{2} \text{ or } t > \frac{|A_1+|A_2|-|R|}{2} \\ q^{2t-|R|} & \frac{|A_1|}{2} \geq t \geq \frac{|R|}{2} \\ q^{|A_1|-|R|} & \frac{|A_2|}{2} \geq t \geq \frac{|A_1|}{2} \\ q^{|A_1+|A_2|-2t-|R|} & \frac{|A_1+|A_2|-|R|}{2} \geq t \geq \frac{|A_2|}{2} \end{cases}, \quad (3.10)$$

which upon using (2.24) leads to

$$S_2^{(A)}(t) = s_{\text{eq}} \begin{cases} 4t & t \leq \frac{|R|}{2} \\ |R| + 2t & \frac{|A_1+|A_2|-|R|}{2} \geq t \geq \frac{|R|}{2} \\ |A_1| + |A_2| & t > \frac{|A_1+|A_2|-|R|}{2} \end{cases}. \quad (3.11)$$

For $|A_1| < |R|$ we find that

$$G(R, K(R); t) = 1, \implies S_2^{(A)}(t) = S_2^{(A_1)}(t) + S_2^{(A_2)}(t). \quad (3.12)$$

The mutual information between A_1 and A_2 is thus nonzero only for the case $|R| < |A_1|$. One can also obtain $S_2^{(A_1 A_2)}$ using the partition function method of appendix A.3, and one finds agreement with (3.11)–(3.12).

An alert reader may notice that the expressions (3.11)–(3.12) coincide with the expressions for the evolution of entanglement entropy after a global quench in holographic systems [21, 22]. We will see below that this is not an accident; the results agree for any number of intervals.

To conclude this subsection, we note that if there exists some B for which $G(A, B; t)$ does not have the value in (3.9) for some A in the light-cone of B , then we can always construct intervals A_1 and A_2 for which $S_2^{(A_1 A_2)}(t)$ deviates from (3.11)–(3.12). See figure 10. Thus $S_2^{(A_1 A_2)}(t)$ agrees with the holographic result for all t and all A_1, A_2 if and only if (3.9) is satisfied.

3.3 Random void distribution and maximal entanglement growth

Consider a region $A = A_1 \cdots A_n$ consisting of n intervals $A_1 = [l_1, r_1], \dots, A_n = [l_n, r_n]$, separated by intervals R_1, \dots, R_{n-1} , as in figure 7. The entanglement entropy $S_2^{(A)}$ can be calculated using the partition function method, and is discussed in detail in appendix A.3. The result can be written as

$$S_2^{(A)}(t) = s_{\text{eq}} \min_{\{\gamma\}(t)} \left[n_\gamma t + \sum_{\{l_i, r_j\} \in \gamma} |l_i - r_j| \right] \quad (3.13)$$

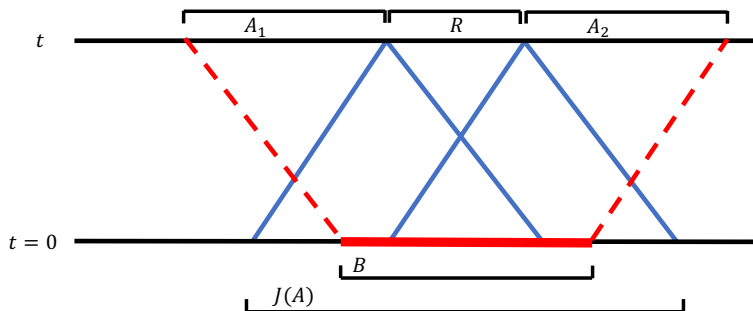


Figure 10. For any B in $J(R)$ such that R is in the light-cone of B , we can construct A_1, A_2 such that $K(R) \equiv D(A_1 R A_2) \cap J(R) = B$. Then if $G(R, B, t)$ does not agree with (3.9), then using (2.24), we see that the resulting $S_2^{(A_1 A_2)}(t)$ will deviate from (3.11)–(3.12), and thus from the holographic expression.

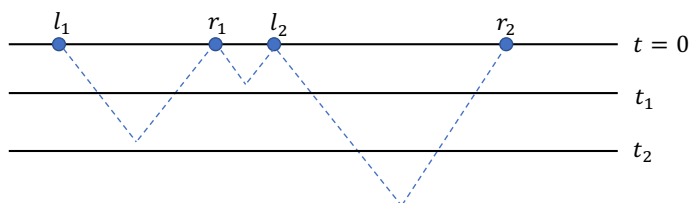


Figure 11. We show an example to explain the notation of equation (3.13). At time t_1 , the only connectable pair is $\{l_2, r_1\}$, whereas at time t_2 , $\{l_2, r_1\}$ and $\{l_1, r_1\}$ are both connectable pairs. The allowed configurations in the set $\{\gamma\}(t_2)$ are $\gamma_1 = \{l_1, r_2, \{l_2, r_1\}\}$, $\gamma_2 = \{l_2, r_2, \{l_1, r_1\}\}$, and $\gamma_3 = \{l_1, r_1, l_2, r_2\}$.

where the set $\{\gamma\}(t)$ and n_γ are defined as follows. Call any pair of left and right endpoints $\{l_i, r_j\}$ from the set of all endpoints $\{l_1, r_1, \dots, l_n, r_n\}$ a connectable pair at time t if $t > |l_i - r_j|/2$. We can then group together the elements of $\{l_1, r_1, \dots, l_n, r_n\}$ into different configurations such that every element is either unconnected to any other point, or part of one connected pair. Only connectable pairs can be connected. $\{\gamma\}(t)$ is the set of all such configurations γ at time t , and n_γ is the number of unconnected points in the configuration γ . Each connected pair $\{l_i, r_j\}$ in a configuration γ contributes $|l_i - r_j|$, while each unconnected point contributes t , so that we get (3.13). The notation is explained further with the help of an example in figure 11.

Extending our discussion in section 3.2 for two intervals, one can show that the same expression (3.13) follows by using only the following elements: (i) sharp light-cone growth; (ii) the random void distribution (3.8) for all initial operators; (iii) large q . The derivation is given in appendix B.

The expression (3.13) can also be shown to be equivalent to the expression of entanglement entropy (in a scaling regime) for holographic systems after a quench. Holographic systems are a certain class of strongly coupled $(1+1)$ -dimensional conformal field theories (CFT) which have a gravity dual. In the large central charge limit, their entanglement entropy can be calculated using classical gravity. More explicitly, in the regime with $t, |A_1|, \dots, |A_n|, |R_1|, \dots, |R_{n-1}|$ large while $|A_i|/t, |R_i|/t$ are fixed, the holographic

entanglement entropy after a global quench has the following form (see e.g. [11])

$$S_{\text{hol}}(t) = \min_{\sigma} \left[\sum_{i=1}^n S_{\text{interval}}(t, |l_i - r_{\sigma(i)}|) \right] \quad (3.14)$$

where

$$S_{\text{interval}}(t, R) = \begin{cases} 2s_{\text{eq}}t & t < \frac{R}{2} \\ s_{\text{eq}}R & t \geq \frac{R}{2} \end{cases} \quad (3.15)$$

and σ are permutations of $\{1, \dots, n\}$, and s_{eq} is the equilibrium entropy density. The contribution we get from a permutation σ at time t on the right-hand side of (3.14) is equal to the contribution we get on the r.h.s. of (3.13) from a configuration $\gamma(\sigma)$ where we first pair each l_i with $r_{\sigma(i)}$, and then connect l_i and $r_{\sigma(i)}$ if they are connectable. The set of $\gamma(\sigma)$ for all σ at time t is a subset of $\{\gamma\}(t)$ defined below (3.13). We can obtain the full set $\{\gamma\}(t)$ if for each choice of σ , in addition to $\gamma(\sigma)$, we include configurations where any number of the connectable pairs of the type $\{l_i, r_{\sigma(i)}\}$ are not connected. But disconnecting a pair of connected points in a configuration while leaving other points unchanged always increases the contribution from that configuration in (3.13), so any element in $\{\gamma\}(t)$ which cannot be obtained as $\gamma(\sigma)$ for any σ at time t gives a larger contribution than some $\gamma(\sigma)$. Thus, the minimum value in (3.13) is the same as the minimum in (3.14).

We can check that a minimal configuration in (3.13) only involves connections between adjacent endpoints. So we get the same result if we restrict the definition of connectable points below (3.13) to adjacent endpoints l_i and r_j such that $t > |l_i - r_j|/2$.

As the number of intervals increases, equation (3.14) (or equivalently (3.13)) gives rise to intricate patterns of time-dependence when the relative sizes of the intervals $|A_i|$ and their separations $|R_i|$ are varied. It is remarkable that these patterns can be reproduced by the extremely simple underlying rules of sharp light-cone growth and the random void distribution. We note, however, that while for the random unitary circuits, S_2 coincides with the von Neumann entropy in the large q limit, this is no indication that this is true for holographic systems in the large c limit.⁷ So the entanglement spectrum of holographic systems cannot be fully approximated by random unitary circuits, and while it is natural to expect that the random void distribution should play some role in holographic systems, it cannot be the full story.

In [11], it was shown using the strong subadditivity condition that the holographic expression (3.14) in fact maximizes the entanglement growth for an arbitrary number of intervals among all $(1+1)$ -dimensional systems with a strict light cone. Thus we find that the random void distribution (3.8) together with sharp light-cone growth maximizes entanglement growth (recall that S_2 is upper-bounded by the von Neumann entropy). In the two-interval case, this statement implies that any system with $S_2(t)$ not equal to the holographic result must have $S_2(t)$ smaller than (3.11) and (3.12). Using (2.24), this implies

⁷ S_2 for a different configuration (two offset intervals in a thermal field double state) was calculated in [17] in a holographic system, and found to be different from the von Neumann entropy. S_2 in this setup can be calculated for random unitary circuits and is found to agree with the von Neumann entropy, but not with S_2 of holographic systems.

that

$$G(A, B; t) \geq G_{\text{RVD}}(A, B; t) \tag{3.16}$$

where $G_{\text{RVD}}(A, B; t)$ is the random void distribution (RVD) expression (3.9).

4 Void formation in two Clifford circuit models

As contrasts to the random void distribution, we now consider the void formation structure of two other examples of unitary circuits: (i) a free propagating model in which entanglement can only be spread, but not created, which may thus be considered a proxy for free theories; (ii) a circuit built from perfect tensors, which may be considered a model for non-chaotic systems, as while it can generate entanglement in certain initial product states, like all Clifford circuits it does not lead to growth of operator entanglement, and also does not have the form of the out-of-time-ordered correlator (OTOC) expected in chaotic systems [8, 9]. Both models are special examples of Clifford circuits [23–25], a class of unitary circuits where under time evolution, a basis operator \mathcal{O}_α transitions to another basis operator.

More explicitly, consider a unitary circuit defined by (3.1)–(3.2) and figure 8, where now each $U_t^{i,i+1} = \tilde{U}$, and \tilde{U} is a (fixed) unitary matrix that evolves each basis operator in $\mathcal{G}_i \otimes \mathcal{G}_{i+1}$ to some basis operator. Under this time-evolution, for a given \mathcal{O}_α , $c_\alpha^\beta(t) = 1$ for a single β , and $c_\alpha^{\beta'}(t) = 0$ for all $\beta' \neq \beta$. Since the evolution is unitary, there is a one-to-one mapping from initial operators \mathcal{O}_α to final operators \mathcal{O}_β .

To study entanglement growth, now instead of generic homogeneous pure product states, we have to consider a more restricted set, as these Clifford circuit models do not generate entanglement in an arbitrary initial pure product state. Furthermore, initial basis operators in the entire system can in general both grow and decrease in size under the action of the circuit U . The set of initial states we look at are again of the form

$$\rho_0 = \frac{1}{q^L} \sum_{a \in I} \mathcal{O}_a \tag{4.1}$$

where the set I consists of q^L basis operators, which are in general no longer just tensor products of powers of Z_i . In each of the models, we will choose ρ_0 such that the end points of all basis operators in the associated I move outwards with $v = 1$ under time-evolution.

As before, the entanglement growth can be obtained from (2.9)–(2.13). Note that for Clifford circuits the off-diagonal terms in (2.10) vanish identically due to the one-to-one mapping between initial and final basis operators. The sharp light-cone growth of operators in I again implies that $S_2^{(A)}(t)$ is given by (2.17) for a single interval.

In Clifford circuits, we cannot have a void distribution like in (3.8) for individual initial operators \mathcal{O}_α , as the probability of going to any final operator is either 0 or 1. Thus the functions $N_A(t)$, $N(A, B; t)$ and $G(A, B; t)$ defined in (2.11)–(2.13) are now the numbers (rather than the expectation values of the numbers) of initial operators in I satisfying a given property. It is instructive to contrast void formation functions $G(A, B; t)$ in these models with those from the random void distribution, and see how they lead to different entanglement growth.

Before discussing the models explicitly, here we summarize some common features, which are also shared by random unitary circuits in the large q limit:

1. While their void formation structure is very different from the random void distribution, we will see the corresponding void formation functions nevertheless satisfy the unitarity constraints (2.20) and (2.29)–(2.31).
2. The unitarity constraint (2.20) is satisfied in the following way. At $t > |A|/2$, take any of the $q^{2t-|A|}$ basis operators in $I \cap X(A)$, say $\tilde{\mathcal{O}}$, where $X(A)$ is the region of length $2t - |A|$ in the center of $J(A)$ (shown in figure 4). Defining \tilde{I} as the set of all initial basis operators in $I \cap J(A)$ which are equal to $\tilde{\mathcal{O}}$ in $X(A)$, one finds that

$$\sum_{\alpha \in \tilde{I}, \beta \text{ trivial in } A} |c_{\alpha}^{\beta}(t)|^2 = 1 . \tag{4.2}$$

In the Clifford circuits, we can interpret this as the fact that any choice of initial operator within the region $X(A)$ is consistent with evolving to a final operator which is equal to the identity in A . Since the total number of operators in $I \cap X(A)$ is equal to $q^{2t-|A|} = G(A, J(A); t)$, this also means that when we fix the part of the initial operator within $X(A)$, the initial operator in the entire region $J(A)$ which can evolve to the identity is fully determined.

It is tempting to speculate that given sharp light-cone growth, equation (4.2) is true in all unitary systems.

4.1 Free propagation model

In this model, the two-site unitary matrices in the circuit are given by [11]

$$\tilde{U}_{a_1 a_2, b_1 b_2} = \delta_{a_1 b_2} \delta_{a_2 b_1} \tag{4.3}$$

which is a discrete version of the quasiparticle models for entanglement growth proposed in [27]. \tilde{U} takes a product state to a product state at all times, but can spread the entanglement to large distances if we consider an initial state

$$|\psi_0\rangle = \dots \otimes \left(\frac{1}{\sqrt{q}} \sum_{n=0}^{q-1} |n\rangle_0 \otimes |n\rangle_1 \right) \otimes \left(\frac{1}{\sqrt{q}} \sum_{n'=0}^{q-1} |n'\rangle_2 \otimes |n'\rangle_3 \right) \dots \tag{4.4}$$

which has short-range entanglement between adjacent pairs of sites.

The evolution of basis operators⁸ in this model has a simple form: the basis operators at different sites evolve independently from each other, and all operators that are initially at an even (odd) site move to the right (left) at speed 1, so that at an odd time t , for an initial operator $\mathcal{O} \equiv \otimes_i O_i$,

$$\mathcal{O}(t) = \otimes_{i=\text{even}} O_{i+t} \otimes_{i=\text{odd}} O_{i-t} \tag{4.5}$$

⁸We use the following conventions to avoid complications due to lattice effects which are not relevant in the continuum limit. All spatial regions we consider have their left endpoint at an even site and their right endpoint at an odd site. We always consider times at which an odd number of layers of unitaries have been applied.

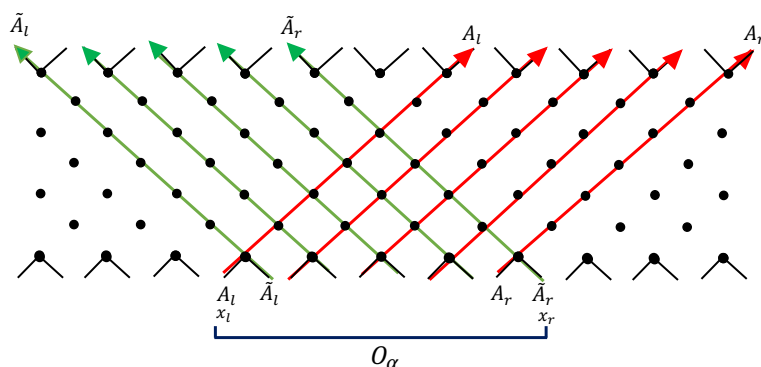


Figure 12. Growth of an operator \mathcal{O}_α in I . Any operator in I has its left endpoint at an even site x_l and its right endpoint at an odd site x_r . The evolution of operators follows (4.5) with operators at even (odd) sites moving right (left), which are shown respectively using red and green. At time t , the left and right endpoints are at $x_l - (t - 1)$ and $x_r + (t - 1)$ respectively (for sufficiently large times we have $t - 1 \approx t$). The resulting operator has a central void of length $2t - L_i$, where L_i is the length of the initial operator, as well as other voids of length 1. The number of sites with non-trivial operators remains fixed as a function of time.

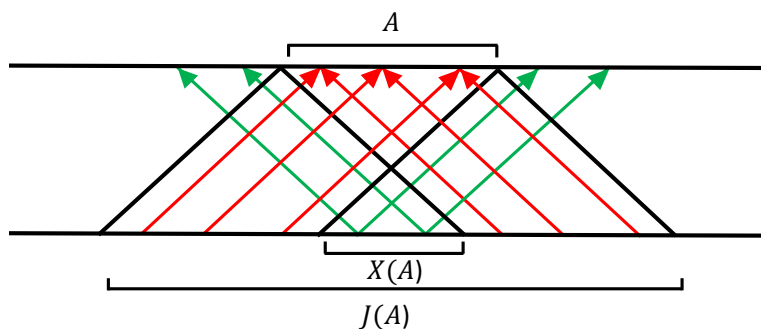


Figure 13. All operators from the region $X(A)$ propagate to \bar{A} . Any operator non-trivial in $J(A) - X(A)$ has a part that propagates to A .

The density matrix for (4.4) is of the form (4.1), with I the set of operators which can have any basis operator O_i at an even site i , but at site $i + 1$ have some fixed \tilde{O}_{i+1} determined by O_i . For example, if $O_i = X^m Z^n$, then $\tilde{O}_{i+1} = X^m Z^{q-n}$. The time evolution of basis operators in I can be readily obtained from (4.5), see figure 12 for an illustration.

In figure 13, we see that all initial operators from sites in $X(A)$ propagate to \bar{A} at time t , while operators on sites in the region $J(A) - X(A)$ propagate to A . An operator \mathcal{O}_α in I becomes trivial in A if and only if it is of the form $\mathcal{O}_\alpha = \tilde{O}_{X(A)} \otimes \mathbf{1}_{J(A)-X(A)}$, where $\tilde{O}_{X(A)}$ can be any operator in $I \cap X(A)$. Thus the statements (2.20) and (4.2) are both true in this model.

Now look at the form of $G(A, B; t)$. Since any initial operator in $J(A) \cap I$ that becomes trivial in A must be contained within $X(A)$, the number of basis operators in I contained in a subset B of $J(A)$ that develop a void in A is given by

$$G_{\text{free}}(A, B; t) = q^{|B \cap X(A)|} . \tag{4.6}$$

Note that this is very different from the form (3.9) from the random void distribution. In particular, while (3.9) depends only on the length of B and not on its position in $J(A)$, (4.6) is sensitively dependent on the position of B in $J(A)$. Moreover, equation (4.6) can be greater than 1 when $|B| < 2|A|$, if B has some overlap with $X(A)$. When applied to the entanglement entropy for two intervals using (2.24), such behavior can lead to $S_2^{(A)}(t)$ behaving non-monotonically in time,⁹ whereas (3.11)–(3.12) are non-decreasing.

One can also see that, as anticipated from (3.16), the expression (4.6) is always greater than (3.9)

$$G_{\text{free}}(A, B; t) \geq G_{\text{RVD}}(A, B; t), \quad \forall A, B. \quad (4.7)$$

Equation (4.7) is equivalent to $|B \cap X(A)| \geq |B| - 2|A|$, which is always true as $|B| - |B \cap X(A)| \leq J(A) - X(A) = 2|A|$ (see figure 13). It can also be checked the constraints (2.29)–(2.31) obtained from unitarity are satisfied in this model.

4.2 Perfect tensor model

In this model, which was previously considered in [12], the Hilbert space at each site has dimension 3 (q below should be understood as being 3), with a basis $\{|0\rangle, |1\rangle, |2\rangle\}$, and \tilde{U}^\dagger acts on the Hilbert space as:

$$\tilde{U}^\dagger |i\rangle \otimes |j\rangle = |-i - j\rangle \otimes |j - i\rangle, \quad (4.8)$$

with addition defined modulo 3. \tilde{U} is a perfect tensor, that is, any balanced bipartition of its indices into inputs and outputs gives a unitary transformation. The perfect tensor model does not generate entanglement in every initial pure product state: for example, states $\dots |0\rangle \otimes |0\rangle \dots$ and $\dots |\psi_3\rangle \otimes |\psi_3\rangle \dots$ remain invariant under the action of U . We will consider an initial state of the form

$$|\psi\rangle = \dots \otimes |\psi_3\rangle_0 \otimes |0\rangle_1 \otimes |\psi_3\rangle_2 \otimes |0\rangle_3 \dots \quad (4.9)$$

where $|\psi_3\rangle \equiv \frac{1}{\sqrt{3}} \sum_{k=0}^2 |k\rangle$, for which I is the set of basis operators with powers of X on even sites and powers of Z on odd sites.

From (4.8) one finds that acting on operators in the basis (2.2), \tilde{U} sends any basis operator on two sites to another basis operator,

$$\tilde{U}^\dagger (b_1 \otimes b_2) \tilde{U} = b'_1 \otimes b'_2, \quad (4.10)$$

and has the following properties:

1. It takes any operator with a power of X on site i and a power of Z on site $i + 1$ to an operator which is non-trivial on both i and $i + 1$.
2. It takes any basis operator non-trivial on a single site to a basis operator non-trivial on both sites.
3. If we know any two of b_1, b_2, b'_1, b'_2 , the other two are fully determined.

⁹Which was well known in the context of the quasiparticle model [21, 22, 28].

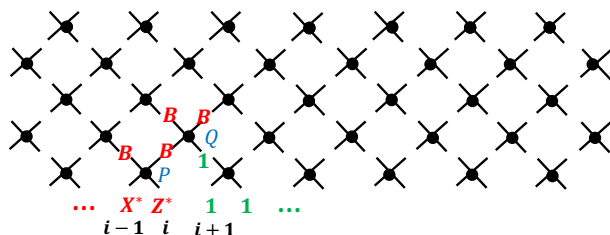


Figure 14. Sharp light-cone growth of operators in the perfect tensor model. The unitary \tilde{U} at point P in the circuit acts on operators at $i - 1$, i , and gives some operator which is nontrivial on both $i - 1$ and i . In the figure, X^* , Z^* refer to any powers of X and Z , and B to any non-trivial operator. \tilde{U} at Q acts on an operator which is non-trivial on site i and trivial on site $i + 1$, giving an operator which is non-trivial on both i and $i + 1$. By repeatedly using the fact that operators nontrivial on a single site evolve to operators non-trivial on two sites, we see that the right endpoint moves to $i + (t - 1)$ at time t .

4. For $b_1 \otimes b_2$ of the form $X^m \otimes Z^n$, the set of b'_1 runs over all one-site basis operators and if we fix b'_1 , then b'_2 , b_1 and b_2 are uniquely determined (and similarly if we fix b'_2 , b'_1 , b_1 and b_2 are uniquely determined).

From items (1) and (2), we see that all operators in I grow outwards with speed 1, as illustrated in figure 14. Thus we have the same form of $S_2^{(A)}(t)$ for a single interval A as we found in the random circuit and free propagation models with the initial states we considered there. Note that (4.9) is a product state, so unlike the free propagation model, the perfect tensor model can generate entanglement in an initial pure product state.

From items (3) and (4), one can show that for any basis operator $\mathcal{O}_\alpha = P_{X(A)} \otimes Q_{J(A)-X(A)} \in J(A) \cap I$, for $\mathcal{O}_\alpha(t)$ to be $\mathbf{1}$ in the region A , $P_{X(A)}$ can be any operator in $I \cap X(A)$, and if we fix $P_{X(A)}$, then $Q_{J(A)-X}$ is uniquely fixed. The basic idea is illustrated in figure 15. This implies (4.2), and also implies (2.20) as $G(A, J(A), t)$ is equal to the number of basis operators in $I \cap X(A)$, which is $q^{2t-|A|}$. Note that an initial operator that becomes trivial in A at time t will in general be non-trivial in the region $J(A) - X(A)$, unlike in the free propagation model.

Further differences from the free propagation model can be seen by looking at the time-evolution of single initial operators in figure 16. The void formation in this model has a fractal structure, similar to the fractal Clifford circuits discussed in [24, 25].¹⁰ The number of non-trivial operators in $J^+(\mathcal{O}_\alpha)$ grows unboundedly with time, in contrast to the situation in the free propagation model shown in figure 12.

We do not have a closed form for the void formation function $G(A, B; t)$ for a general region B , but one can readily check in examples that (3.16) is satisfied. See figure 17. We do not have a general derivation of the constraints (2.29)–(2.31) in this model, but we have checked in a number of examples that they are obeyed.

¹⁰In [25], all Clifford circuits for $q = 2$ are classified into three types: periodic, glider and fractal. The latter two types cannot leave any pure translation-invariant stabilizer state invariant. In the perfect tensor model here with $q = 3$, the operator evolution has a structure similar to fractal Clifford circuits, but as mentioned earlier this model leaves some pure translation-invariant stabilizer states such as $|0\rangle \otimes \dots \otimes |0\rangle$ invariant. The model also has similarities to the fractal Clifford circuit model in [26].

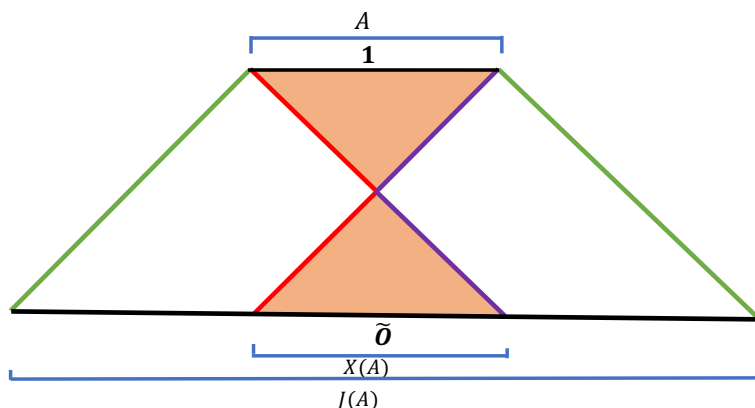


Figure 15. From causality the operators in the past domain dependence of A and future domain dependence of $X(A)$ are fixed by operators in A and $X(A)$ respectively. From the two “Cauchy surfaces”, shown in red and purple, which bound the regions of determined operators, by repeatedly using items (3) and (4) one can uniquely determine the all other operators in the portion of the circuits between the green lines, including the remaining parts of the initial operator in $J(A)$.

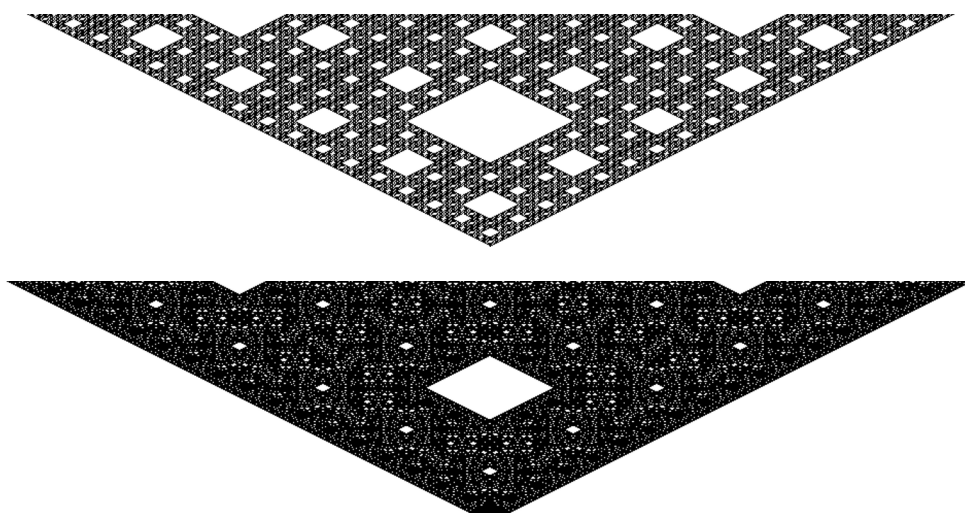


Figure 16. In the top figure, we show the time-evolution of an initial operator from I which is non-trivial on one site in the perfect tensor model. In the bottom figure, we show the time-evolution of an operator non-trivial on 40 sites. Time increases in the upward direction, and the time-evolved operator has non-trivial support on black sites, and the identity operator on white sites. It is evident from the examples that for fixed time, fewer voids are formed in the case where the initial operator is bigger.

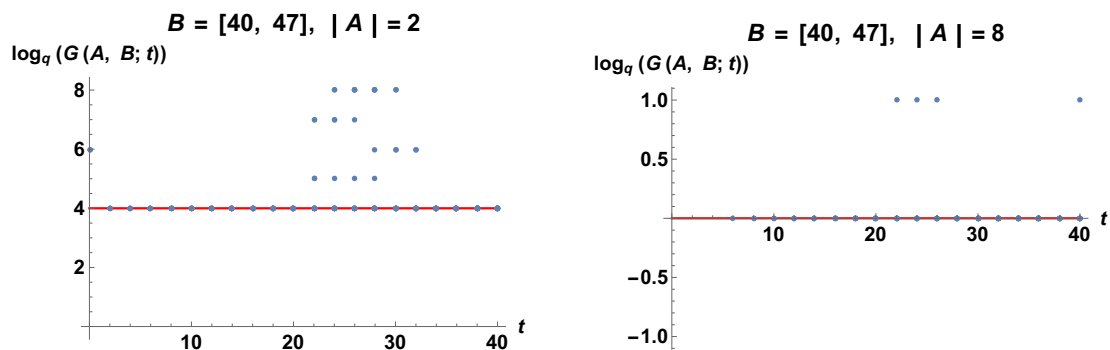


Figure 17. We consider a system with $L = 80$, and sites labelled from $x = 0$ to $x = 79$. We fix the region $B = [40, 47]$, and look at $G(A, B; t)$ for different choices of A . We consider intervals with $|A| = 2 < |B|/2$ on the left, and $|A| = 8 > |B|/2$ on the right, and in both cases we sample over a large number of intervals A and show $\log_q(G(A, B; t))$ as a function of time in a single plot with the blue data points. The value expected from the random void distribution (3.9) is 4 in the first case, and 0 in the second case. We see clearly that $G(A, B; t) \geq G_{\text{RVD}}(A, B; t)$, and there are cases where $G(A, B; t) > G_{\text{RVD}}(A, B; t)$.

5 Conclusions and discussion

In this paper, we examined the implications of void formation in operator evolution for entanglement growth, and showed that it plays an important role in maintaining unitarity of entanglement growth and generation of mutual information and multi-partite entanglement. We showed that the void formation probability for generic operators in random unitarity circuits is given by the random void distribution (3.8). We also showed that the intricate time-dependence of holographic entanglement entropies for an arbitrary number of intervals after a global quench can be understood as a consequence of the very simple rules of sharp light-cone growth and the random void distribution.

Furthermore, we found sharp differences between the void formation properties of random unitary circuits and the non-chaotic circuit models we studied, which suggests that void formation can be used to characterize differences in the operator evolution of chaotic and integrable systems which are not captured by the movement of the operator endpoints alone [29]. It is also an interesting question whether void distribution can be used to distinguish between different classes of chaotic systems. For example, it is conceivable that the random void distribution may only apply to highly chaotic systems.

In our discussion, for simplicity of presentation we have restricted to systems with sharp light-cone growth in the evolution of operators. In general systems, the fronts of operator growth should follow a distribution. For example, in random unitary circuits at finite q , the evolution of an operator exhibits a diffusive front around a butterfly velocity v_B which is smaller than the lightcone velocity v_c [8, 9]. Our discussion, including the random void distribution, can be generalized to these situations, although the story is technically more complicated and will be presented elsewhere.

In our discussion we have focused on the second Renyi entropy, which can be conveniently expressed in terms of the expected number of operators that develop a void in a

certain region. It would be interesting to explore the implications of void formation in operator growth for higher Renyi and von Neumann entropies,¹¹ as well as whether the unitarity of these quantities imposes further constraints on void formation. It is possible these quantities will involve other aspects of void formation, and not just the squared absolute values of the operator-spreading coefficients.

While we have restricted to one spatial dimension, our discussion can be immediately generalized to higher dimensions. In one dimension, a process of void formation separates both the original operator and the full space into disconnected parts. Thus it simultaneously creates “holes” in an operator and breaks it into disjoint pieces. This is not true in higher dimensions, where “hole formation” in an operator and breaking an operator apart are distinct void formation processes, as discussed in figure 2. In particular, it is the latter type of process which contributes to mutual information and multi-partite entanglement among disjoint regions. It is also a natural question whether the “hole formation” and “breaking apart” could follow different probability distributions in higher dimensions.

In this paper, we defined a void as a region of identity operators among regions of nontrivial support of an operator. This definition is only appropriate for a lattice system with a finite one-site Hilbert space at infinite temperature. For finite temperature or continuum systems, a mathematically rigorous definition is tricky. Operationally, one can define a void as the part of an operator which is given by the equilibrium density operator.

It would be interesting to explore the implications of void formation for other observables to see how it affects their behavior in integrable and chaotic systems, and also to see if the non-unitarity in the absence of void formation manifests itself in other observables. For example, one can show that in some simple models without void formation, the out-of-time-ordered correlation functions (OTOCs) will also violate unitarity, although the violation appears less dramatic than that in the entanglement entropy. We will leave the exploration of this question for elsewhere.

It is important to see whether one can find other measures to characterize void formation besides than the probability functions we discussed in this paper. For example, how does one characterize the fractal structure of figure 16? A related question is whether such fractal structure is generic among interacting non-chaotic systems.

Acknowledgments

We would like to thank Netta Engelhardt, Paolo Glorioso, Aram Harrow, Lampros Lamprou, Sam Leutheusser, Raghu Mahajan, Juan Maldacena, Márk Mezei, Adam Nahum, Tibor Rakovszky, Jan Zaanen and Ying Zhao for discussions. This work is supported by the Office of High Energy Physics of U.S. Department of Energy under grant Contract Number DE-SC0012567 and DE-SC0019127.

¹¹Although in some cases, like random unitary circuits in the large q limit, all the entropies are the same.

A Derivations in the random unitary circuits

In this appendix, we first briefly review the partition function method introduced in [8–10] for calculating various observables in random unitary circuits, and then use it to derive the random void distribution (3.8) and the time evolution of entanglement entropy.

A.1 Mapping to a classical Ising partition function

Consider the operator spreading coefficient $|c_\alpha^\beta(t)|^2$ introduced in (2.4), which can be written as a matrix element on four copies of the Hilbert space:

$$\begin{aligned}
 |c_\alpha^\beta(t)|^2 &= \left| \frac{1}{q^L} \text{Tr} \left(\mathcal{O}_\beta^\dagger U^\dagger \mathcal{O}_\alpha U \right) \right|^2 \\
 &= \frac{1}{q^{2L}} (\mathcal{O}_\alpha)_{b_1 a_1} (\mathcal{O}_\alpha^\dagger)_{d_1 c_1} U_{a_1 a_2} U_{b_1 b_2}^* U_{c_1 c_2} U_{d_1 d_2}^* (\mathcal{O}_\beta^\dagger)_{a_1 b_1} (\mathcal{O}_\beta)_{c_1 d_2} \\
 &= \langle \mathcal{O}_{\alpha\uparrow} | (U \otimes U^* \otimes U \otimes U^*) | \mathcal{O}_{\beta\uparrow} \rangle .
 \end{aligned} \tag{A.1}$$

In the last line we have introduced, for any operator \mathcal{O} in the system, “up” and “down” spin states on four copies of \mathcal{H} ,

$$\langle abcd | \mathcal{O}_\uparrow \rangle = \frac{\mathcal{O}_{ab}^\dagger \mathcal{O}_{cd}}{\text{Tr}[\mathcal{O}\mathcal{O}^\dagger]}, \quad \langle abcd | \mathcal{O}_\downarrow \rangle = \frac{\mathcal{O}_{ad}^\dagger \mathcal{O}_{cb}}{\text{Tr}[\mathcal{O}\mathcal{O}^\dagger]}. \tag{A.2}$$

In the case where $\mathcal{O} = \mathbf{1}$, we use the notation $|\uparrow\rangle = |\mathbf{1}_\uparrow\rangle$, $|\downarrow\rangle = |\mathbf{1}_\downarrow\rangle$.

The time evolution operator U for the entire system is a tensor product of random unitaries from the Haar ensemble applied at each time on pairs of sites, as shown in (3.1)–(3.2) and figure 8. As explained in [8–10], after averaging over local unitaries with the Haar measure, one can express (A.1) as a partition function of classical Ising spins on a triangular lattice, as shown in figure 18, with the following specifications:

1. The top layer of the lattice corresponds to time t and the bottom layer to $t = 0$. They are determined respectively by $|\mathcal{O}_{\beta\uparrow}\rangle$ and $|\mathcal{O}_{\alpha\uparrow}\rangle$. If $\mathcal{O}_\alpha = \otimes_i \mathcal{O}_i$, then the spins on the lower boundary are given by $O_{i\uparrow}$ at site i . \mathcal{O}_β similarly fixes the spins on the top boundary.
2. One can see that the interactions along the boundaries are the same for all nontrivial operators, so we can represent any nontrivial operator on the boundaries as B_\uparrow . The rules along the bottom and top boundaries are shown in figure 19 and 20.
3. Lattice points in the bulk of the lattice correspond to locations of local unitaries in figure 8, and on each lattice point lies a spin taking value \uparrow or \downarrow . The interactions among spins in the bulk are specified by the rules of figure 21.
4. The partition function is obtained by summing over all possible configurations of bulk spins, and the weight of a given configuration in the partition function is obtained by multiplying the contributions from each interaction vertices in the bulk and along the boundaries.

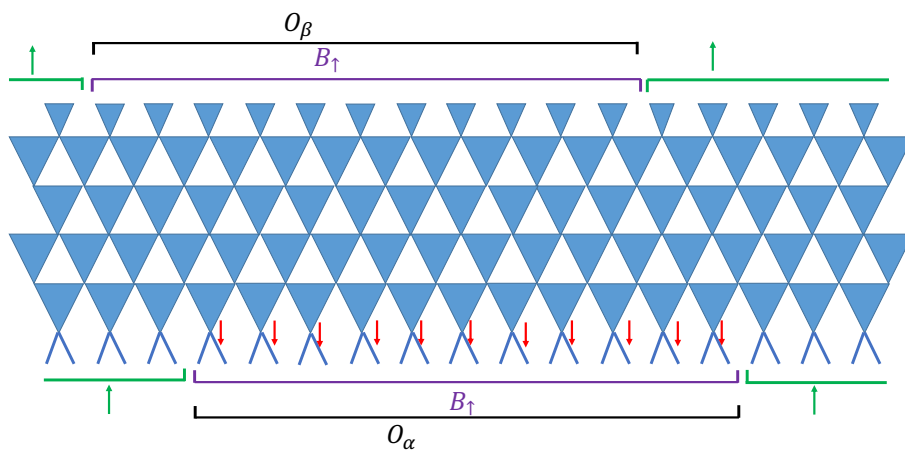


Figure 18. We show the lattice for the partition function corresponding to (A.1) for a particular choice of O_α and O_β , both without voids, and $t = 5$. The boundary conditions, and the spins in the bulk fixed by the boundary conditions through the rules in figure 19, are shown explicitly. The remaining spins can be either up or down.

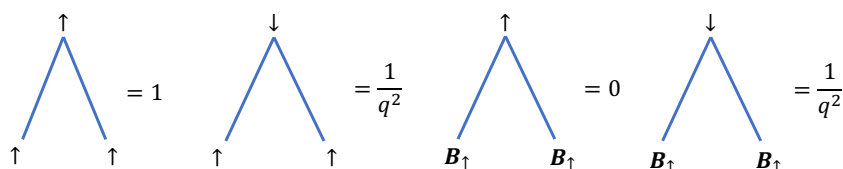


Figure 19. Rules for interactions along the bottom boundary.

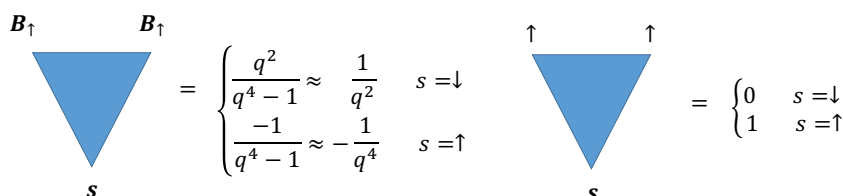


Figure 20. Rules for interactions along the top boundary.

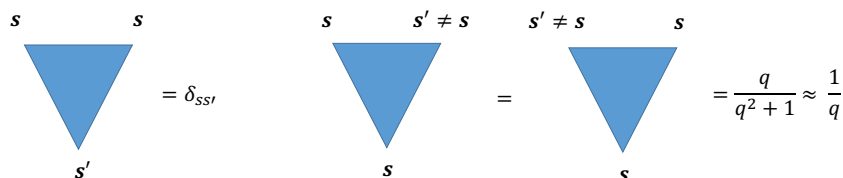


Figure 21. Interactions among the bulk spins, with s, s' taking values in \uparrow, \downarrow . These rules will also apply to the interactions along the top boundary for observables in appendix A.2 and A.3.

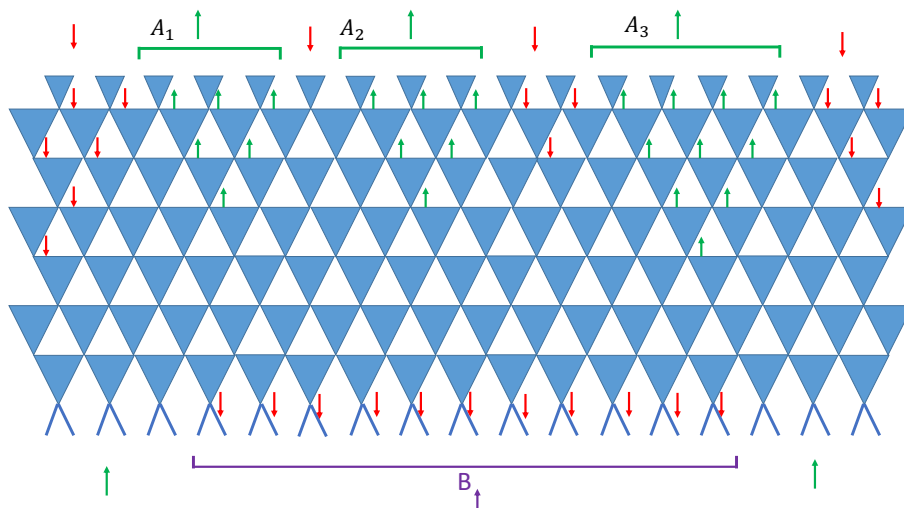


Figure 22. We show the lattice for the partition function corresponding to (A.4) for a case with A consisting of three intervals and $t = 7$. The boundary conditions, and the spins in the bulk fixed by the boundary conditions through the rules in figure 21 and figure 19, are shown explicitly. The remaining spins can be either up or down, but in the large q limit we need to consider a relatively small number of configurations.

A.2 Derivation of the random void distribution

Now let us consider the probability $P_{\mathcal{O}_\alpha}^{(A)}(t) = \sum_{\beta \text{ with void } A} |c_\alpha^\beta(t)|^2$ for a basis operator \mathcal{O}_α to develop a void in region $A = A_1 \cdots A_n$ in random unitary circuits in the large q limit. We will take A to be in the future light cone of \mathcal{O}^α , as otherwise $P_{\mathcal{O}_\alpha}^{(A)}(t)$ is automatically zero.

To find $P_{\mathcal{O}_\alpha}^{(A)}(t)$, it is convenient to consider a slightly different quantity

$$Q_\alpha^{(\bar{A})}(t) = \sum_{\beta \in \bar{A}} |c_\alpha^\beta(t)|^2 \geq P_{\mathcal{O}_\alpha}^{(A)}(t) \quad (\text{A.3})$$

which also includes possible contributions from processes in which \mathcal{O}_α evolves to operators trivial in some disconnected parts of \bar{A} . Such contributions are not in $P_{\mathcal{O}_\alpha}^{(A)}(t)$. Summing (A.1) over all \mathcal{O}_β which have the identity in A we find

$$Q_\alpha^{(\bar{A})}(t) = q^{|\bar{A}|} \langle \mathcal{O}_{\alpha\uparrow} | (U \otimes U^* \otimes U \otimes U^*) | \uparrow \rangle_A | \downarrow \rangle_{\bar{A}} \quad (\text{A.4})$$

where $| \downarrow_{\bar{A}} \rangle = \otimes_{i \in \bar{A}} | \downarrow \rangle_i$ and we have used the fact that

$$| \downarrow \rangle_i = \frac{1}{q} \sum_{a=0}^{q^2-1} | \mathcal{O}_{a\uparrow}^i \rangle \quad (\text{A.5})$$

where the sum over a runs over the complete set of basis operators at site i . Now (A.4) can be calculated with a partition function with boundary conditions as shown in figure 22, with the interactions on the top boundary given by figure 21.

We first consider a basis operator \mathcal{O}_α with no initial voids. In the situation where there exists some A_i for which $t < |A_i|/2$, the rules in figure 21 fix all spins in the past domain

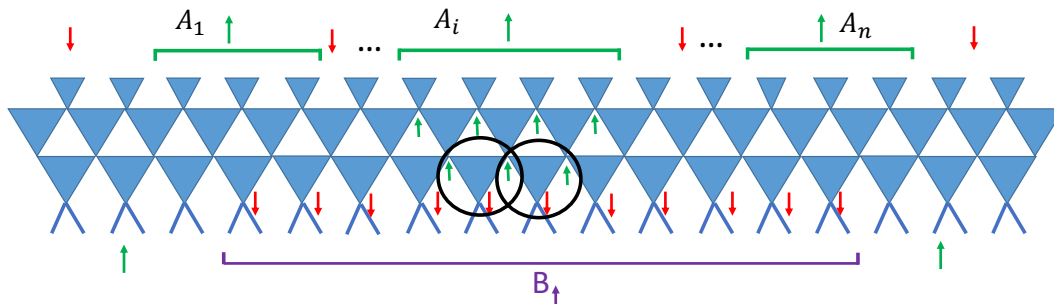


Figure 23. A case with $t < |A_i|/2$ for some A_i . Any configuration in the partition function at such times contributes 0 due to the circled interactions.

of dependence of each A_i to be \uparrow in any configuration in the partition function, while the rules in figure 19 fix the spins attached to \mathcal{O}_α on the lower boundary to be \downarrow . So we get a set of interactions like the circled interactions in figure 23 along the bottom boundary, and hence the contribution from any configuration to (A.4) is zero, and thus $P_{\mathcal{O}_\alpha}^{(A)}(t) = 0$.

We now consider the situation $t > |A_i|/2$ for all i . From the rules of figure 21 and figure 19, in the large q limit, the computation of the partition function corresponding to (A.4) reduces to finding domain walls between up and down spins: for each triangle in the lattice that a domain wall passes through, we get a factor of $\frac{1}{q}$. Starting from the top boundary, a domain wall can either reach the lower boundary or combine with another one to enclose intervals on the top boundary. A domain wall which reaches the lower boundary contributes a factor of q^{-t} , and the shortest domain wall which encloses an interval of length l contributes q^{-l} (which gives the leading contribution in the large q limit).¹²

One possible configuration contributing to (A.4) is shown in figure 24, with the domain walls between the up and down spins enclosing each of the A_i . Since all spins on the lower most bulk layer are \downarrow in this configuration, we get a factor of q^{-L} from the bottom boundary. The domain walls give a factor of $q^{-(|A_1|+\dots+|A_n|)}$. Combing these factors with the prefactor $q^{|\bar{A}|} = q^{L-(|A_1|+\dots+|A_n|)}$ in (A.4), we find the total contribution from this configuration is $q^{-2(|A_1|+\dots+|A_n|)}$. There are other possible domain wall configurations contributing to (A.4), but all these other configurations correspond to the final operators in (A.3) being trivial in some disconnected parts of \bar{A} . An example is given in figure 25. Thus only the configuration of figure 24 contributes to $P_{\mathcal{O}_\alpha}^{(A)}(t)$, and we conclude that

$$P_{\mathcal{O}_\alpha}^{(A)}(t) = q^{-2(|A_1|+\dots+|A_n|)} . \tag{A.6}$$

If \mathcal{O}_α has an initial void, then the domain wall configuration shown in figure 24 still exists, and evaluates to the same value. However, in this case, we also have the possibility that an initial void may evolve into a final void while the disconnected parts of the initial operator evolve independently. Such processes correspond to a new configuration in the partition function for $P_{\mathcal{O}_\alpha}^{(A)}(t)$, shown in figure 26. Clearly when either $|A|$ or $|G|$ is larger

¹²At finite q , it is not sufficient to simply know the length of a domain wall, as a combinatorial factor needs to be included in each configuration to count different possible paths of a given length. But in the large q limit, we can ignore this q -independent combinatorial factor.

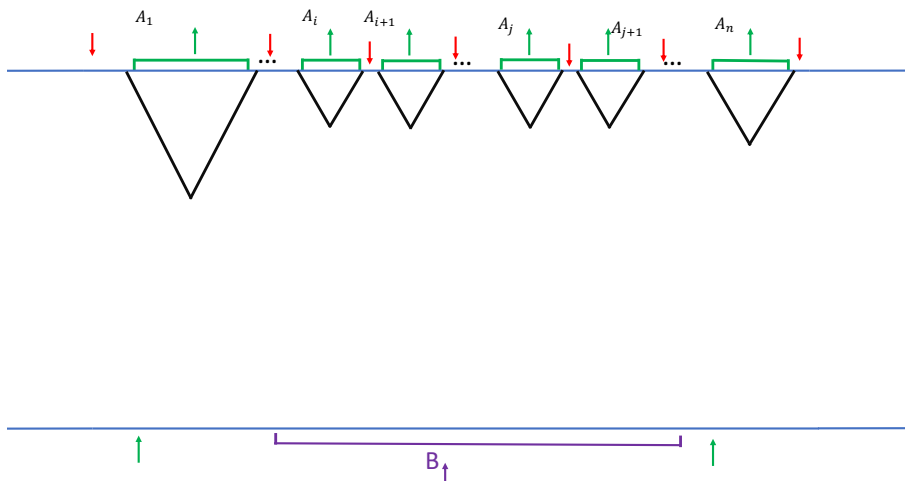


Figure 24. The configuration contributing to $P_{\mathcal{O}_\alpha}^{(A)}(t)$. In this plot to highlight the domain wall structure, we suppress the lattice in the bulk, only showing the domain walls between up and down spins, which are represented by the black lines.

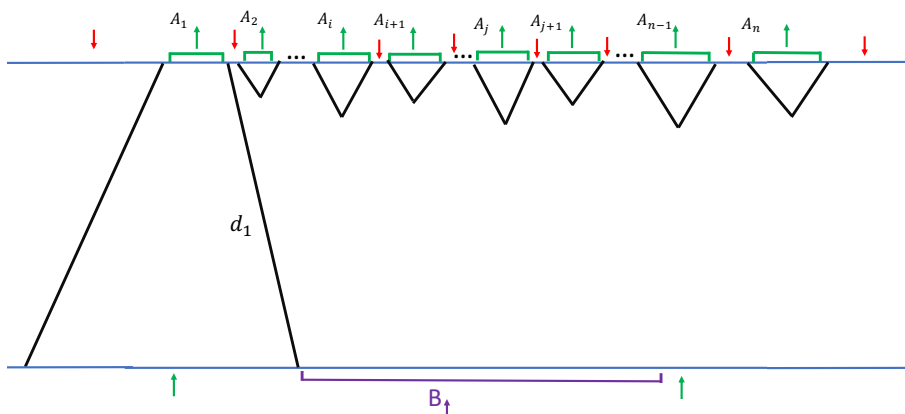


Figure 25. A domain wall configuration that contributes to (A.4), but not to $P_{\mathcal{O}_\alpha}^{(A)}(t)$. It describes processes in which \mathcal{O}_α evolves to operators trivial in the region to the left of A_1 .

than $2t$, this is the only process which can contribute to $P_{\mathcal{O}_\alpha}^{(A)}(t)$. When $|G| < 2t$ and $|A| < 2t$, in general both possibilities exist and compete. The domain wall configuration of figure 24 gives the probability that the initial void closes and opens up a new void, and we have the same value as (A.6). The configuration of figure 26 gives a contribution $q^{-|A|+|G|-2t} = q^{-2|A|}q^{|A|+|G|-2t}$ which dominates when $|G| > 2t - |A|$. Note that for a generic operator (3.6), for \mathcal{O}_α with initial void G , we have $|a_\alpha|^2 \sim q^{-2|G|}$ and thus the overall contribution from such operators is $q^{-|A|-|G|-2t}$ which is subdominant compared with $q^{-2|A|}$ for $|A| < 2t$.

To conclude this discussion, let us note that one can find the probability for an operator of size l_i to evolve into an operator of size l_f (assuming it is allowed by causality, and both operators do not have voids) in the large q limit. This probability is independent of the

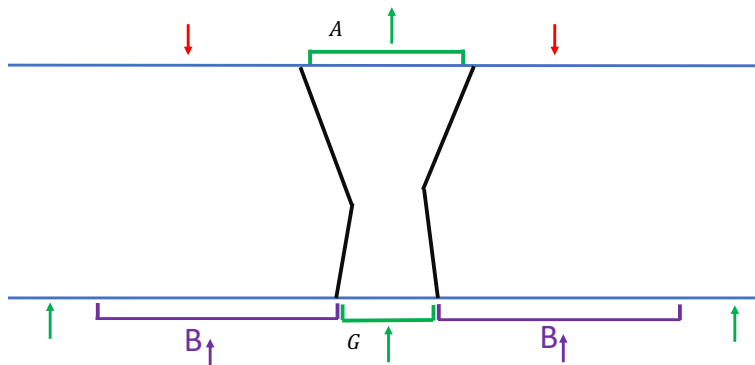


Figure 26. A domain wall configuration that contributes to $P_{\mathcal{O}_\alpha}^{(A)}(t)$ in the case where the initial operator \mathcal{O}_α has a single void in region G , and A is a single interval. This configuration describes processes in which the disconnected parts of \mathcal{O}_α evolve independently to left and right of A . This contribution evaluates to $q^{-|A|+|G|-2t}$, and is bigger than the contribution from figure 24 for $t < (|G| + |A|)/2$.

initial and final operators, and is given by

$$p_{if}(t) = q^{-l_i - l_f - 2t} . \quad (\text{A.7})$$

From here one finds that the probability going to all final operators with length $l_f = l_i + 2t$ allowed by causality is $q^{-l_i - l_f - 2t} q^{2l_f} = q^{l_f - l_i - 2t} = 1$, which leads to the sharp light-cone growth noted in (3.3).

A.3 Derivation of the time-evolution of S_2 in random unitary circuits with large q

Now let us consider the evaluation of $S_2^{(A)}$, which is given by (2.11), which we copy here for convenience

$$e^{-S_2^{(A)}(t)} = \frac{1}{q^{|A|}} \sum_{\alpha \in I} \sum_{\beta \in A} |c_\alpha^\beta(t)|^2 . \quad (\text{A.8})$$

Using (A.1) and summing over all \mathcal{O}_β which are the identity in \bar{A} and all \mathcal{O}_α in I , we find

$$e^{-S_2^{(A)}(t)} = q^L \langle \rho_{0\uparrow} | U \otimes U^* \otimes U \otimes U^* | \downarrow_A \otimes \uparrow_{\bar{A}} \rangle . \quad (\text{A.9})$$

In obtaining (A.9) we again used the fact that on summing over all operators in A , we obtain $|\downarrow_A\rangle$ by using (A.5). Also recall that $\rho_0 = \otimes \prod_i \rho_i$ with each ρ_i given by (2.7). Note that the boundary conditions at top boundary are reversed compared with (A.4). In the large q limit, one has $e^{-\overline{S_2^{(A)}}(t)} = e^{-S_2^{(A)}(t)}$ and all S_n are the same [10], thus the (minus) logarithm of the partition function corresponding to (A.9) gives $\overline{S_2^{(A)}}(t)$ and other entanglement entropies.

The structure of the lattice is the same as in figure 22, but along the top boundary, we have \uparrow spins in \bar{A} and \downarrow spins in A . On the lower boundary, we have $\rho_{i\uparrow}$ at each site. As in last subsection, the evaluation of (A.9) boils down to summing over domain wall configurations. From the fact that each ρ_i is a projector onto a single state in the one-site

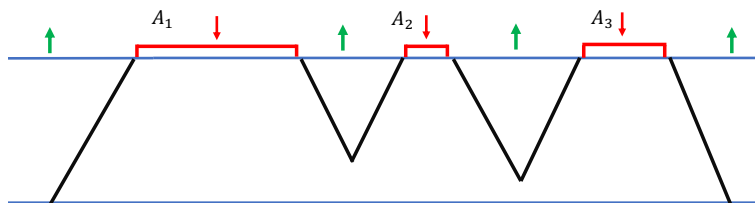


Figure 27. Example of a configuration contributing to the partition function for (A.9). The set γ corresponding to this configuration is $\{l_1, r_3, \{l_2, r_1\}, \{l_3, r_2\}\}$.

Hilbert space, we get a factor of $\frac{1}{q}$ from each site on the lower boundary, irrespective of whether the lowermost bulk layer has an \uparrow or \downarrow spin at that site.¹³ Thus irrespective of where the domain walls end, we get a factor of $\frac{1}{q^L}$ from the lower boundary, cancelling with the prefactor q^L in (A.9). Effectively the bottom boundary does not play any role.

Now consider a region $A = A_1 \cdots A_n$ consisting of n intervals $A_1 = [l_1, r_1], \dots, A_n = [l_n, r_n]$, separated by intervals R_1, \dots, R_{n-1} . Due to the boundary conditions, in each non-zero configuration in the partition function, we have $2n$ starting points of domain walls on the top boundary, from each of the l_i and r_i . As discussed in last subsection, in the large q limit, we only need to specify whether each domain wall starting on the top boundary reaches the bottom boundary, in which case we get a factor of q^{-t} , or joins with another domain wall to enclose an interval of length l , in which case we get a factor of q^{-l} . The latter possibility only exists for $t > l/2$.

Due to the boundary conditions, in order to separate regions of opposite spins, domain walls starting from some left-endpoint l_i can only join with domain walls starting from some right-endpoint r_j , and from the rules of figure 21, domain walls cannot intersect. Let us call a pair of adjacent endpoints l_i and r_j an “allowed pair” at time t if $|l_i - r_j|/2 \leq t$. It is convenient to define a set $\{\gamma\}(t)$ of possible domain wall configurations. An element γ of the set contains some number of “allowed pairs” $\{l_i, r_j\}$ and some number of unpaired points, such that each endpoint appears exactly once. Paired points correspond to domain walls enclosing the region between them, while unpaired points correspond to domain walls reaching the bottom boundary. See figure 27 for an example. The contribution from a γ -configuration evaluates to $(\frac{1}{q})^{f_\gamma}$, where

$$f_\gamma = n_\gamma t + \sum_{\{l_i, r_j\} \in \gamma} |l_i - r_j|, \quad (\text{A.10})$$

and n_γ is the number of unpaired points in γ . In the large q limit, the configuration with minimal $f(\gamma)$ dominates.¹⁴ We thus find

$$S_2^{(A)}(t) = s_{\text{eq}} \min_\gamma f_\gamma, \quad s_{\text{eq}} = \log q. \quad (\text{A.11})$$

¹³If $\rho_i = |\psi\rangle\langle\psi|$, then $\langle abcd|\rho_{i\uparrow}\rangle = \psi_a\psi_b^*\psi_c\psi_d^*$, so $\langle\uparrow|\rho_{i\uparrow}\rangle = \langle\downarrow|\rho_{i\uparrow}\rangle = \frac{1}{q}|\psi|^2|\psi|^2 = 1/q$.

¹⁴In the large q limit we can also ignore any q -independent prefactor, as it contributes an $O(1)$ term to $S_2^{(A)}$ after taking the logarithm.

B Entanglement entropy from the random void distribution

Here we prove that (3.13) follows using only the following rules: (i) sharp light-cone growth, (ii) the random void distribution (3.8), (iii) large q , for A consisting of an arbitrary number of disjoint intervals. Considering (2.11), we can write $N_A(t)$ more explicitly as

$$N_A(t) = N(A, D(Q); t), \quad Q = A_1 R_1 \cdots R_{n-1} A_n \tag{B.1}$$

Before presenting the proof, let us first note the two key elements which are heavily used:

1. At any given time t , if there exists some $|R_i| > 2t$, the two parts of A separated by interval R_i can be independently considered. See figure 28 for an example, where we have a factorized form

$$N(A, D(Q); t) = N(M_1, D(Q_1); t) N(M_2, D(Q_2); t) . \tag{B.2}$$

2. Consider one of the factorized parts, in which $2t$ is greater than all R_i in that part, for example, Q_1 in figure 28. There are possible multiple competing contributions to $N(A_1 A_2, D(Q_1); t)$, which are exhibited in figure 29. One contribution comes from operators which are nontrivial at all sites of $D(Q_1)$, see figure 29(b), which from (3.8) gives

$$N_{(\text{connected})}^A(t) = M_{D(Q_1)}(t) q^{-2|R_1|} = q^{|A_1|+|A_2|-|R_1|-2t} \tag{B.3}$$

where $M_{D(Q_1)}(t) = q^{|D(Q_1)|} = q^{|A_1|+|A_2|+|R_1|-2t}$ is the number of initial basis operators in $D(Q_1)$. There is also a disconnected contribution from figure 29(a), where nontrivial operators in $D(A_1)$ and $D(A_2)$ separated by an initial void evolve independently to region A_1 and A_2 respectively. Note that one may also consider initial operators with a void like in figure 29(c), where the non-trivial parts of the operator are not contained within $D(A_1)$ and $D(A_2)$. Assuming sharp light-cone growth, such an operator cannot give a disconnected contribution in which $\mathcal{O}_1, \mathcal{O}_2$ evolve independently to A_1 and A_2 . However, such an operator can give a connected contribution, which corresponds to situations where the initial void closes and then opens an new void. This contribution is suppressed compared to (B.3), as the phase space for initial operators with a void is suppressed, while the probability for an individual operator to develop a void is the same.

We will prove (3.13) by induction. For two intervals, $n = 2$, we already showed this explicitly in section 3.2 (the discussion leading to (3.11)–(3.12)). Here we give another derivation which connects more directly to the form (3.13). From item 1 above, for $t < |R|/2$,

$$N(A, D(Q), t) = N_{A_1}(t) N_{A_2}(t) \quad \rightarrow \quad S_2^{(A)} = S_2^{(A_1)} + S_2^{(A_2)} . \tag{B.4}$$

For $t > |R|/2$, from item 2 we have two competing contributions: the connected one, which is given by (B.3), and the disconnected one, given by (B.4). Comparing (B.3) and (B.4) in the large q limit leads to the two-interval result from (3.13), with the connected contribution (B.3) corresponding to having the pair (r_1, l_2) in the endpoint configuration γ .

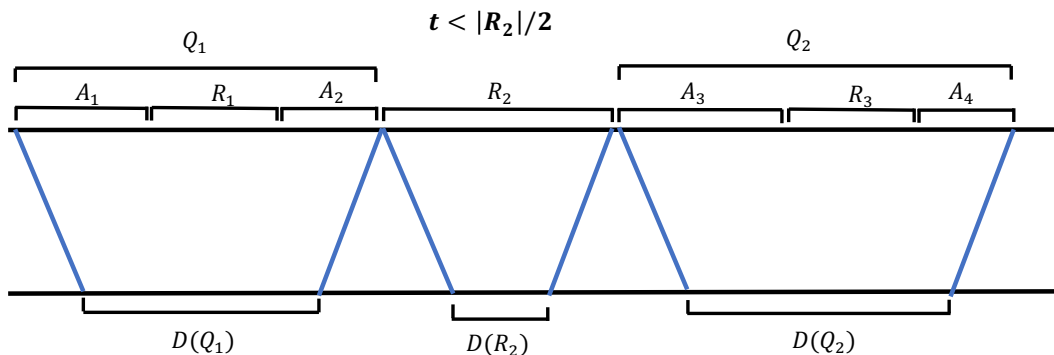


Figure 28. R_2 separates A into two parts with $M_1 = A_1A_2$ and $M_2 = A_3A_4$, which can be independently considered. The initial operator must have identity in region $D(R_2)$ in order to contribute to (B.1), and thus the contributions from two sides of region $D(R_2)$ factorize.

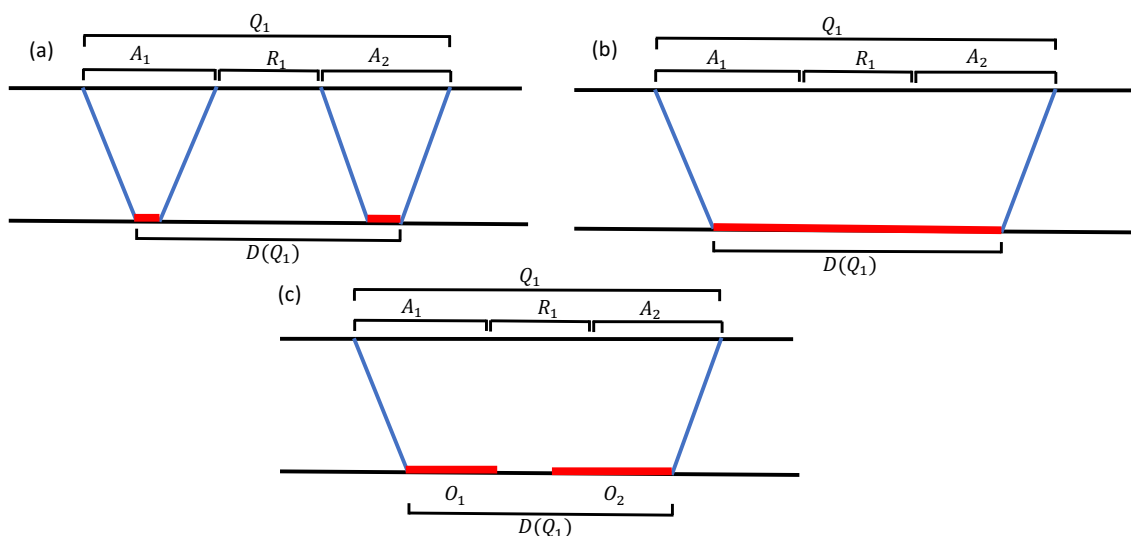


Figure 29. Possible contributions to $N(A_1A_2, D(Q_1); t)$ of figure 28. In the figure, red shaded regions have non-trivial operators at all sites, while the unshaded part of the initial operator is trivial.

Assuming we have (3.13) up to $n = k$ intervals, let us now consider $n = k + 1$.

Let us denote the collection of R_i 's which are greater than $2t$ by \mathcal{R}_{big} . Such R_i 's generate a partition $\mathcal{P}(t)$ of A into m parts (where m can range from 1 to n depending on the time), M_1, \dots, M_m . For the reason stated in figure 28 and item 1 above, the contribution from each M_i factorizes, and we thus have

$$S_2^{(A)}(t) = \sum_{j=1}^m S_2^{(M_j)}(t) . \tag{B.5}$$

The same thing happens to (3.13): if there exists a $|R_i| > 2t$, then the two parts of A separated by interval R_i can be independently minimized as there are no allowed pairing between end points on the left of R_i and those on the right of R_i . Thus equation (B.5) also

applies to (3.13), and each term in the sum of (B.5) is given by (3.13) from our assumption regarding $n \leq k$.

We increase t until $t = r/2$ when the set \mathcal{R}_{big} becomes empty, where r is size of the largest R_i , which we call R_{max} . Slightly before reaching $t = r/2$, from the random void distribution (3.8) initial operators cannot develop a void in region R_{max} , which in the language of (3.13) corresponds to the fact that the two end points of R_{max} cannot be paired with each other. Now when $t > r/2$, from the random void distribution (3.8), there are new contributions coming from initial operators developing a void in region R_{max} , which compete with previously existing ones. Note the new contributions include the connected one for the full region $D(Q)$, but also disconnected ones. See figure 30 for two examples. These new contributions are in one-to-one correspondence with new γ -configurations in (3.13) which come from pairing the two end points of R_{max} . One can also readily check that their respective contributions agree. Thus concludes the proof.

In the above proof, we assumed the random void distribution for all initial operators, an assumption that is not precisely true in random unitary circuits in the large q limit due to the subtlety noted at the end of appendix A.2. Yet we found by the partition function calculation of appendix A.3 that equation (3.13) is true in random unitary circuits in the large q limit, indicating that the cases where $P_{\mathcal{O}_\alpha}^{(A)}(t)$ for initial operators with voids is not given by the random void distribution can be ignored in the calculation of S_2 . For instance, this means that in figure 29(c), the cases in random unitary circuits where the disconnected contribution (due to independent evolution of O_1 and O_2 respectively into A_1 and A_2) is dominant can be ignored. Indeed, on using (A.7) and counting the number of relevant initial and final operators, we can check that the collective disconnected contribution from all initial operators like in figure 29(c) is the same as that from figure 29(a), changing S_2 at most by a q -independent $O(1)$ term which can be neglected in the large q limit.

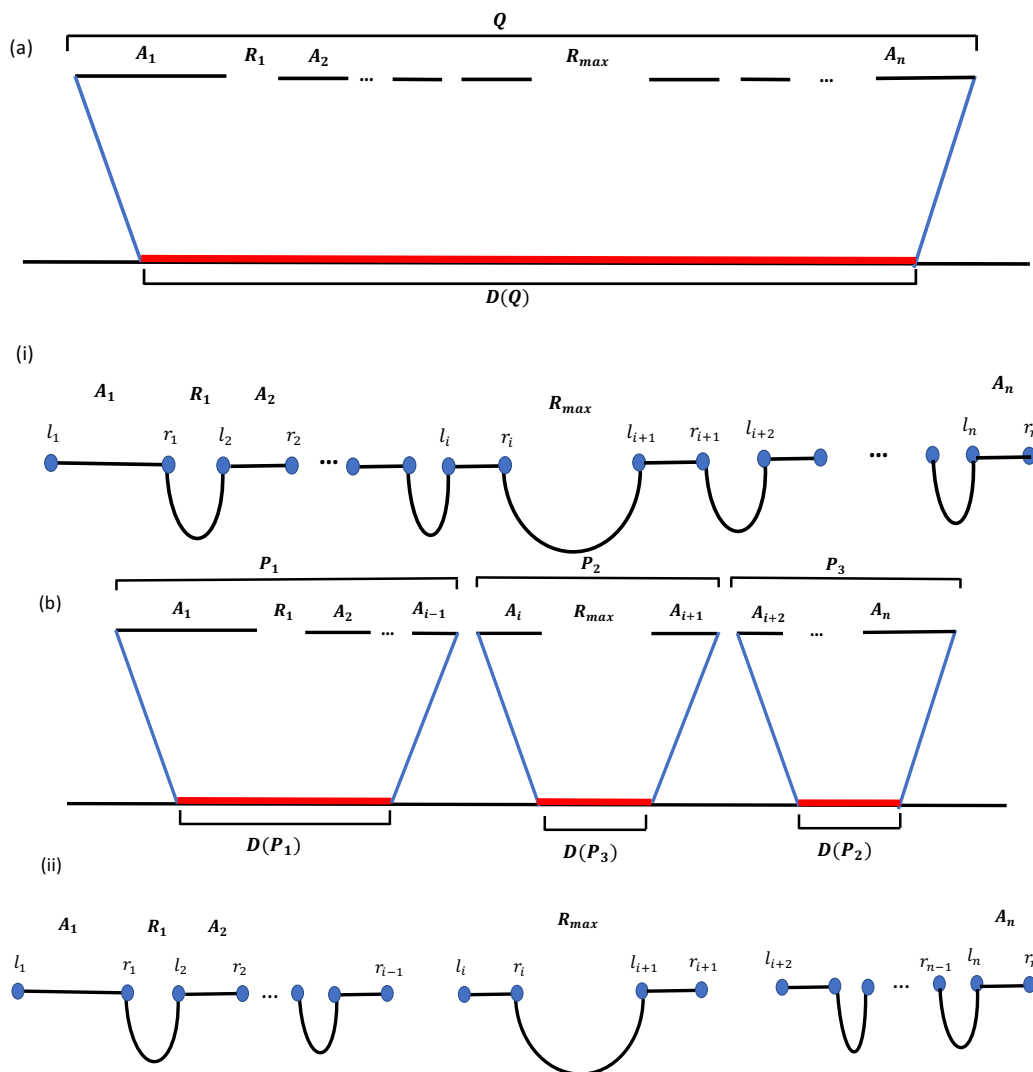


Figure 30. Two sets of initial operators contributing to (2.11), and the corresponding endpoint configurations γ . In (a), initial operators nontrivial everywhere in $D(Q)$ develop a void in $R \equiv R_1 \cdots R_{n-1}$, contributing $q^{|Q|-2t} q^{-2|R|} = q^{|A|+|R|-2t} q^{-2|R|} = q^{|A|-|R|-2t}$ to $N_A(t)$, leading to a possible contribution $S_2^{(A)}(t) = s_{\text{eq}}(|R| + 2t)$. We get the same contribution from γ shown in (i), where l_1 and r_n are the only unpaired points, and all r_m for $m < n$ are paired with l_{m+1} . In (b), initial operators nontrivial everywhere in each $D(P_i)$ go into P_i , forming voids in the R_j 's contained within the P_i 's. This contributes $q^{|P_1|-2t} q^{-2(|R_1|+\dots+|R_{i-2}|)} q^{|P_2|-2t} q^{-2|R_{\text{max}}|} q^{|P_3|-2t} q^{-2(|R_{i+2}|+\dots+|R_{n-1}|)}$ to $N_A(t)$, leading to a possible contribution $S_2^{(A)}(t) = s_{\text{eq}}(|R| - |R_{i-1}| - |R_{i+1}| + 6t)$. We get the same contribution from γ shown in (ii), where $l_1, r_{i-1}, l_i, r_{i+1}, l_{i+2}$, and r_n are the only unpaired points, and all r_m for $m \neq i-1, i+1, n$ are paired with l_{m+1} .

Open Access. This article is distributed under the terms of the Creative Commons Attribution License ([CC-BY 4.0](https://creativecommons.org/licenses/by/4.0/)), which permits any use, distribution and reproduction in any medium, provided the original author(s) and source are credited.

References

- [1] D.A. Roberts, D. Stanford and L. Susskind, *Localized shocks*, *JHEP* **03** (2015) 051 [[arXiv:1409.8180](https://arxiv.org/abs/1409.8180)] [[INSPIRE](#)].
- [2] T. Prosen and I. Piorn, *Operator space entanglement entropy in transverse Ising chain*, *Phys. Rev. A* **76** (2007) 032316 [[arXiv:0706.2480](https://arxiv.org/abs/0706.2480)].
- [3] I. Piorn and T. Prosen, *Operator Space Entanglement Entropy in XY Spin Chains*, *Phys. Rev. B* **79** (2009) 184416 [[arXiv:0903.2432](https://arxiv.org/abs/0903.2432)].
- [4] J. Dubail, *Entanglement scaling of operators: a conformal field theory approach, with a glimpse of simulability of long-time dynamics in 1 + 1d*, *J. Phys. A* **50** (2017) 234001 [[arXiv:1612.08630](https://arxiv.org/abs/1612.08630)] [[INSPIRE](#)].
- [5] C. Jonay, D.A. Huse and A. Nahum, *Coarse-grained dynamics of operator and state entanglement*, [arXiv:1803.00089](https://arxiv.org/abs/1803.00089) [[INSPIRE](#)].
- [6] S.H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **03** (2014) 067 [[arXiv:1306.0622](https://arxiv.org/abs/1306.0622)] [[INSPIRE](#)].
- [7] W.W. Ho and D.A. Abanin, *Entanglement dynamics in quantum many-body systems*, *Phys. Rev. B* **95** (2017) 094302 [[arXiv:1508.03784](https://arxiv.org/abs/1508.03784)] [[INSPIRE](#)].
- [8] A. Nahum, S. Vijay and J. Haah, *Operator Spreading in Random Unitary Circuits*, *Phys. Rev. X* **8** (2018) 021014 [[arXiv:1705.08975](https://arxiv.org/abs/1705.08975)] [[INSPIRE](#)].
- [9] C. von Keyserlingk, T. Rakovszky, F. Pollmann and S. Sondhi, *Operator hydrodynamics, OTOCs, and entanglement growth in systems without conservation laws*, *Phys. Rev. X* **8** (2018) 021013 [[arXiv:1705.08910](https://arxiv.org/abs/1705.08910)] [[INSPIRE](#)].
- [10] T. Zhou and A. Nahum, *Emergent statistical mechanics of entanglement in random unitary circuits*, *Phys. Rev. B* **99** (2019) 174205 [[arXiv:1804.09737](https://arxiv.org/abs/1804.09737)] [[INSPIRE](#)].
- [11] H. Casini, H. Liu and M. Mezei, *Spread of entanglement and causality*, *JHEP* **07** (2016) 077 [[arXiv:1509.05044](https://arxiv.org/abs/1509.05044)] [[INSPIRE](#)].
- [12] P. Hosur, X.-L. Qi, D.A. Roberts and B. Yoshida, *Chaos in quantum channels*, *JHEP* **02** (2016) 004 [[arXiv:1511.04021](https://arxiv.org/abs/1511.04021)] [[INSPIRE](#)].
- [13] E.H. Lieb and D.W. Robinson, *The finite group velocity of quantum spin systems*, *Commun. Math. Phys.* **28** (1972) 251 [[INSPIRE](#)].
- [14] D.A. Roberts and D. Stanford, *Two-dimensional conformal field theory and the butterfly effect*, *Phys. Rev. Lett.* **115** (2015) 131603 [[arXiv:1412.5123](https://arxiv.org/abs/1412.5123)] [[INSPIRE](#)].
- [15] M. Mezei and D. Stanford, *On entanglement spreading in chaotic systems*, *JHEP* **05** (2017) 065 [[arXiv:1608.05101](https://arxiv.org/abs/1608.05101)] [[INSPIRE](#)].
- [16] H. Liu and S.J. Suh, *Entanglement Tsunami: Universal Scaling in Holographic Thermalization*, *Phys. Rev. Lett.* **112** (2014) 011601 [[arXiv:1305.7244](https://arxiv.org/abs/1305.7244)] [[INSPIRE](#)].
- [17] C.T. Asplund, A. Bernamonti, F. Galli and T. Hartman, *Entanglement Scrambling in 2d Conformal Field Theory*, *JHEP* **09** (2015) 110 [[arXiv:1506.03772](https://arxiv.org/abs/1506.03772)] [[INSPIRE](#)].

- [18] S. Xu and B. Swingle, *Locality, Quantum Fluctuations, and Scrambling*, *Phys. Rev. X* **9** (2019) 031048 [[arXiv:1805.05376](#)] [[INSPIRE](#)].
- [19] S. Xu and B. Swingle, *Accessing scrambling using matrix product operators*, *Nature Phys.* **16** (2019) 199 [[arXiv:1802.00801](#)] [[INSPIRE](#)].
- [20] E. Leviatan, F. Pollmann, J.H. Bardarson and E. Altman, *Quantum thermalization dynamics with Matrix-Product States*, [arXiv:1702.08894](#) [[INSPIRE](#)].
- [21] S. Leichenauer and M. Moosa, *Entanglement Tsunami in (1+1)-Dimensions*, *Phys. Rev. D* **92** (2015) 126004 [[arXiv:1505.04225](#)] [[INSPIRE](#)].
- [22] V. Balasubramanian, A. Bernamonti, N. Copland, B. Craps and F. Galli, *Thermalization of mutual and tripartite information in strongly coupled two dimensional conformal field theories*, *Phys. Rev. D* **84** (2011) 105017 [[arXiv:1110.0488](#)] [[INSPIRE](#)].
- [23] D.M. Schlingemann, H. Vogts and R.F. Werner, *On the structure of Clifford quantum cellular automata*, *J. Math. Phys.* **49** (2008) 112104 [[arXiv:0804.4447](#)].
- [24] J. Gutschow, *Entanglement generation of Clifford quantum cellular automata*, *Appl. Phys. B* **98** (2010) 623 [[arXiv:1001.1062](#)].
- [25] J. Gutschow, S. Uphoff, R.F. Werner and Z. Zimboras, *Time asymptotics and entanglement generation of Clifford quantum cellular automata*, *J. Math. Phys.* **51** (2010) 015203 [[arXiv:0906.3195](#)].
- [26] S. Gopalakrishnan and B. Zakirov, *Facilitated quantum cellular automata as simple models with nonthermal eigenstates and dynamics*, *Quant. Sci. Tech.* **3** (2018) P044004 [[arXiv:1802.07729](#)].
- [27] P. Calabrese and J.L. Cardy, *Evolution of entanglement entropy in one-dimensional systems*, *J. Stat. Mech.* **0504** (2005) P04010 [[cond-mat/0503393](#)] [[INSPIRE](#)].
- [28] C.T. Asplund and A. Bernamonti, *Mutual information after a local quench in conformal field theory*, *Phys. Rev. D* **89** (2014) 066015 [[arXiv:1311.4173](#)] [[INSPIRE](#)].
- [29] S. Gopalakrishnan, D.A. Huse, V. Khemani and R. Vasseur, *Hydrodynamics of operator spreading and quasiparticle diffusion in interacting integrable systems*, *Phys. Rev. B* **98** (2018) 220303 [[arXiv:1809.02126](#)] [[INSPIRE](#)].