Eigenstate Correlations, Thermalization and the Butterfly Effect

Amos Chan, Andrea De Luca and J. T. Chalker
Theoretical Physics, Oxford University, Parks Road, Oxford OX1 3PU, United Kingdom
(Dated: October 26, 2018)

We discuss eigenstate correlations for ergodic, spatially extended many-body quantum systems, in terms of the statistical properties of matrix elements of local observables. While the eigenstate thermalization hypothesis (ETH) is known to give an excellent description of these quantities, the butterfly effect implies structure beyond ETH. We determine the universal form of this structure at long distances and small eigenvalue separations for Floquet systems. We use numerical studies of a Floquet quantum circuit to illustrate both the accuracy of ETH and the existence of our predicted additional correlations.

Introduction. Statistical mechanics is one of the pillars of modern physics. There is however a well-known disjuncture between the fundamental laws of classical and quantum mechanics, with reversible time evolution, and the description of generic systems using ensembles defined only by a small number of conserved quantities. Within classical mechanics, a standard supporting argument is the ergodic hypothesis. A chaotic system evolves over time to explore uniformly all states compatible with the conservation laws, so that the microcanonical ensemble is the only possible equilibrium ensemble with fixed energy. For quantum systems, the notion of ergodicity is more problematic and even the definitions of quantum chaos and integrability are under debate [1, 2].

A fruitful direction is to characterise quantum systems in terms of the spectral properties of their Hamiltonian or evolution operator. Here, random matrix theory (RMT) provides an important paradigm [3]. Quantum chaotic systems can be identified, following the Bohigas-Giannoni-Schmidt conjecture [4], from an RMT eigenvalue distribution [5], while the Berry conjecture [6] proposes that their eigenfunctions can be understood as a random superposition of plane waves. Building on these foundations, the eigenstate thermalization hypothesis (ETH) links the properties of eigenvectors for many-body systems to statistical mechanics and the dynamics of equilibration [7–9]. It constitutes a widely accepted expression of the notion of ergodicity for many-body quantum systems [10, 11].

By design, the ETH omits any spatial structure present in the underlying system. Our aim in this paper is to understand universal features of eigenfunction correlations in ergodic many-body systems that follow from spatial structure and lie outside the ETH.

A single-particle counterpart to the questions we address is provided by studies of eigenfunction correlations in the metallic phase [12] or at the Anderson transition [13–16], in models of disordered conductors without interactions. In that case behaviour is controlled by conservation of probability density. By contrast, for interacting systems it is the dynamics of quantum information that determines long-distance correlations.

A direct characterisation of such dynamics is provided by the butterfly effect [17–19]. In the framework of quantum mechanics, this phenomenon concerns the influence of a perturbation induced by the operator $\hat{Y}$ on later measurements of $\hat{X}$. This is quantified by the value of the commutator $[\hat{X}(t), \hat{Y}(0)]$, and an indication of the strength of the effect is given by

$$C(t) = \frac{1}{2}\langle [\hat{X}(t), \hat{Y}(0)]^\dagger [\hat{X}(t), \hat{Y}(0)] \rangle \qquad (1)$$

where $\langle \ldots \rangle$ denotes the thermal average. If the perburbation Y and the measurement X occur at points which are separated in space, the commutator is initially vanishing. More precisely, for short-range interactions in spatially extended lattice models, the Lieb-Robinson bound [20] ensures that $C(t)$ remains exponentially small for a time which grows linearly in the space separation ℓ between the supports of Y and X. In this language, the butterfly effect is that Y necessarily influences $X(t)$ at large t, and $C(t)$ approaches $\langle \hat{Y}\hat{Y}\rangle\langle \hat{X}\hat{X}\rangle$ regardless of the specific choice of $\hat{Y}$ and $\hat{X}$.

An important recent insight [21] is that non-zero $C(t)$ implies correlations in matrix elements of operators beyond those captured by the ETH. We show here that these correlations acquire a specific universal form for pairs of widely-separated local operators in spatially extended systems. Moreover, since the time-scale for propagation of quantum information is long when ℓ is large, these additional correlations may be arbitrarily sharp in energy, in contrast to those of the ETH.

We focus on Floquet systems because they constitute the simplest class. More generally, conservation laws are reflected in correlations at large distances and long times, and the simplest systems are ones with no conserved densities. To escape conservation of energy density, it is necessary to consider evolution with a time-dependent Hamiltonian, and for there to be a fixed evolution operator this time dependence should be periodic.

A convenient way to construct models with time-dependent evolution operators is by using unitary quantum circuits. These have yielded valuable insights into chaotic quantum dynamics both for systems with an evolution operator that is stochastic in time [22–24] and for Floquet systems [25, 26]. In particular, these studies have

confirmed the existence of two regimes for $C(t)$, distinguished by the sign of $v_B|t|-\ell$, where the *butterfly velocity* v_B characterises the speed at which operators spread in space. We use Floquet random unitary circuits in the following for numerical simulations.

ETH and relaxation. We start by recalling the formulation of the ETH and its connection to the autocorrelation function of an observable. Consider a chaotic many-body Floquet system with local interactions and Hilbert space dimension N. Let $\hat{W}$ be the evolution operator for one period, with eigenstates $|\alpha\rangle$ and eigenphases E_α. According to the ETH, the matrix elements of a local Hermitian operator $\hat{X}$ have the form [27]

$$X_{\alpha\beta} \equiv \langle\alpha|\,\hat{X}\,|\beta\rangle = \bar{X}\delta_{\alpha\beta} + N^{-1/2}h(\Delta_{\alpha\beta})R_{\alpha\beta} \qquad (2)$$

where $\Delta_{\alpha\beta} = E_\alpha - E_\beta$ modulo 2π. In this expression, $\bar{X}$ is the value to which the expectation of $\hat{X}$ relaxes at long times [28–30], $h(\omega)$ is a smooth function of ω, and $R_{\alpha\beta}$ are Gaussian random variables with zero mean and unit variance, which are complex and independent for each pair $\alpha > \beta$, and real and independent for each $\alpha = \beta$. Hermiticity of $\hat{X}$ implies that $h(\omega)$ is real and symmetric, and that $R^*_{\alpha\beta} = R_{\beta\alpha}$. Without loss of generality, we consider traceless operators so that $\bar{X} = 0$. It is then useful to define

$$F(\omega) = N^{-1}\Big[\sum_{\alpha\beta}|X_{\alpha\beta}|^2\delta(\Delta_{\alpha\beta}-\omega)\Big]_{\rm av} \qquad (3)$$

with the eigenstate average $[\ldots]_{\rm av}$ either effected by broadening the delta function, or taken over an ensemble of statistically similar systems. The function $F(\omega) \equiv (2\pi)^{-1}[h(\omega)]^2$ characterizes relaxation of the (integer t) autocorrelation function, since

$$\langle X(t)X\rangle = \int_0^{2\pi} {\rm d}\omega\, e^{\imath\omega t}F(\omega) \equiv \langle X^2\rangle f(t)\,, \qquad (4)$$

where the thermal average appropriate for a chaotic Floquet system is the infinite-temperature one, with $\langle\ldots\rangle \equiv N^{-1}\,{\rm Tr}[\ldots]$. Decay of the autocorrelation function on a microscopic relaxation timescale t_R is encoded in $f(t)$, which satisfies $f(0) = 1$ and $f(t) \to 0$ for $t \gg t_R$.

Generic form of four-point correlators. Let $\hat{X}$ and $\hat{Y}$ be local observables acting near the points x and y, with $\ell \equiv |x-y|$. In analogy with Eq. (3), we introduce the correlator of four matrix elements

$$G(\omega_1,\omega_2,\omega_3) = N^{-1}\Big[\sum_{\alpha\beta\gamma\delta} X_{\alpha\beta}Y_{\beta\gamma}X_{\gamma\delta}Y_{\delta\alpha} \\ \delta(\Delta_{\alpha\beta}-\omega_3)\delta(\Delta_{\beta\gamma}-\omega_2)\delta(\Delta_{\gamma\delta}-\omega_1)\Big]_{\rm av}\,. \qquad (5)$$

If the matrix elements of $\hat{X}$ and $\hat{Y}$ were uncorrelated random variables, then only the term $\alpha = \beta = \gamma = \delta$ would survive the average in Eq. (5) and we would have $G(\omega_1,\omega_2,\omega_3) \sim \mathcal{O}(N^{-2})$. Moreover, before averaging we would expect the sum of N^4 terms to yield an $\mathcal{O}(N^{-1})$ contribution with random phase to $G(\omega_1,\omega_2,\omega_3)$. In fact, the sum rule

$$\int_{[-\pi,\pi]^3}{\rm d}\omega_1{\rm d}\omega_2{\rm d}\omega_3\, G(\omega_1,\omega_2,\omega_3) = \langle\hat{X}^2\hat{Y}^2\rangle \qquad (6)$$

shows that $X_{\alpha\beta}Y_{\beta\gamma}X_{\gamma\delta}Y_{\delta\alpha}$ has a coherent $\mathcal{O}(N^{-3})$ component, as well as $\mathcal{O}(N^{-2})$ fluctuations (see also Ref. [21]). We will show that for large ℓ this coherent component is concentrated on the terms in which the two matrix elements of $\hat{X}$ (and also those of $\hat{Y}$) are between pairs of states with almost opposite eigenphase differences. Specifically, the condition for $X_{\alpha\beta}Y_{\beta\gamma}X_{\gamma\delta}Y_{\delta\alpha}$ to make a large contribution to $G(\omega_1,\omega_2,\omega_3)$ is that $(E_\alpha - E_\beta) \approx -(E_\gamma - E_\delta)$, implying also $(E_\beta - E_\gamma) \approx -(E_\delta - E_\alpha)$.

To understand the form of $G(\omega_1,\omega_2,\omega_3)$ we switch to the time domain and consider $C_\ell(t) \equiv \frac{1}{2}\langle|[\hat{X}(t),\hat{Y}]|^2\rangle$, written as

$$C_\ell(t) = \langle\hat{Y}^2\hat{X}^2(t)\rangle - \langle\hat{X}(t)\hat{Y}\hat{X}(t)\hat{Y}\rangle\,. \qquad (7)$$

The second term on the right-hand side is the *out-of-time order correlator* (OTOC) [19, 31]. $C_\ell(t)$ vanishes for short times t and large separations ℓ, while for large times the OTOC is small and $C_\ell(t)$ approaches $\langle\hat{X}^2\rangle\langle\hat{Y}^2\rangle$. Correspondingly, the OTOC has the form $\langle\hat{X}(t)\hat{Y}\hat{X}(t)\hat{Y}\rangle = \langle\hat{X}^2\rangle\langle\hat{Y}^2\rangle k_\ell(\ell/v_B - |t|)$, where $k_\ell(\tau)$ steps between $k_\ell(\tau) = 1$ for τ large and positive, and $k_\ell(\tau) = 0$ for τ large and negative. The width $\Delta\tau \propto \ell^{1/2}$ of the step is set by diffusive processes [23, 24].

We want to connect this behaviour with the statistical properties of matrix elements. We start by introducing a generalised OTOC

$$\langle X(t_3)Y(t_2)X(t_1)Y\rangle \equiv \langle X^2\rangle\langle Y^2\rangle g(t_1,t_2,t_3)\,. \qquad (8)$$

This reduces to the standard OTOC for $t_1 = t_3 \equiv t$ and $t_2 = 0$. More generally, the form of $g(t_1,t_2,t_3)$ for $\ell \gg v_B t_R$ can be deduced by examining two limits. First, consider $\ell - v_B|t_1 - t_2|$ large and positive. Then: (*i*) $[Y(t_2),X(t_1)]$ is small, and (*ii*) averages that are taken approximately simultaneously at x and y factorise. As a consequence, in this regime $\langle X(t_3)Y(t_2)X(t_1)Y\rangle \approx \langle X(t_3)X(t_1)\rangle\langle Y(t_2)Y\rangle$ and so $g(t_1,t_2,t_3) \approx f(t_3-t_1)f(t_2)$, where for simplicity we take the autocorrelation function $f(t)$ of $\hat{X}$ and $\hat{Y}$ to be the same. In the opposite limit, with $\ell - v_B|t_1 - t_2|$ large and negative, $g(t_1,t_2,t_3)$ is small. We can represent the complete function for $\ell \gg v_B t_R$ by writing

$$g(t_1,t_2,t_3) \approx f(t_3-t_1)f(t_2)k_\ell(\ell/v_B-|t_1|)\,. \qquad (9)$$

Corrections to this form are expected only in a time window $|\ell/v_B-|t_1|| \lesssim \Delta\tau$ that is parametrically narrower at large ℓ than the main scale ℓ/v_B.

Clearly $G(\omega_1, \omega_2, \omega_3)$ and $g(t_1, t_2, t_3)$ and are related by Fourier transform. From this and Eq. (9), we find

$$G(\omega_1, \omega_2, \omega_3) \approx F(\omega_2) F(\omega_3) K_\ell(\omega_1 + \omega_3)\,, \qquad (10)$$

where we have introduced the Fourier transform

$$K_\ell(\omega) = \frac{1}{2\pi} \sum_t k_\ell(\ell/v_B - |t|) e^{-i\omega t}\,. \qquad (11)$$

Our discussion of the form of $k_\ell(t)$ implies that $K_\ell(\omega)$ is maximum at $\omega = 0$ and has a width in frequency of order v_B/ℓ. At large ℓ, since $\Delta\tau \ll \ell/v_B$, we can represent $k_\ell(\tau)$ as a step function and obtain the scaling form

$$\lim_{\ell\to\infty} \frac{1}{\ell} K_\ell(u/\ell) = \frac{\sin(u/v_B)}{2\pi u} \qquad (12)$$

dependant only on the butterfly velocity of the model.

Eqns. (10) and (12) constitute our main theoretical results. They apply to a pair of operators acting at points separated by $\ell \gg v_B t_R$. In this limit they show that non-Gaussian correlations of matrix elements, which are not modelled by the ETH, have universal structure in frequency. This structure appears on a much finer scale (v_b/ℓ) than the one $(1/t_R)$ relevant for the Gaussian correlations that are represented by the ETH.

Model. To test these ideas in a computational study, we consider a one-dimensional L-site Floquet unitary circuit [25] generated by Haar-distributed random unitaries, where the quantum states at each site span a q-dimensional Hilbert space. The circuit is defined by a $q^L \times q^L$ Floquet operator $W = W_2 \cdot W_1$, where $W_1 = U_{1,2} \otimes U_{3,4} \otimes \ldots U_{L-1,L}$ and on an open chain $W_2 = \mathbf{1}_q \otimes U_{2,3} \otimes U_{4,5} \otimes \ldots \mathbf{1}_q$. Here each $U_{i,i+1}$ is a $q^2 \times q^2$ random unitary matrix acting on sites i and $i+1$. We note that the circuit can be defined on a closed chain by the replacement $W_2 \to W_2 \otimes U_{1,L}$. This model has no conserved quantities or discrete symmetries. Moreover, many dynamical quantities can be computed analytically for $q \to \infty$ [25]: in this limit the autocorrelation function decays to zero in a single Floquet period ($t_R \to 0$) and the OTOC exhibits a light-cone with $v_B = 2$ and no broadening of the front. At finite q, the Floquet random circuit provides an ideal setting to investigate the general phenomenology of Eqns. (10) and (12).

Numerical simulations. We focus on $q = 2$ where the space of single-site operators is spanned by Pauli operators $\hat\sigma^j_\alpha$, with $\alpha = x, y, z$ and $j = 1, \ldots, L$. We consider system sizes $L = 4, 6, \ldots, 12$ and perform the full diagonalization of the unitary matrix $\hat W$. We first study the statistics of the matrix elements $X_{\alpha\beta}$ and the function $F(\omega)$ for the operator $\hat X = \sigma^1_z$, using a closed chain in order to minimise boundary effects. The short-time behavior of the autocorrelation function can be computed analytically [25] giving $\langle X(t)X\rangle = 1,\ 0,\ 6.67 \times 10^{-3}$ and 3.87×10^{-3}, for $t = 0,\ 1,\ 2$ and 3. Hence in this model t_R is short and $F(\omega)$ is almost constant. For this reason, we neglect the dependence of $h(\omega)$ on ω and compute the combined probability distributions of all off-diagonal matrix elements, and of all diagonal elements. As shown in Fig. 1, ETH gives an outstandingly accurate description for both quantities. (For similar results in a Hamiltonian system, see e.g. Ref [32, 33])

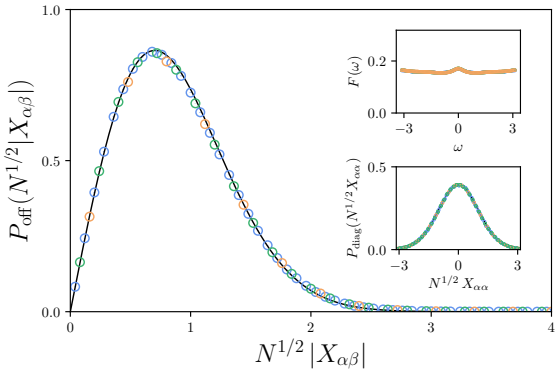

FIG. 1. Comparison of predictions from the ETH with numerical results for Floquet quantum circuits (see text for definitions). Main panel: Scale-collapsed probability distribution $P_{\rm off}(N^{1/2}|X_{\alpha\beta}|)$ of the modulus of off-diagonal elements $X_{\alpha,\beta}$ of a local operator, with $N = 2^L$. Points: data for $L = 6, 8, 10$ in blue, green, yellow respectively. Line: complex Gaussian distribution, as expected from the ETH. Bottom inset: Scale-collapsed probability distribution of the diagonal elements $P_{\rm diag}(N^{1/2}X_{\alpha\alpha})$. Points: data for system sizes as in main panel. Line: real Gaussian distribution, as expected from ETH. Top inset: $F(\omega)$ vs ω for the same system sizes.

Next we turn to the four-point correlators and test the predictions of Eqns. (10) and (12). To maximise the separation $\ell = L-1$ at a given L, we choose two operators acting on sites at opposite ends of an open chain:

$$\hat X = \hat\sigma^1_z\,, \quad \hat Y = \hat\sigma^L_z\,. \qquad (13)$$

An overview of the data for $G(\omega_1, \omega_2, \omega_3)$ is given in Fig 2. It shows the expected narrow maximum near the plane $\omega_1 + \omega_3 = 0$. For a quantitative analysis, we project $G(\omega_1, \omega_2, \omega_3)$ onto two orthogonal lines. First, we have as an identity (taking $\omega_1 + \omega_3 - \omega \bmod 2\pi$)

$$K_\ell(\omega) = \int_{[-\pi,\pi]^3} \mathrm{d}\omega_1 \mathrm{d}\omega_2 \mathrm{d}\omega_3\, \delta(\omega_1 + \omega_3 - \omega) G(\omega_1, \omega_2, \omega_3)\,.$$

Second, we define

$$J(\omega_2) = \int_{[-\pi,\pi]^2} \mathrm{d}\omega_1 \mathrm{d}\omega_3\, G(\omega_1, \omega_2, \omega_3)\,. \qquad (14)$$

From Eq. (10) we expect $J(\omega) \approx F(\omega)$.

Results for both functions are presented in Fig 3, and match excellently the expectations we have described. The data in the main panel show a perfect collapse of the central peak for all accessible system sizes, in agreement

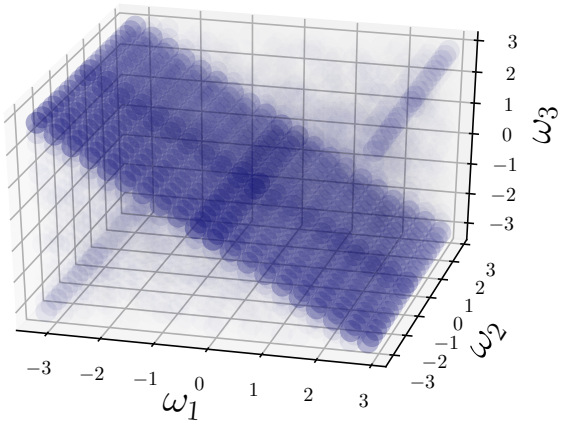

FIG. 2. Histogram of $G(\omega_1, \omega_2, \omega_3)$ as a function of ω_1, ω_2 and ω_3, for $L = 12$ with 20 bins along each axis, averaged over 10^3 realisations. Larger values of $|G(\omega_1, \omega_2, \omega_3)|$ are shown with heavier shading.

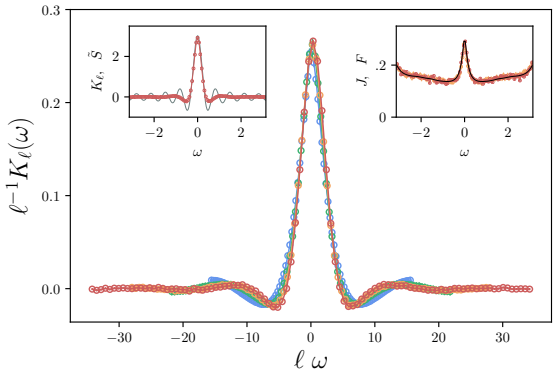

FIG. 3. Main panel: $\ell^{-1}K_\ell(\omega)$ vs $\ell\omega$ for $L = 6, 8, 10, 12$ in blue, green yellow and red respectively. Left inset: $K_\ell(\omega)$ vs ω (red) compared with Fourier transform $\tilde{S}$ of the step function $\Theta(|\ell/v_b - t|)$ (grey) for $L = 12$. Right inset: $J(\omega)$ vs ω for $L = 10, 12$ (yellow and red) compared with $F(\omega)$ vs ω for $\hat{X}$ acting on a site at the end of an open $L = 12$ chain (black).

with the scaling form of Eq. (12). The left inset shows a fit of the central peak in $K_\ell(\omega)$ to the Fourier transform of the step function $\Theta(|\ell/v_B - t|)$. Since t takes only integer values, the fit yields a range of possible values for the butterfly velocity: from $L = 10$, we obtain $v_B \in [1.125, 1.286]$. This range includes the value for a random quantum circuit, $v_B = 2(q^2 - 1)/(q^2 + 1)$ or $v_B = 1.2$ at $q = 2$ [23, 24]. The deviations of the data from the fitting function away from the central peak are due to the diffusive broadening of the step in the OTOC (for these system sizes $\Delta\tau \sim 1$).

The right inset of Fig 3 shows $J(\omega)$ vs ω. In this case data for all system sizes collapse *without* rescaling ω with ℓ, as anticipated from Eq. (10) but in contrast to behaviour for $K_\ell(\omega)$. We expect in addition from Eq. (10) that $J(\omega) = F(\omega)$. In fact, the peak in $J(\omega)$ near $\omega = 0$ is much more pronounced than is shown for $F(\omega)$ in Fig. 1. The discrepancy arises from different choices of boundary conditions: periodic for Fig. 1, open for Fig. 3. Viewed in the time domain, decay of $f(t)$ is slower for an operator at the end of an open chain than with periodic boundary conditions, because its spreading is hindered. The right inset of Fig. 3 also shows that $F(\omega)$ for an operator at the end of an *open* chain has a very similar form to $J(\omega)$.

Finally, and crucially, we show support for our main result in Fig 4. The product form given in Eq. (10) indeed provides a very accurate representation of $G(\omega_1, \omega_2, \omega_3)$.

Discussion. It is natural to ask about the behaviour of other correlators. Within the Floquet model we have described, non-zero correlators must have even numbers of operators acting at each site, since odd powers vanish under the ensemble average. The only two-point correlator is hence $\langle \hat{X}(t)\hat{X}\rangle$. Besides the generalised OTOC, there is a second four-point correlator, with the form $\langle \hat{X}(t_3)\hat{X}(t_2)\hat{Y}(t_1)\hat{Y}\rangle$. This correlator has no long-time structure and therefore no small-frequency features. It

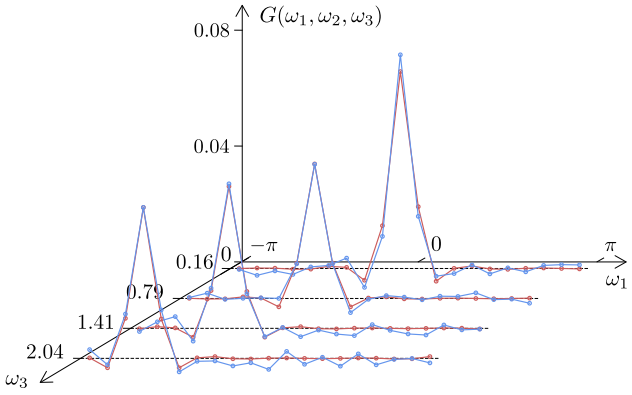

FIG. 4. To test Eq. 10, we plot with offsets $G(\omega_1, \omega_2, \omega_3)$ (blue) and $F(\omega_2)F(\omega_3)K_\ell(\omega_1 + \omega_3)$ (red) vs ω_1 for $L = 12$ at $\omega_2 = -1.73$. Plots for other values of ω_2 and ω_3 show similar agreement.

is captured for large ℓ by the ETH, since it factorises in this limit as $\langle \hat{X}(t_3)\hat{X}(t_2)\rangle\langle \hat{Y}(t_1)\hat{Y}\rangle$. Hence $G(\omega_1, \omega_2, \omega_3)$ is unique in its sharp ω-space structure. This structure implies coherent contributions to $X_{\alpha\beta}Y_{\beta\gamma}X_{\gamma\delta}Y_{\delta\alpha}$ if $(E_\alpha - E_\beta) \approx -(E_\gamma - E_\delta)$.

We expect the phenomenology we have described to be very generic, as it arises from fundamental features of chaotic dynamics in spatially extended systems. In particular, our conclusions will also hold in higher spatial dimensions. In addition, although we have treated the simplified context of Floquet systems, we expect our conclusions to apply with some caveats in the presence of conserved quantities. To be specific, consider a system with a time-independent Hamiltonian $\hat{H}$ and energy as the only conserved density, and examine matrix elements of operators between eigenstates of $\hat{H}$. In this case, the appropriate thermal average at inverse temperature β is

$\langle\cdot\rangle_\beta \equiv \mathcal{N}^{-1}\,\mathrm{Tr}[\,\cdot\, e^{-\beta\hat{H}}]$. A consequence, as for the standard ETH, is that all spectral correlators, including the functions F, G and K_ℓ, as well as v_B, acquire a smooth temperature dependence. Following [34, 35], we expect that for operators $\hat{X}$ and $\hat{Y}$ that do not couple to the conserved charge, our conclusions will hold unchanged. On the other hand, if $\hat{X}$ or $\hat{Y}$ couple to the energy density, both the on-site relaxation in (4) and the OTOC in (7) will present power-law tails. This slow dynamics may affect [36] the separation of scales $\ell \gg v_B t_R$ that we exploited in deriving Eq. (9) and we leave further analysis for future studies.

Acknowledgements. The work was supported in part by the European Union's Horizon 2020 research and innovation programme under the Marie Sklodowska-Curie Grant Agreement No. 794750 (A.D.L.), and in part by EPSRC Grant No. EP/N01930X/1 (A.C. and J.T.C.). We are grateful to S. Gopalakrishnan, A. Nahum and S. Parameswaran for discussions.

[1] F. Haake, *Quantum Signatures of Chaos* (Springer, 2010).
[2] J.-S. Caux and J. Mossel, Journal of Statistical Mechanics: Theory and Experiment **2011**, P02023 (2011).
[3] M. L. Mehta, *Random Matrices* (Academic Press, 2004).
[4] O. Bohigas, M.-J. Giannoni, and C. Schmit, Phys. Rev. Lett. **52**, 1 (1984).
[5] T. Guhr, A. Müller-Groeling, and H. A. Weidenmüller, Physics Reports **299**, 189 (1998).
[6] M. Berry, J. Phys. A **10**, 2083 (1977).
[7] J. M. Deutsch, Phys. Rev. A **43**, 2046 (1991).
[8] M. Srednicki, Phys. Rev. E **50**, 888 (1994).
[9] M. Rigol, V. Dunjko, and M. Olshanii, Nature **452**, 854 (2008).
[10] M. Rigol and M. Srednicki, Phys. Rev. Lett. **108**, 110601 (2012).
[11] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, Advances in Physics **65**, 239 (2016).
[12] D. Thouless, Physics Reports **13**, 93 (1974).
[13] Y. Imry, Y. Gefen, and D. J. Bergman, Phys. Rev. B **26**, 3436 (1982).
[14] J. T. Chalker and G. J. Daniell, Phys. Rev. Lett. **61**, 593 (1988).
[15] J. Chalker, Physica A: Statistical Mechanics and its Applications **167**, 253 (1990).
[16] Y. V. Fyodorov and A. D. Mirlin, Phys. Rev. B **55**, R16001 (1997).
[17] A. Kitaev, in *KITP strings seminar and Entanglement* (2015) http://online.kitp.ucsb.edu/online/entangled15/kitaev/; http://online.kitp.ucsb.edu/online/entangled15/kitaev2/.
[18] Y. Gu, X.-L. Qi, and D. Stanford, Journal of High Energy Physics **2017**, 125 (2017).
[19] J. Maldacena, S. H. Shenker, and D. Stanford, Journal of High Energy Physics **2016**, 106 (2016).
[20] E. H. Lieb and D. W. Robinson, Communications in Mathematical Physics **28**, 251 (1972).
[21] L. Foini and J. Kurchan, (2018), arxiv:1803.10658.
[22] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, Phys. Rev. X **7**, 031016 (2017).
[23] A. Nahum, S. Vijay, and J. Haah, Phys. Rev. X **8**, 021014 (2018).
[24] C. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. Sondhi, Phys. Rev. X **8**, 021013 (2018).
[25] A. Chan, A. De Luca, and J. T. Chalker, (2017), arXiv: 1712.06836.
[26] A. Chan, A. De Luca, and J. T. Chalker, Phys. Rev. Lett. **121**, 060601 (2018).
[27] M. Srednicki, Journal of Physics A: Mathematical and General **32**, 1163 (1999).
[28] G. Biroli, C. Kollath, and A. M. Läuchli, Physical Review Letters **105**, 250401 (2010).
[29] G. Brandino, A. De Luca, R. Konik, and G. Mussardo, Physical Review B, 2012 **85**, 214435 (2011).
[30] H. Kim, T. N. Ikeda, and D. A. Huse, Physical Review E **90**, 052105 (2014).
[31] A. I. Larkin and Y. N. Ovchinnikov, Sov. Phys. JETP **28**, 1200 (1969).
[32] W. Beugeling, R. Moessner, and M. Haque, Phys. Rev. E **91**, 012144 (2015).
[33] D. J. Luitz and Y. B. Lev, Physical review letters **117**, 170404 (2016).
[34] T. Rakovszky, F. Pollmann, and C. von Keyserlingk, Phys. Rev. X **8**, 031058 (2018).
[35] V. Khemani, A. Vishwanath, and D. A. Huse, Phys. Rev. X **8**, 031057 (2018).
[36] A. Dymarsky, (2018), arXiv:1804.08626.