

Eigenstate Correlations, the Eigenstate Thermalization Hypothesis, and Quantum Information Dynamics in Chaotic Many-Body Quantum Systems

Dominik Hahn^{1,*}, David J. Luitz^{2,†} and J. T. Chalker^{3,‡}

¹Max Planck Institute for the Physics of Complex Systems, Nöthnitzer Straße 38, 01187-Dresden, Germany

²Physikalisches Institut, Universität Bonn, Nussallee 12, 53115 Bonn, Germany

³Theoretical Physics, University of Oxford, Parks Road, Oxford OX1 3PU, United Kingdom

 (Received 2 October 2023; revised 25 April 2024; accepted 6 June 2024; published 16 August 2024)

We consider the statistical properties of eigenstates of the time-evolution operator in chaotic many-body quantum systems. Our focus is on correlations between eigenstates that are specific to spatially extended systems and that characterize entanglement dynamics and operator spreading. In order to isolate these aspects of dynamics from those arising as a result of local conservation laws, we consider Floquet systems in which there are no conserved densities. The correlations associated with scrambling of quantum information lie outside the standard framework established by the eigenstate thermalization hypothesis (ETH). In particular, ETH provides a statistical description of matrix elements of local operators between pairs of eigenstates, whereas the aspects of dynamics we are concerned with arise from correlations among sets of four or more eigenstates. We establish the simplest correlation function that captures these correlations and discuss features of its behavior that are expected to be universal at long distances and low energies. We also propose a maximum-entropy ansatz for the joint distribution of a small number n of eigenstates. In the case $n = 2$, this ansatz reproduces ETH. For $n = 4$ it captures both the growth with time of entanglement between subsystems, as characterized by the purity of the time-evolution operator, and also operator spreading, as characterized by the behavior of the out-of-time-order correlator. We test these ideas by comparing results from Monte Carlo sampling of our ansatz with exact diagonalization studies of Floquet quantum circuits.

DOI: [10.1103/PhysRevX.14.031029](https://doi.org/10.1103/PhysRevX.14.031029)

Subject Areas: Condensed Matter Physics,
Quantum Physics, Statistical Physics

I. INTRODUCTION

Although textbook approaches to the thermodynamic equilibrium of quantum systems rely on invoking a weak coupling to a heat bath, it was understood over the course of the past four decades that this is not strictly necessary. In a large class of many-body quantum systems, the interactions between its constituents enable an isolated system to act as its own heat bath and to reach a thermal equilibrium state at long times when starting from most nonequilibrium initial states. While the late-time state remains a pure state, it approaches a *typical state*, representative of the Gibbs ensemble with small fluctuations away from this state for large system size [1]. This phenomenon emerges from the

pseudo-random nature of physical observables in the energy eigenbasis, which was suggested as a criterion for quantum chaos [1] and demonstrated numerically early on [2], with diagonal matrix elements clustering around equilibrium expectation values [3]. Integrable systems can evade this behavior but are not robust in the sense that very weak perturbations suffice to recover thermalization [4].

These observations were subsequently formalized and now constitute the eigenstate thermalization hypothesis (ETH) [1,2,4–13], which can be derived from the assumption that the behavior of the quantum system within a narrow energy window is essentially captured by random matrix theory. This generically leads to Gaussian distributions of the matrix elements of local operators in the energy eigenbasis [8–10,14], with possible exceptions in tails of the distribution for systems with slow particle transport [15–18]. In these Gaussian distributions, the mean values of diagonal matrix elements are fixed to reproduce statistical mechanical averages of observables, while off-diagonal matrix elements have zero mean and an energy structure in their variance that governs the dynamics of autocorrelation functions of the local operators [12,19].

It was pointed out recently on general grounds that this picture cannot be complete [20,21]. If matrix elements of

*Contact author: hahn@pks.mpg.de

†Contact author: david.luitz@uni-bonn.de

‡Contact author: john.chalker@physics.ox.ac.uk

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Open access publication funded by the Max Planck Society.

local operators in the eigenbasis were independent Gaussian random variables, then their mean and variance would determine not only autocorrelation functions but also all higher-order correlators. In particular, the implications for the out-of-time-order correlators (OTOC) [22,23] are in stark contrast to the known behavior of chaotic many-body quantum systems. This leads to the conclusion that there are necessarily correlations between matrix elements, which contain information characterized by higher-order cumulants [20,21,24,25].

More specifically, in quantum systems with local interactions, there are strong bounds on how fast correlations can spread [26], and this limits, for instance, the rate of growth of the entanglement entropy [27,28]. This behavior is reflected in the typically linear growth of the operator entanglement entropy of the time evolution operator of local systems [29,30], and is captured by light-cone structures of out-of-time-order correlators [31]. Our focus here is on the resulting correlations between matrix elements of observables, and related correlations between eigenstates of the time-evolution operator.

A separate potential source of correlations between matrix elements is provided by locally conserved densities. Such correlations were considered early on [32], have been investigated via transport timescales [33], and subsequently observed in larger systems [34,35]. To isolate features arising because of the dynamics of quantum information from those due to locally conserved densities, we focus in the following on Floquet systems, for which time dependence of the Hamiltonian eliminates energy conservation, and in which there are no other local conservation laws.

Recent work has presented a generalized formulation of ETH using free probability theory and numerical tests for higher-order correlations between matrix elements [36,37]. That perspective considers matrix elements of local operators as fundamental objects, revealing the frequency structure of the higher-order free cumulants, particularly the fourth-order free cumulant, which encodes the leading correlations of matrix elements beyond standard ETH.

An alternative perspective, which we adopt here and which was considered, for example, in Ref. [38], is to treat eigenstates of the time-evolution operator, rather than matrix elements of observables, as the relevant set of variables. In particular, the typical time evolution of Renyi entropy in local systems implies nontrivial correlations between eigenstates. Separately, in Ref. [39] a derivation is given of ETH via a study of eigenstates in random Floquet quantum circuits using a field-theoretic approach.

In the present paper, we discuss correlations in chaotic many-body quantum systems that are associated with the dynamics of quantum information scrambling and that can be expected to be universal in spatially extended systems with local interactions. Here we use the term *chaotic* to refer to generic systems whose dynamics is unconstrained

(for example) integrability, many-body localization, or Hilbert space fragmentation. While ETH is formulated as a statement about the statistical properties of matrix elements of operators, we find that it is more transparent to consider instead eigenstates and correlators constructed from them, without reference to particular operators. The first main contribution of our work is to identify the leading-order eigenstate correlator that contains information about this scrambling dynamics, and to discuss its behavior at long distances and low energy differences. The second main contribution is to introduce an ansatz for the joint probability distribution of a small number n of eigenstates of the time-evolution operator. In the case of a pair of eigenstates, this reproduces the Gaussian distribution for matrix elements of local observables that constitutes ETH. Extending this ansatz, we show that a constraint on the joint distribution of four eigenstates is sufficient to capture the essential features of the OTOC and of the operator entanglement entropy of the time-evolution operator. As a third main contribution, we test these ideas by comparing results from exact diagonalization (ED) of the time-evolution operator with those from Monte Carlo (MC) sampling of our ansatz for the joint eigenstate distribution.

The approach that we develop here builds on previous work by one of the present authors and others [21], which argued for the existence of eigenstate correlations on the basis of known behavior of the OTOC and demonstrated their presence using numerical studies of Floquet quantum circuits. Our aim in the following is to establish a general framework for the discussion of these phenomena via the three contributions summarized in the preceding paragraph, none of which were anticipated in Ref. [21].

The ideas and results that we set out in this paper are consistent with, but broadly complementary to, the recent discussion of a generalized version of ETH, formulated to describe the average of products of arbitrary numbers of matrix elements of observables using the language of free probability [36,37]. In particular, our emphasis is different from that of Ref. [37] in two ways. First, we focus on correlations at large distances and small energy separations that we expect to be a universal consequence of the dynamics of quantum information. Second, we find that it is advantageous to consider correlations between eigenstates in place of matrix elements.

The remainder of this paper is organized as follows. In Sec. II we provide a compact overview of our main results. In Sec. III we give details of our calculations, including the microscopic models we use for numerical studies, the determination of parameters in our ansatz for the joint probability distribution of eigenstates, and a summary of the numerical methods used. We develop a treatment of our ansatz based on a perturbative expansion in Sec. IV, and we present numerical results for additional models in Sec. V. We conclude with a summary and outlook in Sec. VI.

II. OVERVIEW

In this section we provide an overview of our results. We introduce the class of model studied and the eigenstate correlators of interest in Sec. II A. We set out the relationship between these correlators and the OTOC in Sec. II B, and indicate generalizations in Sec. II C. We review the sense in which the original form of ETH fails to capture these correlators in Sec. II D, and we propose a representation for them in terms of the joint distribution function for sets of four eigenstates in Sec. II E. We present results for the correlators from exact diagonalization and from Monte Carlo sampling of this distribution in Sec. II F.

A. Models and correlation functions

We start by setting out some essential notation. We consider a one-dimensional Floquet system consisting of L sites, each with a local Hilbert space dimension q , coupled by local interactions. We use W to denote the Floquet operator for the system, which is then a unitary $q^L \times q^L$ matrix that generates evolution through one time period. Defining eigenstates $|a\rangle$ and quasienergies θ_a satisfying $W|a\rangle = e^{-i\theta_a}|a\rangle$, the evolution operator $W(t)$ for an integer number of time steps t has the spectral decomposition

$$W(t) = \sum_a e^{-i\theta_a t} |a\rangle\langle a|. \quad (1)$$

The dynamics can be characterized in terms of correlators of local operators X_α, Y_β, \dots . Here we use upper case letters X, Y, \dots as labels for subsystems on which local operators act, with subscripts α, β, \dots to distinguish different operators acting on a given subsystem. Later we will use \bar{X}, \bar{Y}, \dots to denote the complements of these subsystems. The time evolution in the Heisenberg picture is $X_\alpha(t) = W^\dagger(t)X_\alpha W(t)$, and for a Floquet system it is natural to evaluate correlators using the infinite-temperature density matrix. As we describe in more detail in Sec. III A, the models analyzed in this work consist of local gates drawn from the Haar ensemble. In this case, it is useful to consider an ensemble of realizations of W and to average physical quantities over the ensemble. We indicate this average by $[\dots]_{\text{av}}$. An alternative average is over a Haar distribution of vectors, which we indicate by $[\dots]_0$.

The simplest correlator is the autocorrelation function of a single operator, which has the spectral decomposition

$$q^{-L} \text{Tr}[X_\alpha(t)X_\alpha] = q^{-L} \sum_{ab} |\langle a|X_\alpha|b\rangle|^2 e^{i(\theta_a - \theta_b)t}. \quad (2)$$

In a chaotic many-body quantum system this is expected to decay on a timescale that is microscopic in the sense that it is of order a few Floquet periods [40]. Evidently, its behavior reflects statistical properties of pairs $|a\rangle, |b\rangle$ of eigenstates, which are therefore expected to show features as a function of the quasienergy difference $\theta_a - \theta_b$ that vary

on a scale only a few times smaller than the spectral width 2π [40].

The OTOC has the definition and spectral decomposition

$$\begin{aligned} & q^{-L} \text{Tr}[X_\alpha(t)Y_\beta X_\alpha(t)Y_\beta] \\ &= q^{-L} \sum_{abcd} \langle a|X_\alpha|b\rangle \langle b|Y_\beta|c\rangle \\ & \quad \times \langle c|X_\alpha|d\rangle \langle d|Y_\beta|a\rangle e^{i(\theta_a - \theta_b + \theta_c - \theta_d)t}. \end{aligned} \quad (3)$$

If the subsystems X and Y are separated by a large distance, the main features of the OTOC appear on a large timescale. More specifically, the support of the operator $X_\alpha(t)$ is expected [26] to grow with a butterfly velocity v_B ; the OTOC is constant and nonzero if this support is disjoint from that of Y_β , but falls toward zero when the support of $X_\alpha(t)$ expands to contain that of Y_β . Clearly, behavior of the OTOC reflects statistical properties of sets of four eigenstates, $|a\rangle, |b\rangle, |c\rangle$, and $|d\rangle$, which must show features on a quasienergy scale $2\pi v_B/s$, with s being the spatial separation of the operators, that is much smaller for large s than the spectral width.

Besides the OTOC, a second way to characterize quantum information dynamics is via the spread of entanglement. Consider an initial state $|\psi\rangle$ with low entanglement in the site basis. Although the corresponding density matrix $\rho(t) = W(t)|\psi\rangle\langle\psi|W^\dagger(t)$ remains pure at all t , for a subsystem X that is much smaller than its complement \bar{X} , the reduced density matrix $\rho_X(t) = \text{Tr}_{\bar{X}}\rho(t)$ is expected to evolve toward an equilibrium density matrix. This is probed at the simplest level by considering the purity $\text{Tr}_X[\rho_X(t)]^2$. Since the definition of the purity involves two powers of $W(t)$ and two of $W^\dagger(t)$, its behavior, like that of the OTOC, reflects correlations among sets of four eigenstates. These are characterized by the correlation function defined below in Eq. (7).

Both the OTOC and the purity require choices in their definitions—of the operators denoted by X_α and Y_β in Eq. (3) for the former, and of the initial wave function $|\psi\rangle$ for the latter. This arbitrariness can be eliminated by averaging the OTOC over two complete sets of operators $\{X_\alpha\}$ and $\{Y_\beta\}$ with given supports X and Y , and by considering the operator entanglement entropy of $W(t)$ [29,30,41] in place of the purity of $\rho_X(t)$. Both routes lead to an identical correlator which is defined solely in terms of sets of four eigenstates and the choice of X and Y . We defer discussion of details and present first an alternative argument that singles out the same correlator.

As a starting point, consider the Schmidt decomposition of an eigenstate in terms of tensor products of orthonormal basis states $|i_X\rangle$ and $|i_{\bar{X}}\rangle$ for subsystem X and its complement \bar{X} , which we write as

$$|a\rangle = \sum_{i_X i_{\bar{X}}} [C_X(a)]_{i_X i_{\bar{X}}} |i_X\rangle \otimes |i_{\bar{X}}\rangle. \quad (4)$$

Here $C_X(a)$ is the matrix version of the eigenstate $|a\rangle$, separating states on subsystem X into row indices and states on \bar{X} into column indices. It hence has dimensions $q^{L(X)} \times q^{L(\bar{X})}$, where $L(X) = |X|$, the number of sites in X , and similarly for $L(\bar{X})$.

The problem of constructing correlators from sets of eigenstates is equivalent to one of forming scalars from sets of matrices $C_X(a)$. This can be done by taking the trace of products of an even number of terms, in which the matrices alternate with their Hermitian conjugates. The matrices within such a trace must all refer to a given choice of subsystem X but may refer to multiple eigenstates $|a\rangle, |b\rangle, \dots$. At lowest order this recipe simply yields the quantity

$$\text{Tr}[C_X(a)C_X^\dagger(b)] = \delta_{ab}, \quad (5)$$

which has a value fixed by orthonormality of the eigenstates. At next order it gives

$$M_X(abcd) = \text{Tr}[C_X(a)C_X^\dagger(b)C_X(c)C_X^\dagger(d)]. \quad (6)$$

Such quantities can be represented diagrammatically as shown in Fig. 1.

We now invoke two guiding ideas. One follows from the fact that we want to characterize dynamics in space and time, which suggests that we should consider more than one way of subdividing the system into subsystems, with different alternatives labeled X, Y, \dots . The other follows from the fact that the overall phases of individual eigenstates can be chosen arbitrarily, but physical quantities should be invariant under the transformation $|a\rangle \rightarrow e^{i\phi_a}|a\rangle$. To eliminate the phases ϕ_a , each $C_X(a)$ appearing in a correlator must be accompanied by a Hermitian conjugate $C_Y^\dagger(a)$, referring to the same eigenstate but possibly with a different division into subsystems. Employing both these ideas, we are led to one of the two main quantities of interest in this paper, $M_X(abcd)M_Y^*(abcd)$ and its appropriately normalized ensemble average,

$$F_4(X, Y, \theta) = q^{-L(X, Y)} \left[\sum_{abcd} M_X(abcd)M_Y^*(abcd) \times \delta(\theta - \theta_a + \theta_b - \theta_c + \theta_d) \right]_{\text{av}}. \quad (7)$$

Here and in the following, the argument of the δ function on quasienergy differences is taken modulo 2π . The length $L(X, Y)$ appearing in the normalization is defined by

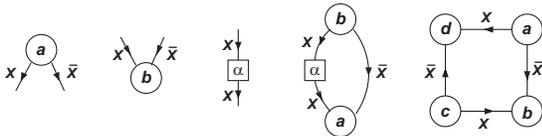


FIG. 1. Ingredients for a diagrammatic representation. From left to right: $C_X(a)$, $C_X(b)^\dagger$, X_a , $\langle a|X_a|b\rangle$, and $M_X(abcd)$.

$$\begin{aligned} L(X, Y) &= 2L - |\bar{X}\backslash\bar{Y}| - |X\backslash Y| \\ &= 2L - |\bar{Y}\backslash\bar{X}| - |Y\backslash X| = L(Y, X), \end{aligned} \quad (8)$$

where notation of the form $|A\backslash B|$ indicates the number of sites that are in subsystem A but not in B . The choice of subscript on $F_4(X, Y, \theta)$ indicates that this quantity characterizes correlations within sets of four eigenstates. The normalization in Eq. (7) is chosen such that the Fourier transform of Eq. (7) with respect to the relative phase θ is unity at $t = 0$, as will be explained below Eq. (11).

Somewhat surprisingly, at this order in powers of the eigenstates, Eq. (7) is the unique outcome of interest from the approach we have sketched for constructing correlators. To see this, consider potential alternatives. Any such alternative should involve two factors, each consisting of a trace over products of Schmidt matrices $C_X(a)$, since each trace carries a subsystem label and we are interested in correlations between a pair of subsystems. Moreover, at this order the two factors together involve four such matrices and four Hermitian conjugates. If these matrices are equally distributed between the two traces, then alternatives to Eq. (7) must all be generated by replacing $M_Y^*(abcd)$ with a similar factor that preserves invariance under changes of eigenstate phases. These can all be obtained from Eq. (7) using the equalities $M_Y(abcd) = M_Y(cdab) = M_{\bar{Y}}(adcb) = M_{\bar{Y}}^*(badc)$ and so are equal to $F_4(X, Y, \theta)$ or $F_4(X, \bar{Y}, \theta)$. Finally, one might regroup matrices under the trace, so that one trace involves six matrices and the other trace involves only two. Then, however, the value of the second trace is fixed via Eq. (5) and is independent of subsystem label, eliminating the spatial dependence of interest.

Among correlators involving only a single X , the lowest-order quantity that is independent of eigenstate phases is $M_X(abba)$ [equal to $M_{\bar{X}}(aabb)$]. From this we define

$$F_2(X, \theta) = q^{-[L+L(X)]} \left[\sum_{ab} M_X(abba)\delta(\theta - \theta_a + \theta_b) \right]_{\text{av}}. \quad (9)$$

The subscript on $F_2(X, \theta)$ indicates that this correlator characterizes correlations between pairs of eigenstates.

The two correlators $F_2(X, \theta)$ and $F_4(X, Y, \theta)$ are the central quantities of interest in the following, together with their counterparts in the time domain, defined by

$$f_2(X, t) = \int_{-\pi}^{\pi} d\theta F_2(X, \theta) e^{i\theta t} \quad (10)$$

and

$$f_4(X, Y, t) = \int_{-\pi}^{\pi} d\theta F_4(X, Y, \theta) e^{i\theta t}. \quad (11)$$

The initial values $f_2(X, 0) = f_4(X, Y, 0) = 1$ follow from completeness of the set of eigenstates and the choices of normalization in Eqs. (7) and (9) [see Eqs. (25) and (26)

for a discussion]. The late-time limits are also system independent. For $f_4(X, Y, t)$ this limit comes (assuming no degeneracies) from terms in Eq. (7) with pairwise equal labels $a = b, c = d$ or $a = d, b = c$. For $f_2(X, t)$ it comes from terms in Eq. (9) with $a = b$. All these terms can be simplified by noting that $C_X(a)C_X^\dagger(a) \equiv \text{Tr}_{\bar{X}}|a\rangle\langle a|$ is the reduced density matrix on subsystem X formed from the eigenstate $|a\rangle$. If $L(X) \ll L$, one expects from ETH that to an excellent approximation $C_X(a)C_X^\dagger(a) = q^{-L(X)}\mathbb{1}_X$, where $\mathbb{1}_X$ is the identity on X . This implies that $\lim_{|t| \rightarrow \infty} f_2(X, t) = q^{-2L(X)}$ and that $\lim_{t \rightarrow \infty} f_4(X, Y, t) = q^{2L-L(X,Y)-L(X)-L(Y)}$ for $L(X), L(Y) \ll L$.

B. Relation to autocorrelation functions of observables and OTOC

As we now discuss, these correlators are related respectively to the autocorrelation function [Eq. (2)] and the OTOC [Eq. (3)] via averages over the operators appearing in the latter two quantities. We begin by stating a key relation between $M_X(abcd)$ and an operator average of matrix elements. Given a subsystem X , choose a complete basis of $q^{2L(X)}$ Hermitian operators $\{X_\alpha\}$ that act on the subsystem and obey the orthonormality condition:

$$q^{-L(X)}\text{Tr}_X[X_\alpha X_\beta] = \delta_{\alpha\beta}. \quad (12)$$

Using the resolution of the identity in the vector space of operators, one finds

$$q^{-L(X)}\sum_\alpha \langle a|X_\alpha|b\rangle \langle c|X_\alpha|d\rangle = M_X(abcd). \quad (13)$$

The operator resolution of the identity and the relation given in Eq. (13) are represented diagrammatically in Fig. 2.

We first apply this to the simpler case of the autocorrelation function. The autocorrelation function averaged over all choices of operator (and over the ensemble of systems) is

$$q^{-2L(X)}\sum_\alpha \left[q^{-L}\text{Tr}[X_\alpha(t)X_\alpha] \right]_{\text{av}} = f_2(X, t). \quad (14)$$

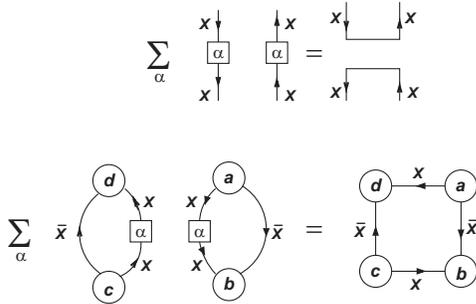


FIG. 2. Diagrammatic representation of (top) the operator resolution of the identity and (bottom) Eq. (13).

The special case $X_\alpha = \mathbb{1}_X$ is the only contribution to this average that survives at late times, giving a value of $\lim_{|t| \rightarrow \infty} f_2(X, t)$ consistent with the discussion above.

Similar arguments apply to the OTOC. In this case we choose two complete sets of operators. Operators in one set act on the subsystem labeled X and are denoted by X_α . Operators in the other set act on the subsystem labeled Y and are denoted by Y_β . Then

$$q^{-[L(X)+L(Y)]}\sum_{\alpha\beta} \langle a|X_\alpha|b\rangle \langle b|Y_\beta|c\rangle \langle c|X_\alpha|d\rangle \langle d|Y_\beta|a\rangle = M_X(abcd)M_Y(bcda). \quad (15)$$

This can be rewritten using $M_Y(bcda) = M_Y^*(abcd)$, and so the average of the OTOC over both sets of operators is

$$q^{-2[L(X)+L(Y)]}\sum_{\alpha\beta} \left[q^{-L}\text{Tr}[X_\alpha(t)Y_\beta X_\alpha(t)Y_\beta] \right]_{\text{av}} = q^{-S(X,Y)}f_4(X, \bar{Y}, t), \quad (16)$$

with $S(X, Y) = L(X) + L(Y) + L - L(X, \bar{Y}) = |X \setminus \bar{Y}| + |Y \setminus \bar{X}|$. Contributions to this average from the special cases $X_\alpha = \mathbb{1}_X$ and/or $Y_\beta = \mathbb{1}_Y$ survive at long times and are responsible for the limiting value given above.

The correlator $f_4(X, Y, t)$ also arises from a discussion of the operator entanglement entropy of the evolution operator. This quantity stems from considering the operator $W(t)$ as a state on a doubled Hilbert space, with components given by the matrix elements $[W(t)]_{i_X i_{\bar{X}} j_Y j_{\bar{Y}}}$. The corresponding reduced operator density matrix, obtained by tracing out the degrees of freedom in the subsystems \bar{X} and \bar{Y} , is

$$[\rho(X, Y, W(t))]_{i_X j_Y, l_X m_Y} = \sum_{ab} [C_X(a)C_X^\dagger(b)]_{i_X l_X} \times [C_Y(b)C_Y^\dagger(a)]_{m_Y j_Y} e^{i(\theta_b - \theta_a)t}. \quad (17)$$

The ensemble-averaged operator purity arising from this reduced density matrix is simply

$$[\text{Tr}[\rho(X, Y, W(t))^2]]_{\text{av}} = q^{L(X,Y)}f_4(X, Y, t). \quad (18)$$

The proportionality between $f_4(X, Y, t)$ and the operator purity of the evolution operator implies a straightforward link to the idea of an entanglement membrane, which has been proposed as a coarse-grained description of entanglement dynamics in chaotic many-body quantum systems [42,43]. For the one-dimensional models we are considering, the entanglement membrane is a curve in space-time, and to discuss the link to operator purity we build on the exposition of Ref. [43]. In outline, coarse-grained features of entanglement dynamics are determined by the line

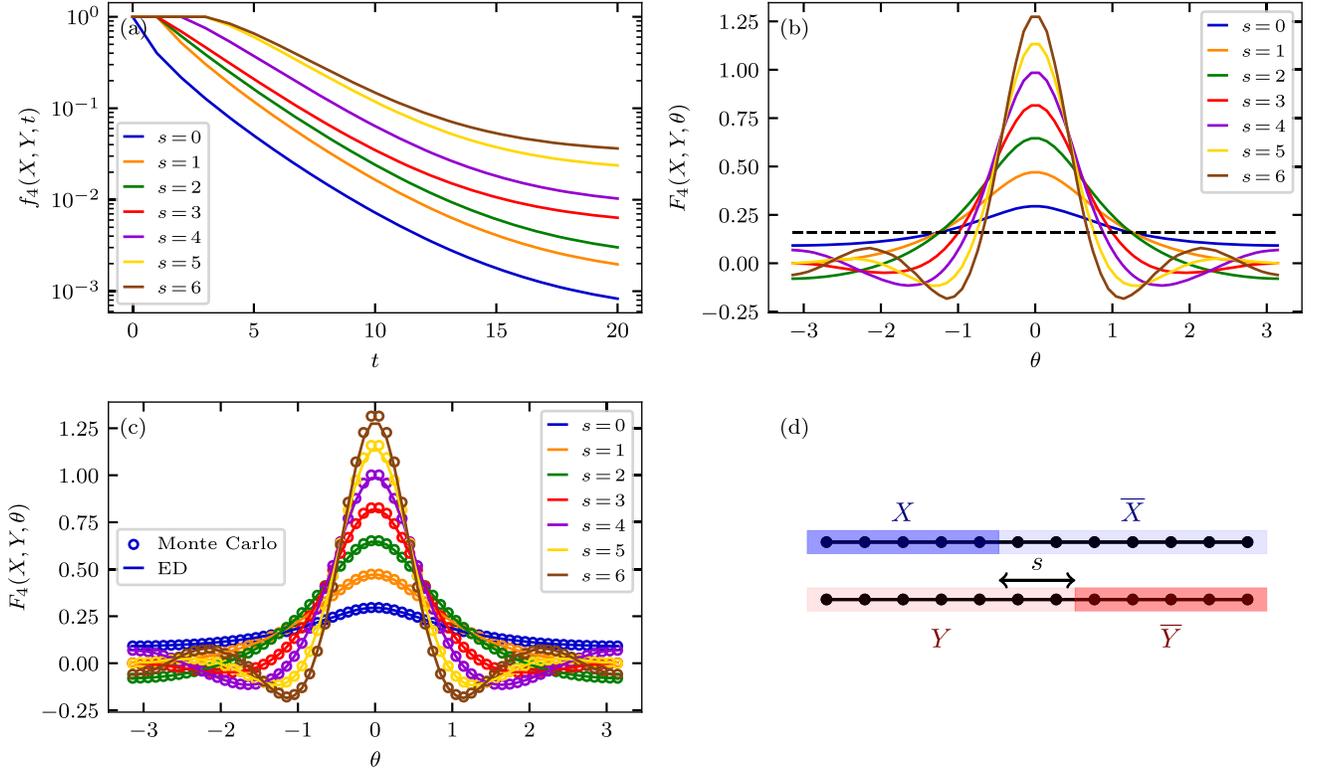


FIG. 3. Overview of main results, calculated for the brickwork circuit model defined in Sec. III A with $L = 12$, $q = 2$ and open boundary conditions. Behavior from ED in (a) and (b), comparison of ED and MC sampling in (c), and geometries of subsystems X and Y in (d). (a) The correlator $f_4(X, Y, t)$ [Eq. (9)] versus t , obtained using ED. Recall [Eq. (16)] that this correlator is proportional to the OTOC $q^{-L} \text{Tr}[X_\alpha(t) \bar{Y}_\beta X_\alpha(t) \bar{Y}_\beta]$ averaged over operators X_α and \bar{Y}_β with support on subsystems X and \bar{Y} , respectively. Decay of the correlator reflects operator spreading, and the onset time for decay increases with s . (b) The correlator $F_4(X, Y, \theta)$ [Fourier transform of $f_4(X, Y, t)$; see Eq. (7)] versus quasienergy difference θ , obtained using ED [contributions to Eq. (7) in which any of the state labels a, b, c , and d are equal have been omitted; they are atypical and carry vanishing weight in the thermodynamic limit]. It has a peak centered on $\theta = 0$ which grows narrower and higher with increasing s , reflecting the short-time plateau in (a). Black dashed line: behavior when the Floquet operator W is modeled using a $q^L \times q^L$ Haar unitary, showing for this structureless case that $F_4(X, Y, \theta)$ is nonzero but θ independent. (c) Comparison of ED results with MC results from the ansatz for the JDF of Eq. (29) fitted to behavior in the geometries of (d), showing excellent agreement between $F_4^{\text{MC}}(X, Y, \theta)$ (open circles from MC sampling) and $F_4(X, Y, \theta)$ (lines from ED) versus θ for various s . (d) Illustration of two ways of dividing the 12-site system with open boundary conditions into subsystems by means of a single spatial cut. In one case the subsystems are labeled X and \bar{X} ; in the other the labels are Y and \bar{Y} . The distance between the spatial cuts in the two cases is denoted by s .

tension $\mathcal{E}(v)$ of this membrane, which is a function of a velocity v . In our notation, $v = s/t$, where s is the distance between the ends of subsystems X and Y , defined in Fig. 3(d). For a fixed choice of X and Y with s large, the operator purity of the evolution operator is proportional to the line tension, and so traces out the function $\mathcal{E}(v)$ as t varies. Hence, the correlators $f_4(X, Y, t)$ and $F_4(X, Y, \theta)$ can be seen as representations of the line tension $\mathcal{E}(v)$.

C. Multitime and multi-quasienergy correlators

An obvious generalization [40] of the OTOC [Eq. (3)] is to introduce three time arguments, by considering the quantity $q^{-L} \text{Tr}[X_\alpha(t+t_2) Y_\beta(t_1) X_\alpha(t) Y_\beta]$. Correspondingly, in the quasienergy domain we have a generalization of the correlator $F_4(X, Y, \theta)$, defined by

$$\begin{aligned}
 F_4(X, Y; \theta, \theta_1, \theta_2) &= q^{-L(X, Y)} \left[\sum_{abcd} M_X(abcd) M_Y^*(abcd) \right. \\
 &\quad \times \delta(\theta - \theta_a + \theta_b - \theta_c + \theta_d) \\
 &\quad \times \delta(\theta_2 - \theta_a + \theta_b) \\
 &\quad \left. \times \delta(\theta_1 - \theta_b + \theta_c) \right]_{\text{av}}, \quad (19)
 \end{aligned}$$

with the Fourier transform

$$\begin{aligned}
 f_4(X, Y; t, t_1, t_2) &= \int_{-\pi}^{\pi} d\theta \int_{-\pi}^{\pi} d\theta_1 \int_{-\pi}^{\pi} d\theta_2 F_4(X, Y; \theta, \theta_1, \theta_2) \\
 &\quad \times e^{i(\theta t + \theta_1 t_1 + \theta_2 t_2)}, \quad (20)
 \end{aligned}$$

which is related to the generalized OTOC by

$$f_4(X, \bar{Y}; t, t_1, t_2) = q^{S(X,Y)-2L(X)-2L(Y)} \times \sum_{\alpha\beta} q^{-L} \text{Tr}[X_\alpha(t+t_2)Y_\beta(t_1)X_\alpha(t)Y_\beta]. \quad (21)$$

The single-quasienergy correlator can be recovered from the multi-quasienergy version using

$$F_4(X, Y, \theta) = \int_{-\pi}^{\pi} d\theta_1 \int_{-\pi}^{\pi} d\theta_2 F_4(X, Y; \theta, \theta_1, \theta_2). \quad (22)$$

For large separation s between the ends of X and \bar{Y} these correlation functions factorize as follows (see Ref. [40]). Consider behavior in the time domain. Let v_B be the butterfly velocity and D the diffusion constant for the spread of an operator front. Then we expect

$$q^{-L} \text{Tr}[X_\alpha(t+t_2)\bar{Y}_\beta(t_1)X_\alpha(t)\bar{Y}_\beta] \approx \begin{cases} q^{-L} \text{Tr}[X_\alpha(t+t_2)X_\alpha(t)] \times \\ \quad \times q^{-L} \text{Tr}[\bar{Y}_\beta(t_1)\bar{Y}_\beta] & s - v_B|t| \gg \sqrt{Dt} \\ 0 & v_B|t| - s \gg \sqrt{Dt} \end{cases}$$

This motivates the approximation,

$$f_4(X, Y; t, t_1, t_2) \cong f_4(X, Y, t) f_2(X, t_2) f_2(Y, t_1), \quad (23)$$

which in the quasienergy domain is

$$F_4(X, Y; \theta, \theta_1, \theta_2) \cong F_4(X, Y, \theta) F_2(X, \theta_2) F_2(Y, \theta_1). \quad (24)$$

For the models we study in this paper, $F_2(X, \theta_2)$ is only weakly dependent on θ , and so the multi-quasienergy correlator carries only limited extra information compared to the single-quasienergy version. For this reason we leave study of $F_4(X, Y; \theta, \theta_1, \theta_2)$ for future work.

D. Existence of correlations beyond ETH

Our objective in the remainder of this work is to find a form for the joint distribution function (JDF) of a small number of eigenstates that reproduces these correlations. We do this using a maximum entropy ansatz with a final form that we build up by considering first individual vectors, then pairs of vectors, and finally sets of four vectors.

To place our approach in context, it is useful to recall (following Refs. [20,40]) the limitations of ETH in its standard formulation when applied to the OTOC. As a starting point, consider the spectral decomposition of the OTOC in terms of operator matrix elements, as displayed in Eq. (3). ETH asserts that matrix elements of the form $\langle a_1|X_\alpha|a_2\rangle$ and $\langle a_3|Y_\beta|a_4\rangle$ appearing in this expression are Gaussian random variables, and are independent apart

from the constraint implied by Hermiticity of the operators X_α and Y_β . The mean values of off-diagonal matrix elements are automatically zero, and those of diagonal matrix elements are zero for traceless operators in the Floquet setting of interest. Finally, the variance of these matrix elements is set by the Hilbert space size and is $\mathcal{O}(q^{-L})$.

Applying these ideas to Eq. (3), the OTOC is given by q^{-L} times a sum of q^{4L} random $\mathcal{O}(q^{-2L})$ terms. Of these, only the q^L terms with $a = b = c = d$ are expected from ETH to have a nonzero average. This would imply an average value for the OTOC of $\mathcal{O}(q^{-2L})$. Treating the remaining terms as independent random variables, one expects $\mathcal{O}(q^{-L})$ fluctuations around this average. In contrast, the true value is $\mathcal{O}(1)$ at short times. To resolve this discrepancy it is necessary that a product of four matrix elements of the form $\langle a|X_\alpha|b\rangle\langle b|Y_\beta|c\rangle\langle c|X_\alpha|d\rangle\langle d|Y_\beta|a\rangle$ has a nonzero average that is $\mathcal{O}(q^{-3L})$ in addition to the $\mathcal{O}(q^{-2L})$ fluctuations captured by the standard version of ETH [40]. These additional correlations are the central concern in this paper and in the generalization of ETH discussed in Refs. [36,37].

A simple demonstration that such correlations must be present, regardless of details of the dynamics, is provided by a sum rule related to the value of the OTOC at $t=0$. From the left-hand side of Eq. (3), assuming for simplicity that the subsystems X and Y do not overlap and using the operator normalization of Eq. (12), we have

$$q^{-L} \text{Tr}[X_\alpha(t)Y_\beta X_\alpha(t)Y_\beta]|_{t=0} = 1. \quad (25)$$

Using this in Eq. (16) with Eq. (11) we have

$$\int_{-\pi}^{\pi} d\theta F_4(X, \bar{Y}, \theta) = 1. \quad (26)$$

This sum rule for $F_4(X, \bar{Y}, \theta)$ is automatically satisfied if eigenstates are Haar-distributed vectors, and in that case $F_4(X, \bar{Y}, \theta)$ is independent of θ . The eigenstate correlations that we are concerned with generate a dependence of $F_4(X, \bar{Y}, \theta)$ on θ but do not alter the fact that, with the normalization of Eq. (7), it has an order of magnitude that is independent of the Hilbert space dimension q^L .

E. Describing correlations beyond ETH

Some constraints on the eigenstate JDF are implied by ETH, which we now consider. ETH specifies statistical properties of both diagonal and off-diagonal matrix elements of local observables between eigenstates, and we discuss the two classes of matrix elements separately.

For a system with a time-independent Hamiltonian, a key part of ETH is that diagonal matrix elements of observables vary smoothly with energy, taking average values compatible with a thermal ensemble at the same energy density, and with fluctuations of a characteristic size that vanishes rapidly as the thermodynamic limit is approached. By contrast, for Floquet systems, statistical properties of diagonal matrix elements of observables are independent of quasienergy. We capture this property of diagonal matrix elements in a Floquet system by taking individual eigenstates to be isotropically distributed vectors in the Hilbert space for the model. We outline in Sec. VI the alternative choice required to model the energy dependence of diagonal matrix elements in Hamiltonian systems within our approach. We denote the isotropic (Haar) distribution for one, two, or four orthonormal vectors by $P_1^{(0)}(a)$, $P_2^{(0)}(a, b)$, and $P_4^{(0)}(a, b, c, d)$, respectively, and set out to modify these distributions in a way that introduces the correlations of interest.

Statistical properties of off-diagonal matrix elements determine the approach to equilibrium and the autocorrelation functions of observables. ETH applied to Floquet systems asserts that these matrix elements are independent Gaussian random variables with a variance that depends only on quasienergy separation. A central idea in our work is that strict independence is incompatible with the correlations implied by the dynamics of quantum information. Instead there are correlations (albeit weak) and these are better handled by considering distributions for eigenstates rather than matrix elements. We take the joint distribution of a pair of eigenstates to have the maximum entropy form

$$P_2(a_1, a_2) = Z_2^{-1} P_2^{(0)}(a_1, a_2) e^{-S_2(a_1, a_2)}, \quad (27)$$

with Z_2 a normalization constant and

$$S_2(a, b) = \sum_X G_2(X, \theta_a - \theta_b) M_X(abba), \quad (28)$$

where the coefficients $G_2(X, \theta)$ act as a Lagrange multipliers and should be chosen to reproduce the behavior of $F_2(X, \theta)$ as determined for a particular system.

Extending this pattern, we take the joint distribution of four eigenstates to have the form

$$P_4(a_1, a_2, a_3, a_4) = Z_4^{-1} P_4^{(0)}(a_1, a_2, a_3, a_4) \times e^{-\sum_{j < k} S_2(a_j, a_k) - S_4(a_1, a_2, a_3, a_4)}, \quad (29)$$

with Z_4 a normalization constant and

$$S_4(a, b, c, d) = \sum_{XY} G_4(X, Y, \theta_a - \theta_b + \theta_c - \theta_d) \times M_X(abcd) M_Y^*(abcd). \quad (30)$$

Here the Lagrange multipliers $G_4(X, Y, \theta)$ should be chosen to reproduce the behavior of $F_4(X, Y, \theta)$.

Two further ingredients are required. One is to establish a practical method for deducing the values of the Lagrange multipliers from information on the correlators. The other is to test the approach by sampling $P_2(a, b)$ or $P_4(a, b, c, d)$ and comparing the results with correlators calculated from exact diagonalisation of W . In this work we focus on the correlator $F_4(X, Y, \theta)$ since it contains the long-distance, low-energy information related to the dynamics of quantum information. Moreover, in the models we study, the lower-order correlator $f_2(X, t)$ decays rapidly in time. This implies that $F_2(X, \theta)$ is approximately independent of quasienergy and so we simply set $G_2(X, \theta)$ to zero in our initial treatment. We return to consideration of a θ -dependent $F_2(X, \theta)$ and nonzero $G_2(X, \theta)$ immediately after our discussion of $F_4(X, Y, \theta)$.

The determination of the Lagrange multipliers from ED data for eigenstate correlator can be seen simply as a fitting problem, but since this involves a high-dimensional parameter space, alternative approaches are desirable. Fortunately, as we describe in Sec. III B, we have been able to find direct and straightforward methods to derive $G_4(X, Y, \theta)$ from $F_4(X, Y, \theta)$ and $G_2(X, Y, \theta)$ from $F_2(X, Y, \theta)$.

F. Results

We implement and test these ideas using an open chain with the Floquet operator defined by a brickwork circuit (see Sec. III A). If nonzero Lagrange multipliers $G_4(X, Y, \theta)$ were included for all choices of subsystem X and Y , their number would be unreasonably large, both as a matter of principle and in practice (there are 2^L different subsystems in total). Since our models have nearest-neighbor interactions, it seems natural to organize subsystems according to the number of cuts necessary to separate them from the full system. For the majority of our work we include nonzero Lagrange multipliers only for subsystems that can be obtained from a full system under open boundary conditions by means of a single spatial cut. We discuss alternative choices of the sets of subsystems X and Y for which the Lagrange multipliers $G_4(X, Y, \theta)$ are nonzero in Sec. III B and Appendix D. As an additional restriction on the choice of nonzero Lagrange multipliers, our fastest method (see Sec. III B 1) for determining their values is effective if the probability distribution of $M_X(abcd)$ is well approximated by a Gaussian, which is the case for the model studied provided the subsystem sizes $L(X)$ and $L(\bar{X})$ are not too small. In order to satisfy this requirement, and in order to limit the total number of Lagrange multipliers under consideration, we include $G_4(X, Y, \theta)$ for all $L - 5$ subsystems X with $L(X) > 2$ that can be obtained from the full system by means of a single spatial cut, and similarly for Y . Taking account of symmetry under interchange of X and Y , this gives

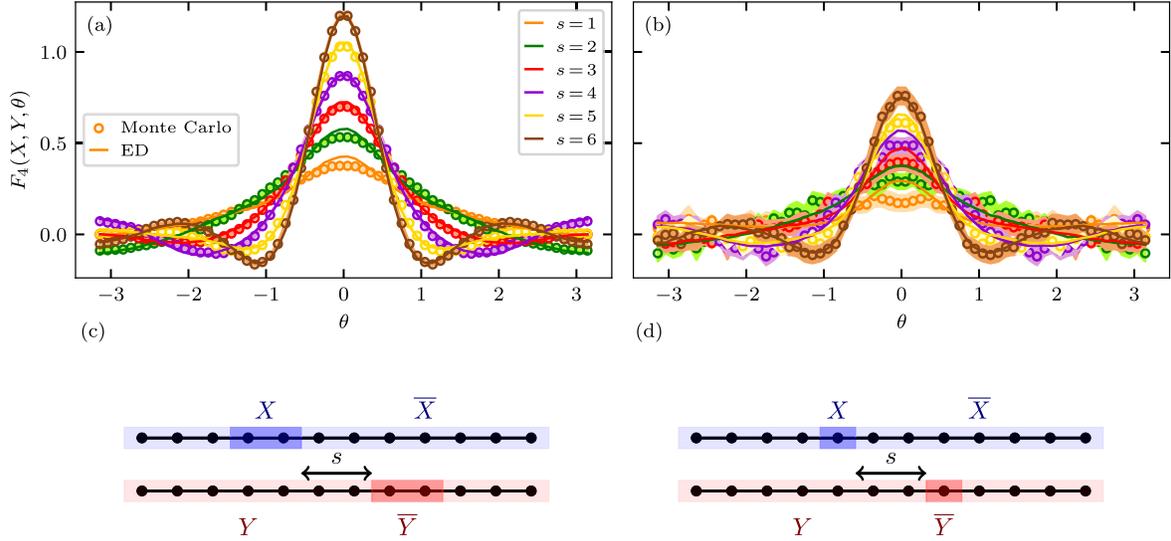


FIG. 4. Test of JDF fitted to behavior in the geometries of Fig. 3(d) but applied to geometries of (c) and (d), respectively. (a),(b) Comparison of data from MC sampling (open circles) and ED (solid lines). (c) Partition used for (a), in which the 12-site system is divided by two spatial cuts into a two-site subsystem X and its complement \bar{X} , or a two-site subsystem \bar{Y} and its complement Y . (d) Partition used for (b), in which the 12-site system is divided by two spatial cuts into a single-site subsystem X and its complement \bar{X} , or a single-site subsystem \bar{Y} and its complement Y . Calculations are for the brickwork circuit model defined in Sec. III A with $L = 12$, $q = 2$ and open boundary conditions.

$(L - 5)(L - 4)/2$ (i.e., 28 for the example studied of $L = 12$) independent Lagrange multipliers. Using as input the values of $F_4(X, Y, \theta)$ for these subsystem choices from ED, we arrive at a final form for the JDF.

We use Monte Carlo sampling of this distribution to compute the correlator [denoted by $F_4^{\text{MC}}(X, Y, \theta)$] with two objectives. First, for the simplest choices of X involving single spatial cuts (and similarly for Y), comparison with ED is a test of our ansatz for the JDF and of our procedure to determine the Lagrange multipliers. Second, it is interesting to see whether this input alone is sufficient to capture correlations more generally. To probe this we compare ED and MC results for the correlator, making choices of X (and also Y) that are defined by more than one spatial cut. This is a test of the extent to which the proposed JDF captures long-distance, low-energy correlations in general. In particular, taking X and Y each to consist of a small number of sites acting as the support for a local observable, we test the implications of the JDF for the (operator-averaged) OTOC.

Some of our principal results are shown in Fig. 3 and discussed in the figure caption. The main conclusions are as follows. (i) As expected from its relation to the OTOC, the correlator $f_4(X, Y, t)$ is time independent at short times and falls off at a timescale that is long if the spatial separation s between subsystems X and \bar{Y} is large. All aspects of this behavior are apparent in Fig. 3(a). (ii) In turn, this implies structure in $F_4(X, Y, \theta)$ at small quasienergies θ , as is visible in Fig. 3(b); the width in quasienergy of this structure decreases with increasing s . (iii) Monte Carlo

sampling of the JDF of Eq. (29), with Lagrange multipliers determined as described in Sec. III B, generates results for $F_4(X, Y, \theta)$ that are in excellent agreement with those from ED, as demonstrated in Fig. 3(c).

Further important results are shown in Fig. 4. Here we examine how well the JDF constructed using correlators for single-cut subsystems can capture correlators for two-cut subsystems. It is apparent from Fig. 4(a) that MC sampling of the JDF generates a moderately good representation of $F_4(X, Y, \theta)$ for the two-site choices of X and Y shown in Fig. 4(c). Equivalently, the JDF determined using information from the geometries of Fig. 3(d) reproduces the main features of the OTOC, as a function of time and spatial separation, for two operators, each supported on two sites in the geometry of Fig. 4(c). A similar picture holds even when the support of the operators is reduced to a single site, as is shown in Fig. 4(b). In this case, statistical errors are enhanced, since in each subsystem there are only one quarter as many operators to average over.

As a final indication of the effectiveness of our approach, we return to the behavior of $F_2(X, \theta)$, which characterizes the correlations that are incorporated by ETH. For simplicity, we consider only the joint distribution of pair of eigenstates, in this way treating $F_2(X, \theta)$ separately from $F_4(X, Y, \theta)$. By determining the Lagrange multiplier $G_2(X, \theta)$ from ED data as described in Sec. III B, we generate and sample from this joint distribution, with results that are shown in Fig. 5. As is evident, our MC data are in excellent agreement with ED results for all values of quasienergy difference θ and all subsystem choices X .

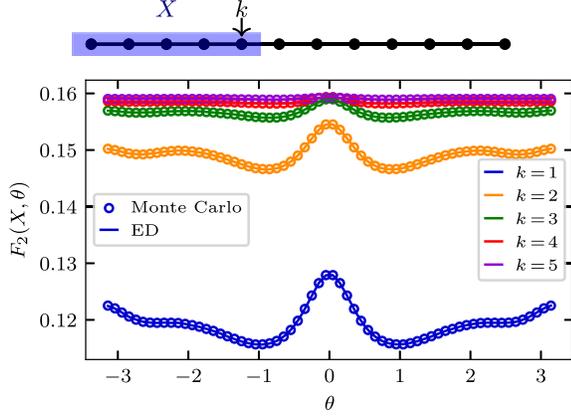


FIG. 5. Numerical results for $F_2(X, \theta)$ in a system with $L = 12$ and $q = 2$ as a function of the left cut position k : ED results (solid lines) versus Monte Carlo sampling (circles). [Delta-function contributions at $\theta = 0$ have been omitted; they are responsible for ensuring that the sum rule $\int_{-\pi}^{\pi} d\theta F_2(X, \theta) = 1$ is satisfied for all X .] The almost perfect agreement between both sets of data indicates that the correlations between pairs of eigenstates $F_2(X, \theta)$, as captured by ETH, may also be represented accurately within our approach.

The fact that deviations are small from the value $F_2(X, \theta) = (2\pi)^{-1} \approx 0.159$ for a pair of Haar-distributed orthogonal unit vectors is justification for our omission of $F_2(X, \theta)$ in our discussion of $F_4(X, Y, \theta)$. A more complete treatment would require the simultaneous inclusion of both $G_2(X, \theta)$ and $G_4(X, Y, \theta)$, following Eq. (29). Some consequences have been discussed previously in Ref. [40] and we do not consider them further here.

The overall aim of MC sampling from our maximum entropy ansatz for the JDF of a small number of eigenstates is to test whether, with a suitable choice of Lagrange multipliers, the JDF reproduces the correlations $F_4(X, Y, \theta)$ measured from ED. This test of our approach is a crucial one, and we believe Fig. 3(c) offers excellent evidence that the JDF can indeed reproduce the required correlations for the model and parameter range investigated there. Further discussion of the determination of Lagrange multipliers, including a treatment of other models, is presented in Sec. V and the Appendixes B and D.

III. MODELS, LAGRANGE MULTIPLIERS, AND NUMERICAL METHODS

In Sec. III A we describe the microscopic models we use to generate the numerical results shown in this paper. In Sec. III B we set out efficient methods to determine the Lagrange multipliers $G_2(X, \theta)$ and $G_4(X, Y, \theta)$ that appear in the JDF for eigenstates. In Sec. III C we give details of the numerical methods used in this paper.

A. Models

We use the brick-wall circuit depicted in Fig. 6 as a simple model of a periodically driven many-body quantum

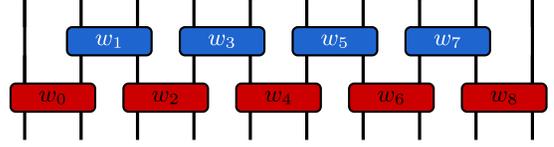


FIG. 6. Unitary time-evolution operator W of the Floquet circuit written as a tensor network.

system with local interactions in one dimension. Such Floquet models have been studied extensively in past work; see, for example, Refs. [40,44,45]. This circuit is defined in terms of two-site gates $w_i \in \mathbb{C}^{q^2 \times q^2}$, and the driving period is decomposed into two parts: In the first half of the period, couplings are active only on even bonds of the system, while in the second half the couplings are active only on odd bonds. Thus, the time evolution operator for the first half period is $W_1 = w_0 \otimes w_2 \otimes w_4 \cdots$ and for the second half period is $W_2 = \mathbb{1} \otimes w_1 \otimes w_3 \otimes \cdots$. The evolution operator over one full period is $W = W_2 W_1$, and for t periods we write $W(t) \equiv W^t$.

In order to define an ensemble of systems, a natural choice would be to draw each unitary matrix w_i independently from the Haar distribution. We find, however, that in this case determination of the Lagrange multipliers $G_4(X, Y, \theta)$ is complicated by effects that we attribute to realizations containing weak links i on which the gate w_i is close to the identity (especially in small systems or with small subsystems). To avoid such weak links, we draw the w_i from a truncated version of the Haar distribution in which all gates with an operator purity above a cutoff are discarded. With operator purity defined as in Eq. (18) (so that the two-site identity operator has a purity of q^4) we take the cutoff to be $0.3 \times q^4$ for local Hilbert space dimension $q = 2$. The consequences of changing or omitting this cutoff are discussed in Sec. V A and Appendix A 2, and results for $q = 3$ are given in Sec. V B.

B. Determining Lagrange multipliers in the JDF

We now discuss the problem of determining the Lagrange multipliers that appear in Eqs. (28) and (30), and that define the JDFs, Eqs. (27) and (29), of a small number of eigenstates. We present a method specific to $G_4(X, Y, \theta)$ in Sec. III B 1 and one specific to $G_2(X, \theta)$ in Sec. III B 2. These are both single-shot methods that make explicit use of information about the probability distributions of $M_X(abcd)$ and $M_X(abba)$, respectively, and for this reason they are particularly efficient. In Sec. III B 3 we outline a third, more generally applicable iterative approach, that is agnostic to the probability distributions involved.

1. Determination of $G_4(X, Y, \theta)$

As noted above, pairwise correlations between eigenstates encoding the decay of autocorrelation functions are

very weak in the Floquet model we consider, in the sense that the correlator $F_2(X, \theta)$ is only weakly dependent on θ , taking values close to that for a Haar-distributed pair of orthogonal vectors. As a simplifying approximation, we therefore set to zero the Lagrange multipliers $G_2(X, \theta)$ [see Eq. (28)] that control these pairwise correlations. We make further choices concerning the set of subsystems X and Y for which Lagrange multipliers $G_4(X, Y, \theta)$ are included in the JDF. Without restrictions there are $2^L - 1$ distinct subsystems: To reduce this number we include only Lagrange multipliers for connected subsystems—those that can be obtained from the full system (which has open boundary conditions) by means of a single cut—and we omit them for all subsystems obtained using multiple cuts.

This choice gives a minimal set of partitions, which turns out to be sufficient to reproduce the main correlations arising from spatial structure and local interactions. As is shown in Fig. 4, it also captures the eigenstate correlations $F_4(X, Y, \theta)$ even for some subsystems pairs X and Y not included in the set. The reason is probably that the most important parameter characterizing the dynamics of information spreading is simply the distance s separating the closest points on the two partitions, and that further possible structure in the partitions is unimportant.

A more general choice is to include Lagrange multipliers for subsystems that can be obtained from the full system by means of either one or two cuts, and we examine this in Appendix D. We find that adding more partitions increases the complexity of calculations without significantly improving the accuracy of MC results for $F_4(X, Y, \theta)$. Systems with periodic rather than open boundary conditions present a new problem, since in this case a minimum of two cuts is required to define a subsystem. We treat this instance in Sec. V C.

Our objective is to find the values of the Lagrange multipliers $G_4(X, Y, \theta)$ for which the JDF reproduces the eigenstate correlator $F_4(X, Y, \theta)$ (known from ED) as accurately as possible. We are able to simplify this task and avoid attacking directly a high-dimensional fitting problem if the quantities $M_X(abcd)$ [Eq. (6)] are Gaussian distributed. Our motivation for treating a model with a truncated Haar distribution of gate unitaries is that in this system $M_X(abcd)$ is well approximated by a Gaussian.

Evidence for this is presented in Fig. 7. Here we consider only the magnitude $|M_X(abcd)|$ since the phase of $M_X(abcd)$ is dependent on the phase convention used for the eigenstates, and we compare the distribution of $|M_X(abcd)|$ with one in which the real and imaginary parts of $M_X(abcd)$ are assumed to be uncorrelated Gaussian variables with equal variances and zero means. An analysis of the dependence of these distributions on the system size L and subsystem size $L(X)$ (see Appendix A) suggests that deviations from a Gaussian vanish in the limit of large L and $L(X)$. Since deviations are significant for small subsystem size, we omit Lagrange multipliers $G_4(X, Y, \theta)$ for

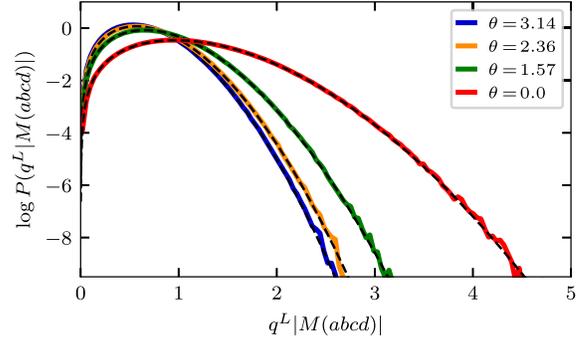


FIG. 7. Probability distribution of $q^L |M_X(abcd)|$ in the Floquet model of Sec. III A (colored data) compared with fitted Gaussian distributions (black dashed lines). Data are for $L = 12$, $q = 2$ and a subsystem X consisting of the four sites closest to the end of an open system, and are shown for the four indicated values of the relative phase $\theta = \theta_a - \theta_b + \theta_c - \theta_d$. Results for other q and X are shown in Appendix A.

subsystems X with $L(X)$ or $L(\bar{X}) \leq 2$. This leaves $(L - 5)^2$ Lagrange multipliers, or $(L - 5)(L - 4)/2$ independent quantities after taking account of symmetry relations.

Similarly, a Gaussian distribution for $M_X(abcd)$ also arises from the Haar distribution for four orthogonal vectors $[P_4^{(0)}(a, b, c, d)$ in the notation of Eq. (29)] in the large $q^L, q^{L(X)}$ limit. In this case, and in contrast to the Floquet model, the covariance is independent of the quasienergy difference θ . In the limit of large q^L and $q^{L(X)}$ it has the simple form

$$[M_X(abcd)M_Y^*(abcd)]_0 \simeq q^{L(X,Y)-4L}. \quad (31)$$

A full expression for Eq. (31), applicable for general q^L and $q^{L(X)}$, can also be obtained in terms of Weingarten functions, but turns out to be unnecessary for the work in this paper, except as noted in Sec. V C.

In order to represent these Gaussian distributions in a compact way, it is convenient to introduce notation in which the $L - 1$ single-cut subsystems X used to define our Lagrange multipliers are labels for basis states in a $(L - 1)$ -dimensional vector space. Then the values of $M_X(abcd)$ for different X are components of an $(L - 1)$ -component column vector \mathbf{M} , so that $[\mathbf{M}]_X = M_X(abcd)$, while the Lagrange multipliers are entries in the $(L - 1) \times (L - 1)$ matrix \mathbf{G}_4 with rows and columns labeled by X and Y , respectively, so that $[\mathbf{G}_4]_{X,Y} = G_4(X, Y, \theta)$. Similarly, the eigenstate correlators $F_4(X, Y, \theta)$ are elements of a matrix F_4 . Then Eq. (30) can be rewritten in the compact form:

$$S_4 = \mathbf{M}^T \mathbf{G}_4 \mathbf{M}^*. \quad (32)$$

The fact that the distributions $P_4(a, b, c, d)$ and $P_4^{(0)}(a, b, c, d)$ are [at large L and $L(X)$] Gaussian for \mathbf{M} suggests that a convenient coordinate system consists of the

components of \mathbf{M} together with additional variables Ω that we do not specify explicitly. We indicate this by writing $P_4(\mathbf{M}, \Omega)$ and $P_4^{(0)}(\mathbf{M}, \Omega)$. The result given in Eq. (31) for the covariance within the Haar distribution implies that

$$\int d\Omega P_4^{(0)}(\mathbf{M}, \Omega) = [\mathcal{Z}_4^{(0)}]^{-1} e^{-\mathbf{M}^T \mathbf{G}_4^{(0)} \mathbf{M}^*},$$

with $[(\mathbf{G}_4^{(0)})^{-1}]_{X,Y} = [[\mathbf{M}]_X [\mathbf{M}^*]_Y]_0 \simeq q^{L(X,Y)-4L}$, (33)

and $\mathcal{Z}_4^{(0)} = \pi^{L-1} / \det \mathbf{G}_4^{(0)}$. This in turn implies that

$$\int d\Omega P_4(\mathbf{M}, \Omega) = [\mathcal{Z}_4]^{-1} e^{-\mathbf{M}^T [\mathbf{G}_4 + \mathbf{G}_4^{(0)}] \mathbf{M}^*},$$
 (34)

and hence that

$$[[\mathbf{M}]_X [\mathbf{M}^*]_Y]_{\text{av}} = [(\mathbf{G}_4 + \mathbf{G}_4^{(0)})^{-1}]_{X,Y}. \quad (35)$$

In addition, we have from Eq. (7)

$$[\mathbf{F}_4]_{X,Y} = (2\pi)^{-1} q^{4L-L(X,Y)} [[\mathbf{M}]_X [\mathbf{M}^*]_Y]_{\text{av}}. \quad (36)$$

Equation (35) allows the determination of the Lagrange multipliers in Eq. (29) in terms of the matrix \mathbf{F}_4 , which is obtained using ED. This method was used to generate Fig. 3(c).

2. Determination of $G_2(X, \theta)$

Next we describe the method we use to determine the Lagrange multipliers $G_2(X, \theta)$, treating explicitly the case of $L - 1$ subsystems X generated by single cuts. A different approach is required to that for $G_4(X, Y, \theta)$ because the probability distribution of $M_X(abba)$ is quite different to that of $M_X(abcd)$ for $a \neq b \neq c \neq d$. Indeed, while (as discussed) $M_X(abcd)$ has a complex Gaussian distribution with zero mean, $M_X(abba)$ is from its definition real and non-negative.

To understand the distribution of $M_X(abba)$ it is useful to start from Eq. (13), which specializes here to

$$M_X(abba) = q^{-L(X)} \sum_{\alpha} |\langle a | X_{\alpha} | b \rangle|^2. \quad (37)$$

With $a \neq b$ we expect from ETH that $\langle a | X_{\alpha} | b \rangle$ for each α is an independent complex Gaussian random variable. From this we can conclude, first, that the two quantities $M_X(abba)$ and $M_{X'}(abba)$ are correlated if the sets $\{X_{\alpha}\}$ and $\{X'_{\alpha}\}$ have operators in common, and second, that statistically independent quantities can be constructed using a transformation that organizes the operators into disjoint sets.

In detail, this transformation is defined recursively by considering the $L - 1$ subsystems in order of increasing size. Again we introduce an $(L - 1)$ -dimensional vector

space, and define the vector \mathbf{M} to have components $[\mathbf{M}]_X = M_X(abba)$. Similarly, we introduce the notation \mathbf{T} for our target vector with statistically independent components $[\mathbf{T}]_X$, and \mathbf{G} with $[\mathbf{G}]_X = G_2(X, \theta)$ for the vector of Lagrange multipliers. To write the transformation we abuse notation and substitute in place of the component label X the value $\ell = L(X)$. Then using Eq. (37) we write

$$\begin{aligned} \mathbf{T}_1 &= \mathbf{M}_1 \\ \text{and } \mathbf{T}_{\ell} &= \mathbf{M}_{\ell} - q^{-1} \mathbf{M}_{\ell-1}, \end{aligned} \quad (38)$$

for $\ell = 2$ to $L - 1$. This can be recast in the matrix form

$$\mathbf{T} = \mathbf{V} \mathbf{M}, \quad (39)$$

where

$$\mathbf{V} = \begin{pmatrix} 1 & 0 & 0 & \dots & 0 \\ -q^{-1} & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & & \\ 0 & 0 & \dots & -q^{-1} & 1 \end{pmatrix}. \quad (40)$$

The effect of this transformation is that Eq. (37) is replaced by $\mathbf{T}_{\ell} = q^{-\ell} \sum'_{\alpha} |\langle a | X_{\alpha} | b \rangle|^2$, where the sum runs over the subset of $n_{\ell} \equiv (q^2 - 1) q^{2(\ell-1)}$ operators X_{α} that act nontrivially at the rightmost site in X [see illustration in Fig. 3(d)]. Since $\langle a | X_{\alpha} | b \rangle$ is Gaussian with mean zero, the variable $s_{\alpha} \equiv |\langle a | X_{\alpha} | b \rangle|^2$ has the distribution $p_{\alpha}(s_{\alpha}) = \sigma_{\alpha} e^{-\sigma_{\alpha} s_{\alpha}}$, where $\sigma_{\alpha} \equiv [|\langle a | X_{\alpha} | b \rangle|^2]_{\text{av}}^{-1}$. We make the approximation that σ_{α} takes the same value σ_{ℓ} for all X_{α} that contribute to a given \mathbf{T}_{ℓ} . This framework applies not only to the true eigenstate distribution under consideration, but also to vectors with a Haar distribution, and in the latter case we denote the value of σ_{α} by $\sigma_{\ell}^{(0)}$. Then

$$[\mathbf{T}_{\ell}]_{\text{av}} = \frac{n_{\ell}}{\sigma_{\ell}} \quad \text{and} \quad [\mathbf{T}_{\ell}]_0 = \frac{n_{\ell}}{\sigma_{\ell}^{(0)}}. \quad (41)$$

With these ingredients in hand, Eq. (28) can be written in the form

$$S_2 = \mathbf{G}^T \mathbf{V}^{-1} \mathbf{T}, \quad \text{so that } \sigma_{\ell} = \sigma_{\ell}^{(0)} + [\mathbf{G}^T \mathbf{V}^{-1}]_{\ell}. \quad (42)$$

Substituting Eq. (41) into Eq. (42) and rearranging, we obtain

$$\mathbf{G}_{\ell} = \sum_{\ell'} \left\{ \frac{n_{\ell'}}{[(\mathbf{V} \mathbf{M})_{\ell'}]_{\text{av}}} - \frac{n_{\ell'}}{[(\mathbf{V} \mathbf{M})_{\ell'}]_0} \right\} \mathbf{V}_{\ell' \ell}. \quad (43)$$

We employ Eq. (43) to determine the Lagrange multiplier $G_2(X, \theta)$ in terms of $[\mathbf{M}]_{\text{av}}$ obtained from ED and the Haar average

$$[M_\ell]_0 = q^{-L}(q^\ell - q^{-\ell}). \quad (44)$$

This approach is used to obtain the data shown in Fig. 5.

Note [from the definition of T_ℓ below Eq. (40) and the discussion following Eq. (41)] that the probability distribution of T_ℓ is consistent with a Gaussian distribution for the matrix elements $\langle a|X_\alpha|b\rangle$, as expected from ETH, with a variance controlled by the Lagrange multipliers $G_2(X, \theta)$.

3. Iterative method for determining Lagrange multipliers

We next describe a straightforward method for determining the Lagrange multipliers without making use of information about the probability distributions of $M_X(abcd)$ and $M_X(abba)$. We treat the case of $G_4(X, Y, \theta)$; the necessary modifications for $G_2(X, \theta)$ are obvious.

Our starting point is a perturbative expansion of Eq. (29) to first order in $G(X, Y, \theta)$, which yields

$$[M_X M_Y^*]_{\text{av}} = [M_X M_Y^*]_0 - \sum_{X'Y'} [K_4]_{XY, X'Y'} G(X', Y', \theta), \quad (45)$$

where $[K_4]_{XY, X'Y'}$ is the connected correlator:

$$[K_4]_{XY, X'Y'} = [M_X M_Y^* M_{X'} M_{Y'}^*]_0 - [M_X M_Y^*]_0 [M_{X'} M_{Y'}^*]_0. \quad (46)$$

Introducing the abbreviations

$$\begin{aligned} [V_4]_{XY} &= [M_X M_Y^*]_0 - [M_X M_Y^*]_{\text{av}} \\ \text{and } [G_4]_{XY} &= G_4(X, Y, \theta), \end{aligned} \quad (47)$$

this gives at first order in perturbation theory

$$G_4 = K_4^{-1} V_4. \quad (48)$$

To go beyond first-order perturbation theory, we define an iterative procedure based on Eq. (48). Let $[M_X M_Y^*]_{\text{MC}}^{(n)}$ denote the MC result obtained with the n th approximant $G_4^{(n)}(X, Y, \theta)$ as Lagrange multiplier, and let $[M_X M_Y^*]_{\text{ED}}$ be the value from ED. Then iterate for $n = 1, 2, \dots$,

$$[V_4^{(n+1)}]_{XY} = [M_X M_Y^*]_{\text{MC}}^{(n)} - [M_X M_Y^*]_{\text{ED}}, \quad (49)$$

with

$$[V_4^{(1)}]_{XY} = [M_X M_Y^*]_0 - [M_X M_Y^*]_{\text{ED}}, \quad (50)$$

and

$$G_4^{(n)} = G_4^{(n-1)} + K_4^{-1} V_4^{(n)}, \quad (51)$$

with

$$G_4^{(0)} = 0. \quad (52)$$

A possible refinement is to replace Eq. (51) with

$$G_4^{(n)} = G_4^{(n-1)} + \alpha K_4^{-1} V_4^{(n)}, \quad (53)$$

where $0 < \alpha \leq 1$ is a real parameter. Small α reduces the risk of overshooting the solution at the expense of a slower convergence rate.

This method is used to produce the data shown in Fig. 14.

C. Numerical methods

In this section, we give details of the ED, MC, and ensemble-averaging procedures used to obtain the data presented in this paper.

1. Exact diagonalization

In order to study eigenstate correlations, we use ED of the Floquet operator W to compute exact eigenvectors and hence averages of $M_X(abcd)M_Y^*(abcd)$ for all choices of our selected subsystems X and Y . We obtain phase-resolved averages by dividing the phase interval $[-\pi, \pi]$ into 64 bins. Each realization of W with L spins generates approximately $q^{4L}/4!$ different quadruples of eigenvectors, and from these we randomly choose up to 10^6 quadruples. We average over between 150 (for $L = 12$) and 1000 (for $L = 10$ and $L = 8$) realizations of W to obtain the ED data in Fig. 3, and over 25 000 realizations for the results shown in Fig. 4. The symmetry relation $F_4(X, Y, \theta) = F_4(X, Y, -\theta)$ allows us to restrict calculations to $\theta \geq 0$. Similarly, in the case of $M_X(abba)$ we take 10^6 tuples out of the approximately $q^{2L}/2$ possibilities for each realization of W and $L = 12$. We average over 1000 realizations for $L = 12$ to obtain the data in Fig. 5.

From these ED results we compute the Lagrange multipliers $G_4(X, Y, \theta)$ and $G_2(X, \theta)$ using the procedures described in Sec. III B. We find that a further symmetrization of the data, using the spatial symmetry under the interchange of X, Y with \bar{X}, \bar{Y} , improves stability.

As discussed in Sec. II D, the coherent contribution of the four-point correlator $\langle a|X_\alpha|b\rangle\langle b|Y_\beta|c\rangle \times \langle c|X_\alpha|d\rangle\langle d|Y_\beta|a\rangle$ is suppressed by a factor q^{-L} in comparison to fluctuations of this quantity. To average out the latter, it is thus necessary to average over $\mathcal{O}(q^{2L})$ samples. In practice, averaging over the operators X_α and Y_β reduces fluctuations, but for large spatial separation s and system size L , the number of required samples scales exponentially with the system size for a fixed support of X and \bar{Y} .

2. Monte Carlo sampling

In order to determine the eigenstate correlator $F_4(X, Y, \theta)$ from the JDF [Eq. (29)] for four eigenstates

$|a\rangle$, $|b\rangle$, $|c\rangle$, and $|d\rangle$, we use Monte Carlo sampling with $e^{-S_4(a,b,c,d)}$ [Eq. (30)] as the weighting term. Similarly, to determine the correlator $F_2(X, \theta)$ from the JDF [Eq. (27)] for two eigenstates $|a\rangle$ and $|b\rangle$, we use Monte Carlo sampling with $e^{-S_2(a,b)}$ [Eq. (28)] as the weighting term.

We follow the Metropolis-Hastings algorithm to obtain a Markov chain of vector quadruples (a, b, c, d) distributed according to P_4 . To generate the next quadruple (a', b', c', d') , we randomly select one of the four vectors a, b, c, d and add a vector $\sqrt{\epsilon}v$ to the chosen state, where v is a Haar-random vector orthogonal to a, b, c, d .

We set $\epsilon = 0.1$ for $L = 12$ and $q = 2$, and $\epsilon = 0.8$ for $L = 8$ and $q = 2$. Since V is unitary, orthonormal vectors retain this property after the transformation.

This choice of update rule with tuning parameter ϵ allows us to perform effective importance sampling, since we propose only relatively small changes to the sample. The new sample is then accepted in the Markov chain with probability

$$P_{\text{accept}} = \min(1, e^{-S_4(a,b,c,d)+S_4(a',b',c',d')}). \quad (54)$$

For each θ we perform 2000 Monte Carlo runs in parallel with up to 10^6 samples per run. The results are obtained by averaging over all runs. In the case of small θ , we find in rare cases (one out of 1000 runs) instabilities toward local maxima of the weighting function during the sampling process. This is visible by tracking $S_4(a, b, c, d)$ or the acceptance rate. This problem can be circumvented by decreasing the size of the update steps, but only at the expense of longer autocorrelation times. As a compromise, we discard runs where S_4 falls below a threshold value of -50 .

To provide an overall test of our form [Eq. (29)] for the JDF, we show in Fig. 8 a comparison of the distributions of $S_4(a, b, c, d)$ obtained respectively from ED and from MC sampling of the JDF. The excellent agreement between the two distributions over a range of values for θ is evidence of the internal consistency of our approach.

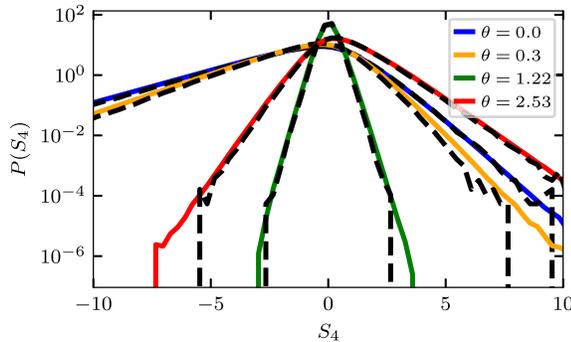


FIG. 8. Probability distribution of $S_4(a, b, c, d)$ [Eq. (30)] for quasienergy differences θ as indicated. Comparison between ED results (black dashed lines) and MC results (solid colored lines). Parameters as in Fig. 3.

IV. PERTURBATIVE TREATMENT OF LAGRANGE MULTIPLIERS

An obvious approach to calculations based on the joint eigenstates distributions of Eqs. (27) and (29) is a perturbative expansion in powers of the Lagrange multipliers $G_2(X, \theta)$ and $G_4(X, Y, \theta)$. In this section we set out a general framework for such an expansion and apply it in several ways. While the expansion does not generate fundamentally new results, it provides a useful perspective that is complementary to the one set out in Secs. II and III.

We use the expansion to provide an alternative justification of the fitting procedure for $G_4(X, Y, \theta)$ to the one described in Sec. III B 1. We also use it to consider Eq. (29) without the simplification employed in Sec. II F of setting $G_2(X, \theta) = 0$. We show that for large q^L and $q^{L(X)}$ that if $F_2(X, \theta)$ is calculated from the JDF for four eigenstates rather than two, the influence of $G_4(X, Y, \theta)$ on the result is small. This means that the predictions of ETH for matrix elements between pairs of eigenstates are only weakly affected by the correlations between sets of four eigenstates $F_4(X, Y, \theta)$ that are introduced with our ansatz for the JDF. Finally, we show that the effect of the Lagrange multipliers $G_2(X, \theta)$ and $G_4(X, Y, \theta)$ on the normalization of eigenstates drawn from the joint distributions is small in large systems.

An exact perturbative expansion requires the evaluation of averages over a Haar distribution of vectors, denoted by $P_1^{(0)}(a)$, $P_2^{(0)}(a, b)$, and $P_4^{(0)}(a, b, c, d)$ in Sec. II E. While the formalism required for this is well developed (see, e.g., Ref. [46]), it is quite cumbersome. Moreover, we require results only at leading order for q^L large. These can be obtained by substituting another distribution in place of the Haar distribution in which each vector component is an independent Gaussian random variable. Specifically, consider a computational basis $\{|k_c\rangle\}$ and denote the overlap of the eigenstate $|a\rangle$ with the basis state $|k_c\rangle$ by $a(k) = |k_c|a\rangle$. Define $S_1(a)$ by

$$S_1(a) = q^L \sum_k |a(k)|^2. \quad (55)$$

We replace $P_4^{(0)}(a_1, a_2, a_3, a_4)$ by

$$P_4^{(G)}(a_1, a_2, a_3, a_4) = \left(\frac{q^L}{\pi}\right)^{4q^L} e^{-\sum_i S_1(a_i)}. \quad (56)$$

Although different vectors drawn from this distribution are not in general exactly orthonormal, orthonormality is recovered in the limit $q^L \rightarrow \infty$. We denote averages with respect to this Gaussian distribution by $[\dots]_G$.

A. Diagrammatic notation

It is useful to employ diagrammatic notation for these Gaussian averages. As in Fig. 1, eigenstates are represented by circles and index contractions following from the definition of $M_X(abcd)$ are indicated by solid lines carrying

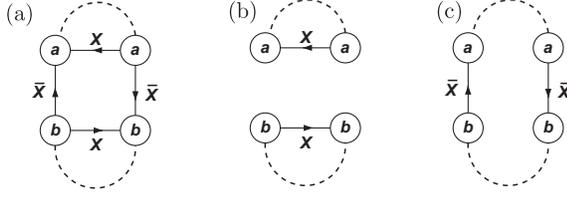


FIG. 9. Diagrammatic representation and evaluation of $[M_X(abba)]_G$. (a) Full diagram, (b) decomposition into loops for a site in subsystem X , (c) decomposition into loops for a site in subsystem \bar{X} .

arrows that run from states $|a\rangle$ toward conjugate states $\langle b|$; these lines carry labels to indicate the subsystem within which the contraction is done. The combination of circles and full lines, with labels for states and subsystems, is fixed by the choice of quantity we average, and diagrams are generated by making all possible Wick pairing of circles. These pairings are represented by dashed lines. The contribution of a diagram is a product of two factors. One factor stems from Eq. (56) and consists of q^{-L} for every dashed line. The other factor arises from sums over the Hilbert space at each site. To evaluate this factor we form closed loops in the diagram consisting alternately of dashed lines and full lines traversed in the direction of the arrows. These full lines must carry the label that corresponds to a subsystem in which the site lies. For each site we have a factor of q for every such closed loop.

For example, the evaluation of $[M_X(abba)]_G$ is shown in Fig. 9. It yields

$$[M_X(abba)]_G = q^{-2L} q^{2L(X)} q^{L(\bar{X})} = q^{L(X)-L}. \quad (57)$$

B. Relating $F_4(X, Y, \theta)$ and $G_4(X, Y, \theta)$

Our first objective is to obtain a relationship between $F_4(X, Y, \theta)$ and $G_4(X, Y, \theta)$ by expanding $e^{-S_4(a,b,c,d)}$ in a power series, averaging with respect to the Gaussian distribution, and then resumming the terms at each order in the expansion that are leading for q^L and $q^{L(X)}$ large. By this means we will recover Eq. (35).

Using the notation of Eq. (32) we require the connected contributions to $[[M]_X[M]_Y[M^T G_4 M^*]^n]_G$ at each order n in perturbation theory that are leading for large system and subsystem sizes. These come from the Wick contractions that generate the largest number of loops in a decomposition of the type illustrated in Fig. 9. These contractions are illustrated for $n = 0, 1$, and 2 in Fig. 10, establishing an obvious pattern for general n . Retaining only these terms we have

$$\begin{aligned} [[M]_X[M^*]_Y]_{av} &\approx \sum_{n=0}^{\infty} (-1)^n [(\mathbf{G}_4^{(0)})^{-1} (\mathbf{G}_4 (\mathbf{G}_4^{(0)})^{-1})^n]_{X,Y} \\ &= [(\mathbf{G}_4^{(0)})^{-1} [\mathbb{1} + \mathbf{G}_4 (\mathbf{G}_4^{(0)})^{-1}]^{-1}]_{X,Y} \\ &= [(\mathbf{G}_4 + \mathbf{G}_4^{(0)})^{-1}]_{X,Y}, \end{aligned} \quad (58)$$

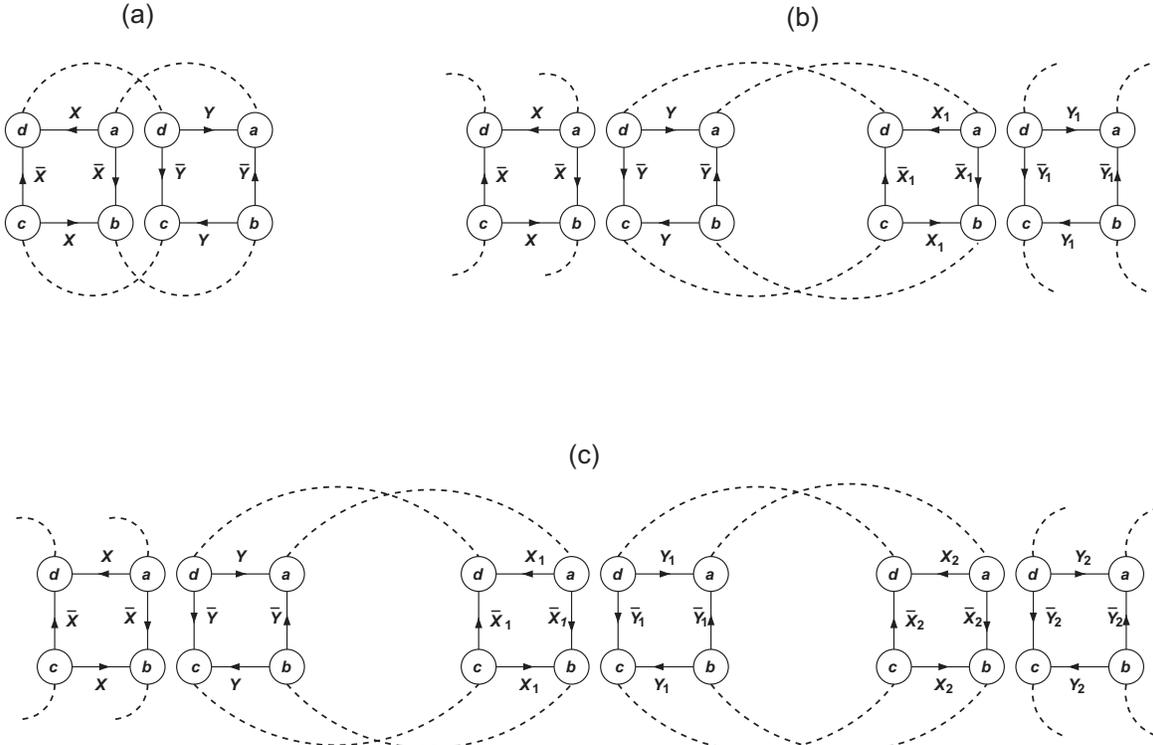


FIG. 10. Leading connected contributions to $[[M]_X[M]_Y[M^T G_4 M^*]^n]_G$ for (a) $n = 0$, (b) $n = 1$, and (c) $n = 2$. The pair of squares in (a) and the leftmost pairs of squares in (b) and (c) represent $[M]_X[M]_Y$; other pairs of squares represent $M^T G_4 M^*$. The dashed lines in (b) and (c) that leave the diagrams on the left are joined to the dashed lines that leave on the right.

where $\mathbf{G}_4^{(0)}$ is as defined following Eq. (33) by the relation $[(\mathbf{G}_4^{(0)})^{-1}]_{X,Y} = [[\mathbf{M}]_X[\mathbf{M}^*]_Y]_G = q^{L(X,Y)-4L}$. For averaging over a Gaussian ensemble, the last relation is exact. [Note that for these terms the factor of $(n!)^{-1}$ arising from the power series expansion of $e^{-S_4(a,b,c,d)}$ is canceled by a combinatorial factor arising in the pairing of terms in Fig. 10.] Comparing Eqs. (35) and (58) we see that this diagrammatic resummation provides an alternative derivation of the main results of Sec. III B 1.

C. Cross-correlations

In our MC studies of the eigenstate JDF [Eq. (29)] we have made two simplifications: one is to omit $G_4(X, Y, \theta)$ when studying $F_2(X, \theta)$, and the other is to omit $G_2(X, \theta)$ when studying $F_4(X, Y, \theta)$. The magnitudes of the resulting errors can be assessed using perturbation theory, as we now discuss.

We begin by considering the effect of $G_4(X, Y, \theta)$ on the value of $F_2(X, \theta)$, or equivalently on the value of $[M_X(abba)]_{\text{av}}$, as follows. To first order in perturbation theory in $G_4(X, Y, \theta)$ we have

$$\begin{aligned} [M_X(abba)]_{\text{av}} &= [M_X(abba)]_G \\ &- \sum_{X'Y'} G_4(X', Y', \theta) [M_X(abba)M_{X'}(abcd)M_{Y'}^*(abcd)]_{G,c} \\ &+ \mathcal{O}([G_4(X', Y', \theta)]^2), \end{aligned} \quad (59)$$

where $[\cdot\cdot\cdot]_{G,c}$ denotes the connected average, defined by

$$\begin{aligned} &[M_X(abba)M_{X'}(abcd)M_{Y'}^*(abcd)]_{G,c} \\ &= [M_X(abba)M_{X'}(abcd)M_{Y'}^*(abcd)]_G \\ &- [M_X(abba)]_G[M_{X'}(abcd)M_{Y'}^*(abcd)]_G. \end{aligned} \quad (60)$$

The diagrams that contribute to this connected average are shown in Fig. 11. Evaluating these diagrams for the representative case $X = X' = Y'$, we obtain

$$\begin{aligned} &[M_X(abba)M_X(abcd)M_X^*(abcd)]_{G,c} \\ &= q^{-3L}(q^{-L(X)} + 2q^{-L(\bar{X})}). \end{aligned} \quad (61)$$

We compare the zeroth-order term, which is

$$[M_X(abba)]_G = q^{L(X)-L}, \quad (62)$$

from Eq. (57), with the first-order term, using Eq. (61) and $G_4(X, X, \theta) \sim q^{2L}$ from Eq. (33), which gives

$$\begin{aligned} &\sum_{X'Y'} G_4(X', Y', \theta) [M_X(abba)M_{X'}(abcd)M_{Y'}^*(abcd)]_{G,c} \\ &\sim q^{-L}(q^{-L(X)} + 2q^{-L(\bar{X})}). \end{aligned} \quad (63)$$

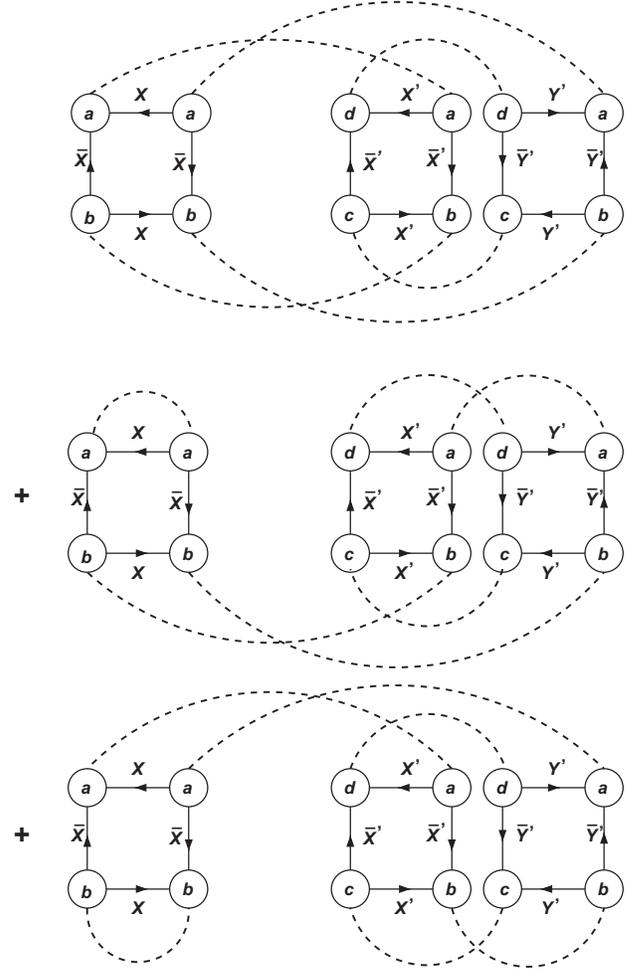


FIG. 11. Contributions to the connected average defined in Eq. (60).

Hence we see that the effect of $G_4(X', Y', \theta)$ on $F_2(X, \theta)$ is small provided that $q^{L(X)}$ and $q^{L(\bar{X})}$ are large.

We have not systematically investigated higher-order terms in Eq. (59), but we expect them individually to be small: Note that although $G_4(X', Y', \theta) \sim q^{2L}$, each such contribution is accompanied by a factor of q^{-4L} from extra dashed lines, as well as diagram-dependent factors from sums over the Hilbert space at each site.

A similar study of the effect of $G_2(X, \theta)$ at first order in perturbation theory on $F_4(X, Y, \theta)$, or equivalently on $[M_{X'}(abcd)M_{Y'}^*(abcd)]_{\text{av}}$, reaches a different conclusion: The influence is not small in q^L or $q^{L(X)}$, but only in powers of $G_2(X, \theta)$. This is not unexpected: The functional form for a correlator involving four eigenstates and depending on three quasienergy differences has been discussed previously [21] and in Sec. II C; it involves factors of both $F_2(X, \theta)$ and $F_4(X, Y, \theta)$ [see Eq. (24)]. We do not pursue this further since we have chosen here to study models in which $F_2(X, \theta)$ lies close to its value for Haar-distributed pairs of eigenstates, and $G_2(X, \theta)$ is therefore small.

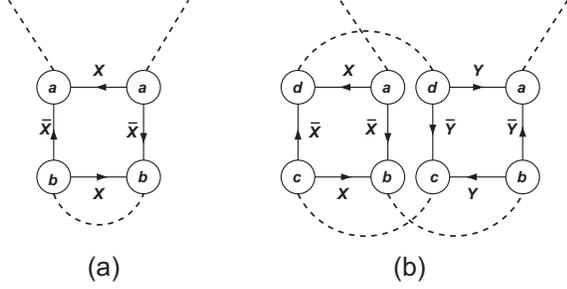


FIG. 12. Contributions to the self-energy Σ at first order. (a) From $G_2(X, \theta)$ and (b) from $G_4(X, Y, \theta)$.

D. Renormalization of propagators and vertices

An alternative perspective is provided by considering the perturbative renormalization of the propagators and vertices appearing in the JDF. We begin with the former: The bare propagator in the theory, represented using dashed lines in the figures, is generated by Eqs. (55) and (56) and carries a factor of q^{-L} . More generally, it acquires a self-energy Σ , and the factor becomes $(q^L + \Sigma)^{-1}$. Our aim is to evaluate Σ at leading order in $G_2(X, \theta)$ and $G_4(X, Y, \theta)$ and compare it to the bare inverse propagator q^L : If it is small under this comparison, then the effects of the vertices on the normalization of eigenstates can be neglected. The diagrams contributing to Σ at this order are shown in Fig. 12. They are diagonal matrices in the site basis, and the magnitudes of their largest entries are, respectively,

$$\Sigma_{kk}^{(i)} = q^{-L} \sum_X G_2(X, \theta) q^{L(X)} \sim q^{2L(X)} \quad (64)$$

[where we have used Eqs. (43) and (44) to estimate $G_2(X, \theta) \sim q^{L+L(X)}$] and

$$\Sigma_{kk}^{(ii)} = q^{-4L} \sum_{X,Y} q^{L(X,Y)} G_4(X, Y, \theta) \sim \mathcal{O}(1) \quad (65)$$

[where we have used an estimate of $G_4(X, Y, \theta)$ given above Eq. (63)]. From this we see that the renormalization of the propagator is indeed small if q^L is large and if $q^{2L(X)} \ll q^L$. In the opposite regime ($q^{2L(X)} > q^L$) we believe that $G_2(X, \theta)$ is small since $F_2(X, \theta)$ approaches the value it takes for Haar-distributed eigenstates, and that $\Sigma_{kk}^{(i)} \ll q^L$ notwithstanding the estimate of Eq. (64).

A similar discussion can be developed of renormalized vertices generated by combining contributions from $S_2(a, b)$ and $S_4(a, b, c, d)$ in Eq. (29) and forming sufficient internal contractions to generate a new effective contribution to either $S_2(a, b)$ or $S_4(a, b, c, d)$. For example, one (out of three possible terms) contributing in this way to $S_2(a, b)$ at first order in both $G_2(X, \theta)$ and $G_4(X, Y, \theta)$ is shown in Fig. 13. We do not pursue this further because this is exactly the same phenomenon as has been discussed from a different perspective in Sec. IV C.

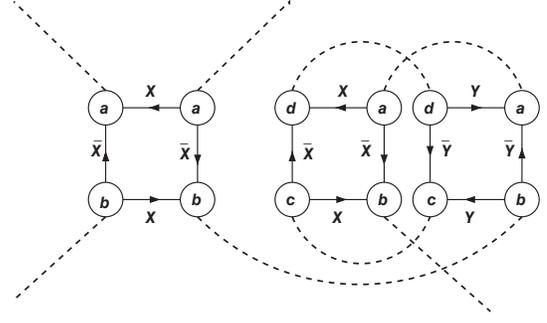


FIG. 13. One of three contributions to the renormalization at first order in both $G_2(X, \theta)$ and $G_4(X, Y, \theta)$ of the vertex that appears in $S_2(a, b)$.

Note that it is not necessary to consider renormalization of $S_2(a, b)$ in powers of $G_2(X, \theta)$ alone, or of $S_4(a, b, c, d)$ in powers of $G_4(X, Y, \theta)$ alone, since these effects (which are not generally small) are covered in full by the approaches described in Secs. IV B and III B.

V. RESULTS FOR FURTHER MODELS

In this section we provide numerical results for brickwork models additional to the one treated in Sec. II. In Sec. VA we omit the cutoff in the operator purity of gates introduced in Sec. III A and present results for gates drawn from a Haar distribution. In Sec. VB we give results with local Hilbert space dimension $q = 3$ and Haar-distributed gates.

A. Results for $q = 2$ with Haar gates

As discussed in Sec. III, the determination of the Lagrange multipliers $G_4(X, Y, \theta)$ in the brickwork model with Haar-distributed gates is complicated by the presence of weak links. Specifically, we find that the simplifying assumption used in Sec. III B 1, namely that the probability distribution of $M_X(abcd)$ is approximately Gaussian, does not hold for Haar-distributed gates. Instead, this distribution exhibits long tails, as we demonstrate in Appendix A 2.

To determine $G_4(X, Y, \theta)$ under these circumstances, we use the iterative fitting procedure introduced in Sec. III B 3, which is not predicated on a particular form for the probability distribution of $M_X(abcd)$. We take all subsystems X and Y that can be obtained using a single cut from a system with open boundary conditions. In contrast to Sec. II, where the requirement of an approximately Gaussian distribution for $M_X(abcd)$ led to the restriction $2 < L(X) < L - 2$, here we include all subsystem sizes $1 \leq L(X) \leq L - 1$. A disadvantage of the iterative fitting procedure is that it requires multiple MC evaluations of $F_4^{\text{MC}}(X, Y, \theta)$, which is slow if the total number of degrees of freedom q^L involved is large; this restricts us to $L = 8$ with $q = 2$. In this case, estimates $F_4^{\text{MC}}(X, Y, \theta)$ are obtained using between 60 and 300 iterations and a step

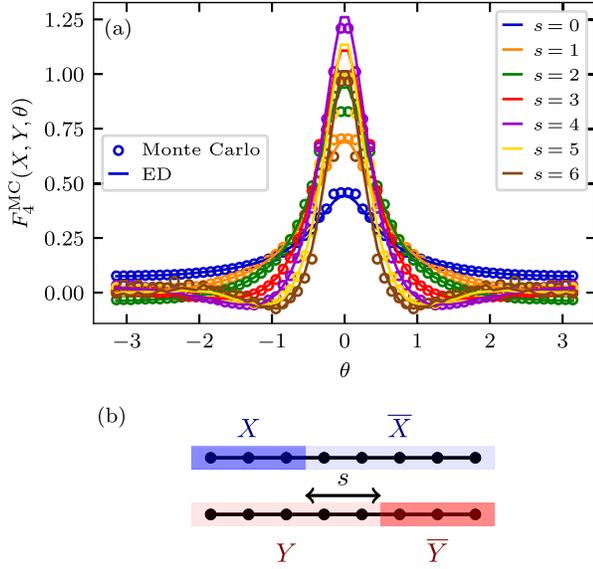


FIG. 14. $F_4^{\text{MC}}(X, Y, \theta)$ (open circles from MC sampling) and $F_4(X, Y, \theta)$ (lines from ED) versus θ for various s , for a brickwork model with Haar-distributed gates, $q = 2$ and $L = 8$.

size $\alpha = 0.2$. The MC results shown in Fig. 14 display very good agreement with ED data.

B. Results for $q=3$ brickwork model

In our study of the brickwork model with $q = 3$ and Haar-distributed gates, we use system size $L = 8$ so that ED calculations are straightforward. The effect of weak links becomes less pronounced with increasing bond dimension q , and as a consequence the single-shot approach to obtain $G_4(X, Y, \theta)$ of Sec. III B 1 becomes more accurate. Conversely, for a given system size the iterative method of Sec. III B 3 is more difficult to apply with increasing q , because a large number of samples is required to obtain accurate estimates for $F_4^{\text{MC}}(X, Y, \theta)$ when the Hilbert space dimension q^L of the system is large (see the discussion of

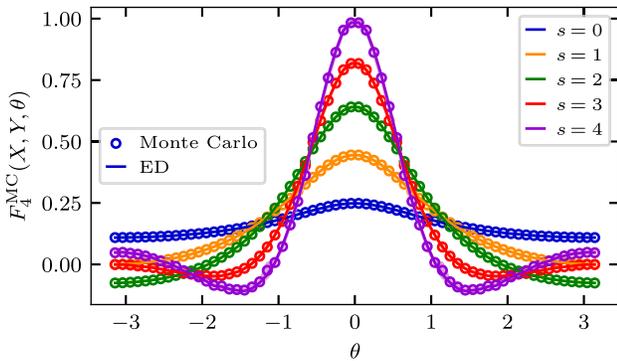


FIG. 15. $F_4^{\text{MC}}(X, Y, \theta)$ (open circles from MC sampling) and $F_4(X, Y, \theta)$ (lines from ED) versus θ for various s , for a brickwork model with Haar-distributed gates, $q = 3$ and $L = 8$.

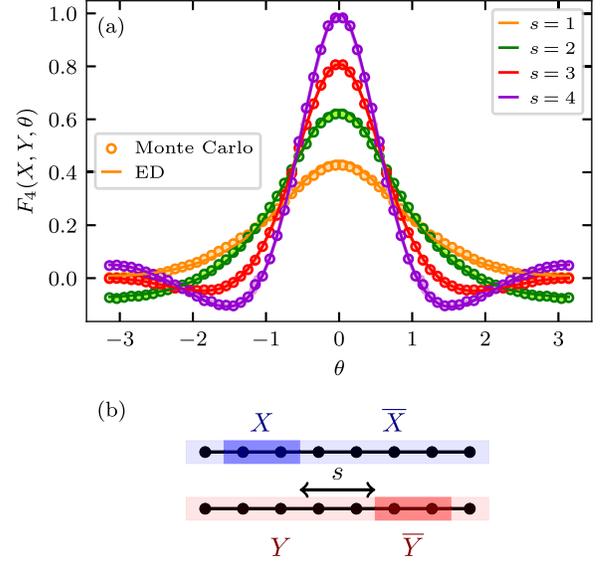


FIG. 16. Test of JDF fitted to behavior in the geometries of Fig. 15(b) but applied to geometries of (b), for the brickwork model with $q = 3$ and $L = 8$. (a) Comparison of data from MC sampling (open circles) and ED (solid lines). (b) Partition used for (a), in which the eight-site system is divided by two spatial cuts into a two-site subsystem X and its complement \bar{X} , or a two-site subsystem \bar{Y} and its complement Y .

Sec. II D and Ref. [40]). We therefore determine Lagrange multipliers for this model using the single-shot approach, taking subsystems X obtained from a system with open boundary conditions by making a single cut, and with $L(X) \geq 2$, $L(\bar{X}) \geq 2$. The MC results shown in Fig. 15 display excellent agreement with ED data.

Furthermore, we test how well the JDF fitted to the geometries of Fig. 14(b) can reproduce the behavior of the OTOC in the geometries of Fig. 16(b). The results shown in Fig. 16(a) display excellent agreement between ED and MC data.

C. Results for $q=2$ brickwork model with periodic boundary conditions

Finally, we show results for the brickwork model for $q = 2$, $L = 12$ using periodic boundary conditions. This leads to a second possible direction for the growth of correlations and two separations between the partitions denoted by s and s' , with the most important separation being the minimum of the two. To simplify the presentation we set $s = s'$.

The results are shown in Fig. 17. In this case, we have used the exact Haar average $G_4^{(0)}$ instead of the leading approximation given in Eq. (31). The solid triangles indicate $F_4^{\text{MC}}(X, Y, \theta)$ calculated Monte Carlo sampling from a JDF that includes Lagrange multipliers $G_4(X, Y, \theta)$ for all connected partitions X, Y generated by two cuts with $L(X), L(Y) > 1$. The results in this case (open circles)

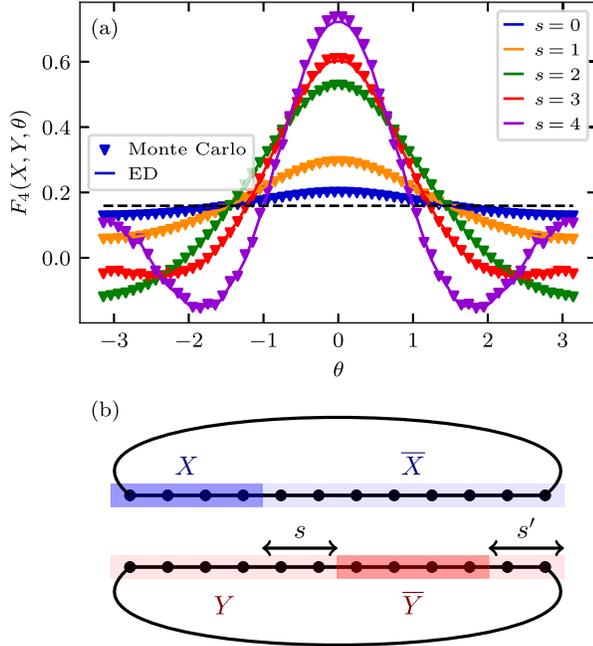


FIG. 17. $F_4^{\text{MC}}(X, Y, \theta)$ (open circles and solid triangles from MC sampling) and $F_4(X, Y, \theta)$ (lines from ED) versus θ for various s , for a brickwork model with Haar-distributed gates and periodic boundary conditions, $q = 2$ and $L = 12$ for the geometries displayed in (b). Partitions in the Lagrange multipliers are generated using two cuts (solid triangles). The results generated from Monte Carlo sampling show very good agreement with the data obtained from exact diagonalization. (b) Partition used for (a). Because of the periodic boundary conditions, two separations s and s' between the partitions exist. We choose $s = s'$.

show very good agreement with $F_4(X, Y, \theta)$ calculated using ED.

VI. SUMMARY AND OUTLOOK

In this work we have analyzed the interplay between the statistical properties of the time-evolution operator for chaotic many-body quantum systems and the quantum information dynamics that these systems display. The eigenstate thermalization hypothesis provides an accurate description of some key aspects, in terms of the probability distribution of matrix elements of observables between eigenstates of the time-evolution operator. In its original form, however, it does not capture the consequences of a finite speed for quantum information spreading. To remedy this, we advocate a change of viewpoint. In place of the matrix elements considered within ETH, we construct correlators from eigenstates. We show that the simplest such correlator capturing signatures of quantum information dynamics at long times and large distances is unique and involves a set of four eigenstates. We also propose an ansatz for the joint probability distribution function of small numbers of eigenstates, in which the values of correlators are

controlled by Lagrange multipliers. We support these general ideas using numerical studies of Floquet quantum circuits, showing firstly that correlators have the expected features, and secondly that it is possible to choose values for the Lagrange multipliers so that our ansatz reproduces these features accurately. We believe our approach is complementary to recent work [36,37] that generalizes ETH using the language of free probability theory.

An advantage of viewing this problem in terms of correlations between eigenstates rather than matrix elements is that even the simplest assumption, of a Haar distribution for eigenstates, yields correlations with the correct order of magnitude. In this instance, however, the correlations are independent of eigenvalues. By contrast, a microscopic model for local quantum dynamics yields a characteristic dependence of correlators on differences in quasienergies, as we have summarized in Fig. 3. We have shown that our ansatz for the joint distribution of eigenstates, together with a suitable choice for the Lagrange multipliers, captures this dependence on eigenvalue differences.

The quantity that characterizes correlations between sets of four eigenstates, denoted $F_4(X, Y, \theta)$ in this paper, is a function not only of the eigenphase difference θ , but also of subsystem choices X and Y . This constitutes a large set of possibilities if there are no restrictions on the way the subsystems are selected. We have focused on the simple set of subsystems that can be obtained from an overall system with open (rather than periodic) boundary conditions by making a single spatial cut, choosing Lagrange multipliers in our ansatz for the JDF so that $F_4(X, Y, \theta)$ is reproduced accurately for these subsystems. In addition, we have shown that the eigenstate correlations imposed in this way are sufficient to reproduce to a good approximation the correlators for some other choices of X and Y . In particular, we have demonstrated that the behavior of the OTOC for operators supported on a few sites of the system (requiring choices of X and Y each involving two cuts) is well approximated by our JDF.

Several obvious directions remain open for future work. Perhaps most importantly, while our discussion in this paper has been restricted to Floquet systems, it would be desirable to extend the approach to systems with a time-independent Hamiltonian. In our context, the significance of such an extension is that diagonal matrix elements of local observables in the basis of Hamiltonian eigenstates are generically functions of energy, a feature absent from Floquet systems. Indeed, ETH is formulated in part to describe this energy dependence. A natural modification of our JDF to impose such an energy dependence is to supplement Eqs. (27) and (29) with an extra factor $e^{-\sum_k S_1(a_k)}$ chosen to bias selected eigenstates toward a prespecified energy shell. An obvious choice is to take

$S_1(a) = \beta \langle a | H | a \rangle$, where H is the Hamiltonian and β is a Lagrange multiplier.

It is worth emphasizing that the correlations induced respectively by $S_1(a)$, $S_2(a, b)$, and $S_4(a, b, c, d)$ are associated with widely separated energy scales. Taking the characteristic strength of local interactions as the unit of energy, $S_1(a)$ selects vectors from a broad energy window, which has a width that increases with system size. In turn, $S_2(a, b)$ generates correlations between pairs of vectors within an energy window with width of order unity. Finally, $S_4(a, b, c, d)$ generates correlations between groups of four vectors with energy or quasienergy differences lying in a narrow energy window, whose width decreases indefinitely with increasing separation between the subsystems X and Y in $F_4(X, Y, \theta)$.

A further direction for future work is to examine correlations between sets of n eigenstates with $n > 4$. This opens up many possibilities, since generalizations of the correlators $F_2(X, \theta)$ and $F_4(X, Y, \theta)$ may involve multiple subsystems in their definition. Restricting to a pair of subsystems X and Y , these higher-order correlators are required, for example, to describe the higher-order Rényi entropies of the operator entanglement for the time-evolution operator. Beyond this, one can ask whether there are physical phenomena exposed only by higher-order correlators.

ACKNOWLEDGMENTS

This work was supported by the Deutsche Forschungsgemeinschaft through the cluster of excellence ML4Q (EXC2004, Project No. 390534769) and by the UK Engineering and Physical Sciences Research Council through Grants No. EP/N01930X/1 and No. EP/S020527/1. We also acknowledge support from the QuantERA II Programme, which has received funding from the European Union's Horizon 2020 research innovation programme (GA 101017733), and from the DFG through the project DQUANT (Project No. 499347025), and from the National Science Foundation under Grant No. NSF PHY-1748958. We thank S. Parameswaran for useful comments.

APPENDIX A: DISTRIBUTION OF $M_X(abcd)$

In this appendix we examine the probability distribution of $M_X(abcd)$, providing further details beyond the information given in Fig. 7. This information is important because one of the two methods that we use for determining the Lagrange multipliers $G_4(X, Y, \theta)$ relies on this distribution having a Gaussian form. In Appendix A 1 we show how the form of the distribution varies with the value of $L(X)$. In Appendix A 2 we investigate the effect on the distribution of the cutoff in the operator entanglement purity of two-qubit gates introduced in Sec. III A.

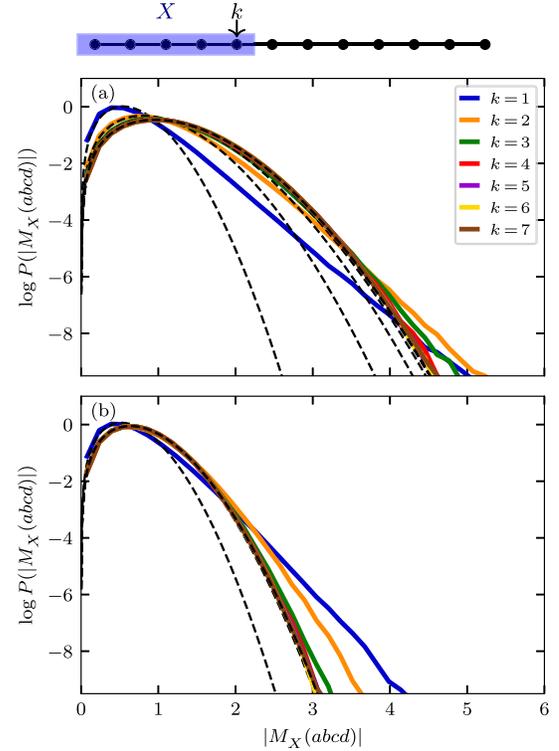


FIG. 18. Probability distribution of $|M_X(abcd)|$ for different subsystems X with $L(X) = k$ and $L = 12$, $q = 2$, and fitted complex Gaussian distributions (dashed black line) as guide for the eye. (a) $\theta = 0.1$ and (b) $\theta = 1.7$. Brickwork model with gates sampled from the distribution defined in Sec. III A and a cutoff $0.3 \times q^4$ for the operator purity. With decreasing partition size $L(X)$, the distribution of $M_X(abcd)$ exhibits non-Gaussian tails.

1. Effect of the subsystem size $L(X)$

The distribution of $M_X(abcd)$ for different partitions X is shown in Figs. 18(a) and 18(b). With increasing partition size $L(X)$, the distribution approaches a Gaussian. This supports our observation that the method we use to determine $G_4(X, Y, \theta)$ becomes more accurate with increasing $L(X)$.

2. Effect of the cutoff in the operator entanglement purity

In Sec. III we presented results for a model (see Sec. III A) that is well adapted to our procedure for finding $G_4(X, Y, \theta)$. This model is designed to have a probability distribution for $M_X(abcd)$ that is close to Gaussian, and has gates drawn from a truncated version of the Haar distribution, with a cutoff on the operator entanglement purity of $c \times q^4$.

In Fig. 19 we examine the effect of the value of this cutoff on the probability distribution of $M_X(abcd)$, considering the range from $c = 0.3$ (the value used in Sec. III) to $c = 1$ (an unrestricted Haar distribution).

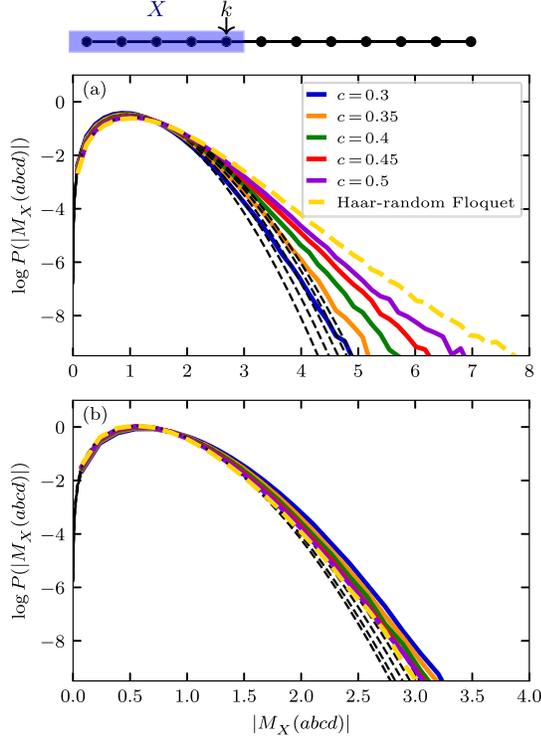


FIG. 19. Probability distribution of $|M_X(abcd)|$ at partition $k = 2$ and $L = 12$, $q = 2$, and a complex Gaussian distribution (dashed black line) as guide for the eye. (a) $\theta = 0.1$ and (b) $\theta = 1.7$. Brickwork model with unitary gates sampled from the distribution defined in Sec. III A and values of the cutoff c as indicated. With increasing cutoff, the distributions show non-Gaussian tails for small relative phases θ .

The distribution shows non-Gaussian tails for small relative phase θ and a Haar-random Floquet model. The tails are suppressed with decreasing cutoff in the operator entanglement purity. We attribute this effect to weak links i on which the gate w_i is close to the identity (especially in small systems or with small subsystems). The effect of such weakly entangling gates on the dynamics of quantum information was studied in more detail in Ref. [47].

APPENDIX B: ACCURACY OF OUR APPROACH WITH INCREASING SYSTEM SIZE

Here we provide complete information on deviations between $F_4^{\text{MC}}(X, Y, \theta)$ and $F_4(X, Y, \theta)$ with Lagrange multipliers $G_4(X, Y, \theta)$ determined using the single-shot procedure described in Sec. III B 1. We compare the results for system sizes $L = 8, 10, 12$ and all possible partitions defined by one cut that have $L(X), L(\bar{X}) > 2$.

The relative deviations are shown in Fig. 20. In all cases the relative deviation is less than 10%, showing the accuracy of our approach. Furthermore, the accuracy improves with increasing system size. We see the largest deviations for small partition sizes $L(X)$ and $L(Y)$. This is expected, as we discuss in Sec. IV: Correction terms in the perturbation theory and contributions from $G_2(X, \theta)$ are most significant when $L(X)$ and $L(\bar{X})$ are small.

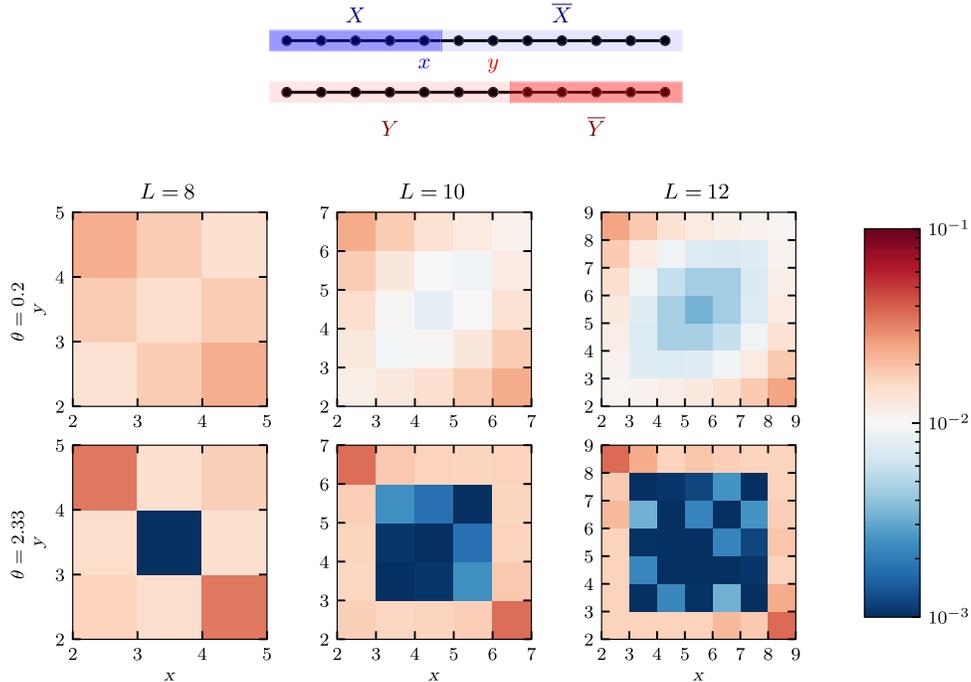


FIG. 20. Relative deviation between $F_4^{\text{MC}}(X, Y, \theta)$ and $F_4(X, Y, \theta)$ for $q = 2$ and $L = 8$ (left-hand column), $L = 10$ (center column), and $L = 12$ (right-hand column), and for $\theta = 0.1$ (top row) and $\theta = 2.03$ (bottom row).

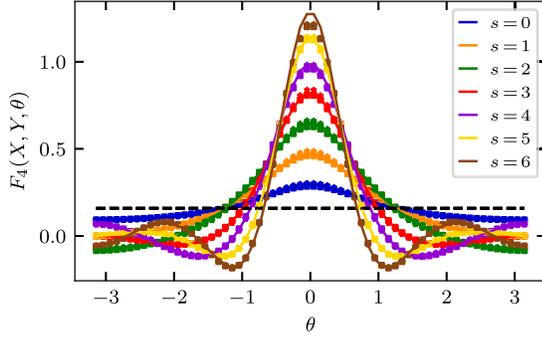


FIG. 21. $F_4(X, Y, \theta)$ (lines from ED) versus θ for various s , and results for single Floquet realizations (squares, stars, and circles) for the circuit model with an additional cutoff in the operator purity, studied in Sec. II. The results of single Floquet realizations are close to the ensemble averages.

APPENDIX C: RESULTS FOR SINGLE REALIZATIONS

In the main text, the eigenstate correlations obtained by exact diagonalization were averaged over multiple realizations of the Floquet circuit to suppress sample-to-sample fluctuations. This is especially important for the $q = 2$ model with Haar gates studied in Sec. VA, where the presence of weak links can dominate the overall dynamics. As we show in this appendix, this is not the case for the Haar model with additional cutoff studied in Sec. II.

In Fig. 21, the averaged result for the eigenstate correlations $F_4(X, Y, \theta)$ together with the results for three different single realizations are shown. The deviations between different realizations of the Floquet circuits are marginal, so we expect that our analysis holds even without ensemble averaging for the circuit analyzed in Sec. II.

APPENDIX D: RESULTS WITH ADDITIONAL LAGRANGE MULTIPLIERS

In the main text, we included Lagrange multipliers $G_4(X, Y, \theta)$ only for partitions X, Y that can be generated by a single cut (apart from a discussion in Sec. VC of connected partitions generated by two cuts in a system with periodic boundary conditions). The purpose of this appendix is to examine the effect of including additional partitions. As an illustration, we include all partitions that can be generated in a system with open boundary conditions using either one or two cuts with $L(X), L(Y), L(\bar{X}), L(\bar{Y}) > 2$.

In Fig. 22 we compare the outcome in this case with results using only partitions generated by means of a single cut. $F_4^{\text{MC}}(X, Y, \theta)$ agrees well in both cases with $F_4(X, Y, \theta)$ obtained from exact diagonalization. However, including more partitions does not reduce the small discrepancies between MC and ED results.

Finally, we investigate how the Lagrange multipliers $G_4(X, Y, \theta)$ are modified after adding additional partitions.

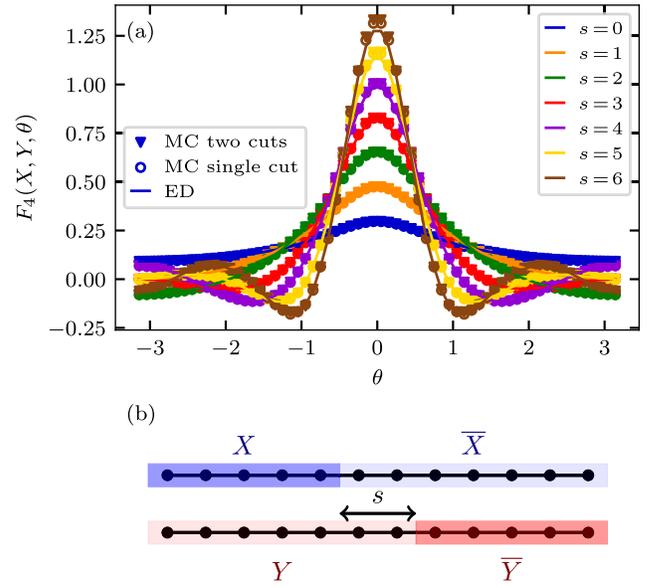


FIG. 22. $F_4^{\text{MC}}(X, Y, \theta)$ including all partitions generated by two cuts with $L(X) > 2$ (solid triangles from MC), $F_4^{\text{MC}}(X, Y, \theta)$ including all partitions generated by a single cut with $L(X) > 2$ (open circles), and $F_4(X, Y, \theta)$ (lines from ED) versus θ for various s . Calculations are for the brickwork circuit model defined in Sec. III A with $L = 12$, $q = 2$ and open boundary conditions.

To present the results it is useful to view $G_4(X, Y, \theta)$ as a matrix, as done in Sec. III B 1, and specify an order for the rows and columns of the matrix. We do this using the index $x = i(L - 5) + (j - 3)$ for a partition with support on the sites $\{i, i + 1, \dots, (i + j) \bmod L\}$. If only partitions generated by a single cut are included, $i = 0$ and $j \geq 3$, leading to a $(L - 5) \times (L - 5)$ matrix. If, on the other hand, partitions generated by two cuts are included, both i and j may be positive, which leads to a $L(L - 5) \times L(L - 5)$ matrix. Restricting to the first $L - 5$ indices allows a direct comparison of the two cases.

In the following, we analyze the componentwise Haar-averaged version of Eq. (32):

$$\mathbf{S}_4(X, Y) = [[\mathbf{M}^T]_X [\mathbf{G}_4]_{X,Y} [\mathbf{M}^T]_Y]_0. \quad (\text{D1})$$

As a first test, we directly compare in Fig. 23 the values of $\mathbf{S}_4(X, Y)$ corresponding to the first $L - 5$ indices. In this case, including more partitions only leads to small modifications of these Lagrange multipliers.

Going further, in Fig. 24 we examine the extra contributions to $\mathbf{S}_4(X, Y)$ that arise after incorporating partitions generated by two cuts. In this figure the Lagrange multipliers related to single-cut generated partitions contribute to the small turquoise square at the lower left-hand corner, and so we are concerned with the remainder of the figure. As is apparent, the values of $\mathbf{S}_4(X, Y)$ arising from most additional partitions are very small in comparison to

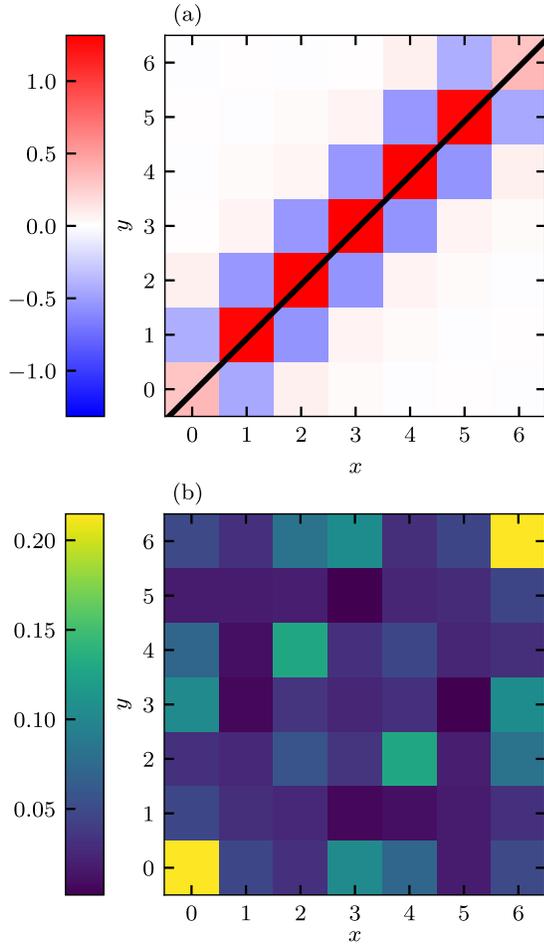


FIG. 23. (a) Comparison of $S_{x,y}^4 = S_{y,x}^4$ including partitions generated by a single cut (above the diagonal) and by two cuts (below the diagonal), for a relative phase $\theta = 0.0$, $L = 12$, $q = 2$. (b) Relative deviation. Adding additional Lagrange multipliers leads to only small modifications.

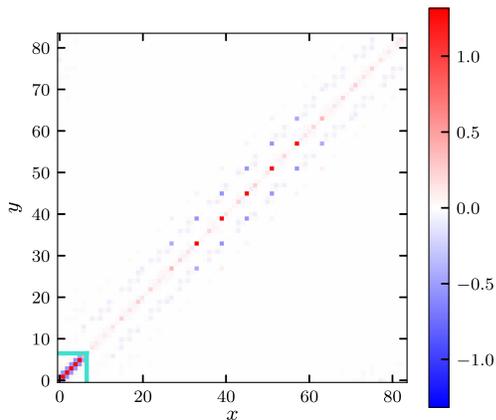


FIG. 24. The complete matrix $S_4(X, Y)$ for the case including partitions generated by two cuts with $L(X) > 2$. The submatrix analyzed in Fig. 23 is indicated as a turquoise square in the lower left-hand corner.

those from single-cut partitions. A few exceptions are also visible: They appear to arise from partitions in which there are two scales for spatial separations, such as (in obvious notation) the partitions $XXXVIIIIXXX$ and $IIIIYYYYIIII$.

-
- [1] A. Peres, *Ergodicity and mixing in quantum theory. I*, *Phys. Rev. A* **30**, 504 (1984).
 - [2] M. Feingold, N. Moiseyev, and A. Peres, *Ergodicity and mixing in quantum theory. II*, *Phys. Rev. A* **30**, 509 (1984).
 - [3] M. Feingold and A. Peres, *Distribution of matrix elements of chaotic systems*, *Phys. Rev. A* **34**, 591 (1986).
 - [4] J. M. Deutsch, *Quantum statistical mechanics in a closed system*, *Phys. Rev. A* **43**, 2046 (1991).
 - [5] M. Srednicki, *Chaos and quantum thermalization*, *Phys. Rev. E* **50**, 888 (1994).
 - [6] M. Rigol, V. Dunjko, and M. Olshanii, *Thermalization and its mechanism for generic isolated quantum systems*, *Nature (London)* **452**, 854 (2008).
 - [7] W. Beugeling, R. Moessner, and M. Haque, *Finite-size scaling of eigenstate thermalization*, *Phys. Rev. E* **89**, 042112 (2014).
 - [8] W. Beugeling, R. Moessner, and M. Haque, *Off-diagonal matrix elements of local operators in many-body quantum systems*, *Phys. Rev. E* **91**, 012144 (2015).
 - [9] R. Mondaini and M. Rigol, *Eigenstate thermalization in the two-dimensional transverse field Ising model. II. Off-diagonal matrix elements of observables*, *Phys. Rev. E* **96**, 012157 (2017).
 - [10] I. M. Khaymovich, M. Haque, and P. A. McClarty, *Eigenstate thermalization, random matrix theory, and behemoths*, *Phys. Rev. Lett.* **122**, 070601 (2019).
 - [11] V. Alba, *Eigenstate thermalization hypothesis and integrability in quantum spin chains*, *Phys. Rev. B* **91**, 155123 (2015).
 - [12] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, *From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics*, *Adv. Phys.* **65**, 239 (2016).
 - [13] F. Borgonovi, F. M. Izrailev, L. F. Santos, and V. G. Zelevinsky, *Quantum chaos and thermalization in isolated systems of interacting particles*, *Phys. Rep.* **626**, 1 (2016).
 - [14] T. Prosen, *Ergodic properties of a generic nonintegrable quantum many-body system in the thermodynamic limit*, *Phys. Rev. E* **60**, 3949 (1999).
 - [15] D. J. Luitz, *Long tail distributions near the many-body localization transition*, *Phys. Rev. B* **93**, 134201 (2016).
 - [16] D. J. Luitz and Y. Bar Lev, *Anomalous thermalization in ergodic systems*, *Phys. Rev. Lett.* **117**, 170404 (2016).
 - [17] M. Serbyn, Z. Papić, and D. A. Abanin, *Thouless energy and multifractality across the many-body localization transition*, *Phys. Rev. B* **96**, 104201 (2017).
 - [18] S. Roy, Y. B. Lev, and D. J. Luitz, *Anomalous thermalization and transport in disordered interacting Floquet systems*, *Phys. Rev. B* **98**, 060201(R) (2018).
 - [19] E. Khatami, G. Pupillo, M. Srednicki, and M. Rigol, *Fluctuation-dissipation theorem in an isolated system of quantum dipolar bosons after a quench*, *Phys. Rev. Lett.* **111**, 050403 (2013).

- [20] L. Foini and J. Kurchan, *Eigenstate thermalization hypothesis and out of time order correlators*, *Phys. Rev. E* **99**, 042139 (2019).
- [21] A. Chan, A. De Luca, and J.T. Chalker, *Eigenstate correlations, thermalization, and the butterfly effect*, *Phys. Rev. Lett.* **122**, 220601 (2019).
- [22] A. I. Larkin and Y. N. Ovchinnikov, *Quasiclassical method in the theory of superconductivity*, *Sov. Phys. JETP* **28**, 1200 (1969), <http://jetp.ras.ru/cgi-bin/e/index/e/28/6/p1200?a=list>.
- [23] S. H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *J. High Energy Phys.* **03** (2014) 067.
- [24] M. Brenes, S. Pappalardi, M. T. Mitchison, J. Goold, and A. Silva, *Out-of-time-order correlations and the fine structure of eigenstate thermalization*, *Phys. Rev. E* **104**, 034120 (2021).
- [25] D. L. Jafferis, D. K. Kolchmeyer, B. Mukhametzhanov, and J. Sonner, *Matrix models for eigenstate thermalization*, *Phys. Rev. X* **13**, 031033 (2023).
- [26] E. H. Lieb and D. W. Robinson, *The finite group velocity of quantum spin systems*, *Commun. Math. Phys.* **28**, 251 (1972).
- [27] P. Calabrese and J. Cardy, *Evolution of entanglement entropy in one-dimensional systems*, *J. Stat. Mech.* (2005) P04010.
- [28] H. Kim and D. A. Huse, *Ballistic spreading of entanglement in a diffusive nonintegrable system*, *Phys. Rev. Lett.* **111**, 127205 (2013).
- [29] T. Zhou and D. J. Luitz, *Operator entanglement entropy of the time evolution operator in chaotic systems*, *Phys. Rev. B* **95**, 094206 (2017).
- [30] 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).
- [31] J. Maldacena, S. H. Shenker, and D. Stanford, *A bound on chaos*, *J. High Energy Phys.* **08** (2016) 106.
- [32] T. Prosen, *Statistical properties of matrix elements in a Hamilton system between integrability and chaos*, *Ann. Phys. (N.Y.)* **235**, 115 (1994).
- [33] A. Dymarsky, *Bound on eigenstate thermalization from transport*, *Phys. Rev. Lett.* **128**, 190601 (2022).
- [34] J. Richter, A. Dymarsky, R. Steinigeweg, and J. Gemmer, *Eigenstate thermalization hypothesis beyond standard indicators: Emergence of random-matrix behavior at small frequencies*, *Phys. Rev. E* **102**, 042127 (2020).
- [35] J. Wang, M. H. Lamann, J. Richter, R. Steinigeweg, A. Dymarsky, and J. Gemmer, *Eigenstate thermalization hypothesis and its deviations from random-matrix theory beyond the thermalization time*, *Phys. Rev. Lett.* **128**, 180601 (2022).
- [36] S. Pappalardi, L. Foini, and J. Kurchan, *Eigenstate thermalization hypothesis and free probability*, *Phys. Rev. Lett.* **129**, 170603 (2022).
- [37] S. Pappalardi, F. Fritzsche, and T. Prosen, *General eigenstate thermalization via free cumulants in quantum lattice systems*, [arXiv:2303.00713](https://arxiv.org/abs/2303.00713).
- [38] Z. D. Shi, S. Vardhan, and H. Liu, *Local dynamics and the structure of chaotic eigenstates*, *Phys. Rev. B* **108**, 224305 (2023).
- [39] Y. Liao and V. Galitski, *Field theory approach to eigenstate thermalization in random quantum circuits*, [arXiv:2210.06480](https://arxiv.org/abs/2210.06480).
- [40] A. Chan, A. De Luca, and J.T. Chalker, *Eigenstate correlations, thermalization, and the butterfly effect*, *Phys. Rev. Lett.* **122**, 220601 (2019).
- [41] T. Prosen and I. Pizorn, *Operator space entanglement entropy in a transverse Ising chain*, *Phys. Rev. A* **76**, 032316 (2007).
- [42] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, *Quantum entanglement growth under random unitary dynamics*, *Phys. Rev. X* **7**, 031016 (2017).
- [43] 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).
- [44] 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).
- [45] S. J. Garratt and J. T. Chalker, *Local pairing of Feynman histories in many-body Floquet models*, *Phys. Rev. X* **11**, 021051 (2021).
- [46] P. W. Brouwer and C. W. J. Beenakker, *Diagrammatic method of integration over the unitary group, with applications to quantum transport in mesoscopic systems*, *J. Math. Phys. (N.Y.)* **37**, 4904 (1996).
- [47] A. Nahum, J. Ruhman, and D. A. Huse, *Dynamics of entanglement and transport in one-dimensional systems with quenched randomness*, *Phys. Rev. B* **98**, 035118 (2018).