

# Temporal Entanglement in Chaotic Quantum Circuits

Foligno, Alessandro; Zhou, Tianci; Bertini, Bruno

DOI:

[10.1103/PhysRevX.13.041008](https://doi.org/10.1103/PhysRevX.13.041008)

License:

Creative Commons: Attribution (CC BY)

*Document Version*

Publisher's PDF, also known as Version of record

*Citation for published version (Harvard):*

Foligno, A, Zhou, T & Bertini, B 2023, 'Temporal Entanglement in Chaotic Quantum Circuits', *Physical Review X*, vol. 13, no. 4, 041008. <https://doi.org/10.1103/PhysRevX.13.041008>

[Link to publication on Research at Birmingham portal](#)

## General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

## Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact [UBIRA@lists.bham.ac.uk](mailto:UBIRA@lists.bham.ac.uk) providing details and we will remove access to the work immediately and investigate.

# Temporal Entanglement in Chaotic Quantum Circuits

Alessandro Foligno<sup>1,2</sup>, Tianci Zhou,<sup>3</sup> and Bruno Bertini<sup>1,2</sup>

<sup>1</sup>*School of Physics and Astronomy, University of Nottingham, Nottingham NG7 2RD, United Kingdom*

<sup>2</sup>*Centre for the Mathematics and Theoretical Physics of Quantum Non-Equilibrium Systems, University of Nottingham, Nottingham, NG7 2RD, United Kingdom*

<sup>3</sup>*Center for Theoretical Physics, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA*



(Received 4 March 2023; revised 21 July 2023; accepted 21 August 2023; published 11 October 2023)

The concept of space evolution (or space-time duality) has emerged as a promising approach for studying quantum dynamics. The basic idea involves exchanging the roles of space and time, evolving the system using a space transfer matrix rather than the time evolution operator. The infinite-volume limit is then described by the fixed points of the latter transfer matrix, also known as influence matrices. To establish the potential of this method as a bona fide computational scheme, it is important to understand whether the influence matrices can be efficiently encoded in a classical computer. Here we begin this quest by presenting a systematic characterization of their entanglement—dubbed temporal entanglement—in chaotic quantum systems. We consider the most general form of space evolution, i.e., evolution in a generic spacelike direction, and present two fundamental results. First, we show that temporal entanglement always follows a volume law in time. Second, we identify two marginal cases—(i) pure space evolution in generic chaotic systems and (ii) any spacelike evolution in dual-unitary circuits—where Rényi entropies with index larger than one are sublinear in time while the von Neumann entanglement entropy grows linearly. We attribute this behavior to the existence of a product state with large overlap with the influence matrices. This unexpected structure in the temporal entanglement spectrum might be the key to an efficient computational implementation of the space evolution.

DOI: [10.1103/PhysRevX.13.041008](https://doi.org/10.1103/PhysRevX.13.041008)

Subject Areas: Quantum Physics,  
Quantum Information,  
Statistical Physics

## I. INTRODUCTION

The first two decades of the new millennium witnessed extraordinary experimental progress in measuring dynamical properties of quantum many-body systems. Experiments are now able to probe, for instance, local relaxation of isolated systems [1–3] and out-of-equilibrium transport [4–8] over surprisingly long timescales. Theoreticians, however, can very rarely provide independent predictions to compare with these experiments, especially concerning dynamics beyond intermediate timescale. Indeed, characterizing a quantum many-body system out of equilibrium, or even simulating its state on a classical computer, remains to date a formidable task.

The situation is slightly more favorable in one dimension, where one can use an extension of the celebrated DMRG algorithm [9,10] to provide a faithful representation

of the time-evolving quantum state [11]. The initial state is represented as a matrix product state (MPS), and a suitable evolution algorithm (e.g., (time-dependent)Density Matrix Renormalization Group [12,13] or Time-Evolution Block-Decimation [14,15]) finds an MPS approximation of the state at time  $t$  for a given level of accuracy. The problem, however, is that the amount of resources required for such an approximation grows exponentially with the entanglement of the state and, in the absence of localization or other ergodicity-breaking mechanisms, the latter builds up very quickly as time elapses. In practice this means that one needs an exponentially growing amount of resources for an accurate representation of the state. This “entanglement barrier” is physical and cannot be avoided whenever one tries to characterize the whole quantum state. The key question, however, is whether or not it is necessary to simulate the evolution of the whole quantum state to compute its experimentally accessible properties, e.g., its correlation functions.

In the course of the past decade several algorithms have been proposed to circumvent the fast entanglement growth of nonequilibrium states [16–26]. The common theme is to sidestep the problem by exploiting the fact that one is

---

*Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.*

typically only interested in the correlation functions of special observables, for instance, those that are *local* in space. A promising one, which motivates our work, is the so-called “folding algorithm” or “transverse folding algorithm” proposed in Ref. [26] (see also Refs. [27–31]), whose main idea is to evolve the system in space, rather than in time. Taking the one-point function in Fig. 1 as an example, this means that one has to contract its tensor-network representation horizontally, by means of an appropriate space transfer matrix rather than vertically using the time evolution operator. The name of the algorithm derives from the fact that this operation becomes much more efficient when considering the “folded representation” of the correlator, i.e., when folding the tensor network around the center as shown in Fig. 1(b), which doubles the local degrees of freedom but keeps the correlations short-ranged. Physically, the vertical column of tensors beneath the observable implements the unitary evolution of the subsystem of interest—the one where the observable acts—while the sections on its two sides encode the nonunitary action exerted on the subsystem by the rest of the system, i.e., the environment. For instance, in the example of Fig. 1 the system is a single spin (or qudit). Inspired by the Feynman-Vernon influence functional approach [32], Ref. [33] proposed to dub “influence matrices” the portions of the tensor network describing the action of the environment. Note that when the environment becomes very

large, the influence matrices become equal to the left and right fixed points of the space transfer matrix  $\mathcal{T}_t$ ; see Fig. 1.

The idea of exchanging space and time to describe infinite systems at finite times proved to be very successful and over the past few years has found interesting applications to the study of spectral statistics and quantum chaos [34–39], entanglement dynamics [40–43], impurity problems [44], and even full-counting statistics of many-body observables [45] and Loschmidt echo [46–48]. When considered as a computational tool for computing correlation functions, however, the folding algorithm has an important limitation: it can only deal with cases where the operator insertions break the translation symmetry in a single spatial point, i.e., one-point functions and, more generally, autocorrelations. In this way one cannot access, for instance, generic two-point functions—such as those needed to compute transport coefficients [49–51]—as they feature two operators separated in both time and space.

Another outstanding question concerns the computational complexity of the folding algorithm. Namely, how hard it is to implement this algorithm on a classical computer for increasingly large times. To answer this question one needs to understand what features of the influence matrices have to be retained to correctly describe expectation values of local operators and what is the amount of resources required to do so. An intuitive estimate can be obtained by studying their entanglement, dubbed “temporal entanglement” [28,52]. Indeed, roughly speaking, if the latter does not grow too fast one can efficiently approximate the influence matrices with matrix product states for arbitrarily high fidelity [55,56]. Following Refs. [28,33], one can argue that temporal entanglement should be small for generic systems. Indeed, the dephasing caused by the environment tends to align corresponding spins in the forward and backward copies (cf. Fig. 1) producing configurations that are diagonal and hence classical. Although plausible, this picture can be proven only in a few special cases. These include certain special chaotic quantum systems—dual-unitary circuits [57]—prepared in a special family of initial states [40,58] and in certain special classes of integrable models [30,59–62]. In generic cases the temporal entanglement is observed to grow in time, even though its growth appears slower than that of spatial entanglement [26,33].

In this work we fill both the aforementioned gaps: (1) We extend the folding algorithm to compute generic two-point functions and (2) we characterize the scaling of temporal entanglement in generic quantum many-body systems.

The main idea for extending the folding algorithm is to embed the two operators in the same system defined on a *timelike surface*, or path,  $\gamma$ , see the illustration in Fig. 2, and evolve it in the orthogonal spacelike direction. In a relativistic field theory one can imagine to implement our construction by boosting to a reference frame where the operators are measured at the same position and then

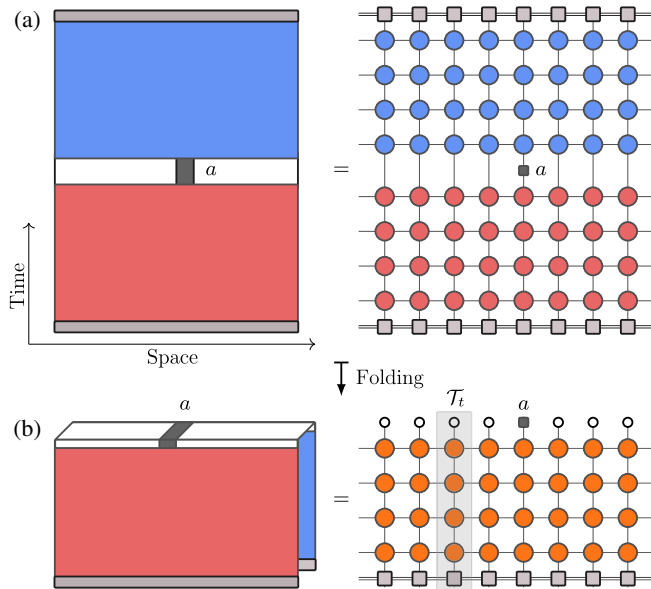


FIG. 1. (a) One-point function of the local operator  $a$  after a quantum quench and its tensor-network representation. Forward and backward time sheets are, respectively, depicted in red and blue. (b) The same one-point function of (a) after folding: in the folded tensor network the number of local degrees of freedom is doubled and the white circles at the top of the tensor network denote a loop. The gray shaded box highlights the space transfer matrix  $\mathcal{T}_t$  (acting from left to right on a lattice of  $t$  sites).

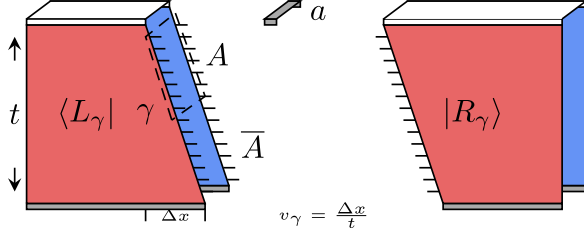


FIG. 2. Generalized influence matrices on a temporal slice  $\gamma$ . The (anti)slope of the path  $\gamma$  is  $v_\gamma = \Delta x/t$ . When  $v_\gamma = 0$ , they correspond to the regular influence matrices. A one-point or two-point function can be evaluated by contracting the left state  $\langle L_\gamma|$ , some relevant operators inserted along the path, and the right state  $|R_\gamma\rangle$ . The temporal entanglement is the larger among the entanglement of  $\langle L_\gamma|$  and that of  $|R_\gamma\rangle$  for a partition  $A, \bar{A}$  on a temporal slice.

use the usual folding algorithm. This setup allows us to treat generic two-point functions, and also gives the option to optimize the evaluation of one-point functions by varying the path on which the influence matrices are evaluated. Note that the extreme case of a timelike surface corresponding with the light cone edge has been considered in Ref. [63] and, for this case, Ref. [62] characterized the complexity of the corresponding influence matrices for integrable dual-unitary circuits.

To characterize the scaling of temporal entanglement, we compute the entanglement entropies of the generalized influence matrices  $\langle L_\gamma|$  and  $|R_\gamma\rangle$  across contiguous bipartitions of  $\gamma$ . We, respectively, denote them by

$$S_{L,A}^{(\alpha)}(\gamma) \quad \text{and} \quad S_{R,A}^{(\alpha)}(\gamma), \quad (1)$$

for subregion  $A$  and Rényi index  $\alpha$ . Our findings are summarized in Table I.

Overall we find that the temporal von Neumann entropy ( $\alpha = 1$ ) always grows linearly in time after a quench from a generic initial state. Nevertheless, we find cases in which Rényi entropies with index  $\alpha > 1$  (higher Rényi entropies from now on) grow sublinearly. In particular, the higher Rényi entropies of vertical influence matrices (the regular ones) in *any* chaotic system are logarithmic in time, while those of *any* influence matrix in a dual-unitary circuit are

bounded by a constant. In these cases the slope of growth of von Neumann entropy is strictly smaller than that of regular state entanglement. These statements are proven analytically for dual-unitary circuits, while in the case of generic circuits they result from the combination of entanglement membrane theory [64,65] and numerical observations.

The observed linear growth of von Neumann entanglement entropy rules out an efficient high-fidelity approximation of the influence matrices via matrix product states [56]. Our findings, however, suggest that there are physically relevant cases where the temporal entanglement spectrum displays a strong separation of scales: There are a few large (at most linearly decaying) Schmidt values and many exponentially small ones. This remarkable structure might be the key for an efficient implementation of the folding algorithm.

In the following subsection, we sketch the key steps to obtain the scalings in Table I and discuss their consequences. A complete description of our setup begins in Sec. II.

### A. Summary of approaches and results

We consider generic quantum many-body systems with local interactions modeled by local brickwork quantum circuits [66]. This is a class of locally interacting systems in discrete time that has recently played a key role in understanding many-body quantum dynamics. The enormous complexity of the latter implies that the theoretical description, or even the mere numerical simulation, of quantum matter out of equilibrium is practically possible only in the short-time regime. Brickwork quantum circuits simplify the picture by imposing strictly local interactions over a finite time step and give rare examples where local observables and information theoretical quantities can be determined at all times. The results obtained in these systems, for instance, through random averaging [64,67–70] and/or space-time duality [40,58,63,71–77], have significantly advanced our understanding of universal properties of the dynamics. Applications include, for instance, operator dynamics and information spreading [64,67–70,78–80], statistical properties of the spectrum [34–39,81–84], and, more broadly, thermalization [58–60,83,85–87]. We also note that quantum circuits are vital tools for experimental

TABLE I. Scaling of the temporal entanglement. We take the second Rényi entropy as a representative of higher Rényi entropies. The maximum is taken over  $L, R$  states and the possible contiguous regions  $A$  on  $\gamma$  for a given slope  $v_\gamma$ . All the circuits have a brickwork architecture (see Sec. II). Only vertical cuts or dual-unitary circuits give sublinear growth in higher Rényi entropy. Italic font denotes analytical results (obtained by membrane picture or exact calculation), bold font denotes numerical evidence.

	$\max[S_{L,A}^{(1)}(\gamma), S_{R,A}^{(1)}(\gamma)]$	$\max[S_{L,A}^{(2)}(\gamma), S_{R,A}^{(2)}(\gamma)]$
Generic circuit ( $v_\gamma \neq 0$ )	$\sim t$	$\sim t$
Generic circuit ( $v_\gamma = 0$ )	$\sim t$	$\lesssim \log t$
Generic dual-unitary circuit	$\sim t$	$\sim 1$

simulation of quantum systems and quantum computation. For instance, they can be used to demonstrate quantum advantage [88–91], to perform randomized benchmarking [92–95], randomized measurements [96–101], shadow tomography [102–104], and, more generally, to study nonequilibrium dynamics of Floquet systems [105,106].

The structure of these circuits look like a Suzuki-Trotter [107,108] approximation of (local) Hamiltonian evolution, but the unitary gates are not necessarily infinitesimal in time or close to the identity: They can be arbitrary unitaries (see the detailed illustration of our setup in Sec. II).

To understand the behavior of  $S_{L/R,A}^{(\alpha)}(\gamma)$  in generic circuits, we take the gates forming the brickwork structure to be independent Haar random matrices. By averaging over the random gates the calculation of entanglement related quantities is mapped into that of the free energy of a statistical mechanical model of emergent spins [64,65]. In particular, we find that the averaged temporal purity,

$$\overline{\exp\left(-S_{L/R,A}^{(2)}(\gamma)\right)}, \quad (2)$$

is the difference of free energies of the same statistical model subjected to different boundary conditions. By minimizing the free energies we find domain-wall configurations that give (cf. Sec. IV)

$$\overline{\exp\left(-S_{L/R,A}^{(2)}(\gamma)\right)} \simeq e^{-v_{\text{TE}}^{(2)} t \log d}, \quad (3)$$

where the linear coefficient  $v_{\text{TE}}^{(2)} \geq 0$  is determined by the line tension  $\mathcal{E}_H(v)$  of the membrane separating the different domains. The line tension is an intrinsic function of the membrane, which, in translational invariant systems (at least after disorder average), only depends on the space-time slope  $v$ . Although the explicit expression is complicated, we have a useful condition,

$$v_{\text{TE}}^{(2)} = 0 \Leftrightarrow \mathcal{E}_H(v_\gamma) = \mathcal{E}_H(0), \quad (4)$$

where  $v_\gamma$  is the antislope [109] of the path  $\gamma$ ; see Fig. 2.

Equation (3) results in a lower bound of the typical growth rates of the temporal entanglement entropies. In particular, we have

$$\begin{aligned} \overline{S_{L/R,A}^{(1)}(\gamma)} &\geq v_{\text{TE}}^{(2)} \log(d)t, \\ \overline{S_{L/R,A}^{(\alpha>1)}(\gamma)} &\geq \frac{1}{2} v_{\text{TE}}^{(2)} \log(d)t. \end{aligned} \quad (5)$$

In fact, following Ref. [69], we argue that this conclusion can be applied to generic Floquet circuits even in the absence of randomness. In this case the entanglement dynamics is still described by an emergent statistical mechanical model and the tension of the associated membrane can be determined perturbatively, dressing the

membrane tension of the Haar random circuit [69]. In practice this means that one can apply Eq. (5) without the average by replacing  $\mathcal{E}_H(v)$  by  $\mathcal{E}(v)$ . Therefore, for generic Floquet circuits and generic paths  $\gamma$ , the temporal entanglement entropies with Rényi index  $\alpha \geq 1$  grow linearly in time [110].

Equation (4), however, also suggests that there are two interesting marginal cases where temporal entanglement entropies can be sublinear.

(I) Constant line tension, i.e.,

$$\mathcal{E}(v) = \text{const.} \quad (6)$$

(II) Vertical path, i.e.,

$$v_\gamma = 0. \quad (7)$$

Condition (I) provides a very stringent constraint. Indeed, invoking general properties of the line-tension function [111], one can conclude that a constant line tension has to be equal to one. This in turn implies a maximal growth rate of the regular *spatial* entanglement after a quantum quench in the circuit. As shown in Refs. [69,112], circuits with this property have to be *dual unitary*.

On the contrary, condition (II) does not involve the line-tension function; it only requires the temporal surface to be vertical (i.e., it holds for regular influence matrices). This means that, intriguingly, the vanishing of the linear coefficient at  $v_\gamma = 0$  should occur for generic circuits.

The two marginal cases (I) and (II) are studied in detail in Secs. V and VI. There we show that in both cases higher Rényi entropies ( $\alpha > 1$ ) display a sublinear growth in time. Nevertheless, their von Neumann entropy ( $\alpha = 1$ ) grows *linearly* (second and third rows of Table I). Namely, one cannot evaluate the scaling of von Neumann entropy via a replica trick as the replica limit does not commute with the large-time limit. Interestingly, a similar discrepancy in the scaling of Rényi entropies was also observed in Refs. [113,114] for the behavior of the “regular” spatial entanglement in circuits with conservation laws.

At the level of entanglement spectrum the mechanism driving the observed sublinear scaling is the same in both cases (I) and (II): the influence matrices have large overlap with a product state of the form  $|\Psi'_A\rangle \otimes |\Psi''_{\bar{A}}\rangle$  on  $A$  and  $\bar{A}$ . Then, an immediate application of Eckart-Young’s theorem [115] implies that the reduced density matrices,

$$\rho_{L,\gamma,A} = \frac{\text{tr}_{\bar{A}}(|L_\gamma\rangle\langle L_\gamma|)}{\| |L_\gamma\rangle\|^2}, \quad \rho_{R,\gamma,A} = \frac{\text{tr}_{\bar{A}}(|R_\gamma\rangle\langle R_\gamma|)}{\| |R_\gamma\rangle\|^2}, \quad (8)$$

have at least one slowly decaying eigenvalue. This eigenvalue determines the slow growth of higher Rényi entropies. Meanwhile, we find exponentially many other eigenvalues of  $\rho_{L/R,\gamma,A}$  that decay exponentially fast with time  $t$ . This produces a linearly growing entanglement

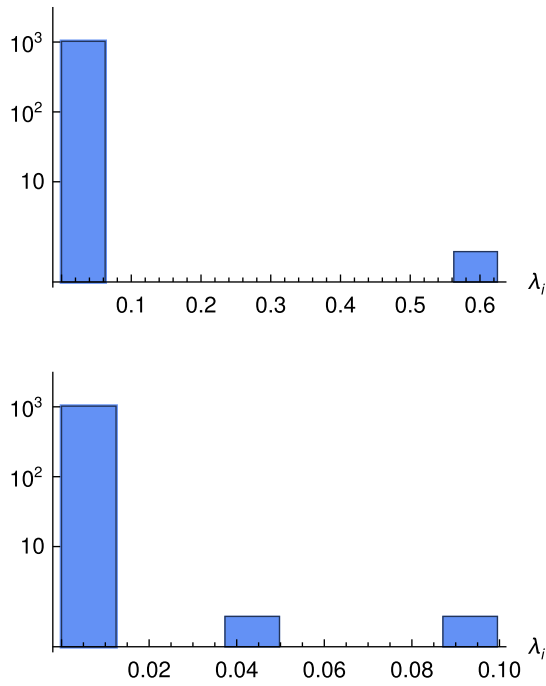


FIG. 3. Schmidt values  $\lambda_i$  for the bipartition of  $\langle L_\gamma \rangle$  in two subsystems of equal size for  $t = 5$ . The y axis reports the number of Schmidt values within a small bin centered on  $\lambda_i$ . The top panel corresponds to the marginal case (I) [cf. Eq. (6)] while the bottom panel corresponds to the marginal case (II) [cf. Eq. (7)].

entropy  $S_{L/R,A}^{(1)}(\gamma)$ . Figure 3 shows the separation of scales in the entanglement spectrum of the temporal state in two representative examples.

In this situation one might be inclined to conclude that the singular states corresponding to the large Schmidt values represent the dominant part of the state. The linear growth of the von Neumann entropy, however, excludes the possibility of constructing a high-fidelity approximation of the state by keeping a polynomial number of Schmidt eigenvectors. Nevertheless, if the objective is to only approximate special observables, for instance, the one-point function  $\text{tr}[\rho_0 a_x(t)]$ , the answer might be different. More generally, it is interesting to ask how much physically relevant information can be extracted faithfully from the first few Schmidt eigenstates in cases with such a strong separation of scales. We leave these questions to future research.

The rest of this paper is laid out as follows. In Sec. II we introduce the precise setting considered in this work. In Sec. III we discuss the folding algorithm and explain its extension to nonvertical cuts. In Sec. IV we determine the scaling of temporal entanglement in generic quantum circuits using the entanglement-membrane approach of Refs. [64,65,69]. The two marginal cases with sublinear growth of higher Rényi entropies are analyzed in Secs. V and VI. In particular, in Sec. V we discuss the scaling of temporal entanglement in dual-unitary circuits evolving from generic initial states, while in Sec. VI we study the scaling of regular temporal entanglement, i.e., of the

influence matrix on the vertical cut, in generic circuits. In Sec. VII we compare the growth of temporal entanglement and that of regular state entanglement. Our conclusions and final remarks are reported in Sec. VIII.

## II. SETTING

We consider the quantum dynamics generated by local quantum circuits acting on a chain of  $2L$  qudits ( $d$  internal states) placed at half integer positions. These circuits have nearest-neighbor interactions, and are often dubbed “brickwork” quantum circuits. The operator performing one step of evolution alternatively evolves the even and odd sublattices:

$$\mathbb{U} = \mathbb{U}_1 \mathbb{U}_2. \quad (9)$$

Here we introduced

$$\mathbb{U}_1 = \bigotimes_{x \in \mathbb{Z}_L} U_{x,x+1/2}, \quad \mathbb{U}_2 = \bigotimes_{x \in \mathbb{Z}_L+1/2} U_{x,x+1/2}, \quad (10)$$

with  $U_{x,x+1/2}$  acting nontrivially, as the  $d^2 \times d^2$  unitary matrix  $U$ , only on the qudits at positions  $x$  and  $x + 1/2$ . The matrix  $U$  is known as “local gate” and specifies the local interactions. Local gates can in principle be different at each space-time point, i.e.

$$U \mapsto U(t, x), \quad (11)$$

representing a disordered system undergoing aperiodic quantum dynamics. In contrast, in the special case where all local gates coincide, the quantum circuit constitutes a clean (two-site shift invariant), periodically driven system.

A useful property of the local gate, which we use later to identify quantum circuits with similar dynamical features, is its *entangling power*. Roughly speaking, the latter is a measure of the ability of the gate to entangle two qubits [116]. Normalizing it to be in  $[0, 1]$ , the entangling power can be expressed as [116]

$$p(U) = \frac{d^4 + d^2 - \text{tr}(U^R U^{R\dagger})^2 - \text{tr}[(US)^R (US)^{R\dagger}]^2}{d^2(d^2 - 1)}. \quad (12)$$

Here  $U^R$  indicates the gate obtained by rotating the original gate  $U$  by a right angle

$$\langle lj|U^R|ki \rangle = \langle kl|U|ij \rangle, \quad i, j, k, l = 0, \dots, d-1, \quad (13)$$

and  $S$  denotes the SWAP gate,

$$\langle ij|S|lk \rangle = \delta_{i,k} \delta_{j,l}. \quad (14)$$

Here we are interested in the evolution of the system for  $t > 0$ . At  $t = 0$  the system is prepared in a generic “pair-product” state,

$$|\Psi_0\rangle = \frac{1}{d^{L/2}} \bigotimes_{x=1}^L \left( \sum_{i,j=0}^{d-1} m_{ij} |i\rangle \otimes |j\rangle \right), \quad (15)$$

where  $\{|i\rangle\}_{i=0}^{d-1}$  is a basis of the configuration space of a single qudit—the “local” Hilbert space. The matrix  $m$ , with elements  $m_{ij}$ , fulfills

$$\text{tr}[mm^\dagger] = d, \quad (16)$$

which ensures that  $|\Psi_0\rangle$  is normalized to one. We consider general pair-product states, rather than simple product states, to keep the staggered structure of the brickwork quantum circuit. Note that a product state is recovered by the choice

$$m_{ij} \propto \delta_{i,i_0} \delta_{j,j_0}, \quad i_0, j_0 \in \{0, \dots, d-1\}, \quad (17)$$

while generically one can think of Eq. (15) as a product state which has been subject to half a step of evolution.

The evolution in a quantum circuit can be conveniently illustrated using a tensor-network-inspired graphical representation [117]. In particular, depicting the components of the local gate and the initial state matrix as

$$U_{(k,l);(i,j)} = \begin{array}{c} k \quad l \\ \diagup \quad \diagdown \\ \square \\ \diagdown \quad \diagup \\ i \quad j \end{array}, \quad m_{ij} = \begin{array}{c} i \quad j \\ \diagup \quad \diagdown \\ \bullet \\ \diagdown \quad \diagup \end{array}, \quad (18)$$

we can represent the state of the system at time  $t$  as follows:

$$|\Psi_t\rangle = \frac{1}{d^{L/2}} \begin{array}{c} \text{[Diagram of a brickwork quantum circuit with red squares and legs]} \end{array}, \quad (19)$$

where we considered  $t = 3$ . As illustrated in the above diagram, we depicted the periodic boundary conditions by connecting left and right boundaries, and used the convention that when legs of different tensors are joined together the index of the corresponding local space is summed over. Moreover, we drop the indices to represent the full vector rather than its components. We will use this convention whenever it does not lead to confusion.

Let us consider the evolution of the reduced density matrix of a finite region  $A$ . Representing it diagrammatically, we have

$$\rho_A(t) = \frac{1}{d^{|A|}} \begin{array}{c} \text{[Diagram of a brickwork quantum circuit with red and blue squares]} \end{array}, \quad (20)$$

where we took  $A = \{1, 3/2, 2, 5/2\}$  and introduced a diagrammatic representation for the Hermitian conjugate of the local gate,

$$U^\dagger = \begin{array}{c} \diagdown \quad \diagup \\ \square \\ \diagup \quad \diagdown \end{array}, \quad (21)$$

and the complex conjugate of the initial state matrix,

$$m^* = \begin{array}{c} \bullet \\ \diagdown \quad \diagup \end{array}. \quad (22)$$

Using this representation we can depict the unitarity of the local gate with the following diagrammatic relations

$$\begin{array}{c} \text{[Diagrammatic relation showing unitarity of the local gate]} \end{array} = \begin{array}{c} \text{[Diagrammatic relation showing unitarity of the local gate]} \end{array} = \begin{array}{c} \text{[Diagrammatic relation showing unitarity of the local gate]} \end{array}. \quad (23)$$

To simplify the diagrams it is convenient to fold them in two. In particular, folding the blue part of the circuit underneath the red one, we can represent the reduced density matrix in Eq. (20) as follows,

$$\rho_A(t) = \frac{1}{d^{|A|/2}} \begin{array}{c} \text{[Diagram of a folded brickwork quantum circuit with orange squares]} \end{array}, \quad (24)$$

where we introduced the *double gate*,





$$\mathcal{T}_x = \text{[Diagram]}, \quad \mathcal{T}_x^{(a)} = \text{[Diagram]}, \quad \mathcal{T}_x^{(b)} = \text{[Diagram]}, \quad (39)$$

$$|R_x\rangle = \text{[Diagram]}, \quad (43)$$

so that the diagram (36) can be written as

$$C_{ab} = \text{tr}[\mathcal{T}_1 \cdots \mathcal{T}_{x_2-1} \mathcal{T}_{x_2}^{(b)} \mathcal{T}_{x_2+1} \cdots \mathcal{T}_{x_1-1} \mathcal{T}_{x_1}^{(a)} \mathcal{T}_{x_1+1} \cdots \mathcal{T}_L], \quad (40)$$

where we consider the generic case of nontranslational invariant circuits. We remark that the space transfer matrices in Eq. (39) are matrix product operators (MPOs) with finite bond dimension  $\chi = d^2$ .

The next step is to note that unitarity can simplify products of transfer matrices. To illustrate this point, let us write down the product of  $2t_2$  transfer matrices  $[\mathcal{T}_x$  in Eq. (39)]. In diagrams it takes the following form:

$$\text{[Diagram]} \quad (41)$$

Unitarity [cf. Eq. (30)] allows us to cancel all the gates above the red dashed lines and propagate the bullets to the legs crossing the dashed lines. We therefore have the following rank-1 decomposition,

$$\mathcal{T}_{y_1} \cdots \mathcal{T}_{y_{2t_2}} = |R_{y_1}\rangle \langle L_{y_{2t_2}}|, \quad \forall y_j, \quad (42)$$

where we introduced the following vectors on the folded time lattice,

$$\langle L_x| = \text{[Diagram]} \quad (44)$$

This means that, for  $L > x_2 - x_1 + 2t_2$ , Eq. (51) can be written as

$$C_{ab} = \langle L_{x_1-1} | \mathcal{T}_{x_1}^{(a)} \mathcal{T}_{x_1+1} \cdots \mathcal{T}_{x_2-1} \mathcal{T}_{x_2}^{(b)} | R_{x_2+1} \rangle. \quad (45)$$

This representation sheds light on the physical interpretation of the two vectors  $\langle L_x|$  and  $|R_x\rangle$ . These objects encode the effect of the rest of the system on the subsystem of size  $x_2 - x_1$  where  $a_x$  and  $b_x$  act. Since their role is analogous to that of the influence functional of Feynman and Vernon [32], they have been dubbed “influence matrices” [33]. Note that in the translational invariant case one can use Eq. (42) to show that  $\langle L_x|$  and  $|R_x\rangle$  are the unique *fixed points*, i.e., eigenvectors corresponding to eigenvalue one, of the space transfer matrix  $\mathcal{T}$  (which is  $x$  independent in translational invariant circuits).

The representation (45) is the main instrument of the folding algorithm. Assuming that one can find an efficient MPS representation for the influence matrices (see Sec. III B), Eq. (45) gives a way to compute two-point functions as matrix elements of an MPO—the product of  $x_2 - x_1 + 1$  space transfer matrices—between two MPSs. Since the bond dimension of the MPO is bounded by  $d^{2(x_2-x_1+1)}$ , this operation can be performed efficiently for small distances  $x_2 - x_1$ . On the other hand, the computation becomes rapidly unfeasible when the distance increases. This represents a serious limitation as, for instance, two-point functions for arbitrary distances fulfilling Eq. (34) are needed to compute transport coefficients in linear response [49–51]. To circumvent this problem we propose an alternative method for contracting the diagram in Eq. (36): Instead of contracting it

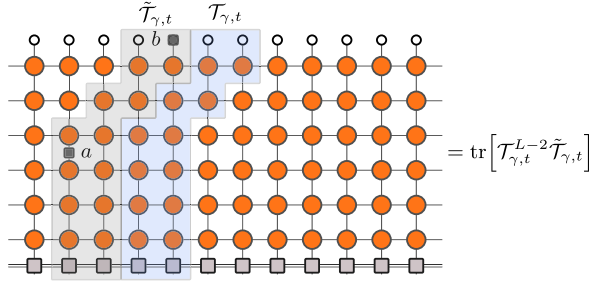


FIG. 4. Folded tensor-network representation of a dynamical two-point function of two operators  $a$  and  $b$  after a quantum quench. The tensor network can be contracted using nonvertical transfer matrices that follow the timelike path  $\gamma$  and propagate in the spacelike direction orthogonal to  $\gamma$ . The two different transfer matrices used are highlighted in the boxes.

in the space direction, we contract it in a more general spacelike direction such that the two points lie on the same timelike surface; see the macroscopic-scale illustration in Fig. 2; an example in terms of the associated tensor network is depicted instead in Fig. 4. The only relevant macroscopic feature of the timelike surface is its the space-time slope  $v_\gamma$ . A more precise lattice definition is given in Sec. III A, while in Sec. III B we discuss the computational complexity of encoding the influence matrices in an MPS.

### A. Generalized folding algorithm

For a precise definition of the generalized folding algorithm it is useful to distinguish between two different regimes:

- (I)  $0 \leq x_2 - x_1 \leq t_2 - t_1$ ,
- (II)  $t_2 - t_1 < x_2 - x_1 \leq t_2 + t_1$ .

Note that regime (II) only arises out of equilibrium: the equilibrium correlation in Eq. (33) exists only in the regime (I). Moreover, regime (I) is also the only regime arising for nonequilibrium one-point functions.

#### 1. Regime (I)

In this regime there exists a path  $\tilde{\gamma}$  connecting  $a$  and  $b$  that is entirely contained in the causal light cone emanating

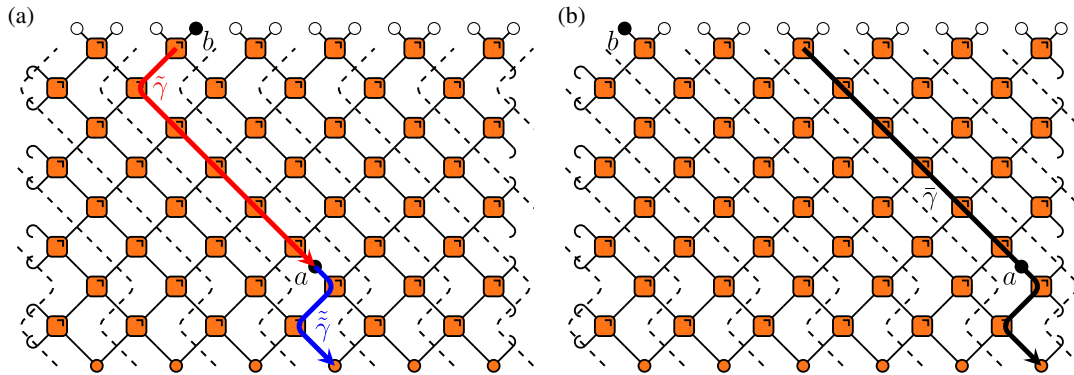


FIG. 5. Diagrammatic representation of the path to contract correlation functions. Panels (a) and (b) depict examples of paths used in the contraction of the correlation functions in the regimes I and II, respectively. Dashed lines help to identify the transfer matrix and time slices to cut open the diagram.

from  $a$ ; i.e., it goes from  $a$  to  $b$  without ever “turning back.” We call this kind of paths *timelike paths*, because all the space-time points they reach are causally connected.

To specify  $\tilde{\gamma}$  we start from the gate below  $b$  and move down in discrete jumps; see Fig. 5(a). At each jump we reach one neighboring gate: either the one at southeast or the one at southwest. Using the variable  $\tilde{\gamma}_i = \pm$  to keep track of whether on the  $i$ th step we jump on the left or on the right, we can represent the path by means of the following sequence,

$$\tilde{\gamma} = \{\tilde{\gamma}_1, \dots, \tilde{\gamma}_N\}, \quad (46)$$

where  $N = 2(t_2 - t_1) - 1$  is the length of the path. For instance,

$$\tilde{\gamma} = \{-, +, +, +, +\} \quad (47)$$

is the path depicted in red in Fig. 5(a). The path  $\tilde{\gamma}$  can be extended to a path  $\gamma$  that reaches the initial state by concatenating it with another timelike path  $\tilde{\tilde{\gamma}}$  from  $(x_1, t_1)$  to  $(y, 0)$  for some  $y \in [x_1 - t_1, x_1 + t_1]$ . The total length of the path  $\gamma$ , i.e., the total number of jumps, is then  $2t_2$ . For instance, in the example of Fig. 5 one can consider

$$\gamma = \{-, +, +, +, +\} \circ \{+, -, +\}, \quad (48)$$

where  $\circ$  denotes the composition operator. The average slope of a given path  $\gamma$  is given by

$$v_\gamma = \frac{1}{|\gamma|} \sum_{i=1}^{|\gamma|} \gamma_i, \quad (49)$$

where  $|\gamma|$  is the length of the path. As mentioned before,  $v_\gamma$  is the only bit of information required for a coarse grained description of the path.

Since the path  $\gamma$  does not turn back, we can use it to “slice” the diagram of the correlation function. Namely, we

subdivide it in a number of timelike slices by cutting the bonds in a direction parallel to  $\gamma$  (see the black dashed lines in Fig. 5) and connect them with suitably defined transfer matrices. In particular, for the configuration in Fig. 5(a) the transfer matrices are given by

$$\mathcal{T}_{\gamma,x} = \text{[Diagram]}, \quad \mathcal{T}_{\gamma,x}^{(ab)} = \text{[Diagram]}. \quad (50)$$

In this way we can write Eq. (36) as

$$C_{ab} = \text{tr}[\mathcal{T}_{\gamma,1} \cdots \mathcal{T}_{\gamma,x_1-1} \mathcal{T}_{\gamma,x_1}^{(ab)} \mathcal{T}_{\gamma,x_1+1} \cdots \mathcal{T}_{\gamma,L}]. \quad (51)$$

This expression can again be simplified using the unitarity of the gates. In particular, we again have

$$\mathcal{T}_{\gamma,y_1} \cdots \mathcal{T}_{\gamma,y_2} = |R_{\gamma,y_1}\rangle \langle L_{\gamma,y_2}|, \quad \forall y_j, \quad (52)$$

where we introduced the generalized or “boosted” influence matrices:

$$\langle L_{\gamma,x}| = \text{[Diagram]}, \quad (53)$$

$$|R_{\gamma,x}\rangle = \text{[Diagram]}. \quad (54)$$

Therefore, for  $L > 2t_2$ , we find

$$C_{ab} = \langle L_{\gamma,x_1-1} | \mathcal{T}_{\gamma,x_1}^{(ab)} | R_{\gamma,x_1+1} \rangle. \quad (55)$$

As opposed to Eq. (45), this expression can always be efficiently contracted if  $\langle L_{\gamma,x}|$  and  $|R_{\gamma,x}\rangle$  admit an efficient MPS representation.

## 2. Regime (II)

In this regime there is no timelike path connecting  $a$  and  $b$ . This means that we cannot embed both  $a$  and  $b$  in the same “thin” transfer matrix as done in Eq. (55). The best strategy in this case is to slice the diagram (36) using transfer matrices corresponding to the path  $\tilde{\gamma} = \gamma_{\text{lc}} \circ \tilde{\gamma}$ , where

$$\gamma_{\text{lc}} = \{+, \dots, +\} \quad (56)$$

is the fastest path allowed by causality (i.e., on the edge of the light cone) and  $\tilde{\gamma}$  is an arbitrary timelike path between the initial state and  $a$  [cf. Fig. 5(b)]. Repeating the above analysis we find that for  $L > x_2 - x_1 + t_2 + t_1$  the correlations can be expressed as

$$C_{ab} = \langle L_{\tilde{\gamma},x_1-1} | \mathcal{T}_{\tilde{\gamma},x_1}^{(a)} \mathcal{T}_{\tilde{\gamma},x_1+1} \cdots \mathcal{T}_{\tilde{\gamma},x_2-t_2+t_1-1} \mathcal{T}_{\tilde{\gamma},x_2-t_2+t_1}^{(b)} | R_{\tilde{\gamma},x_2-t_2+t_1+1} \rangle, \quad (57)$$

where we introduced

$$\mathcal{T}_{\gamma,x}^{(a)} = \text{[Diagram]}, \quad \mathcal{T}_{\gamma,x}^{(b)} = \text{[Diagram]}. \quad (58)$$

We see that the expression (57) involves the product of

$$n = x_2 - x_1 - t_2 + t_1 + 1 \quad (59)$$

transfer matrices, which means  $0 \leq n \leq 2t_1$ . This has two immediate implications. First, the representation (57) gives an advantage over Eq. (45) because it involves fewer transfer matrices. Second, when both  $x_2 - x_1$  and  $t_1$  are large the contraction of Eq. (57) becomes inefficient.

## B. Generalized temporal entanglement

In extreme summary, the upshot of the previous subsection is that an efficient representation of the generalized

influence matrices does indeed lead to an efficient computational scheme for the calculation of correlation functions in many physically relevant cases [118]. This motivates us to investigate whether an efficient representation of the generalized influence matrices is possible. In particular, here we assess whether these objects admit an efficient MPS representation by computing their entanglement. This is the fundamental question to which the rest of this paper is devoted.

The entanglement of the influence matrices is computed in three steps.

- (i) We define reduced density matrices corresponding to an arbitrary nondisjoint bipartition  $A\bar{A}$  of the lattice along the path  $\gamma$ :

$$\rho_{H,\gamma,A} = \text{tr}_{\bar{A}} \frac{|H_{\gamma,x}\rangle\langle H_{\gamma,x}|}{\| |H_{\gamma,x}\rangle \|^2}, \quad H = L, R. \quad (60)$$

- (ii) We compute their Rényi entropies,

$$S_{H,A}^{(\alpha)}(\gamma) := S^{(\alpha)}(\rho_{H,\gamma,A}), \quad H = L, R, \quad \alpha \in \mathbb{R}, \quad (61)$$

where we introduced the function

$$S^{(\alpha)}(\rho) := \frac{1}{1-\alpha} \log \text{tr}[\rho^\alpha]. \quad (62)$$

- (iii) We maximize them over all possible bipartitions  $A\bar{A}$  where  $A$  is a contiguous region.

Before proceeding we note that

$$\rho_{L,\gamma,A}(W) = \rho_{R,\bar{\gamma},A}(W'), \quad (63)$$

where we highlighted the dependence on the double gate (25), introduced

$$W' = \begin{array}{|c|} \hline \text{L} \\ \hline \end{array}, \quad (64)$$

and denoted by  $\bar{\gamma} = \{-\gamma_1, \dots, -\gamma_{2l_2}\}$  the mirror image of the path  $\gamma$  with respect to the vertical line passing through  $b$ .

In the following we will use this relation to focus only on the entanglement properties of one of  $\langle L_{\gamma,x} |$  and  $|R_{\gamma,x}\rangle$ : the remaining case can be easily inferred from Eq. (63) upon replacing  $W$  with  $W'$ . Therefore, from now on we will only look at the entanglement of  $\langle L_{\gamma,x} |$ , and, to lighten the notation, we set

$$\rho_{L,\gamma,A} \mapsto \rho_{\gamma,A}, \quad S_{L,A}^{(\alpha)}(\gamma) \mapsto S_A^{(\alpha)}(\gamma). \quad (65)$$

Moreover, we also drop the dependence of  $\langle L_{\gamma,x} |$  on the point  $x$  at which it is computed, i.e.,

$$\langle L_{\gamma,x} | \mapsto \langle L_\gamma |. \quad (66)$$

## IV. TEMPORAL ENTANGLEMENT IN GENERIC UNITARY CIRCUITS

In this section we specify the unitary gates in Eq. (18) to be (independent) Haar random matrices. We consider the temporal entanglement of the state in Eq. (53) for a typical realization of the disorder and in the long time limit. In our analysis we focus on initial states in product form; i.e., we take  $m$  as in Eq. (17). Indeed, we expect that the choice of the initial state, as long as it is short-range entangled, does not affect the general scaling of entanglement in a random circuit.

The use of the Haar random unitaries follows from the philosophy of random matrix theory. By dispensing with all system-specific details, these strongly chaotic gates allow for analytic calculations while retaining the universal properties of entanglement in strongly interacting systems. Recently, there have been various applications of random unitary circuits to explain aspects of quantum chaos and other nonequilibrium features of generic quantum systems; see, for instance, Refs. [64,67,82,119–125] and the review [66] for a more comprehensive list of references.

The (Rényi) entanglement in a random unitary circuit is described by a statistical mechanical model written in terms of permutation degrees of freedom [65,111,126]. The von Neumann entropy is at the replica limit of the model. The permutations originate from pairings of the unitary evolution with its time reversal. To be more specific, let us consider the example of the  $n$ th Rényi entropy, with  $\mathbb{N} \ni n \geq 2$ , of  $\rho_A(t)$ : the regular density matrix reduced to a subregion  $A$  [cf. Eq. (20)]. In each copy of the time-evolved reduced density matrix  $\rho_A(t)$ , there is one forward and one backward time sheet [cf. Eq. (24)]. Therefore, in total, there are  $n$  forward and  $n$  backward time sheets. When performing random averaging over the gates, each copy of a given gate and its Hermitian conjugate are paired in a fashion similar to the Wick theorem of the free fields. The boundary conditions for the Rényi entropies are domain walls between different types of pairings. If we view the pairings as spin degrees of freedom, the effective statistical mechanical model describing the entanglement is in the ordered phase, and the domain wall continues to exist in the bulk, possibly splitting in a cascade of more elementary domain walls. The (Rényi) entanglement entropy is given by the free energy of these generically interacting domain walls. All these microscopic details can be encoded in a coarse grained line tension of the domain wall, which gives rise to the growth rate of entanglement in the long time limit.

In the upcoming subsections we obtain the general scaling of temporal entanglement in three steps

- (1) We show that the boundary conditions to evaluate the purity of  $\langle L_\gamma |$  correspond to domain walls in the statistical problem (Sec. IVA).
- (2) Averaging over the random unitary gates, we show that the minimal-energy configurations are those

where the domains penetrate in the bulk. Minimizing the free energies by means of the line-tension formalism we find a linear growth of temporal entanglement (Sec. IV B).

- (3) Recalling the arguments of Ref. [69] we infer that the domain-wall picture can be applied also to a single realization of random circuit (without averaging) or, equivalently, to systems without randomness (Sec. IV C).

A technical note: in the upcoming calculations we consider the state  $\langle L_\gamma |$ , with the (anti)slope  $v \geq 0$ ; see Fig. 7. Indeed, the  $v \geq 0$  condition gives rise to nontrivial domain-wall configurations. In the case  $v < 0$  our analysis can be applied to  $|R_\gamma\rangle$ .

### A. Boundary conditions for temporal entanglement

The expression of the  $n$ th Rényi entropy contains  $n$  copies of the forward and backward evolution by the circuit. The pairings emerge naturally on the boundary when contracting copies of these circuits to evaluate Rényi entropies, with or without random averaging.

Let us illustrate this idea in the example of the purity  $\text{tr}[\rho_A^2(t)]$  of the quantum state  $\rho_A(t)$ ; see Fig. 6(a). To form the reduced density matrix  $\rho_A(t)$ , we take partial trace in each copy of  $\rho(t)$ . The partial trace operation is denoted as a contraction of the corresponding indices from the forward and backward copies of the circuit. Multiplying two copies of  $\rho_A(t)$  and taking the trace, we obtain the SWAP contraction in region A. In this quantity, there are two copies of the unitary gate  $U$  and two copies of  $U^*$ . There are two ways to contract them, which we denote as  $\mathbb{1}$  and (12) permutations:

$$\mathbb{1}: \overbrace{U \otimes U^* \otimes U \otimes U^*}, \quad (12): \overbrace{U \otimes U^* \otimes U \otimes U^*}. \quad (67)$$

The top boundary thus has a domain-wall boundary condition between  $\mathbb{1}$  and (12) permutations.

Temporal entanglement is defined for an “operator state,” namely a state in the folded space. Therefore, the ket itself involves a forward and a backward evolution: see, e.g., Eq. (53). The permutation boundary conditions are the same if we were to consider an operator state on a spatial slice, which have been computed explicitly in Ref. [80]. For completeness, we repeat the derivation for the second Rényi entropy for the operator state  $\langle L_\gamma |$ . The boundary conditions involve permutations in the symmetric group  $S_4$ . The purity of a subregion on the temporal slice is

$$\frac{\text{tr}_A(\text{tr}_{\bar{A}}(|L_\gamma\rangle\langle L_\gamma|)^2)}{\langle L_\gamma | L_\gamma \rangle^2}. \quad (68)$$

Here we choose the initial product state in the diagrams to be normalized to 1, which is different from the  $m$  state in Eq. (18). To be consistent with the random circuit literature,

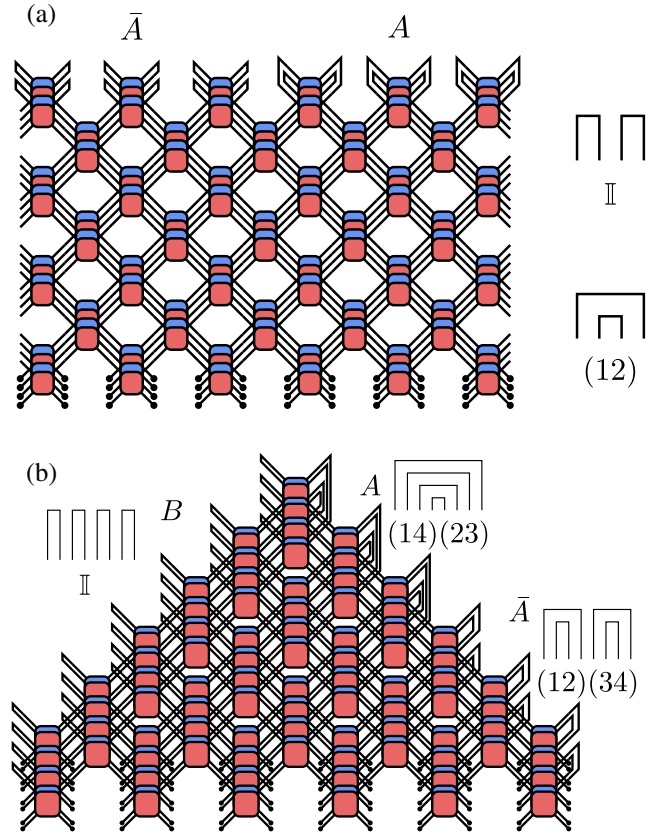


FIG. 6. Permutation boundary conditions for (a) the state purity  $\text{tr}[\rho_A^2(t)]$  and (b) the operator purity  $\text{tr}_A[\text{tr}_{\bar{A}}(|L_\gamma\rangle\langle L_\gamma|)^2]$ . The dots at the bottom represent a generic product initial state. Red gates are forward evolution  $u$ ; blue gates are backward evolution  $u^*$ . (a) The top boundary conditions implement the partial trace, the matrix multiplication of  $\rho_A$  with itself and the trace in  $\bar{A}$ . Right: region A has boundary condition  $\mathbb{1}$ , region  $\bar{A}$  has boundary condition (12). (b) The top boundary conditions are permutation elements in  $S_4$ . They are, respectively, given by  $\mathbb{1}$ , (12)(34), and (14)(23). We consider a general contiguous partition. The ratio of size  $|A|$  and the total size  $|A| + |\bar{A}|$  is set to be  $r \in [0, 1]$ .

we choose to normalize the boundary condition as shown in Fig. 6 for upward pointing legs. For downward pointing legs, we use permutations normalized as the loop state as in Eq. (27). The temporal Rényi-2 entropy thus has two terms,

$$S_A^{(2)}(\gamma) = -\log(\text{tr}_A[\text{tr}_{\bar{A}}(|L_\gamma\rangle\langle L_\gamma|)^2]) + 2\log(\langle L_\gamma |), \quad (69)$$

where the second term is twice the Rényi-2 entropy of  $A\bar{A}$ . The boundary conditions for the first term is shown in Fig. 6(b).

### B. Entanglement in terms of domain-wall line tension: Disorder average

In Fig. 6, we see that pairings (permutations) emerge naturally as boundary conditions when evaluating entanglement-related quantities. In fact, in quantum chaotic

systems these pairings are also the dominant degrees of freedom in the interior of the multilayer unitary evolution. One simple way to introduce pairings in the bulk is through random averaging of the gate over Haar ensemble. Indeed, the latter are the only degrees of freedom surviving the average.

Taking the purity diagram in Fig. 6(a) as an example, each four-layer gate after random averaging can only give a tensor of  $\mathbb{1}$  or  $(12)$  as the output at its bottom legs. Hence we can label the gate with “spin” variables taking values in  $\mathbb{1}$  or  $(12)$ . The  $\mathbb{1}$  and  $(12)$  can form contiguous domains connecting to the  $\mathbb{1}$  and  $(12)$  on the boundary. We label a general domain wall between a pairing  $\sigma$  on the left and a pairing  $\mu$  on the right as  $\sigma^{-1}\mu$ . The domain wall between  $\mathbb{1}$  and  $(12)$  is thus  $\mathbb{1}^{-1}(12) = (12)$ . Because of constraints from unitarity and locality of the interactions, this  $(12)$  domain wall can only wander within the light cone and cannot branch. The entanglement is the free energy, or tension, of the domain wall. Since disorder fluctuations are negligible over large enough scales, the system is asymptotically translationally invariant and the domain wall macroscopically should be a straight line. Using  $v$  to denote the inverse of the domain wall’s slope, we can write the Rényi entropy at leading order as

$$-\log \overline{\text{tr}[\rho_A^2(t)]} \simeq s_{\text{eq}} \min_v \mathcal{E}_H(v)t, \quad (70)$$

where  $\overline{\cdot}$  denotes the average over Haar random gates,  $\mathcal{E}_H(v)$  is the line tension of a domain wall in the Haar random circuit [127],

$$\mathcal{E}_H(v) = \frac{\log \frac{d^2+1}{d} + \frac{1+v}{2} \log \frac{1+v}{2} + \frac{1-v}{2} \log \frac{1-v}{2}}{\log d}, \quad (71)$$

and  $s_{\text{eq}}$  is the infinite temperature equilibrium entropy  $\log d$  (we recall that  $d$  is the local Hilbert space dimension). For a generic product initial state, the domain-wall end point at the bottom is not fixed and the Rényi entropy is obtained by minimizing over different slopes.

Since the random circuit is left-right symmetric after disorder averaging, the minimum in Eq. (70) is taken at  $v = 0$ , i.e., a vertical line; see Fig. 7(a). This gives a linear growth where the line tension is the entanglement growth rate, which is called entanglement velocity [128].

Let us now consider the entanglement of the operator state  $\langle L_\gamma \rangle$ . In particular, we consider the following averaged version of it:

$$\bar{S}_A^{(2)}(\gamma) := -\log \overline{\text{tr}_A[\text{tr}_{\bar{A}}(|L_\gamma\rangle\langle L_\gamma|^2)]} + \log \overline{(\langle L_\gamma | L_\gamma \rangle)^2}. \quad (72)$$

As noted above, this quantity involves permutations in  $S_4$ . In this case the leading contribution is again given by suitable domain walls; however, differently from before

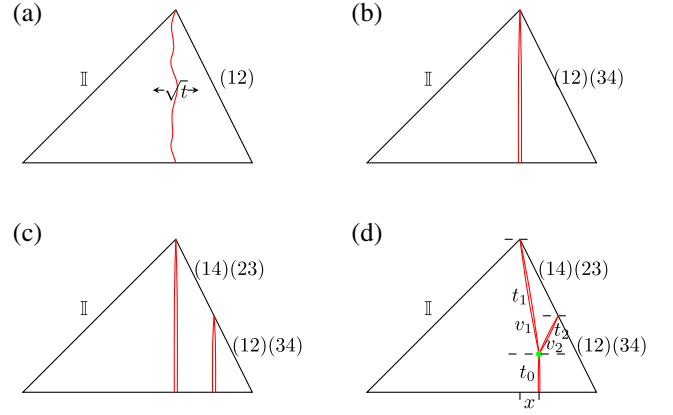


FIG. 7. Domain-wall configurations for puritylike diagrams. (a) Domain wall for purity  $\text{tr}(\rho_A^2)$ . The single domain wall [a transposition  $(12)$ ] undergoes a random walk down to the bottom boundary. The size fluctuation is  $\sqrt{t}$ , which is subleading to  $t$  and ignored in other panels. (b) Domain walls for  $(\text{tr}\rho_A^2)^2$ . The two transpositions are independent. (c) Configurations when the two sets of domain wall do not interact. (d) The two sets of domain wall meet in a middle time slice. The three segments have time duration  $t_1$ ,  $t_2$ , and  $t_0$ , and antislopes  $v_1$ ,  $v_2$ , and  $0$  respectively. The antislope of the right edge of the triangle is  $v$ . The green vertex represents the fusion. It can have a nontrivial weight, but only affects the free energy by an  $\mathcal{O}(1)$  amount.

these domain walls will involve multiple elementary transpositions.

For instance, let us consider the second term on the right-hand side of Eq. (72):

$$F_2(t) := -\log \overline{(\langle L_\gamma | L_\gamma \rangle)^2}. \quad (73)$$

The boundary condition is  $\mathbb{1}$  on subsystem  $B$  and  $(12)(34)$  on subsystem  $A\bar{A}$  [cf. boundary conditions of the first term in Fig. 6(b)]. Hence it has two commutative domain walls at the intersection of  $B$  and  $A\bar{A}$ . These two domain walls are independent; each of them go vertically down to the bottom [Fig. 7(b)] and give a contribution of  $\mathcal{E}_H(0)t$ . And thus the second term corresponds to twice the state Rényi entropy:

$$F_2(t) \simeq 2s_{\text{eq}}\mathcal{E}_H(0)t. \quad (74)$$

The boundary conditions of the first term in Eq. (69) contain four domain walls. There are two commutative transpositions  $(14)(23)$  at the tip of the diagram (cut between  $B$  and  $A\bar{A}$ ), and other two transpositions  $(13)(24)$  at the entanglement cut between  $A$  and  $\bar{A}$ . The two transpositions  $(14)(23)$  alone are independent, so are  $(13)(24)$ . If the two sets do not meet at an intermediate time slice before reaching the bottom boundary, we can separately minimize their free energies. The equilibrium configuration is that all the domain walls go down vertically [Fig. 7(c)], giving total free energy,

$$F_{1,\text{vert}}(t) := -\log \left\{ \text{tr}_A [\text{tr}_{\bar{A}} (|L_\gamma\rangle\langle L_\gamma|^2)] \right\} \\ \simeq 2s_{\text{eq}}\mathcal{E}_H(0)t + 2s_{\text{eq}}\mathcal{E}_H(0)(1-r)t, \quad (75)$$

where  $r \in [0, 1]$  is the ratio of size  $|A|$  and total size  $|A| + |\bar{A}|$ . The difference with  $F_2(t)$  in Eq. (74) is then  $\mathcal{E}_H(0)t$ : This gives us the following upper bound,

$$\bar{S}_A^{(2)}(\gamma) \lesssim 2(1-r)\mathcal{E}(0)t. \quad (76)$$

If the two sets meet in the middle, the domain walls can fuse to different permutations according to the group multiplication rules. For example, if (13)(24) and (14)(23) completely fuse together, the domain wall becomes (13)(24)  $\times$  (14)(23) = (12)(34). We see that the number of domain walls reduces to two, thus reducing the energy cost. For this reason a configuration like the one displayed in Fig. 7(d) can compete for the minimal free energy.

We now set up the minimization problem assuming the two sets to have (anti)slope  $v_1$  and  $v_2$  before they meet in the middle. After the fusion, the resulting domain wall (12)(34) is composed by two independent components. They cost free energy  $2\mathcal{E}_H(0)t_0$  for the remaining duration of  $t_0$ . The vertex can have a nontrivial weight, but, as long as it is not zero, it only brings in a  $\mathcal{O}(1)$  correction and can be neglected when considering the leading order free energy in the long time limit. The geometry is depicted in Fig. 7(d). Therefore, we can write the free energy as

$$F_{1,Y}(t, t_0) \simeq 2s_{\text{eq}}[\mathcal{E}_H(v_1)t_1 + \mathcal{E}_H(v_2)t_2 + \mathcal{E}_H(0)t_0], \quad (77)$$

where  $t_0$ ,  $t_1$ , and  $t_2$  are the duration of the two sets of domain walls and satisfy the geometric relations. The subscript  $Y$  denotes the merging configuration.

$$v_1 = \frac{x}{t_1}, \quad v_2 = \frac{x - rv_\gamma t}{t_2}. \quad (78)$$

To parametrize time, we set

$$t_0 = r_0 t, \quad t_1 - t_2 = rt, \quad t_1 + t_0 = t, \quad (79)$$

where  $r \in [0, 1]$  is the ratio of size  $|A|$  with respect to  $|A| + |\bar{A}|$ , and  $r_0 \in [0, 1-r]$  is the portion of the merged domain wall.  $r_0 = 0$  corresponds to merge at the bottom;  $r_0 = 1-r$  corresponds to taking  $v_1 = v_\gamma$ , i.e., merge immediately when available. To minimize  $F_1(t, t_0)$ , we first fix  $t_0$  and vary  $x$ . The implicit  $x$  derivative gives

$$\mathcal{E}'_H\left(\frac{x}{t_1}\right) + \mathcal{E}'_H\left(x - \frac{v_\gamma r t}{t_2}\right) = \mathcal{E}'_H(v_1) + \mathcal{E}'_H(v_2) = 0. \quad (80)$$

For a reflection symmetric system the physical solution is

$$v_1 = -v_2; \quad (81)$$

namely, the two sets of domain walls in Fig. 7(d) meet symmetrically from left and right toward each other.

The minimization with respect to  $t_0$  depends on the explicit form of the line-tension function. In particular, an explicit calculation for the case of the random circuit line tension  $\mathcal{E}_H$  is carried out in Appendix A. The resulting expression of  $\min_{t_0} F_{1,Y}(t, t_0)$  is piecewise continuous in  $v$  and depends on  $d$ . Nevertheless, it is still linear in  $t$  [cf. Eq. (A6)]. Combining with the linear bound in Eq. (76), we conclude that in a Haar random circuit,

$$\bar{S}_A^{(2)}(\gamma) \simeq s_{\text{eq}}v_{\text{TE,H}}^{(2)}t. \quad (82)$$

where the temporal Rényi entanglement velocity  $v_{\text{TE,H}}^{(2)} > 0$ . The explicit expressions of  $v_{\text{TE,H}}^{(2)}$  for Haar random circuit can be found in Eq. (A12).

### C. Typical circuit without averaging

In this subsection, we argue that the line-tension formalism discussed above can be applied to the calculation of temporal entanglement in generic chaotic circuits without introducing disorder averaging. The recipe is to replace  $\mathcal{E}_H(v)$  with a “dressed” line-tension function  $\mathcal{E}(v)$  characterized by the following general properties [111],

$$\mathcal{E}(v) \geq |v|, \quad \mathcal{E}''(v) \geq 0, \quad \mathcal{E}(v) = \mathcal{E}(-v), \quad (83)$$

where the last one follows from the parity symmetry of the system.

The basic arguments follow Ref. [69], where the concept of line-tension function is generalized to nonrandom circuits and we will briefly recall them here. The key observation is that the pairings between a unitary gate and its Hermitian conjugate continue to dominate the configuration sum, or “path integral,” that determines the Rényi entropies. Indeed, these are real positive quantities that do not suffer from phase cancellation. A single domain wall, such as (12), will be dressed by nonpairing degrees of freedom, but only perturbatively to have an  $\mathcal{O}(1)$  width.

Our problem is slightly more complicated than the one discussed in Ref. [69] because the contributions generating temporal entanglement involve composite domain walls [for instance, a domain wall (123) can appear]. Nevertheless, the two sets of domain walls (14)(23) and (13)(24) have the same dressed line-tension function  $\mathcal{E}(v)$  when they do not interact. Indeed, (14)(23) can be mapped to (12)(34) by relabelling the third and fourth copies of the unitary and its complex conjugate—it is a symmetry of the multireplica dynamics if we look at the patch of (14)(23) alone. The symmetry no longer holds when the two sets of domain walls meet each other and interact. This process, however, only dresses the interaction vertex in Fig. 7(d), which introduces an order  $\mathcal{O}(1)$  correction to the free energy. Below the interaction vertex, the two sets of domain walls

fuse to (12)(34), which again has the same line-tension function  $\mathcal{E}(v)$ . This justifies the use of a single line-tension function  $\mathcal{E}(v)$  to characterize the scaling of the temporal Rényi-2 in Eq. (69).

Specifically, following the steps discussed in the previous section, we obtain

$$S_A^{(2)}(\gamma) \simeq s_{\text{eq}} v_{\text{TE}}^{(2)} t, \quad (84)$$

where the temporal Rényi entanglement velocity is determined by

$$v_{\text{TE}}^{(2)} = \min \left[ \min_{r_0 \in [0, r]} \mathcal{F}(r_0), 2(1-r)\mathcal{E}(0) \right]. \quad (85)$$

The parameters  $r_0$  and  $r$  are defined in Eq. (79) and

$$\mathcal{F}(r_0) = 2[\mathcal{E}[v_1(r_0)] - \mathcal{E}(0)](1-r_0) \quad (86)$$

$$+ 2\mathcal{E}[v_1(r_0)](1-r_0-r). \quad (87)$$

In this case, the remaining minimization over  $r_0$  cannot be performed explicitly as the minimum depends on the precise form of the line tension. However, we can show that for  $\mathcal{E}(0) > 0$ ,

$$\min_{r_0 \in [0, r]} \mathcal{F}(r_0) = 0 \Leftrightarrow \mathcal{E}(v_\gamma) = \mathcal{E}(0), \quad (88)$$

which implies generic linear growth of the temporal Rényi-2 entropy apart from marginal cases.

We begin to prove this property by noting that the two terms in Eqs. (86) and (87) are both non-negative,

$$\begin{aligned} [\mathcal{E}[v_1(r_0)] - \mathcal{E}(0)](1-r_0) &\geq 0, \\ \mathcal{E}[v_1(r_0)](1-r_0-r) &\geq 0, \end{aligned} \quad (89)$$

in the relevant range  $r_0 \in [0, r]$ , because of the convexity and parity of line-tension function. Indeed, these two properties imply that the function is either constant or has a unique local minimum in  $v = 0$ , i.e.,

$$\mathcal{E}(v) \geq \mathcal{E}(0) > 0, \quad \forall v. \quad (90)$$

In fact, the above inequality indicates that Eq. (89) can both be zero for generic  $r$  only if  $r_0 = r$  and

$$\mathcal{E}(v_\gamma) = \mathcal{E}(0). \quad (91)$$

Noting that the reversed implication is obvious, we conclude the proof.

In fact, due to the general properties (83) of the line-tension function, Eq. (91) admits solution only in two cases: either constant line-tension function [Eq. (6)] or

$v_\gamma = 0$  [Eq. (7)]. These are the two marginal cases mentioned in Sec. IA.

## V. TEMPORAL ENTANGLEMENT IN CHAOTIC DUAL-UNITARY CIRCUITS

In this section we consider the first of the two marginal cases identified in Eq. (91): the one in which the line-tension function is constant. As discussed in Sec. IA [cf. the discussion around Eq. (6)], this situation can only be realized when the local gates forming the time evolution operator in Eq. (10) are dual unitary [57]. In terms of our diagrammatic representation, the dual-unitarity condition means that the gates fulfill

$$\begin{array}{c} \text{---} \\ \text{---} \end{array} \begin{array}{c} \text{---} \\ \text{---} \end{array} = \begin{array}{c} \text{---} \\ \text{---} \end{array}, \quad (92)$$

in addition to the standard unitarity conditions (23). Without additional fine-tuning, the gates fulfilling Eqs. (23) and (92) are quantum chaotic [35,57,71].

Imposing the condition (92) enables one to make a number of exact statements concerning dynamics and spectral properties of the quantum circuit [34,35,38,40,58,62,63,71–73,75,76,78,79,85,86]. In particular, dual-unitary circuits have been shown to admit a class of “solvable” initial states for which one can compute exactly the full time evolution of any local subsystem [58]. For solvable initial states the generalized influence matrices  $\langle L_\gamma \rangle$  and  $|R_\gamma\rangle$  [cf. Eqs. (54) and (53)] take the following product form:

$$\langle L_\gamma \rangle = \langle \bigcirc \rangle^{\otimes |\gamma|}, \quad |R_\gamma\rangle = |\bigcirc\rangle^{\otimes |\gamma|}, \quad (93)$$

where  $|\bigcirc\rangle$  is the loop state of Eq. (27). This form immediately implies a strictly vanishing temporal entanglement.

Here, however, we are interested in the behavior of temporal entanglement for *generic*, nonsolvable, initial states. Specifically—recalling that for the family of initial states Eq. (15) considered here the solvable instances correspond to the cases where the matrix  $m$  is unitary [77]—we consider the case where  $m$  is *not* unitary.

Since the time evolution in a chaotic system should not depend on the initial configuration, one might expect that the behavior of solvable states is somewhat representative of the generic situation. Namely, that the temporal entanglement is always small for dual-unitary circuits. In fact, as we now discuss, this intuition turns out to be incorrect: Even though higher temporal Rényi entropies are bounded by a sublinear function of time (in agreement with our entanglement-membrane analysis of the previous section), the von Neumann temporal entanglement entropy is always linear for non-fine-tuned dual-unitary circuits. In the upcoming subsections we show these facts by



analyzing separately the cases of higher Rényi entropies and von Neumann entropy.

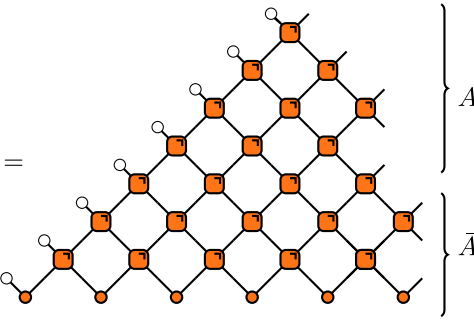
For simplicity, in the main text we consider paths  $\gamma = A\bar{A}$  with *constant* slope  $v_\gamma$ . This means that the slope is the same in both  $A$  and  $\bar{A}$ . This assumption is lifted in Appendixes B and C where we present the most general form of our results.

### A. Bound on temporal higher Rényi entropies

In this subsection we show that higher temporal Rényi entropies are *sublinear* in time, in agreement with the entanglement membrane analysis of the previous section. More precisely, we prove the following bound,

$$S_A^{(\alpha>1)}(\gamma) \leq \frac{\alpha}{\alpha-1} \log \left[ \frac{d^{\tau_A} \mathcal{P}(\tau/2)}{\mathcal{P}(\tau_{\bar{A}}/2)} \right], \quad (94)$$

where  $A, \bar{A}$  correspond to a contiguous bipartition of the  $2t$  legs of the influence matrix,



$$\langle L_\gamma | = \left. \begin{array}{c} \text{Lattice structure} \\ \text{---} A \\ \text{---} \bar{A} \end{array} \right\}, \quad (95)$$

$\tau_A$  ( $\tau_{\bar{A}}$ ) is the number of up-pointing legs in  $A$  ( $\bar{A}$ ) fulfilling

$$\tau_A + \tau_{\bar{A}} = (1 + v_\gamma)t \equiv \tau, \quad (96)$$

while

$$\mathcal{P}(t) = \text{tr}[\rho_{[0,\infty)}(t)^2] \quad (97)$$

is the purity of the “regular” reduced density matrix  $\rho_{[0,\infty)}(t)$  [cf. Eq. (24)] corresponding to a half-infinite subsystem  $[0, \infty)$  with open boundary conditions.

To prove Eq. (94) we proceed in two steps, which are detailed in Appendix B.

*Step 1. General bound on higher Rényi entropies.*—We take advantage of the unitarity of the gates and of the Eckart-Young theorem [115] to bound the temporal Rényi entropies in terms of the norm of the state  $\langle L_\gamma |$ :

$$S_A^{(\alpha)}(\gamma) \leq \frac{\alpha}{\alpha-1} \log \left( \frac{\langle L_{\bar{A}} | L_{\bar{A}} \rangle}{\langle L_\gamma | L_\gamma \rangle} \right). \quad (98)$$

*Step 2. Dual-unitary case.*—Specializing the treatment to the dual-unitary case we can relate  $\langle L_\gamma | L_\gamma \rangle$  to the spatial purity [cf. Eq. (97)]. In particular, we find

$$\langle L_\gamma | L_\gamma \rangle = d^r \mathcal{P}(\tau/2). \quad (99)$$

Plugging it into Eq. (98) we obtain Eq. (94).

The physical interpretation of Eq. (94) is immediate: in dual-unitary circuits the growth of higher temporal Rényi entropies is controlled by that of spatial purity. If the initial state is solvable, then the purity is minimized to  $d^{-2t}$  and the temporal entanglement is zero. For more general, nonsolvable states the purity is no longer strictly  $d^{-2t}$ , but—since dual-unitary circuits maintain a maximal entanglement velocity [77]—it can only acquire subexponential corrections. This implies that all higher temporal Rényis are *sublinear* in time.

To make further progress we introduce the following assumption.

*Assumption 1.*—For any generic dual-unitary circuit evolving from a non solvable state, we have

$$\mathcal{P}(t) \simeq \frac{Ct}{d^{2t}}, \quad (100)$$

where  $\simeq$  denotes the leading order in the asymptotic expansion for large times and  $C > 0$  a time independent constant.

The scaling in Eq. (100) can be proven by averaging each dual-unitary gate of the circuit over random single qubit rotations of its legs (see Ref. [77] and Appendix D). Thus, we expect it to hold for typical dual-unitary circuits: This is in agreement with our numerical investigations, as shown in Fig. 8 for some representative examples.

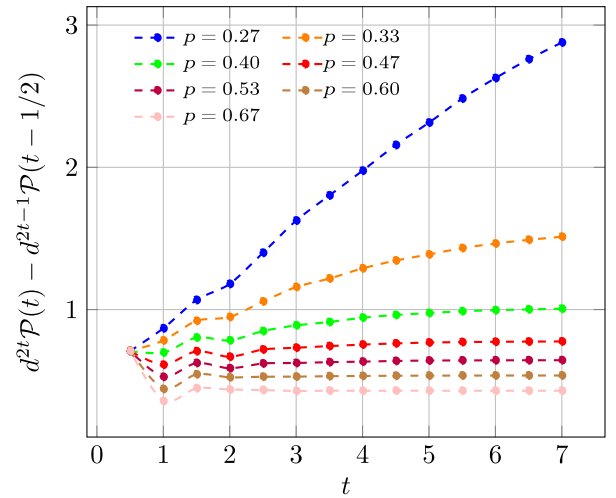


FIG. 8. Evolution of  $d^{2t}\mathcal{P}(t) - d^{2(t-1)}\mathcal{P}(t-1)$  as a function of  $t$  for homogeneous dual-unitary circuits with different entangling power  $p$  [cf. Eq. (12)]. If the scaling (100) holds, this quantity should saturate to the constant  $C/2$ . We considered  $d = 2$  and parametrized the gates as described in Appendix E. Note that  $C$  grows upon decreasing  $p$ : this is consistent with the fact that at the noninteracting point  $p = 0$  the purity decays with an exponent  $\lambda < 2 \log d$ .

As discussed in Appendix D we expect the case  $C = 0$  to hold only for solvable initial states, for which  $\mathcal{P}(t) = d^{-2t}$  and the temporal entanglement is identically 0 for any bipartition [58]. The discussion in Appendix D, however, also shows that these states are “unstable” from the point of view of purity scaling: for any arbitrary small perturbation of a solvable state one has Eq. (100) with  $C > 0$  and the behavior of temporal entanglement is the one we discuss here.

Using Assumption 1 and considering an appropriate scaling of the bipartition,

$$r \equiv \frac{|A|}{2t} < 1, \quad (101)$$

it is immediate to show that the higher Rényi entropies saturate to a constant [129],

$$S_A^{(\alpha)}(\gamma) \simeq \frac{\alpha}{\alpha-1} \log(1-r). \quad (102)$$

A direct numerical test of Eq. (102) is not straightforward as we have only access to short times. Therefore, we can only consider gates for which the asymptotic form Eq. (100) is attained early. With this restriction the bound appears convincingly obeyed. For instance, in Fig. 9 we consider a comparison between Eq. (102) and the four gates of Fig. 8 with higher entangling power.

## B. Linear growth of temporal entanglement entropy

In Sec. VA, we showed that the higher Rényi entropies are bounded by a constant for any partition with ratio  $r < 1$ . However, since Rényi entropies are nonincreasing functions of the Rényi index, this result only provides a lower bound for the temporal entanglement entropy, i.e.,

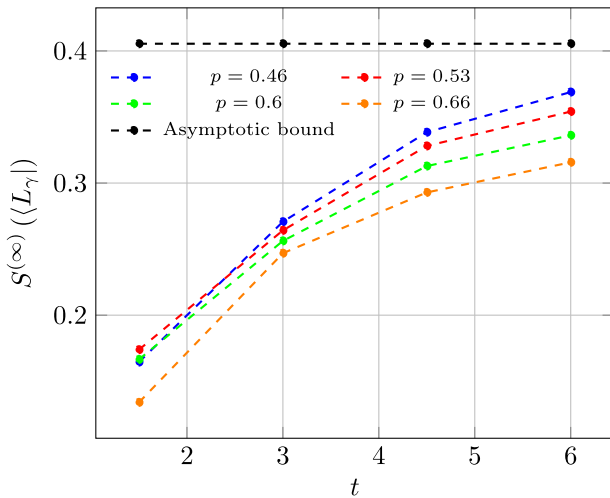


FIG. 9. Growth of  $S^{(\infty)}$  for the state  $L_\gamma$ , taken with  $v_\gamma = 1$ , for gates of various entangling power  $p$ , compared with the asymptotic bound given by Eq. (102), for a fixed value of  $r = 1/3$ .

$$S_A(\gamma) = \lim_{\alpha \rightarrow 1} S_A^{(\alpha)}(\gamma). \quad (103)$$

In this subsection, we show that  $S_A(\gamma)$  in a typical dual-unitary circuit grows linearly in time for nonsolvable initial states.

Denoting again by  $r$  the ratio between the number of legs in region  $A$  and the total [cf. Eq. (101)], we can bound  $S_A(\gamma)$  from above and below,

$$s(r)t \log d \leq S_A(\gamma) \leq s(r)t \log d + \mathcal{O}(\log t), \quad (104)$$

with the same function:

$$s(r) = \begin{cases} (1 + v_\gamma)r^2 & r \in [0, \frac{2}{v_\gamma+3}] \\ \frac{4(1-r)[(2+v_\gamma)r-1]}{1+v_\gamma} & r \in (\frac{2}{v_\gamma+3}, 1]. \end{cases} \quad (105)$$

The  $\log(t)$  margin in Eq. (104) is subleading with respect to the linear scaling of  $s(r)t \log d$  and therefore the latter determines the long timescaling of  $S_A(\gamma)$ .

In the derivation of the upper bound, we only use Assumption 1. For the lower bound, we additionally employ Assumption 2.

*Assumption 2.*—The membrane picture of entanglement holds for the second Rényi entropy of the state in Eq. (119).

This assumption is in line with general expectations from the membrane theory [69] and can be verified numerically. A representative example is reported in Fig. 10. We see that, even though there are strong deviations for short times, the numerical results seem to approach the membrane theory predictions as time increases (see Appendix C 3 for a more thorough discussion of the validity of this assumption).

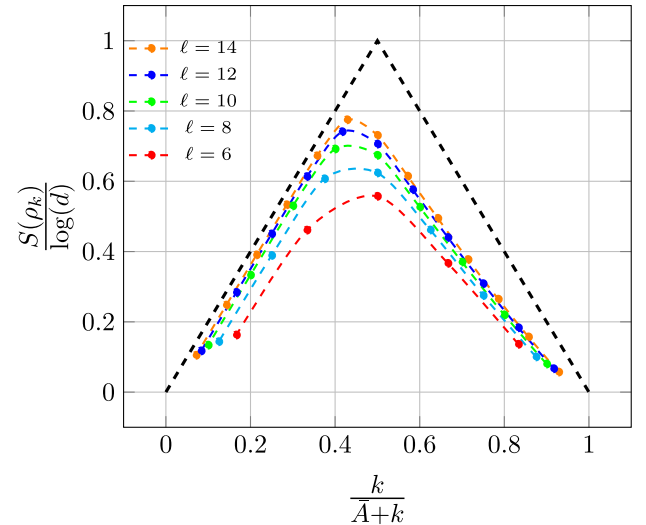


FIG. 10. Slope of the entanglement entropy of the matrices  $\rho_k$  as a function of  $k$ , for various values of  $\ell \equiv \bar{A} + k$  accessible numerically. The dashed line represents the asymptotic profile of the curve according to Eq. (123) as per Assumption 2. The path chosen is the one at the edge of the lightcone.

We sketch the proof in three steps.

*Step 1. Reduction.*—First we observe that the entanglement of a state, for a given bipartition, is unaffected by the action of unitary matrices acting locally on the two separate bipartitions. Thanks to this observation and the dual unitarity of the gates, we can consider a simplified version of our state  $\langle L_\gamma | \mapsto \langle L_{\gamma'} |$  where the new path  $\gamma'$  corresponds to the edges of the light cones of the two bipartitions (see the detailed derivation in Appendix C 1). This is easily understood graphically by looking at the diagram below and noting that the area in red corresponds to matrices that, when viewed horizontally, are unitary. Therefore, removing them will not affect its entanglement.

$$\langle L_\gamma | = \text{Diagram} \mapsto \text{Diagram} \quad (106)$$

After this operation we end up with some bullet states on the top of region  $A$ , which do not entangle with any other part of the system: They can also be removed without affecting the result.

To sum up, as far as the entanglement is concerned, we can reduce  $\langle L_\gamma |$  to the following state,

$$\langle L_{\gamma'} | = \text{Diagram} \quad (107)$$

in which region  $A$  has  $\tau_A = (1 + v_\gamma)|A|/2$  sites and region  $\bar{A}$  has  $|\bar{A}|$  sites.

*Step 2. Lower and upper bounds.*—We define  $\tau_A + 1$  orthogonal projectors in region  $A$ :

$$\begin{aligned} P_0 &= |\circ\rangle\langle\circ| \otimes |\circ\rangle\langle\circ| \otimes \cdots |\circ\rangle\langle\circ| \otimes \mathbb{1}_{\bar{A}}, \\ P_1 &= |\circ\rangle\langle\circ| \otimes |\circ\rangle\langle\circ| \otimes \cdots (\mathbb{1}_{d^2} - |\circ\rangle\langle\circ|) \otimes \mathbb{1}_{\bar{A}}, \\ P_k &= |\circ\rangle\langle\circ|^{\otimes \tau_A - k} \otimes (\mathbb{1}_{d^2} - |\circ\rangle\langle\circ|) \mathbb{1}_{d^2}^{\otimes k-1} \otimes \mathbb{1}_{\bar{A}}, \\ P_{\tau_A} &= (\mathbb{1}_{d^2} - |\circ\rangle\langle\circ|) \otimes \mathbb{1}_{d^2} \otimes \cdots \mathbb{1}_{d^2} \otimes \mathbb{1}_{\bar{A}}. \end{aligned} \quad (108)$$

In words, the projector has three different actions, which we highlight by different colors in the following graphical equation:

$$\langle L_{\gamma'} | P_k = \text{Diagram} \quad (109)$$

The  $k$ th projector  $P_k$  keeps the bottom  $k - 1$  sites (blue) intact, projects each of the top  $\tau_A - k$  sites (red) to a bullet state and the  $k$ th site (green) to the orthogonal complement of the bullet state.

One can easily verify that the projectors are orthogonal and form a complete basis, i.e.,

$$P_i P_j = \delta_{ij} P_i, \quad \sum_{k=0}^{\tau_A} P_k = \mathbb{1}_A \otimes \mathbb{1}_{\bar{A}}. \quad (110)$$

These projectors decompose  $|L_{\gamma'}\rangle$  into  $\tau_A + 1$  states, which are orthogonal in  $A$ . Namely,

$$\text{tr}_A(P_i |L_{\gamma'}\rangle \langle L_{\gamma'}| P_j) = 0, \quad i \neq j. \quad (111)$$

The reduced density matrix,

$$\rho_{\bar{A}} = \frac{1}{\langle L_{\gamma'} | L_{\gamma'} \rangle} \text{tr}_A(|L_{\gamma'}\rangle \langle L_{\gamma'}|), \quad (112)$$

is then written as a classical mixture of  $\tau_A + 1$  reduced density matrices,

$$\rho_{\bar{A}} = \sum_{k=0}^{\tau_A} P_k \rho_k, \quad (113)$$

where the classical probability is

$$p_k = \frac{\text{tr}(P_k |L_{\gamma'}\rangle \langle L_{\gamma'}|)}{\langle L_{\gamma'} | L_{\gamma'} \rangle}, \quad (114)$$

and the reduced density matrices are

$$\rho_k = \frac{\text{tr}_A(P_k |L_{\gamma'}\rangle \langle L_{\gamma'}|)}{\text{tr}(P_k |L_{\gamma'}\rangle \langle L_{\gamma'}|)}. \quad (115)$$

The concavity lower bound and mixing upper bound of  $S(\rho_{\bar{A}})$  confine the von Neumann entropy to the following interval:

$$\sum_{k=0}^{\tau_A} p_k S(\rho_k) \leq S_A(\gamma) \leq \sum_{k=0}^{\tau_A} p_k S(\rho_k) - \sum_{k=0}^{\tau_A} p_k \log p_k. \quad (116)$$

The Shannon entropy of the classical probability  $p_k$  is at most  $\log(\tau_A + 1) \sim \mathcal{O}(\log t)$ . Therefore, we conclude that

$$\sum_{k=1}^{\tau_A} p_k S(\rho_k) \leq S_A(\gamma) \leq \sum_{k=1}^{\tau_A} p_k S(\rho_k) + \mathcal{O}(\log t), \quad (117)$$

where we also removed  $k = 0$  from the summation since  $S(\rho_0) = 0$ .

*Step 3. Evaluation of  $\sum_{k=1}^{\tau_A} p_k S(\rho_k)$ .*—We evaluate  $p_k$  in Appendix C 2 using Assumption 1. The asymptotic expression reads as



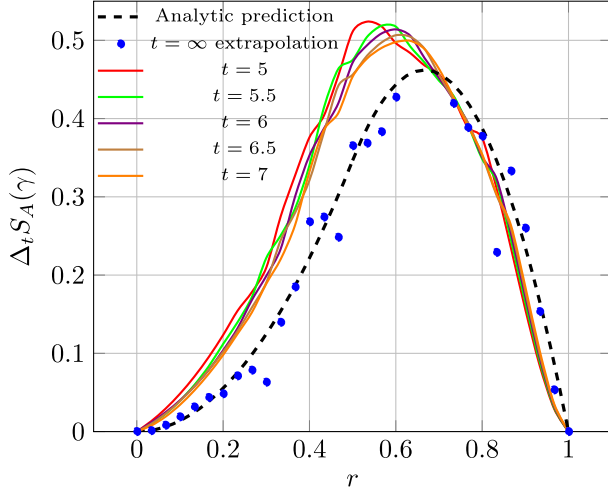


FIG. 13. Slope of the entanglement entropy for the state  $\langle L_\gamma |$ , obtained by taking finite differences of  $S_A(\gamma)$  for two subsequent time steps. Given the discrete nature of the states, only some rational values of  $r$  are allowed at each time, so we interpolated between those in order to take the difference. In blue, we show an extrapolation of these data in the limit  $t \rightarrow \infty$ , which we ultimately compare with the asymptotic prediction, in black [obtained from Eq. (105)]. The extrapolation is attained by observing that, due to the logarithmic form of the corrections to  $S_A(\gamma)$ , for large enough  $t$  we have  $\Delta_t S_A(\gamma) \simeq A + B/t$ , where  $A$  is the desired asymptotic value. Then, we performed a linear fit of the data in  $1/t$  to estimate  $A$ .

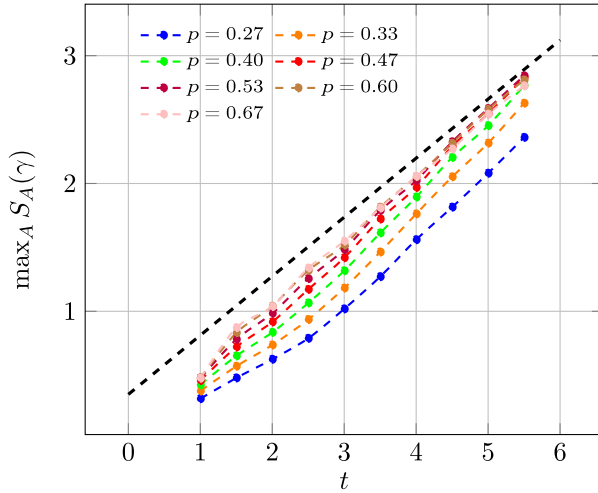


FIG. 14. Maximal entanglement of the influence matrix on the diagonal path  $\gamma_{lc}$  [cf. Eq. (56)] versus the length of the path  $|\gamma_{lc}|$  for different values of the entangling power  $p$ . The entanglement reported is the maximum attained among all the possible non-disjoint bipartitions. The gates are parametrized as explained in Appendix E. The asymptotic growth seems to be independent of the entangling power. The initial  $p$ -dependent transient is larger for smaller values of  $p$ . This is consistent with the fact that for  $p = 0$  the gates are SWAPS and  $\langle L_{\gamma_{lc}} |$  is a product state for all initial states Eq. (15). The black dashed line corresponds to the theoretical prediction of the growth, as in Eq. (105), plus an arbitrary constant chosen for convenience.

## VI. TEMPORAL ENTANGLEMENT OF THE VERTICAL STATE

In this section, we consider the second marginal case of Eq. (91) in which the timelike surface is vertical ( $v_\gamma = 0$ ) and the chaotic quantum circuit is arbitrary. Namely, we look at the scaling in  $|\gamma| = 2t$  of the temporal entanglement of the original influence matrix for generic circuits.

We find that higher Rényi entropies grow logarithmically in time:

$$S_A^{(\alpha)}(\gamma) \sim \log(t), \quad \alpha > 1. \quad (124)$$

We begin by showing the sublinear growth via a direct application of the Eckart-Young strategy employed in Appendix B 1. Specifically, we use the upper bound,

$$\begin{aligned} S_A^{(\alpha)}(\gamma) &\leq \frac{\alpha}{1-\alpha} \log \frac{[\langle L_\gamma | (|\Psi_A\rangle \otimes |\Psi_{\bar{A}}\rangle)]^2}{\langle L_\gamma | L_\gamma \rangle \langle \Psi_A | \Psi_A \rangle \langle \Psi_{\bar{A}} | \Psi_{\bar{A}} \rangle} \\ &:= \frac{2\alpha}{1-\alpha} \log r_t, \end{aligned} \quad (125)$$

by means of the overlap of the state  $\langle L_\gamma |$  and a factorized state  $\langle \Psi_A | \otimes \langle \Psi_{\bar{A}} |$ .

To find a product state with large overlap we employ the membrane theory. Specifically, we consider the state

$$\langle \Psi_A | \otimes \langle \Psi_{\bar{A}} | = \langle L_{\gamma/2} | \otimes \langle L_{\gamma/2} |, \quad (126)$$

which is depicted in Fig. 15(b). Assuming the circuit to be Haar random, the norm of these states are determined by the line tension at  $v = 0$  as follows,

$$\begin{aligned} \overline{\langle L_\gamma | L_\gamma \rangle} &\sim \exp[-\mathcal{E}_H(0) \log(d)t], \\ \overline{\langle \Psi_A | \Psi_A \rangle} &\sim \exp[-\mathcal{E}_H(0) \log(d)t/2], \\ \overline{\langle \Psi_{\bar{A}} | \Psi_{\bar{A}} \rangle} &\sim \exp[-\mathcal{E}_H(0) \log(d)t/2], \end{aligned} \quad (127)$$

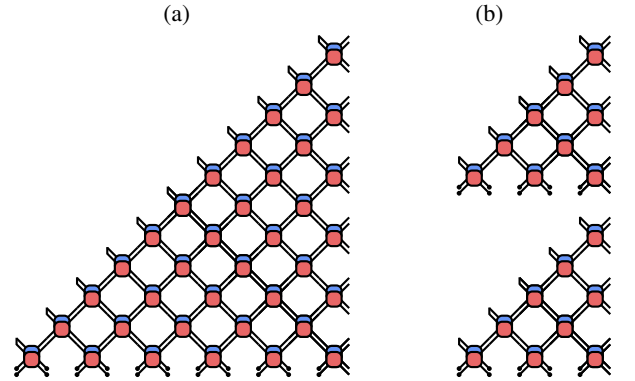


FIG. 15. (a) The vertical state. (b) A tensor product state on  $A$  and  $\bar{A}$ . The state is written as a tensor product of two vertical states defined on a time lattice with half of the sites.

where we recall [cf. Eq. (71)]

$$\mathcal{E}_H(0) = \frac{\log(d^2 + 1) - \log 2d}{\log d}. \quad (128)$$

The average of the overlap is

$$\overline{\langle L_\gamma | (|\Psi_A\rangle \otimes |\Psi_{\bar{A}}\rangle)} \sim \exp[-\mathcal{E}_H(0) \log(d)t]. \quad (129)$$

The estimation in Eq. (129) relies upon evaluating the random averaging in Fig. 16(c), where the region in which the two states differ is only populated by the permutation  $\mathbb{1}$ . Thus, the minimal free-energy configuration continues to have a domain wall going vertically down. Combining Eqs. (127) and (129) we find

$$\bar{r}_t = \frac{|\overline{\langle L_\gamma | (|\Psi_A\rangle \otimes |\Psi_{\bar{A}}\rangle)}|}{\sqrt{\langle L_\gamma | L_\gamma \rangle \langle \Psi_A | \Psi_A \rangle \langle \Psi_{\bar{A}} | \Psi_{\bar{A}} \rangle}} = O(t^\alpha), \quad (130)$$

The exponent  $\alpha$  can be found by studying the subleading contributions from the random walk of the domain wall. All the domain walls in Fig. 16 are subject to the noncrossing condition at the right boundary. If we view from bottom to top, this is the random walk that first hit  $x = 0$  (the

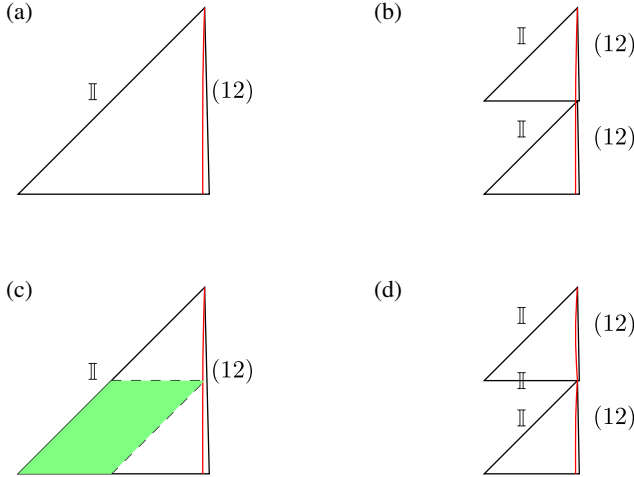


FIG. 16. The domain-wall analysis that produces Eqs. (127) and (129). (a)  $\overline{\langle L_\gamma | L_\gamma \rangle}$ : the dominant configuration is one domain wall going down vertically at equilibrium. (b)  $\overline{\langle \Psi_A | \otimes \langle \Psi_{\bar{A}} | (|\Psi_{\bar{A}}\rangle \otimes |\Psi_A\rangle)}$ : the two parts of the product state factorize so there are two independent vertical domain walls with half the size of (a). (c)  $\overline{\langle L_\gamma | (|\Psi_A\rangle \otimes |\Psi_{\bar{A}}\rangle)}$ : the green region is where  $\langle L_\gamma |$  and  $\langle \Psi_{\bar{A}} | \otimes \langle \Psi_A |$  differ. It has only one copy of  $u \times u^*$ , which, upon averaging, generates a patch of  $\mathbb{1}$ . The domain wall will avoid the green region and goes down vertically. (d) The green region after average provides the  $\mathbb{1}$  boundary conditions for the two dashed lines in (c) (see main text for the bottom rim of the upper triangle; also see Appendix F). The top triangle has a pinned domain wall ending at the rightmost point, the bottom triangle still host a free random walk.

coordinate of the right boundary) for  $t = 0$  (the final time when viewing from bottom to top). The probability distribution for this process is known as the Lévy-Smirnov distribution and reads as

$$p(x) = \frac{x}{t^{3/2}} e^{-x^2/t}. \quad (131)$$

The three independent averages inside the square root in the numerator of Eq. (130) [one in Fig. 16(a), the other two in Fig. 16(b)] correspond to a free boundary condition at the bottom, each of which contributes a polynomial factor  $t^{-1/2}$  (integrate the Lévy-Smirnov distribution in  $x$ ). For the average of the overlap, the green region in Fig. 16(c) represents the missing part in  $\langle \Psi_A | \langle \Psi_{\bar{A}} |$  compared with  $\langle L_\gamma |$ . It can only produce  $\mathbb{1}$ , which becomes the boundary condition of the lower triangle and bottom rim of the top triangle in Fig. 16(d) (the boundary condition for the bottom rim after random averaging is  $|\mathbb{O}\rangle$  tensor product with the vectorized density matrix of the initial state in the folded space. Once we project this state in the space spanned by  $\mathbb{1}$  and (12), as prescribed by the Haar average of the gates, it becomes  $|\mathbb{O}\mathbb{O}\rangle$ , i.e., a  $\mathbb{1}$  boundary). The lower triangle contributes a  $t^{-1/2}$  factor as we argued above. Instead, because of the  $\mathbb{1}$  boundary condition at the bottom rim, the domain wall in the top triangle is penalized by a factor of  $1/d$  when it further moves to the left. Thus the domain wall is pinned to a slope of  $v = 0$ . We end up with a  $t^{-3/2}$  factor in the Lévy-Smirnov distribution for the pinned domain wall. Putting it all together, we have

$$\bar{r}_t \simeq \frac{t^{-3/2} t^{-1/2}}{\sqrt{t^{-1/2} t^{-1/2} t^{-1/2}}} = t^{-5/4}. \quad (132)$$

These random wall arguments can be made more precise by solving a set of recursive relations of the averaged terms in Eq. (130); see Appendix F. The prediction (132) is compared with exact solution of the recursive relations in Fig. 17. The power-law decay of  $\bar{r}_t$  suggests that also  $r_t$  in Eq. (125) should decay as a power law, leading to Eq. (124).

On the other hand, a direct numerical evaluation is still compatible with a linear growth in time of  $S_A(\gamma)$ ; see Fig. 18. Interestingly, we see that for certain choices of gates the growth of temporal entanglement entropy is slower than the lower bound for dual-unitary circuits (see Sec. VB). This indicates that dual-unitary circuits do not produce an extremal temporal entanglement growth.

## VII. TEMPORAL VERSUS SPATIAL ENTANGLEMENT

Having argued that temporal entanglement grows linearly after a quench in generic quantum circuits, the natural question is whether its growth is faster or slower than that of “spatial entanglement,” i.e., regular state entanglement.

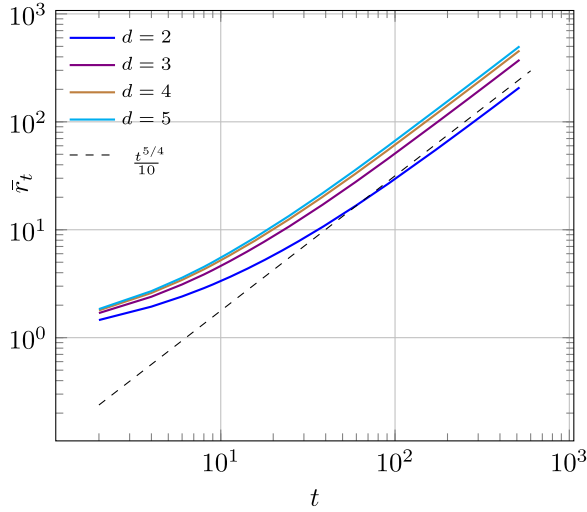


FIG. 17. Polynomial decay of  $\bar{r}_t$ . According to Eq. (F29) we show the asymptotic expected behavior  $\propto t^{-5/4}$  as a dotted line. In the calculation we considered sites of local dimension  $d$  and an initial product state.

This question can be addressed precisely in the case of dual-unitary circuits. Indeed, for these circuits we have that state entanglement grows at the maximal possible speed for generic initial states [77], i.e.,

$$S_{\text{sp}}(t) \simeq 2t \log d. \quad (133)$$

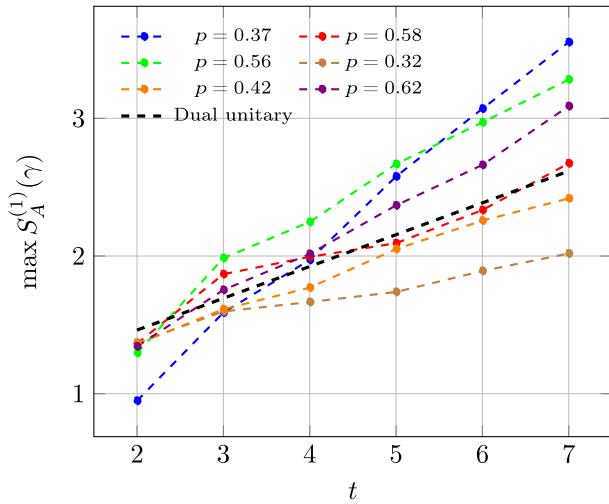


FIG. 18. Growth of the entanglement entropy for the vertical cut state  $\langle L_\gamma \rangle$ , given random choices of the dual-unitary gate (kept constant in space and time) with entangling power  $p$ . We considered the contiguous bipartition  $A\bar{A}$  of  $\gamma$  yielding the maximum entanglement. The entangling power is computed according to Eq. (5) in Ref. [130], which has been normalized by a factor  $(d+1)/(d-1)$  in order to have  $p \in [0, 1]$ . The black line represents the growth of  $S_A^{(1)}(\gamma)$  in dual-unitary circuits [Eq. (105)].

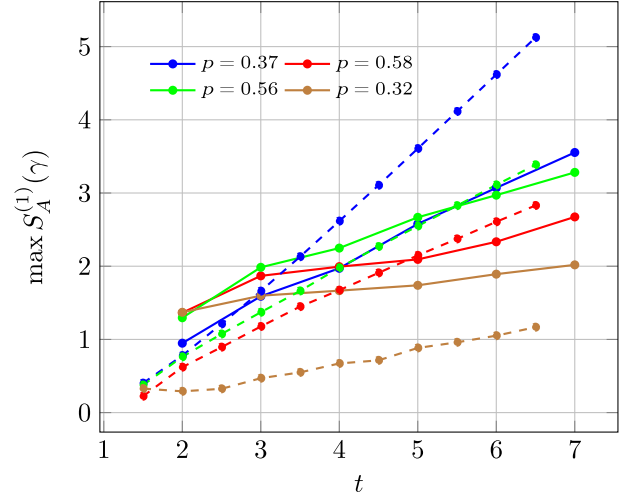


FIG. 19. Comparison between the growth of spatial and temporal entanglement along the vertical path for generic unitary gates. The continuous lines report temporal entanglement of unitary gates with different entangling power (normalized such that  $p \in [0, 1]$ ) while the dashed lines report the corresponding state entanglement. The plot suggests a faster asymptotic growth for the spatial entanglement.

On the other hand, we can use our asymptotic result of Sec. VB to see that

$$S_A(\gamma) \lesssim \max_r S(r)t = \frac{(1+v_\gamma)t}{(2+v_\gamma)} \log d \leq \frac{2t}{3} \log d, \quad (134)$$

where in the first step we computed the maximum of Eq. (105) and in the second we used that it is monotonic in  $v_\gamma$ .

Comparing Eqs. (133) and (134) we see that the temporal entanglement is lower than the spatial entanglement for every path  $\gamma$ . Our numerical investigations suggest that, for small enough  $v_\gamma$ , temporal entanglement grows slower than spatial entanglement also in generic quantum circuits. For instance, in Fig. 19 we report a comparison between the entanglement of the vertical state ( $v_\gamma = 0$ ) and that of the regular time-evolving state for different times: We see that the former has a consistently smaller growth rate for all the gates considered. When the slope of the path is increased, however, the growth of temporal entanglement appears to match that of state entanglement. See, for instance, the comparison between spatial entanglement and temporal entanglement of the diagonal path ( $v_\gamma = 1$ ) reported in Fig. 20.

## VIII. DISCUSSION

In this work we studied spacelike propagation approaches to quantum nonequilibrium dynamics. The main idea is to compute the time evolution of relevant observables by exchanging the roles of space and time. For

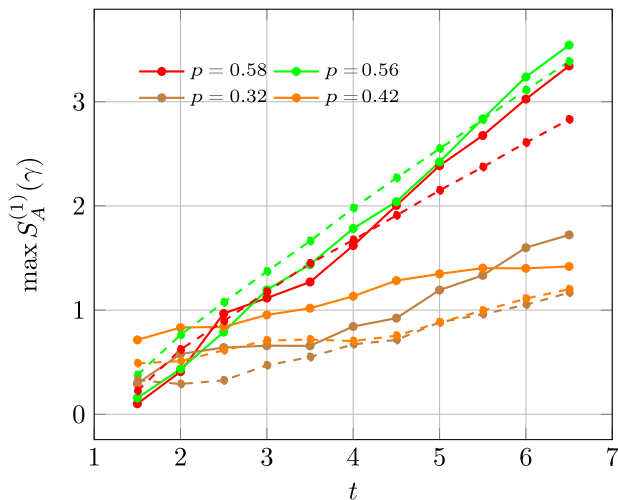


FIG. 20. Comparison between the growth of spatial and temporal entanglement along the diagonal path for generic unitary gates. The continuous lines report temporal entanglement of unitary gates with different entangling power (normalized such that  $p \in [0, 1]$ ) while the dashed lines report the corresponding state entanglement.

large enough volumes, the “dual system” living in the time direction—also known as space-time swapped system [43,45]—reaches (left and right) stationary states dubbed “influence matrices” [33]. Since in many-body systems characterizing a stationary state is often easier and more efficient than characterizing a time-evolving state at intermediate times, spacelike propagation approaches are particularly promising and are attracting an increasing amount of attention [26–30,34,40–45,57–62,87,131,132].

Here we studied whether these ideas can be used to devise an efficient computational scheme to access correlation functions of local operators in generic systems. Our starting point has been the “folding algorithm” of Ref. [26], which uses the above idea to compute autocorrelation functions of local operators in one-dimensional quantum systems. The algorithm represents the time-evolving expectation value as a two-dimensional tensor network and proceeds by embedding the local operator in a system on the vertical time lattice which is then evolved in the space direction. To also access two-point functions between causally connected operators, we generalized the folding algorithm by considering propagation in a generic spacelike direction, i.e., in any direction in the two-dimensional space-time forming an angle  $\alpha$  smaller than  $\pi/4$  with the space direction. The idea is to consider the system on the lattice along a timelike slice, or path, connecting the two points and evolve it in the orthogonal spacelike direction.

We then investigated the efficiency of the generalized folding algorithm by computing the scaling in time of the temporal entanglement, i.e., the entanglement of the influence matrices [28]. Performing a comprehensive investigation in chaotic quantum circuits, we showed that for

generic spacelike evolutions (or states on timelike slice) the entanglement of the influence matrices grows linearly in time, preventing an efficient classical storing. However, we also showed that the volume-law scaling of temporal entanglement is much more subtle than one might expect due to the nontrivial structure of the temporal entanglement spectrum. Indeed, we found physically relevant cases where it separates into a few large Schmidt values (decaying at most polynomially in time) and many small ones (decaying exponentially). This means that the growth of temporal entanglement cannot be characterized via a replica trick.

More specifically, we identified two cases where all temporal Rényi entropies with index larger than one grow sublinearly in time: (i) standard space evolution (the one of the original folding algorithm of Ref. [26], where the timelike surface is vertical) in *generic quantum circuits* and (ii) any spacelike evolution in dual-unitary circuits. This phenomenon is very similar in nature to the sub-ballistic scaling of spatial Rényi entropies observed in circuits with diffusive conservation laws [113,114]. As in the latter case, the time-evolving state has large overlap with a product state over a spatial bipartition of the system; in our case, the influence matrices have large overlap with a product state (which we identified) over a temporal bipartition of the system. This means that the reduced density matrix has a small number of slowly decaying eigenvalues controlling the scaling of higher Rényi entropies.

On the other hand, we showed that the von Neumann temporal entanglement entropy grows *linearly* in time in both the cases (i) and (ii), but it has a strictly smaller rate of growth compared to regular state entanglement. Specifically, while for generic circuits we argued for a linear growth based on the absence of physical constraints and we characterized it numerically, for dual-unitary circuits we were able to provide a closed form expression for the slope of growth. This expression is always nonzero and smaller than the slope of growth of state entanglement. We stress that a strictly positive growth rate of temporal entanglement in dual-unitary circuits is particularly surprising because—due to their maximally fast dephasing [58,83]—these systems are expected to be the chaotic system generating the lowest temporal entanglement [33]. We also emphasize that, to the best of our knowledge, this is the first analytical account of the noncommutativity of replica and large-time limit generating different scalings of Rényi entropies.

Combined with the results of Ref. [62], our findings suggest that the behavior of temporal entanglement after a quantum quench is a dynamical chaos indicator; i.e., it discriminates between integrable and chaotic dynamics. Indeed, while Ref. [62] provided evidence for a generic sublinear scaling of temporal entanglement in integrable models, here we showed that it grows linearly in chaotic systems (modulo some genericity assumption on the initial state). This scenario is in agreement with the



characterization put forward in Ref. [133], which proposed volume-law spatiotemporal entanglement as the defining feature of quantum chaotic systems. From this point of view, temporal entanglement seems to behave similarly to the operator space entanglement of local operators [16,134–136]—another conjectured dynamical chaos indicator [71,72,137,138].

Our work opens several directions for future research. An obvious one is to understand whether it is possible to exploit our findings on the structure of the temporal entanglement spectrum to devise efficient computational schemes. In particular, the fact that influence matrices have a large product-state component might be used to extract information on the large-time dynamics of certain special observables.

Another compelling question is to confirm our numerical observation that the von Neumann entropy of the standard influence matrix grows linearly in time for generic circuits, but its growth is slower than that of regular state entanglement. Because of the noncommutativity of large-time and replica limits this cannot be achieved by a direct application of the entanglement membrane approach. Indeed, in this case the membrane approach can only describe higher Rényi entropies and not von Neumann: one cannot perform the analytic continuation. One possible strategy is to use the approach developed here for dual-unitary circuits: decompose the reduced density matrix as a convex combination and use data processing inequality and convexity of the von Neumann entropy to bound it.

Finally, a further avenue for future research is to assess the performance of our generalized folding algorithm in nonergodic systems, like nearly integrable ones, where the temporal entanglement grows slowly. This could provide a very efficient way to extract numerically linear transport coefficients and, more generally, characterize nonlinear transport in such systems. For instance, it could be applied to the characterization of anomalous transport in integrable systems with non-Abelian charges [139].

### ACKNOWLEDGMENTS

We thank Lorenzo Piroli, Pavel Kos, and Alessio Lerose for helpful discussions and valuable comments on the manuscript. T. Z. acknowledges discussion with and comments from Adam Nahum and Dmitry Abanin on the initial preprint. This work has been supported by the Royal Society through the University Research Fellowship No. 201101 (A. F. and B. B.) and by the National Science Foundation under Grant No. NSF PHY-1748958 (B. B. and T. Z.). T. Z. is supported as a postdoctoral researcher from NTT Research Award No. AGMT DTD 9.24.20 and the Massachusetts Institute of Technology. We acknowledge the accommodation of the KITP program “Quantum Many-Body Dynamics and Noisy Intermediate-Scale Quantum Systems” in which part of the work took place.

### APPENDIX A: MINIMIZATION OF THE FREE-ENERGY TERM FOR HAAR RANDOM CIRCUIT

Here we explicitly carry out the minimization of  $F_{1,Y}(t, t_0)$  with respect to  $t_0$  using the random circuit line tension in Eq. (71). We recall that, setting  $v_1 = -v_2$  in Eq. (77), we have

$$F_{1,Y}(t, r_0) \simeq 4s_{\text{eq}}t \left[ \mathcal{E}_H(v_1(r_0)) \left[ 1 - r_0 - \frac{r}{2} \right] + \mathcal{E}_H(0) \frac{r_0}{2} \right], \quad (\text{A1})$$

where

$$v_1(r_0) = \frac{rv_\gamma}{2 - r - 2r_0} \in \left[ \frac{r}{2 - r - 2r_0} v_\gamma, v_\gamma \right], \quad (\text{A2})$$

and we use  $r_0 = t_0/t \in [0, r]$ . The expression of the free energy in this case is

$$F_{1,Y}(t, r_0) - F_2(t) \simeq 2s_{\text{eq}}t \left[ \{ \mathcal{E}_H(v_1(r_0)) - \mathcal{E}_H(0) \} (1 - r_0) + \mathcal{E}_H(v_1(r_0)) (1 - r - r_0) \right]. \quad (\text{A3})$$

We now solve this final minimization using the explicit random circuit line tension  $\mathcal{E}_H$ . We set the  $r_0$  derivative of the above expression to zero,

$$\mathcal{E}_H(0) - 2\mathcal{E}_H(v_1) + 2v_1\mathcal{E}'_H(v_1) = 0, \quad (\text{A4})$$

and find that the equation is solved for  $v_1 = v_d$ , where

$$v_d := \frac{d-1}{\sqrt{d^2+1}}. \quad (\text{A5})$$

The derivative  $\partial_{r_0} F_{1,Y}(t, r_0)$  is negative for  $v_1 \in [0, v_d]$  and positive for  $v_1 \in (v_d, 1]$ . So  $v_1 = v_d$  is the minimal for  $v_1 \in [0, 1]$ . However, since  $v_1 \in [(r/2 - r)v_\gamma, v_\gamma]$ , depending on the choice of  $r$ , the free energy falls into three cases:

$$\begin{aligned} & \min_{r_0 \in [0, r]} F_{1,Y}(t, r_0) - F_2(t) \\ & \simeq 2s_{\text{eq}}t \begin{cases} r[\mathcal{E}_H(v_\gamma) - \mathcal{E}_H(0)] & v_\gamma \in [0, v_d] \\ r[\mathcal{E}_H(v_d) \frac{v_\gamma}{v_d} - \mathcal{E}_H(0) \frac{v_d+v_\gamma}{2v_d}] & v_\gamma \in [v_d, \frac{2-r}{r} v_d] \\ (2-r)\mathcal{E}_H(\frac{r}{2-r} v_\gamma) - \mathcal{E}_H(0) & v_\gamma \in (\frac{2-r}{r} v_d, 1]. \end{cases} \end{aligned} \quad (\text{A6})$$

To obtain  $\bar{S}_A^{(2)}(\gamma)$ , we further compare this minimum with the free energy of decoupled configurations in Fig. 7, i.e.,  $2s_{\text{eq}}t(1-r)\mathcal{E}(0)$  from Eq. (76).

We first note that

$$(2-r)\mathcal{E}_H\left(\frac{r}{2-r}v_\gamma\right) - \mathcal{E}_H(0) \geq (1-r)\mathcal{E}_H(0), \quad (\text{A7})$$

due to  $\mathcal{E}(v) \geq \mathcal{E}(0)$ . So for the case of  $v_\gamma \in ((2-r/r)v_d, 1]$  the Y-shaped configuration can never dominate. Physically, the minimal for the Y shape here corresponds to taking  $r_0 = 0$ , which represents two tilted sets of domain walls meeting at the very bottom. Its free energy can always be lowered if all the domain walls go down vertically.

Then we compare the expression for  $v_\gamma \in [v_d, (2-r)v_d/r]$  and  $2s_{\text{eq}}t(1-r)\mathcal{E}(0)$ . Setting

$$r\left[\mathcal{E}_H(v_d)\frac{v_\gamma}{v_d} - \mathcal{E}_H(0)\frac{v_d+v_\gamma}{2v_d}\right] \leq (1-r)\mathcal{E}_H(0), \quad (\text{A8})$$

gives  $v_\gamma \leq (2-r)v'_d/r$ , where

$$v'_d := \frac{1}{2} \frac{\log \frac{d^2+1}{2d}}{\text{arctanh}(v_d)}. \quad (\text{A9})$$

We have  $v'_d \leq v_d$  and  $\lim_{d \rightarrow \infty} (v'_d - v_d) = 0$ . For  $v_d < (2-r)v'_d/r$  to hold, we require  $r \leq 2v'_d/(v_d + v'_d)$ .

Finally, the case of  $v_\gamma \in [0, v_d)$ . For the Y-shaped configuration to dominate, we should have

$$r[\mathcal{E}_H(v_\gamma) - \mathcal{E}_H(0)] \leq (1-r)\mathcal{E}_H(0). \quad (\text{A10})$$

This requires  $r \leq \mathcal{E}_H(0)/\mathcal{E}_H(v_\gamma)$ .

In summary, we conclude that Haar random circuits have

$$\bar{S}_A^{(2)}(\gamma) \simeq s_{\text{eq}} v_{\text{TE,H}}^{(2)} t, \quad (\text{A11})$$

where

$$v_{\text{TE,H}}^{(2)} = \begin{cases} 2r[\mathcal{E}_H(v_\gamma) - \mathcal{E}_H(0)] & r \leq \frac{\mathcal{E}_H(0)}{\mathcal{E}_H(v_\gamma)} v_\gamma < v_d \\ 2(1-r)\mathcal{E}_H(0) & r > \frac{\mathcal{E}_H(0)}{\mathcal{E}_H(v_\gamma)} v_\gamma < v_d \\ 2r\left[\mathcal{E}_H(v_d)\frac{v_\gamma}{v_d} - \mathcal{E}_H(0)\frac{v_d+v_\gamma}{2v_d}\right] & r \leq \frac{2v'_d}{v_d+v'_d} v_\gamma \geq v_d \\ 2(1-r)\mathcal{E}_H(0) & r > \frac{2v'_d}{v_d+v'_d} v_\gamma \geq v_d. \end{cases} \quad (\text{A12})$$

## APPENDIX B: RÉNYI ENTROPIES IN DUAL-UNITARY CIRCUITS

In this appendix we present the detailed calculations leading to the bounds on temporal higher Rényi entropies discussed in Sec. VA.

### 1. Upper bound on temporal Rényi entropies for generic quantum circuits

In this subsection we bound  $S_A^{(\alpha)}(\gamma)$  in terms of the norm of the state  $\langle L_\gamma |$ . We begin by writing the Schmidt decomposition of the state  $\langle L_\gamma |$  between the region A and the rest  $\bar{A}$ . Namely,

$$\frac{\langle L_\gamma |}{\sqrt{\langle L_\gamma | L_\gamma \rangle}} = \sum_{r=1}^{\min(d^{|A|}, d^{|\bar{A}|})} \Lambda_r \langle A_r |_A \otimes \langle B_r |_{\bar{A}}, \quad (\text{B1})$$

where  $\{|A_r\rangle_A\}$  and  $\{|B_r\rangle_{\bar{A}}\}$  are orthogonal states, while the Schmidt values  $\{\Lambda_r\}$  fulfill

$$0 \leq \Lambda_r \leq \dots \leq \Lambda_{r-1}, \quad \sum_{r=1}^{\min(d^{|A|}, d^{|\bar{A}|})} \Lambda_r^2 = 1. \quad (\text{B2})$$

The integer

$$n = \min\{r | \Lambda_r = 0\} \quad (\text{B3})$$

is referred to as the Schmidt rank of the state.

Next, we invoke the Eckart-Young theorem [115] to bound from below the largest Schmidt value. To this end we first recall the statement of the theorem

*Theorem 1 (Eckart-Young).*—The scalar product of an unnormalized state  $|\Phi_n\rangle$  of Schmidt rank  $n$  over the bipartition  $B\bar{B}$  and a normalized state  $|\Phi_k\rangle$  with rank  $k < n$  fulfills the following lower bound,

$$|\langle \Phi_n | \Phi_k \rangle| \leq \sqrt{\sum_{j=1}^k \Lambda_j^2}, \quad (\text{B4})$$

where  $\{\Lambda_r\}$  are the Schmidt values of  $|\Phi_n\rangle$ . The state saturating the bound is unique up to a global phase and reads as

$$|\Phi_k^*\rangle = \frac{1}{\sqrt{\sum_{j=1}^k \Lambda_j^2}} \sum_{r=1}^k \Lambda_r |a_r\rangle_B \otimes |b_r\rangle_{\bar{B}}, \quad (\text{B5})$$

where  $\{|a_r\rangle_B\}$  and  $\{|b_r\rangle_{\bar{B}}\}$  are sets of orthogonal states.

This formulation of the Eckart-Young theorem can be directly proven using the von Neumann trace inequality [140].

Using Theorem 1 we have that the largest Schmidt value  $\Lambda_1$  of any state  $|\Phi_n\rangle$  fulfills

$$\Lambda_1 \geq |\langle \Phi_n | \Phi_1 \rangle|, \quad (\text{B6})$$

for any normalized product state  $|\Phi_1\rangle$ . Specializing the theorem to our case, we consider a bipartition of the  $2t$  sites in  $\tau_A$  on the top and  $\tau_{\bar{A}} = 2t - \tau_A$  on the bottom halves.

In particular, we fix

$$\frac{\tau_A}{2t} \equiv r, \quad (\text{B7})$$

and consider the following product state in this bipartition,

$$\langle \tilde{L} | = \langle \bigcirc |^{\otimes \tau_A} \otimes \langle L_{\gamma_{\bar{A}}} |, \quad (\text{B8})$$

where  $\gamma_{\bar{A}}$  is the second part of the path  $\gamma$ , which comprises  $\tau_{\bar{A}}$  steps. Using only the unitarity of the gates, it is immediate to see that the scalar product of the state with  $\tau_A |\bigcirc\rangle$  states leads to a cancellation of the first  $\tau_A$  diagonal rows. Namely,

$$\begin{aligned} \langle L_{\gamma} | \left( |\bigcirc\rangle^{\otimes \tau_A} \right) &= \text{Diagram 1} \\ &= \text{Diagram 2} \\ &= \langle L_{\gamma_{\bar{A}}} | \end{aligned} \quad (\text{B9})$$

So that we find

$$\Lambda_1 \geq \frac{\langle L_{\gamma} | \tilde{L} \rangle}{\sqrt{\langle L_{\gamma} | L_{\gamma} \rangle \langle \tilde{L} | \tilde{L} \rangle}} = \sqrt{\frac{\langle L_{\gamma_{\bar{A}}} | L_{\gamma_{\bar{A}}} \rangle}{\langle L_{\gamma} | L_{\gamma} \rangle}}. \quad (\text{B10})$$

This gives

$$\max_A S_A^{(\infty)}(\gamma) = -\log \Lambda_1^2 \leq \log \frac{\langle L_{\gamma_{\bar{A}}} | L_{\gamma_{\bar{A}}} \rangle}{\langle L_{\gamma} | L_{\gamma} \rangle}. \quad (\text{B11})$$

Next, we use the known inequality [141],

$$S^{(\alpha)}(\rho) \leq \frac{\alpha}{\alpha-1} S^{(\infty)}(\rho), \quad \alpha > 1, \quad (\text{B12})$$

fulfilled by the function in Eq. (62), to obtain Eq. (98).

## 2. Norm of $\langle L_{\gamma} | L_{\gamma} \rangle$ for dual-unitary circuits

Here we compute  $\langle L_{\gamma} | L_{\gamma} \rangle$  in the special case of dual-unitary circuits. Using the dual-unitarity relations for double gates,

$$\text{Diagram 1} = \text{Diagram 2} = \text{Diagram 3}, \quad (\text{B13})$$

one can easily show that

$$\begin{aligned} \langle L_{\gamma} | L_{\gamma} \rangle &= \text{Diagram 1} \\ &= \text{Diagram 2} \\ &= \langle L_{\gamma_{1c}} | L_{\gamma_{1c}} \rangle := \mathcal{N}_{\tau}, \end{aligned} \quad (\text{B14})$$

where we introduced the diagonal path [cf. Eq. (56)],

$$\gamma_{1c} = \{+, +, \dots, +\}, \quad (\text{B15})$$

with length  $\tau = (1 + v_{\gamma})t$  and denoted by  $\mathcal{N}_{\tau}$  the norm of  $\langle L_{\gamma_{1c}} |$ .

We now observe that the latter quantity is directly related to spatial entanglement. Indeed, computing the purity of the regular density matrix  $\rho_A(t)$  [cf. Eq. (24)] and choosing  $A = [t - x, \infty]$ , we find

$$\mathcal{P}(t) = \frac{1}{d^{2t}} \mathcal{N}_{2t}. \quad (\text{B16})$$

Combining this equation with Eqs. (B14), (B11), and (B12) we recover the bound in Eq. (94).

## APPENDIX C: LINEAR GROWTH OF TEMPORAL ENTANGLEMENT ENTROPY IN DUAL-UNITARY CIRCUITS

In this appendix we present the detailed calculations leading to the bound on temporal entanglement entropy discussed in Sec. V B.

### 1. Reduction

Consider a generic bipartition of a state  $L_{\gamma}$ ,  $\gamma = \gamma_A \circ \gamma_{\bar{A}}$ :

$$\langle L_{\gamma} | = \text{Diagram 1} \quad (\text{C1})$$

Now we observe that, since the entanglement is invariant under local unitary transformations, the entanglement between  $A$  and  $\bar{A}$  is not changed by the transformation,

$$\langle L_\gamma | \mapsto \langle L_\gamma | (U_A^\dagger \otimes U_{\bar{A}}^\dagger), \quad (\text{C2})$$

for any unitary matrices  $U_A$  and  $U_{\bar{A}}$  acting, respectively, only in  $A$  and  $\bar{A}$ . We consider the transformations  $U_A$  and  $U_{\bar{A}}$  removing the largest number of gates; in the example shown in Eq. (C1), this corresponds to

$$[U_{\bar{A}}]_{i,j} = \begin{array}{c} i_1 - j_1 \\ i_2 - j_2 \\ i_3 - j_3 \\ i_4 - j_4 \end{array}, \quad [U_A]_{i,j} = \begin{array}{c} i_1 - j_1 \\ i_2 - j_2 \\ i_3 - j_3 \\ i_4 - j_4 \\ i_5 - j_5 \end{array}, \quad (\text{C3})$$

where  $i_a/j_a$  correspond to the  $a$ th digit of  $i/j$  in base  $d^2$ . The corresponding  $\langle L_{\gamma'} |$  state has the following form,

$$\langle L_{\gamma'} | = \begin{array}{c} \text{Diagram of a triangular lattice with paths } \gamma'_A \text{ (black) and } \gamma'_{\bar{A}} \text{ (red) highlighted.} \end{array}, \quad (\text{C4})$$

where we highlighted the new paths  $\gamma'_A$  and  $\gamma'_{\bar{A}}$  forming the edge of  $\langle L_{\gamma'} |$ . This new state has now effectively  $\tau_A = |A|(1 + v_A)/2$  sites in the bipartition  $|A|$ , since the remaining product bullet states are disentangled with the rest.

## 2. Evaluation of $p_k$

Let us evaluate  $\langle L_{\gamma'} | P_k | L_{\gamma'} \rangle$  in order to compute Eq. (114). We are considering states  $\langle L_{\gamma'} |$  as the one shown in Eq. (C4), corresponding to a path (we ignore the bullet states disentangled from the rest),

$$\gamma' = \underbrace{\{+, +, \dots, +\}}_{\tau_A} \circ \underbrace{\{-, -, \dots, -\}}_{|\bar{A}|(1-v_{\bar{A}})/2} \circ \underbrace{\{+, +, \dots, +\}}_{|\bar{A}|(1+v_{\bar{A}})/2}, \quad (\text{C5})$$

of total length  $\tau_A + |\bar{A}|$ .

Graphically, it is easy to see that, using the unitarity of the gates, any scalar product of the type  $\langle L_{\gamma'} | (|\circ\rangle^{\otimes x})$ , where the  $x$  bullet states are applied from the top, deletes the first  $x$  main diagonal of the state  $\langle L_{\gamma'} |$ :

$$\langle L_{\gamma'} | (|\circ\rangle^{\otimes x}) = \begin{array}{c} \text{Diagram of a triangular lattice with } x \text{ bullet states at the top.} \end{array} = (\text{C6})$$

$$= \langle L_{\gamma \setminus x} |, \quad (\text{C7})$$

where with  $\gamma \setminus x$  we indicate the path  $\gamma$ , where the first  $x$  jumps have been deleted. Using the definition of the projectors  $P_k$  in Eq. (108), we then find

$$\langle L_{\gamma'} | P_k | L_{\gamma'} \rangle = \langle L_{\gamma' \setminus \tau_A - k} | L_{\gamma' \setminus \tau_A - k} \rangle - \langle L_{\gamma' \setminus (\tau_A - k + 1)} | L_{\gamma' \setminus (\tau_A - k + 1)} \rangle, \quad (\text{C8})$$

for  $k > 0$ , and

$$\langle L_{\gamma'} | P_0 | L_{\gamma'} \rangle = \langle L_{\gamma' \setminus \tau_A} | L_{\gamma' \setminus \tau_A} \rangle. \quad (\text{C9})$$

Finally, using Eq. (B14) and the shape of  $\gamma'$  in Eq. (C5), we find

$$\begin{aligned} \langle L_{\gamma'} | P_k | L_{\gamma'} \rangle &= \mathcal{N}_{|\bar{A}|(1+v_{\bar{A}})/2+k} - \mathcal{N}_{|\bar{A}|(1+v_{\bar{A}})/2+k-1}, \\ \langle L_{\gamma'} | P_0 | L_{\gamma'} \rangle &= \mathcal{N}_{|\bar{A}|(1+v_{\bar{A}})/2}. \end{aligned} \quad (\text{C10})$$

Using again Eq. (B14), we also have

$$\langle L_{\gamma'} | L_{\gamma'} \rangle = \mathcal{N}_{(1+v_{\gamma})t}, \quad (\text{C11})$$

where we used the fact that

$$|A| \frac{1+v_A}{2} + |\bar{A}| \frac{1+v_{\bar{A}}}{2} = \frac{1+v_{\gamma}}{2} 2t. \quad (\text{C12})$$

Finally, using Assumption 1, we find the asymptotic scaling,

$$\langle L_{\gamma'} | P_k | L_{\gamma'} \rangle \sim C > 0, \quad (\text{C13})$$

$$\langle L_{\gamma'} | P_0 | L_{\gamma'} \rangle \sim C |\bar{A}| \frac{1+v_{\bar{A}}}{2} > 0, \quad (\text{C14})$$

$$\langle L_{\gamma'} | L_{\gamma'} \rangle \sim C(1+v)t, \quad (\text{C15})$$

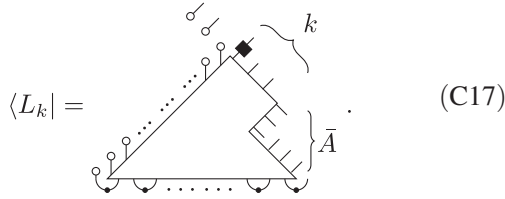
which, putting back in the definition of  $p_k$  in Eq. (114), gives

$$p_k = \begin{cases} \frac{1}{(1+v_\gamma)t} & k \neq 0 \\ \frac{|\bar{A}|(1+v_{\gamma\bar{A}})}{2(1+v_\gamma)t} & k = 0 \end{cases}; \quad (\text{C16})$$

in the main text we considered a coarse grained constant slope for the path  $\gamma$ , which corresponds to the substitution, in Eq. (C16),  $v_{\gamma_A} = v_{\gamma_{\bar{A}}} = v_\gamma$ , that leads to Eq. (118).

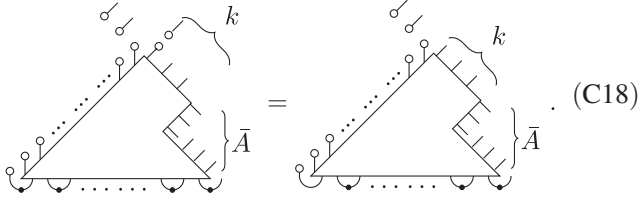
### 3. Entropy of the state $|L_k\rangle$ via membrane approach

In this appendix we use the entanglement membrane approach to compute the second Rényi entropy of the state in Eq. (119) which we call  $|L_k\rangle$  and repeat the expression here:



$$\langle L_k | = \quad (\text{C17})$$

The projector (black box) is  $\mathbb{1}_{d^2} - |\circ\rangle\langle\circ|$ . By using the dual-unitary property, the action of the  $|\circ\rangle\langle\circ|$  is equivalent to replacing the left boundary state via a solvable EPR state:

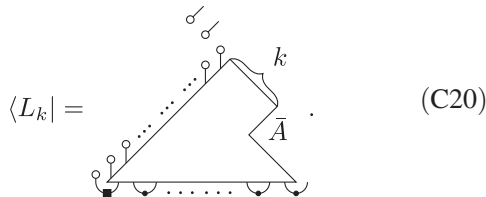


$$= \quad (\text{C18})$$

Therefore this is the part to be projected out. We decompose the initial state on the bottom left of the diagram into a component of the solvable state (EPR state) and a remainder term (box state):

$$\text{EPR state} = \frac{1}{d} \text{EPR state} + \text{box state}. \quad (\text{C19})$$

From Eq. (C18),  $\langle L_k |$  can be simplified to



$$\langle L_k | = \quad (\text{C20})$$

We then closely follow the discussion of Sec. IV B. The evaluation follows exactly as in Eq. (69) when the state  $\langle L_\gamma |$  is replaced by  $\langle L_k |$ ; namely,

$$S^{(2)}(|L_k\rangle) = F_1(x) - F_2(x), \quad (\text{C21})$$

with

$$F_1(x) = -\log(\text{tr}_A[\text{tr}_{\bar{A}}(|L_k\rangle\langle L_k|)^2]), \quad (\text{C22})$$

and

$$F_2(x) = -2 \log \text{tr}(|L_k\rangle\langle L_k|). \quad (\text{C23})$$

Here  $A$  is the subsystem formed by the top  $\tau_A$  sites (of which the top  $\tau_A - k$  sites are decoupled product states though, so it is as if  $A$  contained only  $k$  sites for what concerns entanglement calculations).

The calculation follows the same lines as the one outlined in Sec. IV B with one main difference: since the square state



$$(\text{C24})$$

in the bottom left corner of  $|L_k\rangle$  is orthogonal to the loop state  $|\circ\rangle$ , its fourfold copy is orthogonal to the identity permutation in  $S_4$ . This means that the optimal domain-wall configurations are not those reported in Fig. 7 but, instead, look like those reported in Fig. 22. This gives

$$F_1(x) = F_2(x) + 2 \min(k, |\bar{A}|) \log d, \quad (\text{C25})$$

where we used that for dual-unitary circuits the line tension is equal to one. Plugging in Eq. (C21) and using the monotonicity in  $\alpha$  of the Rényi entropies we arrive at Eq. (123).

In Fig. 21, we checked the validity of Eq. (123), by comparing the maximum value of  $S^{(1)}(\rho_k)$  as a function of  $\ell \equiv k + |\bar{A}|$  (i.e., the number of sites defining the corresponding Hilbert space) for  $v_\gamma = 1$ . From the membrane theory [cf. Eq. (123)] we expect an asymptotic growth of the peak to be equal to

$$\sup_{\substack{k, \\ k+|\bar{A}|=\text{const}}} \min(k, |\bar{A}|) \log(d^2) = \frac{k + |\bar{A}|}{2} \log(d^2) = \ell \log(d). \quad (\text{C26})$$

This prediction agrees with the data for the higher values of the entanglement power  $p$ . For lower values of the entangling power we expect the asymptotic form to arise at larger system sizes. The growth rate of the sup in Eq. (C26) is actually a necessary and sufficient condition

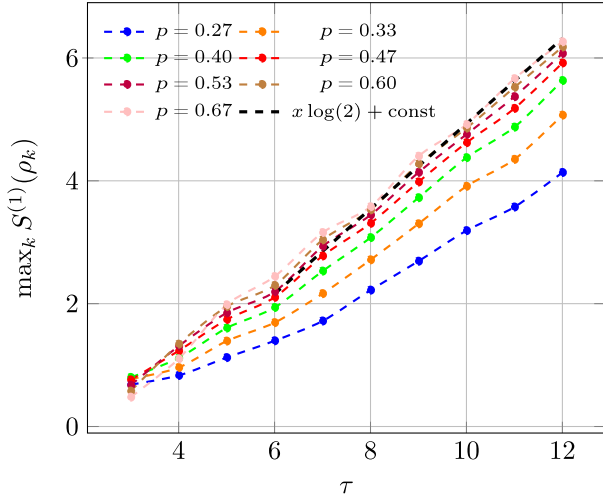


FIG. 21. Entanglement entropy for the matrix  $\rho_k$ , obtained by maximizing the value over all possible choices of  $k$ , keeping  $\tau = |\bar{A}| + k$  fixed, for generic dual-unitary gates of different entangling power, local Hilbert space dimension  $d = 2$ , and  $v_\gamma = 1$ .

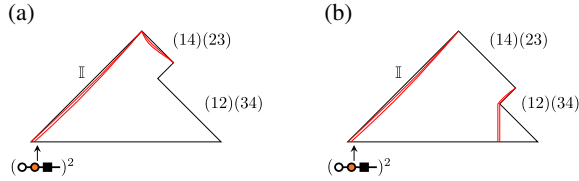


FIG. 22. Domain-wall configuration giving the leading contribution to Eq. (C22). (a) When  $k < |A|$ , the two domain walls (14)(23) split into four domain walls (12)(34)  $\times$  (13)(24). Two of them go to the left to contract with the orthogonal states. The other two (12)(34) go to the right and cancel with the two domain walls at the interactions of  $A$  and  $\bar{A}$ . (b) The domain walls at the tip do not split. They go to the left to contract with the orthogonal states.

for the validity of Eq. (123). This is because, using the positivity of mutual information one has:

$$|S(\rho_k^{(\ell)}) - S(\rho_{k-1}^{(\ell)})| \leq \log(d^2), \quad (\text{C27})$$

where we stressed for clarity that here we consider a density matrix with  $\ell$  sites and a bipartition (on which entanglement is defined) with  $k, \ell - k$  sites. Combining Eq. (C27) with Eq. (C26), one must have

$$S(\rho_k^{(\ell)}) = \min(k, \ell - k) \log(d^2) + o(k) + o(\ell), \quad (\text{C28})$$

which is indeed Eq. (123).

#### APPENDIX D: ASYMPTOTIC BEHAVIOR OF $\mathcal{P}(t)$ UNDER RANDOM DUAL-UNITARY GATES

In this appendix we characterize the asymptotic behavior of the purity in dual-unitary circuits with random local

gates. In particular, following Ref. [77] we consider local gates of the form

$$u_+(\tau, x) \otimes u_-(\tau, x) \cdot U \cdot v_+(\tau, x) \otimes v_-(\tau, x), \quad (\text{D1})$$

where  $U$  is a fixed two-site dual-unitary gate and  $u_\pm, v_\pm$  are random single-site matrices  $\in U(d)$  distributed independently in the space-time. In this setting, Ref. [77] proved that if

$$p(U) \geq \frac{d^2 - 1}{d^2} \left( 1 - \frac{1}{\sqrt{2d + 2}} \right) \quad (\text{D2})$$

[cf. Eq. (12)], then

$$d^x \mathbb{E}_u[\mathcal{P}(x/2)] \leq A + Bx, \quad A, B \geq 0, \quad (\text{D3})$$

where  $\mathbb{E}_u[\cdot]$  is the average over  $\mathbf{u}$  drawn from the full  $U^{\otimes 4L_t}(d)$  group (Haar average).

Here we want to show that for nonsolvable states  $d^x \mathbb{E}_u[\mathcal{P}(x/2)]$  is bounded by a linearly growing function also from below. Namely,

$$d^x \mathbb{E}_u[\mathcal{P}(x/2)] \geq Cx. \quad (\text{D4})$$

Defining the convenient auxiliary quantity,

$$\mathcal{M}_x = d^x \mathcal{P}(x/2) - d^{x-1} \mathcal{P}(x/2 - 1/2), \quad (\text{D5})$$

our goal is to show

$$\lim_{x \rightarrow \infty} \mathbb{E}_u[\mathcal{M}_x] = \mathbb{E}_u[\mathcal{M}_\infty] > 0. \quad (\text{D6})$$

This proves Assumption 1 in the random dual-unitary setting.

First, we note that  $\mathbb{E}_u[\mathcal{M}_x]$  can be related to the function  $\mathcal{Q}_x$ —introduced in Eq. (85) of Ref. [77]—as follows:

$$\mathbb{E}_u[\mathcal{M}_x] = d^x \frac{c - 1}{\sqrt{d^2 - 1}} \mathcal{Q}_x. \quad (\text{D7})$$

Here the parameter  $c$  is defined in terms of the initial state matrix  $m$  [cf. Eq. (15)]:

$$c = \frac{1}{d} \text{tr}[(mm^\dagger)^2] \in [1, d]. \quad (\text{D8})$$

In particular, the value  $c = 1$  corresponds to a unitary initial state matrix  $m$ , i.e., a solvable initial state. In this case, noting that  $\mathcal{Q}_x$  is finite for  $c = 1$ , Eq. (D7) gives  $\mathcal{M}_x = 0$ . This is the expected result for solvable states: the norm of  $\langle L_\gamma \rangle$  is equal to one and its increment  $\mathcal{M}_r$  is zero. From now on we consider  $c > 1$  and argue that in this case  $\mathbb{E}_u[\mathcal{M}_x]$  is always strictly larger than zero.

We begin by noting that a direct application of Eq. (86) of Ref. [77] gives

$$\mathbb{E}_u[\mathcal{M}_x] = \mathbb{E}_u[\mathcal{M}]_{x_0} + \sum_{i=x_0+1}^x \mathcal{S}_i. \quad (\text{D9})$$

Here we introduced

$$\mathcal{S}_x = d^x \frac{c-1}{\sqrt{d^2-1}} \mathcal{R}_x, \quad (\text{D10})$$

where  $\mathcal{R}_x$  is the function defined in Eq. (87) of Ref. [77].

Next, using Eq. (103) of Ref. [77] we conclude

$$|\mathcal{S}_x| \leq Aa^x, \quad (\text{D11})$$

where  $p$  is the entangling power of  $U$  [cf. Eq. (12)] and we introduced

$$\lambda = (1-p)^2 + \frac{p^2}{d^2-1}, \quad (\text{D12})$$

$$a = \frac{d+c}{d+1} d\lambda, \quad (\text{D13})$$

$$A = \frac{(c-1)^2 d+1}{(d+c)^2 d-1} \sqrt{\frac{d^2-1}{d^2 \lambda^3}}. \quad (\text{D14})$$

For high enough values of the entangling power,

$$p > \bar{p}(d) = \frac{d^2-1}{d^2} \left(1 - \frac{1}{\sqrt{2d+2}}\right), \quad (\text{D15})$$

it is immediate to verify that  $a < 1$  for any value of  $c$  in the range  $[1, d]$ , which allows us to find an upper bound for  $\mathbb{E}_u[\mathcal{M}_x]$ . Namely,

$$\begin{aligned} \mathbb{E}_u[\mathcal{M}_x] &\leq \mathbb{E}_u[\mathcal{M}_{x_0}] + \sum_i |\mathcal{S}_i| \\ &\leq \mathbb{E}_u[\mathcal{M}_{x_0}] + \frac{A}{1-a} a^{x_0+1}. \end{aligned} \quad (\text{D16})$$

This bound was first presented in Ref. [77]. Our goal here is to bound  $\mathbb{E}_u[\mathcal{M}_x]$  also from below, showing that it is always strictly larger than 0. To this end we combine Eqs. (D9), (D11), and the triangle inequality to write

$$\begin{aligned} \mathbb{E}_u[\mathcal{M}_x] &\geq \mathbb{E}_u[\mathcal{M}_{x_0}] - \sum_i |\mathcal{S}_i| \\ &\geq \mathbb{E}_u[\mathcal{M}_{x_0}] - A \frac{a^{x_0+1} - a^{x+1}}{1-a}. \end{aligned} \quad (\text{D17})$$

For  $p$  fulfilling the bound (D15) one has  $a < 1$ . This implies that, if we find an  $x_0$  such that

$$\mathbb{E}_u[\mathcal{M}_{x_0}] > A \frac{a^{x_0+1}}{1-a}, \quad (\text{D18})$$

then

$$\lim_{x \rightarrow \infty} \mathbb{E}_u[\mathcal{M}_x] = \mathbb{E}_u[\mathcal{M}_\infty] > 0. \quad (\text{D19})$$

In order to get some intuition it is useful to consider two limiting cases. First, we fix the values of  $x$ ,  $p$ ,  $d$ , and restrict ourselves to a neighborhood of the solvable case, which corresponds to  $c = 1$ . We choose the neighborhood to be small compared to the other parameters, so that we can treat everything perturbatively around the lowest nontrivial order of the solvable case:

$$c \in [1, 1 + \epsilon], \quad \epsilon \ll \frac{1}{x}, \quad 1 - a. \quad (\text{D20})$$

In this situation, it is easy to see that

$$\mathbb{E}_u[\mathcal{M}_x] = O(c-1), \quad A = O[(c-1)^2], \quad (\text{D21})$$

which immediately imply the validity of condition (D18) if  $c \neq 1$ . This shows that some properties of the solvable case are not stable under perturbations.

The other useful limit is  $d \gg 1$ , which makes the expressions (D14) and (D13) much easier to handle. We consider  $\mathbb{E}_u[\mathcal{M}_4]$  at leading order in  $d$ , expanded at the first relevant order for  $1-p$ . We consider gates with entangling power close to the one of the Hadamard gate, or larger, meaning that

$$1-p \lesssim O(d^{-1}). \quad (\text{D22})$$

The asymptotic expression at the lowest relevant order is

$$\begin{aligned} \mathbb{E}_u[\mathcal{M}_4] &= (c-1)[1+c(1-p)] > (c-1) \\ &= O(d^0). \end{aligned} \quad (\text{D23})$$

Expanding Eqs. (D14) and (D13), we find

$$A = \frac{(c-1)^2 d^3}{(d+c)^2 [(1-p)^2 d^2 + 1]^{3/2}} \lesssim O(d^3), \quad (\text{D24})$$

$$a = \frac{d+c}{d} \frac{(1-p)^2 d^2 + 1}{d} = O(d^{-1}). \quad (\text{D25})$$

Putting together everything in condition (D18), we see the left-hand side is  $O(d^0)$  and the right-hand side is  $O(d^{-2})$ , so the condition is respected. Finally, to address the general case, we compute  $\mathbb{E}_u[\mathcal{M}_{x_0}]$  numerically for a high value of  $x_0$ , showing that the bound (D18) holds for a nontrivial interval of entangling powers.

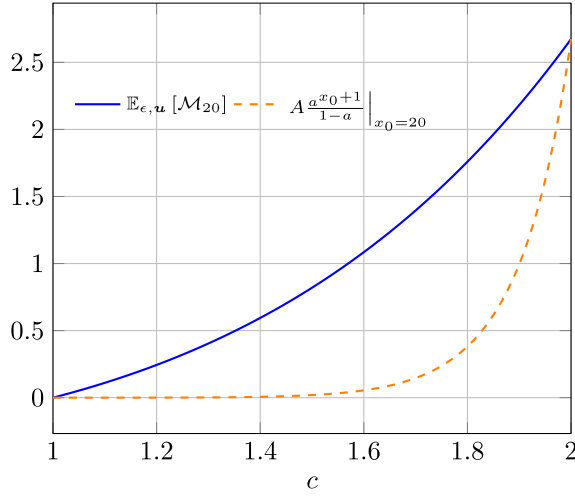


FIG. 23.  $\mathbb{E}_u[\mathcal{M}_{x_0}]$  and  $A(a^{x_0+1}/1-a)$  versus  $c$  for  $x_0 = 20$ ,  $p = \bar{n}$ , and  $d = 2$ . Note the linear and quadratic growth around  $c = 1$  according to Eq. (D21).

In particular, we consider the case  $d = 2$ . In this case, the values of the entangling power  $p$  fulfilling the bound (D15) are given by

$$\frac{2}{3} > p > \bar{p}(2) \approx 0.4438, \quad (\text{D26})$$

where we used that for  $d = 2$  the maximal attainable value of  $p$  is  $2/3$  [77].

Computing  $\mathbb{E}_u[\mathcal{M}_x]$  for  $x = 20$ , we verify the inequality (D18) holds for any  $c$  if we pick  $p$  in the interval

$$\frac{2}{3} > p \geq \bar{n} \approx 0.47660548. \quad (\text{D27})$$

Note that, to verify the equality for  $p$  close to  $\bar{p}(d)$  one would need to consider arbitrarily large values of  $x_0$ . Indeed the denominator  $(1-a)^{-1}$  diverges at  $p = \bar{p}(d)$ . As an example, we show in Fig. 23 the value of  $Aa^{x_0+1}/(1-a)$  versus  $\mathbb{E}_u[\mathcal{M}_{x_0}]$ , for  $x_0 = 20$  as a function of  $c$ .

### APPENDIX E: PARAMETRIZATION OF DUAL-UNITARY GATES FOR NUMERICAL EXPERIMENTS

To produce the data presented in plots involving dual unitary gates (i.e., Figs. 8–14), we parametrized the gates as in Eq. (D1), with fixed one-site unitaries  $u_{\pm}, v_{\pm}$ :

$$u_+ = \begin{pmatrix} 0.204 - 0.971i & -0.108 - 0.068i \\ 0.125 + 0.0254i & -0.524 + 0.842i \end{pmatrix}, \quad (\text{E1})$$

$$u_- = \begin{pmatrix} -0.279 - 0.921i & 0.238 + 0.132i \\ -0.272 + 0.017i & -0.649 + 0.710i \end{pmatrix}, \quad (\text{E2})$$

$$v_+ = \begin{pmatrix} -0.025 - 0.367i & -0.921 - 0.127i \\ 0.908 - 0.202i & 0.005 + 0.368i \end{pmatrix}, \quad (\text{E3})$$

$$v_- = \begin{pmatrix} 0.380 - 0.321i & 0.436 + 0.750i \\ 0.807 + 0.318i & 0.260 - 0.424i \end{pmatrix}, \quad (\text{E4})$$

and two-site dual-unitary given by

$$U(p) = \begin{pmatrix} e^{-iJ(p)} & 0 & 0 & 0 \\ 0 & 0 & -ie^{iJ(p)} & 0 \\ 0 & -ie^{iJ(p)} & 0 & 0 \\ 0 & 0 & 0 & e^{-iJ(p)} \end{pmatrix}, \quad (\text{E5})$$

where

$$J(p) = \arcsin\left(\sqrt{1 - \frac{3p}{2}}\right), \quad (\text{E6})$$

and  $p \in [0, 1]$ . Using the definition of entangling power in Eq. (12) one can immediately verify that

$$p[U(p)] = p. \quad (\text{E7})$$

### APPENDIX F: RECURRENCE RELATION FOR $\langle L_{\gamma} L_{\gamma} \rangle$

In this appendix we consider the Haar averages in Eq. (130). A doubly folded averaged unitary is projected on two-dimensional local vector space, spanned by the normalized vectors

$$|\circ_2\rangle = \frac{|1\rangle}{d}, \quad |\square_2\rangle = \frac{|(12)\rangle}{d}, \quad (\text{F1})$$

where  $|(12)\rangle, |1\rangle$  refer to the permutation vectors represented in Fig. 6(a).

Moreover, a Haar averaged unitary gate fulfills the following relations:

$$\begin{aligned} & \text{Gate with 4 legs} = \text{Purple square gate}, \\ & \text{Purple square gate with 2 white circles} = \frac{d}{d^2+1} \left( \text{White circle} + \text{White square} \right), \\ & \text{Purple square gate with 2 white squares} = \text{White circle} + \text{White square}. \end{aligned} \quad (\text{F2})$$

Here, we choose a specific normalization for the initial state such that



$$\begin{array}{c} \circ \\ | \\ \bullet \end{array} = \begin{array}{c} \square \\ | \\ \bullet \end{array} := 1. \quad (\text{F3})$$

We can define the quantity

$$\mathcal{A}_{x,y} = d^{x+y}, \quad (\text{F4})$$

with  $y \leq x + 1$ . It is immediate to see that

$$\overline{\langle L_{\gamma_{t+1}} | L_{\gamma_{t+1}} \rangle} = \mathcal{A}_{t,t+1}. \quad (\text{F5})$$

Using Eqs. (F2) and (F3), one finds the following recurrence relations:

$$\mathcal{A}_{x,y} = \begin{cases} \frac{d^2}{d^2+1} (\mathcal{A}_{x-1,y-1} + \mathcal{A}_{x,y-1}) & y = x + 1 \\ \frac{d^2}{d^2+1} (\mathcal{A}_{x-1,y} + \mathcal{A}_{x,y-1}) & 0 < y \leq x \\ \left(\frac{2d^2}{d^2+1}\right)^x & y = 0. \end{cases} \quad (\text{F6})$$

The treatment for the average  $\overline{\langle L_{\gamma} | L_{\gamma/2} \rangle \otimes | L_{\gamma/2} \rangle}$  is similar: the average corresponds to the diagram

$$d^{2t-1} = d^t d^{t-1}. \quad (\text{F7})$$

The top diagram can be expressed again with recursive relation, defining the quantity  $\mathcal{B}_{x,y}$ :

$$\mathcal{B}_{x,y} = d^{x+y}, \quad (\text{F8})$$

which fulfills

$$\mathcal{B}_{x,y} = \begin{cases} \frac{d^2}{d^2+1} (\mathcal{B}_{x-1,y-1} + \mathcal{B}_{x,y-1}) & y = x \\ \frac{d^2}{d^2+1} (\mathcal{B}_{x-1,y} + \mathcal{B}_{x,y-1}) & y < x \\ 1 & y = 0. \end{cases} \quad (\text{F9})$$

Equation (130) can then be expressed in terms of these quantities as (assuming a generic bipartition of the temporal state at time  $t$  in  $t = t_1 + t_2$ )

$$\bar{r}_t = \frac{\mathcal{B}_{t_1,t_1} \sqrt{\mathcal{A}_{t_2,t_2+1}}}{\sqrt{\mathcal{A}_{t,t+1} \mathcal{A}_{t_1,t_1+1}}}. \quad (\text{F10})$$

Interestingly, we can map the recurrence relation (F6) in a different problem. First, we slightly change normalization by defining

$$\mathcal{A}_{x,y} = \tilde{\mathcal{A}}_{x,y} \left(\frac{d^2}{d^2+1}\right)^{x+y}, \quad (\text{F11})$$

then the quantity  $\tilde{\mathcal{A}}_{x,y}$  can be thought of as the number of paths connecting the two black dots in the following grid, without crossing the dashed line  $y = x + 2$  and in the minimum number of steps:

$$\quad (\text{F12})$$

Each path gains a weight  $1 + 1/d^2$  every time it touches the top boundary and a factor 2 for every crossed red link of the bottom boundary. To compute the asymptotic scaling of this quantity we can ignore the  $1 + 1/d^2$  weight, which does not change the scaling for  $d$  large enough, since the number of paths touching the top boundary exactly  $p$  times is exponentially suppressed in  $p$  with respect to the total number of paths, which balances the

$$\left(\frac{d^2+1}{d^2}\right)^p \quad (\text{F13})$$

weight (this holds for  $d > 1$ ). This statement can be made more precise using Theorem 2 of Ref. [142]. It is possible to show then that the number of paths touching the boundary  $p$  times is the following (the convention is to

set a binomial coefficient to 0 if the top argument is lower than the bottom one, or the latter is  $< 0$ ):

$$\binom{x+y-p}{x} - \binom{x+y-p}{x+2}, \quad (\text{F14})$$

so that, considering the appropriate weight for these paths, and setting  $x = y - 1 = t$  (we ignore the weights for the red links on the bottom for the sake of this argument), the total is

$$\sum_{p=0}^t \left[ \binom{2t+1-p}{t} - \binom{2t+1-p}{t+2} \right] \left(1 + \frac{1}{d^2}\right)^p. \quad (\text{F15})$$

Studying the asymptotic scaling of this sum (using the Stirling formula and approximating the sum with an integral, expanded around the maximum), we can see the scaling at leading order in  $t$  is unaffected as long as  $d > 1$ . We can call  $a_n$  the number of paths connecting  $(x, y)$  to  $(n, 0)$ , then we can write

$$\tilde{A}_{x,y} \sim \sum_{n=0}^x (a_n - a_{n+1}) 2^n, \quad (\text{F16})$$

with

$$a_n = \begin{cases} \binom{x+y-n}{x-n} - \binom{x+y-n}{x+3} & 0 \leq n \leq x, n \leq y-3 \\ \binom{x+y-n}{y} & 0 \leq n, n \geq y-2 \\ 0 & n < 0, n > x. \end{cases} \quad (\text{F17})$$

In particular, we can rewrite Eq. (F10) as

$$\sum_{n=0}^x (a_n - a_{n+1}) 2^n = a_0 + \sum_{n=1}^x a_n 2^{n-1}, \quad (\text{F18})$$

and find an asymptotic expression for  $a_n$  using Stirling's formula [we take  $(x, y) = (t, t+1)$ ]:

$$a_n 2^{n-1} \sim \frac{1}{2} f\left(\frac{n}{t}\right),$$

$$f(z) = \frac{2^{zt}}{\sqrt{2\pi t}} \frac{(2-z)^{(2-z)t}}{(1-z)^{(1-z)t}} \left[ \frac{6}{t} + 2z - z^2 + O\left(\frac{z}{t}\right) \right]. \quad (\text{F19})$$

Finally, we can estimate the sum in Eq. (F18) with an integral, computed with the saddle point approximation:

$$\sum_{n=1}^t a_n 2^{n-1} = O\left(t \int_0^\infty f(z) dz\right). \quad (\text{F20})$$

We expand  $\log[f(z)]$  around its minimum  $z_0$ , at the leading orders in  $t$ :

$$z_0 = \sqrt{\frac{2}{t}} - \frac{11}{4t} + o\left(\frac{1}{t}\right),$$

$$\frac{d \log(f)}{dz} \Big|_{z=z_0} = O\left(\frac{1}{\sqrt{t}}\right),$$

$$\frac{d^2 \log(f)}{dz^2} \Big|_{z=z_0} = -t + o(t),$$

$$\log[f(z_0)] = t \log(4) - \log(t) + \frac{1}{2} \left[ \log\left(\frac{4}{\pi}\right) - 1 \right]. \quad (\text{F21})$$

Using the saddle point approximation, for large  $t$  we find

$$\int_0^\infty f(z) dz \sim \frac{4^t}{t^{-3/2}} \frac{2}{\sqrt{e\pi}} \int_{-\sqrt{2}}^\infty e^{-x^2/2} dx, \quad (\text{F22})$$

which gives

$$\mathcal{A}_{t,t+1} = O\left(\left(\frac{2d^2}{d^2+1}\right)^{2t} t^{-1/2}\right). \quad (\text{F23})$$

We can compute  $\mathcal{B}_{t,t}$  with a similar approach: in this case we need to consider the paths connecting a point  $(x, y)$  to the origin without crossing the line  $x = y + 1$ . As before, we approximate this quantity by ignoring the weights obtained touching the top boundary  $1 + 1/d^2$ . We can then write

$$\mathcal{B}_{x,y} = O\left(\left(\frac{d^2}{d^2+1}\right)^{x+y} \sum_{n=0}^x b_n \left(\frac{d^2+1}{d^2}\right)^n\right), \quad (\text{F24})$$

where  $b_n$  is the number of paths connecting  $(x, y)$  to  $(n, 1)$ :

$$b_n = \begin{cases} \binom{x+y-n-1}{y-1} - \binom{x+y-n-1}{x+1} & 0 \leq n \leq x, n \leq y-2 \\ \binom{x+y-n-1}{y-1} & 0 \leq n \leq x, n \geq y-1 \\ 0 & n < 0, n > x. \end{cases} \quad (\text{F25})$$

As before, we estimate the sum (F24) with an integral, in the case  $(x, y) = (t, t)$ :

$$\mathcal{B}_{t,t} = O\left(t \int_0^\infty g(z) dz\right),$$

$$g\left(\frac{n}{t}\right) \sim b_n \left(\frac{d^2+1}{d^2}\right)^n, \quad (\text{F26})$$

$$g(z) = \left(\frac{d^2}{d^2+1}\right)^{zt} \frac{1}{t\sqrt{2\pi t}} \frac{(2-z)^{(2-z)t}}{(1-z)^{(1-z)t}} [4 + 2z + O(z^2)].$$

The difference with the previous case is that  $g'(z) \neq 0$  in the domain we are interested in, as long as  $d > 1$ . The maximum value attained is  $g(z=0)$ ; thus we can estimate the integral expanding  $\log(g)$  around  $z=0$  at first order,

$$\int_0^\infty g(z)dz \sim \frac{g^2(0)}{g'(0)} = O(4^t t^{-5/2}), \quad (\text{F27})$$

finding the scaling

$$\mathcal{B}_{t,t} = O\left(\left(\frac{2d^2}{d^2+1}\right)^{2t} t^{-3/2}\right). \quad (\text{F28})$$

Plugging Eqs. (F23)–(F28) into Eq. (F10), we find

$$\bar{r}_t = O(t^{-5/4}). \quad (\text{F29})$$

- 
- [1] I. Bloch, J. Dalibard, and W. Zwerger, *Many-Body Physics with Ultracold Gases*, *Rev. Mod. Phys.* **80**, 885 (2008).
- [2] T. Kinoshita, T. Wenger, and D. S. Weiss, *A Quantum Newton's Cradle*, *Nature (London)* **440**, 900 (2006).
- [3] T. Langen, S. Erne, R. Geiger, B. Rauer, T. Schweigler, M. Kuhnert, W. Rohringer, I. E. Mazets, T. Gasenzer, and J. Schmiedmayer, *Experimental Observation of a Generalized Gibbs Ensemble*, *Science* **348**, 207 (2015).
- [4] M. Schemmer, I. Bouchoule, B. Doyon, and J. Dubail, *Generalized Hydrodynamics on an Atom Chip*, *Phys. Rev. Lett.* **122**, 090601 (2019).
- [5] P. N. Jepsen, J. Amato-Grill, I. Dimitrova, W. W. Ho, E. Demler, and W. Ketterle, *Spin Transport in a Tunable Heisenberg Model Realized with Ultracold Atoms*, *Nature (London)* **588**, 403 (2020).
- [6] N. Malvania, Y. Zhang, Y. Le, J. Dubail, M. Rigol, and D. S. Weiss, *Generalized Hydrodynamics in Strongly Interacting 1D Bose Gases*, *Science* **373**, 1129 (2021).
- [7] D. Wei, A. Rubio-Abadal, B. Ye, F. Machado, J. Kemp, K. Srakaew, S. Hollerith, J. Rui, S. Gopalakrishnan, N. Y. Yao, I. Bloch, and J. Zeiher, *Quantum Gas Microscopy of Kardar-Parisi-Zhang Superdiffusion*, *Science* **376**, 716 (2022).
- [8] M. K. Joshi, F. Kranzl, A. Schuckert, I. Lovas, C. Maier, R. Blatt, M. Knap, and C. F. Roos, *Observing Emergent Hydrodynamics in a Long-Range Quantum Magnet*, *Science* **376**, 720 (2022).
- [9] S. R. White, *Density Matrix Formulation for Quantum Renormalization Groups*, *Phys. Rev. Lett.* **69**, 2863 (1992).
- [10] S. R. White, *Density-Matrix Algorithms for Quantum Renormalization Groups*, *Phys. Rev. B* **48**, 10345 (1993).
- [11] U. Schollwöck, *The Density-Matrix Renormalization Group in the Age of Matrix Product States*, *Ann. Phys. (Amsterdam)* **326**, 96 (2011).
- [12] A. J. Daley, C. Kollath, U. Schollwöck, and G. Vidal, *Time-Dependent Density-Matrix Renormalization-Group Using Adaptive Effective Hilbert Spaces*, *J. Stat. Mech.* (2004) P04005.
- [13] S. R. White and A. E. Feiguin, *Real-Time Evolution Using the Density Matrix Renormalization Group*, *Phys. Rev. Lett.* **93**, 076401 (2004).
- [14] G. Vidal, *Efficient Classical Simulation of Slightly Entangled Quantum Computations*, *Phys. Rev. Lett.* **91**, 147902 (2003).
- [15] G. Vidal, *Efficient Simulation of One-Dimensional Quantum Many-Body Systems*, *Phys. Rev. Lett.* **93**, 040502 (2004).
- [16] T. Prosen and M. Žnidarič, *Is the Efficiency of Classical Simulations of Quantum Dynamics Related to Integrability?*, *Phys. Rev. E* **75**, 015202(R) (2007).
- [17] J. Haegeman, J. I. Cirac, T. J. Osborne, I. Pižorn, H. Verschelde, and F. Verstraete, *Time-Dependent Variational Principle for Quantum Lattices*, *Phys. Rev. Lett.* **107**, 070601 (2011).
- [18] J. Haegeman, C. Lubich, I. Oseledets, B. Vandereycken, and F. Verstraete, *Unifying Time Evolution and Optimization with Matrix Product States*, *Phys. Rev. B* **94**, 165116 (2016).
- [19] E. Leviatan, F. Pollmann, J. H. Bardarson, D. A. Huse, and E. Altman, *Quantum Thermalization Dynamics with Matrix-Product States*, [arXiv:1702.08894](https://arxiv.org/abs/1702.08894).
- [20] B. Kloss, Y. B. Lev, and D. Reichman, *Time-Dependent Variational Principle in Matrix-Product State Manifolds: Pitfalls and Potential*, *Phys. Rev. B* **97**, 024307 (2018).
- [21] C. D. White, M. Zaletel, R. S. K. Mong, and G. Refael, *Quantum Dynamics of Thermalizing Systems*, *Phys. Rev. B* **97**, 035127 (2018).
- [22] M. Žnidarič, *Nonequilibrium Steady-State Kubo Formula: Equality of Transport Coefficients*, *Phys. Rev. B* **99**, 035143 (2019).
- [23] C. Krumnow, J. Eisert, and O. Legeza, *Towards Overcoming the Entanglement Barrier When Simulating Long-Time Evolution*, [arXiv:1904.11999](https://arxiv.org/abs/1904.11999).
- [24] T. Rakovszky, C. W. von Keyserlingk, and F. Pollmann, *Dissipation-Assisted Operator Evolution Method for Capturing Hydrodynamic Transport*, *Phys. Rev. B* **105**, 075131 (2022).
- [25] C. von Keyserlingk, F. Pollmann, and T. Rakovszky, *Operator Backflow and the Classical Simulation of Quantum Transport*, *Phys. Rev. B* **105**, 245101 (2022).
- [26] M. C. Bañuls, M. B. Hastings, F. Verstraete, and J. I. Cirac, *Matrix Product States for Dynamical Simulation of Infinite Chains*, *Phys. Rev. Lett.* **102**, 240603 (2009).
- [27] A. Müller-Hermes, J. I. Cirac, and M. C. Bañuls, *Tensor Network Techniques for the Computation of Dynamical Observables in One-Dimensional Quantum Spin Systems*, *New J. Phys.* **14**, 075003 (2012).
- [28] M. B. Hastings and R. Mahajan, *Connecting Entanglement in Time and Space: Improving the Folding Algorithm*, *Phys. Rev. A* **91**, 032306 (2015).
- [29] M. Sonner, A. Leroise, and D. A. Abanin, *Characterizing Many-Body Localization via Exact Disorder-Averaged Quantum Noise*, *Phys. Rev. B* **105**, L020203 (2022).
- [30] M. Frías-Pérez and M. C. Bañuls, *Light Cone Tensor Network and Time Evolution*, *Phys. Rev. B* **106**, 115117 (2022).
- [31] A. Leroise, M. Sonner, and D. A. Abanin, *Overcoming the Entanglement Barrier in Quantum Many-Body Dynamics via Space-Time Duality*, *Phys. Rev. B* **107**, L060305 (2023).
- [32] R. Feynman and F. Vernon, *The Theory of a General Quantum System Interacting with a Linear Dissipative System*, *Ann. Phys. (N.Y.)* **24**, 118 (1963).

- [33] A. Leroose, M. Sonner, and D. A. Abanin, *Influence Matrix Approach to Many-Body Floquet Dynamics*, *Phys. Rev. X* **11**, 021040 (2021).
- [34] B. Bertini, P. Kos, and T. Prosen, *Exact Spectral Form Factor in a Minimal Model of Many-Body Quantum Chaos*, *Phys. Rev. Lett.* **121**, 264101 (2018).
- [35] B. Bertini, P. Kos, and T. Prosen, *Random Matrix Spectral Form Factor of Dual-Unitary Quantum Circuits*, *Commun. Math. Phys.* **387**, 597 (2021).
- [36] A. Flack, B. Bertini, and T. Prosen, *Statistics of the Spectral Form Factor in the Self-Dual Kicked Ising Model*, *Phys. Rev. Res.* **2**, 043403 (2020).
- [37] S. J. Garratt and J. T. Chalker, *Local Pairing of Feynman Histories in Many-Body Floquet Models*, *Phys. Rev. X* **11**, 021051 (2021).
- [38] F. Fritzsche and T. Prosen, *Eigenstate Thermalization in Dual-Unitary Quantum Circuits: Asymptotics of Spectral Functions*, *Phys. Rev. E* **103**, 062133 (2021).
- [39] S. J. Garratt and J. T. Chalker, *Many-Body Delocalization as Symmetry Breaking*, *Phys. Rev. Lett.* **127**, 026802 (2021).
- [40] B. Bertini, P. Kos, and T. Prosen, *Entanglement Spreading in a Minimal Model of Maximal Many-Body Quantum Chaos*, *Phys. Rev. X* **9**, 021033 (2019).
- [41] M. Ippoliti and V. Khemani, *Postselection-Free Entanglement Dynamics via Spacetime Duality*, *Phys. Rev. Lett.* **126**, 060501 (2021).
- [42] M. Ippoliti, T. Rakovszky, and V. Khemani, *Fractal, Logarithmic, and Volume-Law Entangled Nonthermal Steady States via Spacetime Duality*, *Phys. Rev. X* **12**, 011045 (2022).
- [43] B. Bertini, K. Klobas, V. Alba, G. Lagnese, and P. Calabrese, *Growth of Rényi Entropies in Interacting Integrable Models and the Breakdown of the Quasiparticle Picture*, *Phys. Rev. X* **12**, 031016 (2022).
- [44] J. Thoenniss, M. Sonner, A. Leroose, and D. A. Abanin, *Efficient Method for Quantum Impurity Problems Out of Equilibrium*, *Phys. Rev. B* **107**, L20115 (2023).
- [45] B. Bertini, P. Calabrese, M. Collura, K. Klobas, and C. Rylands, *Nonequilibrium Full Counting Statistics and Symmetry-Resolved Entanglement from Space-Time Duality*, arXiv:2212.06188 [Phys. Rev. Lett. (to be published)].
- [46] B. Pozsgay, *The Dynamical Free Energy and the Loschmidt Echo for a Class of Quantum Quenches in the Heisenberg Spin Chain*, *J. Stat. Mech.* (2013) P10028.
- [47] L. Piroli, B. Pozsgay, and E. Vernier, *From the Quantum Transfer Matrix to the Quench Action: The Loschmidt Echo in XXZ Heisenberg Spin Chains*, *J. Stat. Mech.* (2017) 023106.
- [48] L. Piroli, B. Pozsgay, and E. Vernier, *Non-Analytic Behavior of the Loschmidt Echo in XXZ Spin Chains: Exact Results*, *Nucl. Phys.* **B933**, 454 (2018).
- [49] G. Mahan, *Many-Particle Physics* (Plenum Press, New York, 1990).
- [50] A. Altland and B. Simons, *Condensed Matter Field Theory* (Cambridge University Press, Cambridge, England, 2010).
- [51] B. Bertini, F. Heidrich-Meisner, C. Karrasch, T. Prosen, R. Steinigeweg, and M. Žnidarič, *Finite-Temperature Transport in One-Dimensional Quantum Lattice Models*, *Rev. Mod. Phys.* **93**, 025003 (2021).
- [52] Not to be confused with the “timelike entanglement” studied in the context of conformal field theory and AdS/CFT correspondence; see Refs. [53,54].
- [53] K. Doi, J. Harper, A. Mollabashi, T. Takayanagi, and Y. Taki, *Pseudoentropy in dS/CFT and Timelike Entanglement Entropy*, *Phys. Rev. Lett.* **130**, 031601 (2023).
- [54] K. Doi, J. Harper, A. Mollabashi, T. Takayanagi, and Y. Taki, *Timelike Entanglement Entropy*, *J. High Energy Phys.* **05** (2023) 052.
- [55] F. Verstraete and J.I. Cirac, *Matrix Product States Represent Ground States Faithfully*, *Phys. Rev. B* **73**, 094423 (2006).
- [56] N. Schuch, M.M. Wolf, F. Verstraete, and J.I. Cirac, *Entropy Scaling and Simulability by Matrix Product States*, *Phys. Rev. Lett.* **100**, 030504 (2008).
- [57] B. Bertini, P. Kos, and T. Prosen, *Exact Correlation Functions for Dual-Unitary Lattice Models in 1 + 1 Dimensions*, *Phys. Rev. Lett.* **123**, 210601 (2019).
- [58] L. Piroli, B. Bertini, J.I. Cirac, and T. Prosen, *Exact Dynamics in Dual-Unitary Quantum Circuits*, *Phys. Rev. B* **101**, 094304 (2020).
- [59] K. Klobas, B. Bertini, and L. Piroli, *Exact Thermalization Dynamics in the, “Rule 54” Quantum Cellular Automaton*, *Phys. Rev. Lett.* **126**, 160602 (2021).
- [60] K. Klobas and B. Bertini, *Exact Relaxation to Gibbs and Non-Equilibrium Steady States in the Quantum Cellular Automaton Rule 54*, *SciPost Phys.* **11**, 106 (2021).
- [61] A. Leroose, M. Sonner, and D.A. Abanin, *Scaling of Temporal Entanglement in Proximity to Integrability*, *Phys. Rev. B* **104**, 035137 (2021).
- [62] G. Giudice, G. Giudici, M. Sonner, J. Thoenniss, A. Leroose, D. A. Abanin, and L. Piroli, *Temporal Entanglement, Quasiparticles, and the Role of Interactions*, *Phys. Rev. Lett.* **128**, 220401 (2022).
- [63] S. Gopalakrishnan and A. Lamacraft, *Unitary Circuits of Finite Depth and Infinite Width from Quantum Channels*, *Phys. Rev. B* **100**, 064309 (2019).
- [64] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, *Quantum Entanglement Growth under Random Unitary Dynamics*, *Phys. Rev. X* **7**, 031016 (2017).
- [65] T. Zhou and A. Nahum, *Emergent Statistical Mechanics of Entanglement in Random Unitary Circuits*, *Phys. Rev. B* **99**, 174205 (2019).
- [66] M. P. A. Fisher, V. Khemani, A. Nahum, and S. Vijay, *Random Quantum Circuits*, *Annu. Rev. Condens. Matter Phys.* **14**, 335 (2023).
- [67] C. W. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. L. Sondhi, *Operator Hydrodynamics, OTOCs, and Entanglement Growth in Systems without Conservation Laws*, *Phys. Rev. X* **8**, 021013 (2018).
- [68] T. Rakovszky, F. Pollmann, and C. W. von Keyserlingk, *Diffusive Hydrodynamics of Out-of-Time-Ordered Correlators with Charge Conservation*, *Phys. Rev. X* **8**, 031058 (2018).
- [69] T. Zhou and A. Nahum, *Entanglement Membrane in Chaotic Many-Body Systems*, *Phys. Rev. X* **10**, 031066 (2020).
- [70] V. Khemani, A. Vishwanath, and D. A. Huse, *Operator Spreading and the Emergence of Dissipative Hydrodynamics under Unitary Evolution with Conservation Laws*, *Phys. Rev. X* **8**, 031057 (2018).

- [71] B. Bertini, P. Kos, and T. Prosen, *Operator Entanglement in Local Quantum Circuits I: Chaotic Dual-Unitary Circuits*, *SciPost Phys.* **8**, 67 (2020).
- [72] B. Bertini, P. Kos, and T. Prosen, *Operator Entanglement in Local Quantum Circuits II: Solitons in Chains of Qubits*, *SciPost Phys.* **8**, 68 (2020).
- [73] P. W. Claeys and A. Lamacraft, *Maximum Velocity Quantum Circuits*, *Phys. Rev. Res.* **2**, 033032 (2020).
- [74] B. Bertini and L. Piroli, *Scrambling in Random Unitary Circuits: Exact Results*, *Phys. Rev. B* **102**, 064305 (2020).
- [75] C. Jonay, V. Khemani, and M. Ippoliti, *Triunitary Quantum Circuits*, *Phys. Rev. Res.* **3**, 043046 (2021).
- [76] T. Zhou and A. W. Harrow, *Maximal Entanglement Velocity Implies Dual Unitarity*, *Phys. Rev. B* **106**, L201104 (2022).
- [77] A. Foligno and B. Bertini, *Growth of Entanglement of Generic States under Dual-Unitary Dynamics*, *Phys. Rev. B* **107**, 174311 (2023).
- [78] A. Chan, A. De Luca, and J. T. Chalker, *Solution of a Minimal Model for Many-Body Quantum Chaos*, *Phys. Rev. X* **8**, 041019 (2018).
- [79] I. Reid and B. Bertini, *Entanglement Barriers in Dual-Unitary Circuits*, *Phys. Rev. B* **104**, 014301 (2021).
- [80] H. Wang and T. Zhou, *Barrier from Chaos: Operator Entanglement Dynamics of the Reduced Density Matrix*, *J. High Energy Phys.* **12** (2019) 020.
- [81] A. J. Friedman, A. Chan, A. De Luca, and J. T. Chalker, *Spectral Statistics and Many-Body Quantum Chaos with Conserved Charge*, *Phys. Rev. Lett.* **123**, 210603 (2019).
- [82] A. Chan, A. De Luca, and J. T. Chalker, *Spectral Statistics in Spatially Extended Chaotic Quantum Many-Body Systems*, *Phys. Rev. Lett.* **121**, 060601 (2018).
- [83] P. Kos, T. Prosen, and B. Bertini, *Thermalization Dynamics and Spectral Statistics of Extended Systems with Thermalizing Boundaries*, *Phys. Rev. B* **104**, 214303 (2021).
- [84] B. Bertini, P. Kos, and T. Prosen, *Exact Spectral Statistics in Strongly Localized Circuits*, *Phys. Rev. B* **105**, 165142 (2022).
- [85] P. W. Claeys and A. Lamacraft, *Ergodic and Nonergodic Dual-Unitary Quantum Circuits with Arbitrary Local Hilbert Space Dimension*, *Phys. Rev. Lett.* **126**, 100603 (2021).
- [86] R. Suzuki, K. Mitarai, and K. Fujii, *Computational Power of One- and Two-Dimensional Dual-Unitary Quantum Circuits*, *Quantum* **6**, 631 (2022).
- [87] K. Klobas and B. Bertini, *Entanglement Dynamics in Rule 54: Exact Results and Quasiparticle Picture*, *SciPost Phys.* **11**, 107 (2021).
- [88] F. Arute, K. Arya, R. Babbush, D. Bacon, J. C. Bardin, R. Barends, R. Biswas, S. Boixo, F. G. S. L. Brandao, D. A. Buell *et al.*, *Quantum Supremacy Using a Programmable Superconducting Processor*, *Nature (London)* **574**, 505 (2019).
- [89] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, *Characterizing Quantum Supremacy in Near-Term Devices*, *Nat. Phys.* **14**, 595 (2018).
- [90] A. Deshpande, B. Fefferman, M. C. Tran, M. Foss-Feig, and A. V. Gorshkov, *Dynamical Phase Transitions in Sampling Complexity*, *Phys. Rev. Lett.* **121**, 030501 (2018).
- [91] G. Muraleedharan, A. Miyake, and I. H. Deutsch, *Quantum Computational Supremacy in the Sampling of Bosonic Random Walkers on a One-Dimensional Lattice*, *New J. Phys.* **21**, 055003 (2019).
- [92] Y. Liu, M. Otten, R. Bassirianjahromi, L. Jiang, and B. Fefferman, *Benchmarking Near-Term Quantum Computers via Random Circuit Sampling*, [arXiv:2105.05232](https://arxiv.org/abs/2105.05232).
- [93] E. Magesan, J. M. Gambetta, and J. Emerson, *Characterizing Quantum Gates via Randomized Benchmarking*, *Phys. Rev. A* **85**, 042311 (2012).
- [94] E. Magesan, J. M. Gambetta, and J. Emerson, *Scalable and Robust Randomized Benchmarking of Quantum Processes*, *Phys. Rev. Lett.* **106**, 180504 (2011).
- [95] T. Proctor, K. Rudinger, K. Young, M. Sarovar, and R. Blume-Kohout, *What Randomized Benchmarking Actually Measures*, *Phys. Rev. Lett.* **119**, 130502 (2017).
- [96] T. Brydges, A. Elben, P. Jurcevic, B. Vermersch, C. Maier, B. P. Lanyon, P. Zoller, R. Blatt, and C. F. Roos, *Probing Rényi Entanglement Entropy via Randomized Measurements*, *Science* **364**, 260 (2019).
- [97] A. Elben, J. Yu, G. Zhu, M. Hafezi, F. Pollmann, P. Zoller, and B. Vermersch, *Many-Body Topological Invariants from Randomized Measurements in Synthetic Quantum Matter*, *Sci. Adv.* **6**, eaaz3666 (2020).
- [98] A. Elben, B. Vermersch, M. Dalmonte, J. I. Cirac, and P. Zoller, *Rényi Entropies from Random Quenches in Atomic Hubbard and Spin Models*, *Phys. Rev. Lett.* **120**, 050406 (2018).
- [99] H. Pichler, G. Zhu, A. Seif, P. Zoller, and M. Hafezi, *Measurement Protocol for the Entanglement Spectrum of Cold Atoms*, *Phys. Rev. X* **6**, 041033 (2016).
- [100] B. Vermersch, A. Elben, L. M. Sieberer, N. Y. Yao, and P. Zoller, *Probing Scrambling Using Statistical Correlations between Randomized Measurements*, *Phys. Rev. X* **9**, 021061 (2019).
- [101] B. Vermersch, A. Elben, M. Dalmonte, J. I. Cirac, and P. Zoller, *Unitary  $n$ -Designs via Random Quenches in Atomic Hubbard and Spin Models: Application to the Measurement of Rényi Entropies*, *Phys. Rev. A* **97**, 023604 (2018).
- [102] S. Aaronson, *Shadow Tomography of Quantum States*, [arxiv:1711.01053](https://arxiv.org/abs/1711.01053).
- [103] H.-Y. Huang, R. Kueng, and J. Preskill, *Predicting Many Properties of a Quantum System from Very Few Measurements*, *Nat. Phys.* **16**, 1050 (2020).
- [104] M. Ohliger, V. Nesme, and J. Eisert, *Efficient and Feasible State Tomography of Quantum Many-Body Systems*, *New J. Phys.* **15**, 015024 (2013).
- [105] N. Keenan, N. Robertson, T. Murphy, S. Zhuk, and J. Goold, *Evidence of Kardar-Parisi-Zhang Scaling on a Digital Quantum Simulator*, *Quantum Inf.* **9**, 72 (2023).
- [106] A. Morvan, T. Andersen, X. Mi, C. Neill, A. Petukhov, K. Kechedzhi, D. Abanin, A. Michailidis, R. Acharya, F. Arute *et al.*, *Formation of Robust Bound States of Interacting Microwave Photons*, *Nature (London)* **612**, 240 (2022).
- [107] H. F. Trotter, *On the Product of Semi-Groups of Operators*, *Proc. Am. Math. Soc.* **10**, 545 (1959).

- [108] M. Suzuki, *General Theory of Fractal Path Integrals with Applications to Many-Body Theories and Statistical Physics*, *J. Math. Phys. (N.Y.)* **32**, 400 (1991).
- [109] The dimension of slope in space time is time/distance = 1/velocity.
- [110] Here we assumed reflection symmetry of the gate [hence,  $\tilde{\mathcal{E}}(v) = \tilde{\mathcal{E}}(-v)$ ]. In the absence of this property Eq. (4) takes a different form, but our conclusion about linear growth of temporal entanglement continues to apply.
- [111] 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).
- [112] T. Zhou and A. W. Harrow, *Maximal Entanglement Velocity Implies Dual Unitarity*, *Phys. Rev. B* **106**, L201104 (2022).
- [113] T. Rakovszky, F. Pollmann, and C. W. von Keyserlingk, *Sub-ballistic Growth of Rényi Entropies due to Diffusion*, *Phys. Rev. Lett.* **122**, 250602 (2019).
- [114] Y. Huang, *Dynamics of Rényi Entanglement Entropy in Diffusive Qudit Systems*, *IOP SciNotes* **1**, 035205 (2020).
- [115] C. Eckart and G. Young, *The Approximation of One Matrix by Another of Lower Rank*, *Psychometrika* **1**, 211 (1936).
- [116] S. A. Rather, S. Aravinda, and A. Lakshminarayanan, *Creating Ensembles of Dual Unitary and Maximally Entangling Quantum Evolutions*, *Phys. Rev. Lett.* **125**, 070501 (2020).
- [117] J. I. Cirac, D. Pérez-García, N. Schuch, and F. Verstraete, *Matrix Product States and Projected Entangled Pair States: Concepts, Symmetries, Theorems*, *Rev. Mod. Phys.* **93**, 045003 (2021).
- [118] This is always true for two-point functions on equilibrium states [cf. (33)] (which encode linear response coefficients [51]) and one-point functions. For nonequilibrium two-point functions this is the case away from the regime  $t_2 - t_1 < x_2 - x_1 \leq t_2 + t_1$ .
- [119] A. Nahum, S. Vijay, and J. Haah, *Operator Spreading in Random Unitary Circuits*, *Phys. Rev. X* **8**, 021014 (2018).
- [120] M. Žnidarič, *Exact Convergence Times for Generation of Random Bipartite Entanglement*, *Phys. Rev. A* **78**, 032324 (2008).
- [121] A. W. Harrow and R. A. Low, *Random Quantum Circuits Are Approximate 2-Designs*, *Commun. Math. Phys.* **291**, 257 (2009).
- [122] J. Emerson, Y. S. Weinstein, M. Saraceno, S. Lloyd, and D. G. Cory, *Pseudo-Random Unitary Operators for Quantum Information Processing*, *Science* **302**, 2098 (2003).
- [123] B. Skinner, J. Ruhman, and A. Nahum, *Measurement-Induced Phase Transitions in the Dynamics of Entanglement*, *Phys. Rev. X* **9**, 031009 (2019).
- [124] Y. Li, X. Chen, and M. P. A. Fisher, *Measurement-Driven Entanglement Transition in Hybrid Quantum Circuits*, *Phys. Rev. B* **100**, 134306 (2019).
- [125] A. Chan, R. M. Nandkishore, M. Pretko, and G. Smith, *Unitary-Projective Entanglement Dynamics*, *Phys. Rev. B* **99**, 224307 (2019).
- [126] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, *Quantum Entanglement Growth under Random Unitary Dynamics*, *Phys. Rev. X* **7**, 031016 (2017).
- [127] Here we use  $\mathcal{E}_H$  to denote the line tension for the average purity decay (annealed average). The expression is exact. The line tension for the average second Rényi entanglement (quenched average) has  $\mathcal{O}(d^{-4})$  correction on top of this; see Ref. [65].
- [128] H. Kim and D. A. Huse, *Ballistic Spreading of Entanglement in a Diffusive Nonintegrable System*, *Phys. Rev. Lett.* **111**, 127205 (2013).
- [129] This holds for  $r$  strictly smaller than 1. For  $r = 1 - A/r^\alpha$ , logarithmic violations may occur.
- [130] P. Zanardi, C. Zalka, and L. Faoro, *Entangling Power of Quantum Evolutions*, *Phys. Rev. A* **62**, 030301(R) (2000).
- [131] B. Bertini, K. Klobas, and T.-C. Lu, *Entanglement Negativity and Mutual Information after a Quantum Quench: Exact Link from Space-Time Duality*, *Phys. Rev. Lett.* **129**, 140503 (2022).
- [132] K. Klobas and T. Prosen, *Space-like Dynamics in a Reversible Cellular Automaton*, *SciPost Phys. Core* **2**, 10 (2020).
- [133] N. Dowling and K. Modi, *An Operational Metric for Quantum Chaos and the Corresponding Spatiotemporal Entanglement Structure*, [arXiv:2210.14926](https://arxiv.org/abs/2210.14926).
- [134] T. Prosen and I. Pižorn, *Operator Space Entanglement Entropy in a Transverse Ising Chain*, *Phys. Rev. A* **76**, 032316 (2007).
- [135] I. Pižorn and T. Prosen, *Operator Space Entanglement Entropy in  $xy$  Spin Chains*, *Phys. Rev. B* **79**, 184416 (2009).
- [136] 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**, 234001 (2017).
- [137] V. Alba, J. Dubail, and M. Medenjak, *Operator Entanglement in Interacting Integrable Quantum Systems: The Case of the Rule 54 chain*, *Phys. Rev. Lett.* **122**, 250603 (2019).
- [138] V. Alba, *Diffusion and Operator Entanglement Spreading*, *Phys. Rev. B* **104**, 094410 (2021).
- [139] V. B. Bulchandani, S. Gopalakrishnan, and E. Ilievski, *Superdiffusion in Spin Chains*, *J. Stat. Mech.* (2021) 084001.
- [140] L. Mirsky, *A Trace Inequality of John von Neumann*, *Mon. Hefte Math.* **79**, 303 (1975).
- [141] H. Wilming, M. Goihl, I. Roth, and J. Eisert, *Entanglement-Ergodic Quantum Systems Equilibrate Exponentially Well*, *Phys. Rev. Lett.* **123**, 200604 (2019).
- [142] M. Z. Spivey, *Enumerating Lattice Paths Touching or Crossing the Diagonal at a Given Number of Lattice Points*, *Electron. J. Comb.* **19**, P24 (2012).