ABSTRACT

Title of Dissertation: CONSTRUCTING AN ERGODIC THEORY OF QUANTUM INFORMATION DYNAMICS

Amit Vikram Anand
Doctor of Philosophy, 2024

Dissertation Directed by: Professor Victor Galitski
Department of Physics

The ergodic theory of classical dynamical systems, originating in Boltzmann's ergodic hypothesis, provides an idealized description of how the flow of information within energy surfaces of a classical phase space justifies the use of equilibrium statistical mechanics. While it is an extremely successful mathematical theory that establishes rigorous foundations for classical chaos and thermalization, its basic assumptions do not directly generalize to quantum mechanics. Consequently, previous approaches to quantum ergodicity have generally been limited to model-specific studies of thermalization, or well-motivated but imprecise general conjectures.

In this Dissertation, we develop a general theoretical framework for understanding how the energy levels of a quantum system drive the flow of quantum information and constrain the applicability of statistical mechanics, guided by two prominent conjectures. The first of these, the Quantum Chaos Conjecture (QCC), aims to characterize which quantum systems may thermalize, by postulating a connection between ergodicity or chaos and the statistical properties of random matrices. The second, the Fast Scrambling Conjecture (FSC), is concerned with how fast a quantum system may thermalize, and posits a maximum speed of thermalization in a sufficiently "local" many-body system.

This Dissertation is divided into three main parts. In the first part, Theory of Quantum Dynamics and the Energy Spectrum, we tackle these conjectures for a general isolated quantum system through results that may be understood as new formulations of the energy-time uncertainty principle. For QCC, we introduce precise quantum dynamical concepts of ergodicity and quantitatively establish their connections to the statistics of energy levels, deriving random matrix statistics as a special consequence of these dynamical notions. We subsequently build on one of these connections to derive an energy-time uncertainty principle that accounts for the full structure of the spectrum, introducing sufficient sensitivity for many-body systems. The resulting quantum speed limit allows us to prove a precise formulation of FSC from the mathematical properties of the energy spectrum. In doing so, we generalize QCC beyond the statistics of random matrices alone, and FSC beyond requirements of locality, establishing precise versions of these statements for the most general quantum mechanical Hamiltonian.

In the second part, Quantum Systems Beyond the Chaotic-Integrable Dichotomy, we demonstrate the need for the aforementioned precise formulations of these conjectures, by showing that looser formulations can be readily violated in "maximally" chaotic or integrable systems that would be most expected to satisfy them. Finally, in the third part, Experimental Probes of Many-Body Quantum Ergodicity, we develop tools to experimentally probe the structure of energy levels associated with ergodic dynamics, and demonstrate a generalization of these probes to open systems in an experiment with trapped ions.

CONSTRUCTING AN ERGODIC THEORY OF QUANTUM INFORMATION DYNAMICS

by

Amit Vikram Anand

Dissertation submitted to the Faculty of the Graduate School of the
University of Maryland, College Park in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
2024

Advisory Committee:
Professor Victor Galitski, Chair/Advisor
Professor Paulo Bedaque
Professor Alexey Gorshkov
Professor Christopher Jarzynski
Professor Nicole Yunger Halpern

Acknowledgments

I thank my advisor, Victor Galitski, for introducing me to the field of quantum chaos, whose importance I was previously unfamiliar with, but which I now consider to be the home of some of physics' most compelling questions. I am particularly grateful for having been given the support and freedom to explore new ways of understanding the central issues of the field, which form a core part of this Dissertation, and for his encouragement of collaborations.

I also thank Paulo Bedaque, Alexey Gorshkov, Christopher Jarzynski, and Nicole Yunger Halpern for agreeing to serve on my Advisory Committee. Christopher Jarzynski's Spring 2020 course on Nonequilibrium Statistical Mechanics was particularly useful for the research presented here, especially for its introductory treatment of ergodicity and mixing.

Collaborations have helped me grow in ways that would've been hard to match in isolation. Yunxiang Liao helped me throughout my first foray into "many-body quantum chaos", and has always been available for detailed discussions on the topic. The significant flavor of quantum information in this Dissertation is in no small part due to a collaboration with Peter Zoller and team: Lata Kh Joshi, Andreas Elben, and Benoít Vermersch, that introduced me to its key role in statistical mechanics; I am particularly grateful to Lata and Andreas for numerous illuminating discussions over these years that helped me acquire a broader perspective on the field. More recently, we have been finding interesting connections to mathematical results through several collaborations with Laura Shou, and to quantum algorithms with Abu Musa Patoary. Further, the opportunity to work with Christopher Monroe's trapped ion lab, with Kate S. Collins, Arinjoy De, and William Morong, has provided a direct connection of the ideas in this Dissertation to experiments and a sense for the real-world complications that

can arise in this context.

My approach to physics has also been shaped by mentors during my undergraduate years. I thank L. Sriramkumar for my first long-term research experience in theoretical physics, my undergraduate honours thesis, and advice on choosing a focus in the early stages of my Ph.D. Working with Chethan Krishnan led me to appreciate the value of taking a formal approach, bordering on mathematics, to clarify fundamental questions involving quantum systems. I am grateful to T. Padmanabhan for emphasizing the conceptual clarity of a Hilbert-space approach over a Lagrangian one (despite the latter's apparent invariance properties) for certain quantum problems, and most of all for getting me interested in theoretical physics in the first place through his popular science writings, long before we worked together.

Other than direct collaborators, I must also acknowledge several valuable discussions on different aspects of physics over the years. Gautam Nambiar has been a constant resource for discussing almost anything in physics since we joined graduate school. Alireza Parhizkar has frequently been open to discussing different aspects of gravity whenever I felt some nostalgia for curved spacetime. I am also grateful to Christopher Baldwin, Richard Barney, Andrey Grankin, Naren Manjunath, Masoud Mohammadi-Arzanagh, Subhayan Sahu, Jeet Shah, Michael Winer, and Christopher White for other technical discussions related to and outside my research. Since my undergrad days, Shoy Ouseph and Pranav T. M. have been part of several conversations, arguments, and debates about physics, respectively keeping me tied to more abstract theory and more concrete experiments than whatever I may be thinking about at any given moment.

I thank Sagar Airen, Sharoon Austin, Saurav Das, Sanket Doshi, Saurabh Kadam, Sohitri Ghosh, Arghya Sadhukhan, Deepak Sathyan, Arpith Shanbhag, and Srivatsa Tata for various interesting conversations, many of them also for helping me settle in at UMD in my first year and some later on. I also thank Amitava Banerjee for letting me add his physical copy of Sinai's *Introduction to Ergodic Theory*,

a major influence on Chap. 2, to my own collection of books.

The pandemic lockdown of 2020 began shortly after I began my research in quantum chaos. In that time, as I was growing more familiar with the field, the scarcity of direct interactions made it difficult to absorb the overall zeitgeist of the field, particularly the kind of perspectives not found in papers. This has been increasingly remedied in the last couple of years, and I am grateful to several people for sharing their direct perspectives on different issues, of which the following constitute a proper subset: Namit Anand, Marin Bukov, Bassam Fayad, Wen Wei Ho, Manas Kulkarni, Jorge Kurchan, David Long, Vadim Oganesyan, Anatoli Polkovnikov, Tomaž Prosen, Pedram Roushan, Dibyendu Roy, Lea Santos, Dries Sels, Rafał Świętek, and Lev Vidmar. I also particularly thank Arul Lakshminarayan, Vaibhav Madhok, Prabha Mandayam, and their research groups, for informative interactions during a recent visit back to IIT Madras.

Finally, I am grateful to my parents Usha and Anand for several things, of which encouraging me to pursue my dreams and facilitating everything I wanted or needed to do to get to this point are two of many.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

BV	Balazs-Voros
COE	Circular Orthogonal Ensemble
CSE	Circular Symplectic Ensemble
CUE	Circular Unitary Ensemble
DFT	Discrete Fourier Transform
ETH	Eigenstate Thermalization Hypothesis
FSC	Fast Scrambling Conjecture
GOE	Gaussian Orthogonal Ensemble
GSE	Gaussian Symplectic Ensemble
GUE	Gaussian Unitary Ensemble
KAM	Kolmogorov-Arnold-Moser
KS	Kolmogorov-Sinai
ML	Margolus-Levitin
MT	Mandelstam-Tamm
OTOC	Out-of-Time-Ordered Correlator
PSFF	Partial Spectral Form Factor
QCC	Quantum Chaos Conjecture
RMT	Random Matrix Theory
SFF	Spectral Form Factor
SYK	Sachdev-Ye-Kitaev
TFD	Thermo-Field Double
TR	Time-Reversal
WD	Wigner-Dyson

List of Publications

This Dissertation is based on the following publications:

[1] Amit Vikram and Victor Galitski. "Dynamical quantum ergodicity from energy level statistics." *Phys. Rev. Res. 5, 033126 (2023)*. *arXiv:2205.05704 [quant-ph]*. (for Chapter 2 and Appendix A).

[2] Amit Vikram and Victor Galitski. "Exact universal bounds on quantum dynamics and fast scrambling." *Phys. Rev. Lett. 132, 040402 (2024)*. *arXiv:2212.14021 [quant-ph]*. (for Chapter 3 and Chapter 6).

[3] Amit Vikram, Laura Shou, and Victor Galitski. "Proof of a universal speed limit on fast scrambling in quantum systems." *arXiv:2404.15403 [quant-ph] (2024)*. (for Chapter 4).

[4] Yunxiang Liao, Amit Vikram, and Victor Galitski. "Many-body level statistics of single-particle quantum chaos." *Phys. Rev. Lett. 125, 250601 (2020)*. *arXiv:2005.08991 [cond-mat.stat-mech]*. (for Chapter 5).

[5] Laura Shou, Amit Vikram, and Victor Galitski. "Spectral anomalies and broken symmetries in maximally chaotic quantum maps." *arXiv:2312.14067 [quant-ph] (2023)*. (for Chapter 7 and Appendix D).

[6] Lata Kh Joshi, Andreas Elben, Amit Vikram, Benoît Vermersch, Victor Galitski, and Peter Zoller. "Probing many-body quantum chaos with quantum simulators." *Phys. Rev. X 12, 011018 (2022)*. *arXiv:2106.15530 [quant-ph]*. (for Chapter 8 and Appendix E).

The following publication, not directly a part of this Dissertation, is closely related to Chapter 7:

[7] Abu Musa Patoary, Amit Vikram, Laura Shou, and Victor Galitski. "Chaotic roots of the modular multiplication dynamical system in Shor's algorithm." *Accepted in Phys. Rev. Res. (22 March 2024)*. *arXiv:2306.16446 [quant-ph] (2023)*.

A version of Appendix C.2 is planned to be included in ongoing work, related to Refs. [2,3]:

[8] Victor Galitski*, Laura Shou*, and Amit Vikram*, "Mathematical note on quantum scrambling bounds", in preparation.
* - alphabetical order.

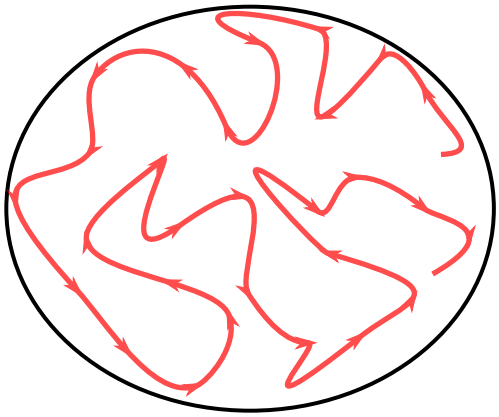
In addition, Chapter 9 and Appendix F are largely based on an experimental collaboration with Christopher Monroe's trapped ion group (Kate S. Collins, Arinjoy De, William Morong) and the authors of [6], implementing the protocol of Ref. [6]. Sec. 9.3 and Fig. 9.1 adapt contributions from Kate S. Collins and Lata Kh Joshi; Sec. 9.1 also incorporates inputs from Lata Kh Joshi on an initial draft.

Chapter 1: Introduction

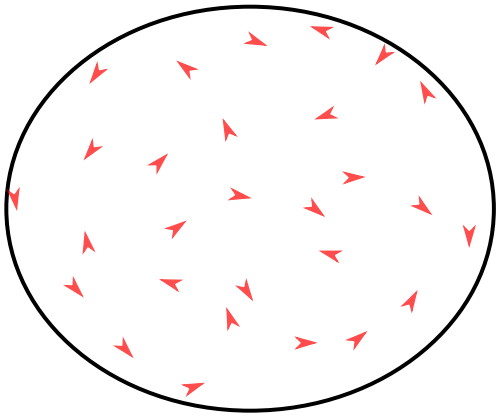
1.1 Historical motivation

Nearly a century after the laws of quantum mechanics were (relatively) set in stone, as a deceptively simple map to linear algebra, we still don't fully understand what these laws are capable of. When quantum systems have too many degrees of freedom, the linear algebra quickly becomes intractable for semiconductor-based classical computers, as well as our presumably more refined neuron-based ones. It would seem that the order of the day is to piece together better ways of viewing quantum dynamics in large systems that will help us improve this situation. This would not be dissimilar to the trajectory of classical mechanics: After Newton wrote down the second law of mechanics $F = ma$ in the late 17th century, the following centuries saw the development of increasingly sophisticated formulations of this same law that could tackle larger and more complex systems (e.g., constraints, complicated geometries, etc.). A number of key ideas include the action principle, Lagrangian and Hamiltonian mechanics, and the idea of invariants under coordinate and canonical transformations [9]. Once these formulations were in place, one could begin to ask and answer questions that would seem beyond reach given just vanilla $F = ma$, for example: What states of the system can we access from a given set of initial conditions? Are there conserved quantities? What does it mean for a system to be integrable (or solvable)? Which systems have no closed form solutions, nor even good numerical approximations?

1.1.1 Boltzmann's classical ergodic hypothesis



(a) An ergodic trajectory in a schematic phase space.



(b) The microcanonical ensemble in phase space (a uniform-looking distribution obtained by retaining just the arrows on the ergodic trajectory, imitating time-averaging over discrete steps).

Figure 1.1: Boltzmann's ergodic hypothesis and the microcanonical ensemble: a schematic. A finite time-average over a trajectory leads to the appearance of a microcanonical distribution for sufficiently coarse-grained observables (pictorially represented by observing just the "arrows" in the trajectory); the distribution becomes increasingly refined for a longer time-average, and approaches the exact microcanonical ensemble for an infinite time-average.

With questions of this nature, we enter the realm of classical ergodic theory [10–12] (see also Chap. 2.2.1 for a more technical discussion). Historically, the field grew out of Boltzmann's attempts to justify the microcanonical ensemble in statistical mechanics, an argument that is worth recapitulating here in a more modern form. Given a region of phase space $\mathscr{P}$ and an initial state $x \in \mathscr{P}$, the trajectory $x(t)$ is ergodic in $\mathscr{P}$ if it equally explores all regions of $\mathscr{P}$ over infinite time. In other words, we want the overall fraction of time spent by $x(t)$ in any region $B \subseteq \mathscr{P}$ to be proportional to the fraction of the phase space volume (measured by μ) of $\mathscr{P}$ occupied by B:

$$\lim_{(t_2-t_1)\to\infty} \frac{1}{t_2-t_1} \int_{t_1}^{t_2} \mathrm{d}t\ \overline{\Theta}(x(t)\in B) = \frac{\mu(B)}{\mu(\mathscr{P})}. \tag{1.1}$$

The right hand side is the microcanonical distribution in $\mathscr{P}$, which is by definition an equal distribution with respect to the phase space volume. We have also introduced the boolean indicator function (also

called the Iverson bracket):

$$\overline{\Theta}(X) = \begin{cases} 1, & \text{if } X \text{ is true,} \\ 0, & \text{if } X \text{ is false.} \end{cases} \tag{1.2}$$

A pictorial depiction of this relation is shown in Fig. 1.1.

Eq. (1.1) is connected to statistical mechanics in two ways [13]. It ensures that time averages of all quantities in an ergodic trajectory agree with those computed with the microcanonical ensemble. Further, if the phase space is partitioned into thermodynamic macrostates B_k with fixed values of "accessible" macroscopic observables, it ensures that the system spends the largest amount of time in the macrostate with the largest entropy, i.e., phase space volume $\mu(B_k)$, which is usually the thermal state $B_k = B_{\text{th}}$. Here, the thermal macrostate B_{th} refers to the region of $\mathscr{P}$ in which all "accessible" macroscopic observables are sufficiently close (in a sense determined by the desired level of accuracy, e.g., to leading order in the number of particles) to their thermal expectation values (obtained in the microcanonical ensemble). We expect that $\mu(B_{\text{th}})/\mu(\mathscr{P}) \to 1$ for a thermodynamically large many-particle system, and therefore, only a vanishing fraction of times is spent outside the thermal macrostate:

$$\lim_{(t_2-t_1)\to\infty} \frac{1}{t_2 - t_1} \int_{t_1}^{t_2} \mathrm{d}t\ \overline{\Theta}(x(t) \notin B_{\text{th}}) \to 0. \tag{1.3}$$

Boltzmann's ergodic hypothesis therefore guarantees the emergence of a time-averaged microcanonical ensemble, as well as thermalization for almost all times if the thermal macrostate has enough entropy. We note that *permanent* thermalization for all times $t \to \infty$ is impossible for an individual ergodic trajectory as it must visit all macrostates B_k infinitely often. More generally, Poincaré recurrences [10–12] ensure that even a non-ergodic trajectory must visit arbitrarily small neighborhoods of every point $x(t)$ on the trajectory infinitely often, making permanent thermalization impossible if *any* portion of the trajectory lies outside the thermal macrostate B_{th}. For an N-particle system, these recurrence time scales are expected to grow exponentially with N and can quickly exceed the ob-

served lifetime of the universe, leading to the appearance of permanent thermalization in practically accessible timescales.

Is ergodicity necessary for statistical mechanics?— All the same, it is important to qualify this idealized picture with a comment on its practicality. It is generally incredibly difficult to prove the ergodicity of almost all trajectories in a given $\mathscr{P}$ in anything more than the simplest systems [11]. Further, one often does not need the full power of the ergodic hypothesis, because the many-particle phase space is large, and one does not need the ideal microcanonical ensemble in every region of $\mathscr{P}$ for the thermalization of the few-particle observables that one can access [14]. This is instead believed to follow from the properties of few-particle observables themselves — they can attain their thermal expectation value in most of the trajectory even if the system is not ergodic. One also does not need the system to spend equal amounts of time throughout the phase space just for it to spend more time in the thermal state than anywhere else.

Thus, ergodicity is frequently too strong a requirement from a practical standpoint [14]. Strictly speaking, the primary role of ergodicity is to determine the dynamical structure of the system in phase space without reference to specific observables (which can, e.g., guarantee thermalization for all observables whether accessible or not), while the properties of certain observables themselves are sufficient for ensuring statistical mechanics holds for accessible few-particle observables. This classical distinction between ergodicity and the statistical mechanics of *accessible* observables is important to stress, because we will find that there is a similar, but more complicated, distinction in our treatment of quantum ergodicity in time-independent unitary systems (see Sec. 2.7).

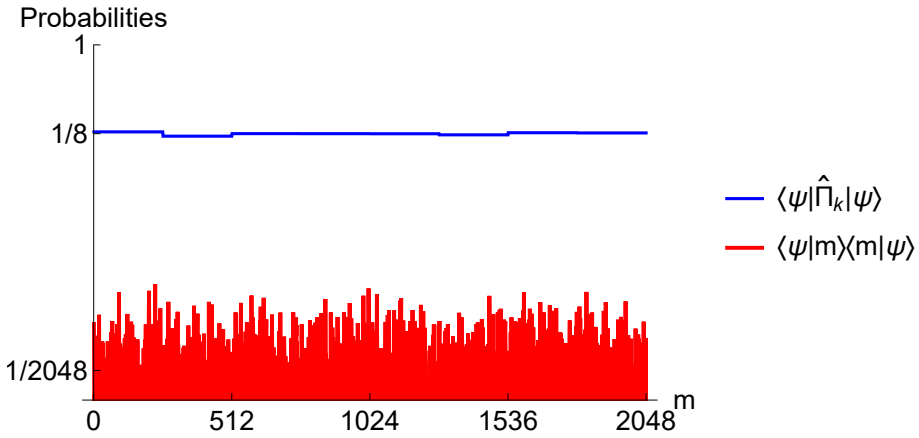


Figure 1.2: An illustration of the typicality of statistical mechanics in quantum systems (log-linear). Generic random states appear to be equally distributed with respect to a set of orthogonal coarse-grained projectors, representing the microcanonical ensemble. The overlaps of the random state $|\psi\rangle$ with pure orthonormal basis states $|m\rangle$, as well as 8 projectors $\hat{\Pi}_k$ of dimension $\mathrm{Tr}[\hat{\Pi}_k] = 256$ (corresponding to sets of 256 consecutive $|m\rangle$s) are shown in red and blue respectively in a Hilbert space of dimension $D = 2048$. This corresponds to Eq. (1.4) if the $|m\rangle$ are taken to be the $|E_n\rangle$. The fluctuations in $\langle\psi|\hat{\Pi}_k|\psi\rangle$ are seen to be much smaller relative to their typical value of $\mathrm{Tr}[\hat{\Pi}_k]/D = 1/8$ [representing the subleading terms $o(\ldots)$], than those in $|\langle m|\psi\rangle|^2$ (fluctuating around $1/D = 2048$).

1.1.2 Typicality and von Neumann's quantum ergodic theorem

Now, we turn to quantum systems with a Hilbert space $\mathscr{H}$ of finite dimension $\dim(\mathscr{H}) = D$. We will work in natural units for Planck's constant and Boltzmann's constant[1], $\hbar = k_B = 1$. If a classical limit exists, the finite Hilbert space dimension corresponds to a finite measure $\mu(\mathscr{P}) < \infty$ of the classical phase space (loosely because pure states occupy a phase space volume $\sim h^f = (2\pi)^f$ for a system with f degrees of freedom by semiclassical arguments [15]; see also App. A.5). We emphasize that this assumption does not preclude a macroscopic or thermodynamic limit $D \to \infty$ (or classically, $f \to \infty$); however, much like the disconnect between Boltzmann's ergodic hypothesis and the statistical mechanics of accessible observables, such a limit can make strict conditions on the entire Hilbert space (or phase space) highly complicated to verify, and impractical for "easily" accessible observables.

[1]Despite brief mentions of black holes and quantum gravity, particularly in Chaps. 3 and 4, the gravitational constant G and speed of light c will not play a direct role in this Dissertation.

The first (known) steps towards understanding ergodicity in quantum mechanics were taken by von Neumann [16]. One such step is the observation that every Hamiltonian quantum system conserves the entire probability distribution $\{|\langle E_n|\psi(t)\rangle|^2\}_n$ of energy in any time evolving state $|\psi(t)\rangle$, identifying disjoint subsets of $\mathscr{H}$ each with a given probability distribution of energy; the system is then ergodic in these fixed-probability-distribution subsets as long as the E_n are rationally incommensurate [see also the discussion after Eq. (2.10)]. However, this condition on the E_n is trivial to ensure with arbitrarily weak perturbations to the system (there are irrational numbers arbitrarily close to any real number E_n), and does not necessarily determine the behavior of observables.

Another key observation of Ref. [16] is that to represent measurements of macroscopic observables, one should probe a quantum system with a complete set of orthogonal projectors $\hat{\Pi}_k$, which can be measured simultaneously but can each "coarse-grain" over several pure states (i.e., project onto subspaces of the Hilbert space, each with dimension ≥ 1). These should be thought of as the quantum analogue of partitioning the phase space $\mathscr{P}$ into cells. We will use such projectors extensively in Chaps. 3 and 4, to represent observables that can only gather information about specific subsystems, such as few-particle observables.

Von Neumann's formulation of the "quantum ergodic theorem" is based on such a choice of projectors $\hat{\Pi}_k$, and is a far-reaching precursor to more modern ideas related to "random matrix theory" and the "eigenstate thermalization hypothesis", which we will encounter in Sec. 1.2.1. The intuition behind this result is easily summarized as follows: Almost all randomly chosen vectors in a Hilbert space look uniformly distributed after coarse-graining, as depicted in Fig. 1.2. In other words, randomness is *sufficient* for microcanonical behavior.

Quantitatively, keeping the energy eigenbasis $|E_n\rangle$ fixed, Ref. [16] samples a complete set of projectors $\hat{\Pi}_k$ from a Haar distribution, which is invariant under unitary transformations $\hat{\Pi}_k \to \hat{U}^\dagger \hat{\Pi}_k \hat{U}$. The expectation value of each $\hat{\Pi}_k$ in any energy eigenstate is proportional to its dimension $\mathrm{Tr}[\hat{\Pi}_k]$ to

leading order (in the limit of infinite projector dimension, corresponding to macroscopic observables):

$$\langle E_n|\hat{\Pi}_k|E_n\rangle = \frac{\mathrm{Tr}[\hat{\Pi}_k]}{\mathrm{Tr}[\hat{\mathbb{1}}]} + o(\ldots), \tag{1.4}$$

with probability 1 in the Haar ensemble (see Refs. [16, 17] for a more precise statement, called "normal typicality"). The right hand side represents the microcanonical distribution; the subleading term $o(\ldots)$ is comparable to $(\sqrt{\mathrm{Tr}[\hat{\Pi}_k]})/\,\mathrm{Tr}[\hat{\mathbb{1}}]$ for the Haar distribution, but its precise behavior is not important for our present discussion (however, it will resurface in Chaps. 3, 4, 6 and more significantly in Chap. 8). This statement already has the basic conceptual nuggets of the (diagonal) eigenstate thermalization hypothesis (ETH) [18–22], which is frequently invoked in present-day studies of statistical mechanics. The primary conceptual difference is that ETH supposes that a specific, *accessible* few-body observable (which may be of physical interest) behaves as if sampled from this Haar distribution in narrow energy windows, and further accounts for a macroscopic energy-dependence of the distribution [21].

Eq. (1.4) has an important dynamical consequence: For any normalized initial state $|\psi\rangle$, as long as the energy levels E_n are non-degenerate, the time-averaged expectation value of a randomly chosen $\hat{\Pi}_k$ is microcanonical with probability 1, to leading order in the projector dimension:

$$\lim_{(t_2-t_1)\to\infty} \frac{1}{t_2-t_1}\int_{t_1}^{t_2} \mathrm{d}t\ \langle\psi(t)|\hat{\Pi}_k|\psi(t)\rangle = \sum_n |\langle E_n|\psi\rangle|^2 \langle E_n|\hat{\Pi}_k|E_n\rangle = \frac{\mathrm{Tr}[\hat{\Pi}_k]}{\mathrm{Tr}[\hat{\mathbb{1}}]} + o(\ldots). \tag{1.5}$$

The resemblance to Eq. (1.1) is evident. In fact, a stronger statement can be made if the differences $E_n - E_m$ are also non-degenerate, using the subleading nature of off-diagonal elements $\langle E_n|\hat{\Pi}_k|E_n\rangle = o(\ldots)$ for $n \neq m$, relative to the microcanonical distribution. For almost all randomly chosen projectors

$\hat{\Pi}_k$, every state $|\psi\rangle$ thermalizes to the microcanonical distribution at almost all times:

$$\langle\psi(t)|\hat{\Pi}_k|\psi(t)\rangle = \frac{\mathrm{Tr}[\hat{\Pi}_k]}{\mathrm{Tr}[\hat{\mathbb{1}}]} + o(\ldots), \quad \forall t \in \mathbb{R} \setminus W_{\mathrm{th}}, \tag{1.6}$$

where the set of times W_{th} in which thermalization does not occur occupies a vanishing fraction of all times,

$$\lim_{(t_2-t_1)\to\infty} \frac{1}{t_2-t_1} \int_{t_1}^{t_2} \mathrm{d}t\ \overline{\Theta}(t \in W_{\mathrm{th}}) = 0. \tag{1.7}$$

In other words, the system spends most of its time in thermal equilibrium for any initial state, provided the observables behave as if they are sampled from the Haar distribution. This is a quantum analogue of Eq. (1.3).

It is also worth noting a related set of statements [23,24] (called "canonical typicality"[2]), which essentially state that given a choice of projectors $\hat{\Pi}_k$, a pure state $|\phi\rangle$ sampled from the Haar distribution of normalized states (invariant under unitary transformations $|\phi\rangle \to \hat{U}|\phi\rangle$) has expectation values corresponding to the microcanonical distribution:

$$\langle\phi|\hat{\Pi}_k|\phi\rangle = \frac{\mathrm{Tr}[\hat{\Pi}_k]}{\mathrm{Tr}[\hat{\mathbb{1}}]} + o(\ldots), \tag{1.8}$$

with probability 1. The intuition behind this is similar to Eq. (1.4) (see Fig. 1.2). Further, a time evolving $|\phi(t)\rangle$ continues to satisfy Eq. (1.8) at almost all times, with probability 1 in the Haar ensemble, leading to thermalization as in Eq. (1.7) — this requires no constraints on the energy levels.

The two statements collectively establish the following properties of any quantum system:

1. Normal typicality: Almost all randomly chosen sets of observables thermalize in every state, given mild non-degeneracy conditions on the energy levels and spacings.

[2]Strictly speaking, canonical typicality assumes more than just a set of projectors, namely all pure states corresponding to one subsystem in a factorization of the system into subsystems. We will deal with this stronger version in Chap. 8, and allude to its connection to projectors in Chap. 4.

2. Canonical typicality: Almost all randomly chosen states thermalize every set of observables.

We should, however, stress that neither is a direct counterpart to Boltzmann's ergodic hypothesis. Normal typicality states instead that the "thermal equilibrium" region occupies nearly the entirety of every ergodic subset of the Hilbert space, if defined for a typical randomly chosen observable, and does not directly address the dynamics of states. It establishes that "statistical mechanics is likely to be valid" for most observables in all systems, and says nothing of physical observables. Even with its "upgrade" to ETH, it only supposes that a physically accessible observable "looks like" such a typical observable.

Canonical typicality, however, may seem at first sight to be saying the same thing as Boltzmann: almost all quantum states are ergodic and thermalize to the microcanonical distribution with respect to every single observable, as per Eq. (1.8). Are we to conclude from this that all quantum systems, with no exceptions whatsoever, should be considered trivially ergodic? We will argue that such a conclusion would not be well-supported even with the most heuristic quantum-to-classical correspondence.

The key distinction is that one should not draw a direct correspondence between quantum states $|\phi\rangle$ and a classical state $x \in \mathscr{P}$. Instead, the best classical analogy to $|\phi\rangle$, or more precisely its density operator $\hat{\rho} = |\phi\rangle\langle\phi|$, is a (probability) *distribution* $\rho(x \in \mathscr{P})$ on the classical phase space. For example, it is these quantities that satisfy the respective Liouville equations [9, 15, 25]:

$$\begin{aligned} \frac{\mathrm{d}\rho}{\mathrm{d}t} &= \{H, \rho\}_{\mathrm{PB}}, \quad \text{for classical distributions,} \\ \frac{\mathrm{d}\hat{\rho}}{\mathrm{d}t} &= -i[\hat{H}, \hat{\rho}], \quad \text{for quantum density operators,} \end{aligned}$$

with $\{\cdot,\cdot\}_{\mathrm{PB}}$ denoting the classical Poisson bracket. The classical analogy to canonical typicality is then the following statement: generic probability distributions $\rho(x \in \mathscr{P})$ appear to be uniform on $\mathscr{P}$ after

some coarse-graining[3], whose intuition is again captured quite well by Fig. 1.2.

Von Neumann's developments, and modern refinements, still leave some questions unanswered: can we make any nontrivial statements about quantum ergodicity, in the sense of the dynamics of quantum states, without such typicality arguments that average over ensembles, and without assuming that we can access only a set of unspecified observables as in ETH? Answering this question will be tantamount to addressing the quantum analogue of ergodic theory.

1.1.3 What do we mean by an ergodic theory of quantum information dynamics?

The typicality results briefly reviewed in Sec. 1.1.2, while powerful, are still statistical arguments about typical observables, and do not directly address the dynamics of an individual quantum system in a nontrivial way. In the distinction between ergodic theory and the statistical mechanics of accessible observables, these results are more directly connected to the latter. But are there compelling reasons to think that there are more nontrivial results that are closer to ergodic theory in quantum mechanics, beyond von Neumann's observations?

The idea of effective theories has, in retrospect, largely insulated us from what might have otherwise been a pressing need to build a quantum ergodic theory from first principles. If one is interested in describing a certain class of macroscopic systems, even some of their universal features, it often suffices to write an effective theory that captures some essential aspects of those systems, or make a hypothesis such as ETH for accessible observables (which itself is frequently too strong). On the other hand, there are fundamental open problems that do not rely on accessible observables: for example, a particularly mysterious connection between the distribution of energy levels of supposed "ergodic/chaotic" quantum systems and random matrices (reviewed in Sec. 1.2), or the question of

[3]For a somewhat more formal construction, we can split the phase space into cells C_k, and allow $\rho(x \in C_k)$ to be a randomly chosen constant for each cell, subject to $\sum_k \rho(C_k) = 1$. An inessential aspect of this analogy is that classical distributions are more properly represented by highly entangled mixed states (see App. A.5), but the intuition for pure states is similar.

fundamental speed limits on quantum processes associated with statistical mechanics (reviewed in Sec. 1.3). The answers to such questions must depend directly on the kind of dynamics possible in a quantum system, rather than a specific choice of observables, and are particularly suited to motivate and test a quantum "ergodic theory".

Mathematically, there is a formal definition of ergodic theory in terms of notions that are tailored to fit classical dynamics, specifically focusing on the invariance of the measure μ. For example, the classic text *Ergodic Theory* by Cornfeld, Fomin, and Sinai [12] begins with "Ergodic theory studies motion in a measure space...[which] will turn out to be the phase space of a dynamical system". Or, with a more technical flavor, *Introduction to Ergodic Theory* by Sinai [11] offers "it may be considered that ergodic theory studies the category of measure spaces in which the morphisms are measure preserving transformations." But these notions appear to be less useful for the quantum dynamics of individual systems. The problem is not that a measure-centric approach does not generalize to quantum mechanics: the Hilbert space of normalized states can be endowed with a Haar measure (invariant under unitary transformations of states, including time evolution), and as noted in Refs. [26, 27] one can partly define a quantum measure algebra based on the entanglement properties (purities and overlaps) of pairs of mixed states [see also Eq. (2.73) and App. A.5]. However, both in classical and quantum mechanics, such methods only show sensitivity to the existence, degeneracy, and rational independence of eigenvalues [10–12, 16] (partly due to averaging over non-zero measure regions), and lack sufficient detail to tackle the distributions of energy eigenvalues.

In this Dissertation, our route towards "Constructing an Ergodic Theory of Quantum Information Dynamics" takes its more physical starting point as the energy levels themselves, which are invariants of quantum systems under unitary transformations. In fact, it is not an entirely misleading picture to think of the energy eigenstates of a Hamiltonian quantum system as quantum analogues of Boltzmann's invariant energy surfaces in phase space. The unique problem in quantum mechanics is that

this picture is too trivial: each eigenstate evolves with a phase factor $|E_n\rangle e^{-iE_n t}$, with fixed expectation values for each observable, and not much else happens. All the interesting physics is in the superposition of these evolving states, and how the E_n affect these superpositions. If we maintain sufficient analytical control over these superpositions (rather than statistically averaging over ensembles of nonzero measure), we may expect to have enough resolution to relate quantitative details of the energy levels to quantum dynamics as necessary for the open problems reviewed in Secs. 1.2 and 1.3.

Further, rather than viewing unitary time evolution as (Haar) measure-preserving transformations on the Hilbert space, we also expect it to be more useful to emphasize unitary time evolution as transformations that preserve the inner products of states. This has two advantages. Firstly, inner products are directly measurable invariants in quantum mechanics, in contrast to Hilbert space measures. Further, there is a direct physical correspondence between quantum mechanical inner products and classical phase space measures, e.g. via the correspondence between the overlaps of quantum states and classical probability distributions discussed in Sec. 1.1.2 (with quantitative details in Chap. 2 and App. A.5). Thus, instead of merely *applying* the classical ergodic theory of measure-preserving transformations to quantum mechanical Hilbert spaces, we shall likely be better placed to address nontrivial aspects of quantum dynamics if we base our approach on quantum mechanical invariants such as inner products, which better capture the *physical* role of the invariants (rather than their formal mathematical role as measures) in classical ergodic theory.

With this perspective, we can attempt to formulate something of an analogue to the opening statements of the ergodic theory texts [11, 12] quoted above. We expect a "natural" quantum ergodic theory to be a collection of results that describe how invariants in linear algebra (e.g., the set of eigenvalues of a linear transformation that implements dynamics, such as energy levels) determine the orbits of vectors (quantum states) in terms of invariant quantities defined through inner products (probabilities and overlaps of the vector with states that are unambiguously specified relative to the

initial state, or relative to the energy levels). Or in briefer language that focuses on the physics:

How energy level distributions affect the observable-independent dynamics of quantum states. (1.9)

This Dissertation sets in place a few such results, specifically tackling Conjecture 1.1 in Sec. 1.2 and Conjecture 1.2 in Sec. 1.3, with a hope for more to come.

1.2 Aspects of the energy spectrum: Quantum chaos and random matrix theory

The biggest physical clue for the role of the energy eigenvalues E_n comes from what is called the "quantum chaos conjecture" [28]. This is not a conjecture in the mathematical sense [29] (it is easy enough to find counter-examples to most literal interpretations of the conjecture), but should be regarded as a guiding principle that tells us where to look for more precise concepts concerning quantum ergodicity. We will find it convenient to club this conjecture with a closely related statement for integrable systems, called the Berry-Tabor conjecture. Collectively, they state that

Conjecture 1.1 (Quantum chaos conjecture)**.** *The statistical distributions of the energy levels of quantum systems typically show the following patterns:*

1. *The energy levels of "chaotic" quantum systems are distributed similar to the eigenvalues of a random matrix, in sufficiently small energy windows of an "ergodic" subspace [30–33].*

2. *The energy levels of "integrable" quantum systems are distributed similar to a Poisson (locally uncorrelated) distribution in sufficiently small energy windows [34].*

These statements were originally motivated from numerical studies of typical quantizations of chaotic and integrable classical systems. In fully quantum systems without a classical limit, the labels "chaotic", "ergodic" and "integrable" don't have any agreed-upon meanings, and usually serve as a

proxy for how "easy" it is to solve for the dynamics of a given system. An alternate approach is to adopt these spectral signatures as (heuristic) *definitions* [35] of "quantum chaos" or "integrability"; this semantic choice still leaves us with little inkling of what precise forms of quantum dynamics correspond to these words, and is one that we will not follow here.

Instead, our primary focus will be directly addressing the connection between the energy spectrum and quantum dynamics, as motivated in Sec. 1.1.3. Given that the associated spectra are much better studied than the dynamics, an important step for our developments is to understand what these spectra look like. We work towards this goal in the remainder of this Section, with Fig. 1.3 providing a quick visual summary.

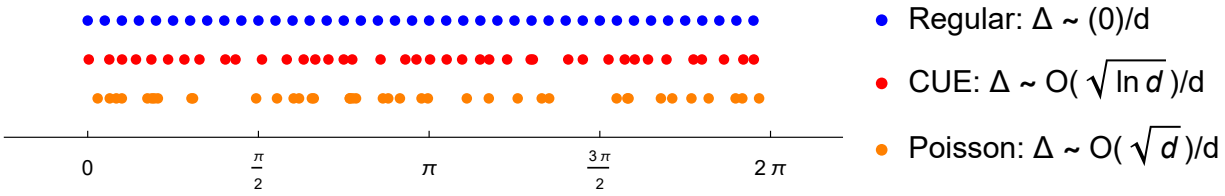


Figure 1.3: A depiction of spectral fluctuations for different distributions of energy levels in $[0, 2\pi)$, representing an energy window of $d = 40$ energy levels (here, with $D = d$): regular $E_n = 2\pi n/d$ (blue), a random CUE matrix (red), and an uncorrelated [Poisson] distribution (orange). Here, $\Delta_n = E_n - 2\pi n/d$, and Δ represents the typical size of the Δ_n (or deviation from regularity of a spectrum, see also Chap. 2). These should be compared to the analytical expressions in Eq. (1.33). The key takeaway is that a random matrix spectrum looks considerably closer to having regularly spaced energy levels than a Poisson one.

1.2.1 Haar random matrices

In Sec. 1.1.2, we saw that Haar distributions over operators $\hat{A}$ (specifically, projectors $\hat{A} = \hat{\Pi}_k$), which show invariance under unitary transformations (i.e., $\hat{A}$ and $\hat{U}\hat{A}\hat{U}^\dagger$ have the same probability), lead to quantum thermalization behaviors with probability 1. The eigenvalues of $\hat{A}$ played no role in this argument, mainly because projectors by definition have a fixed set of eigenvalues $\{0, 1\}$ that cannot be changed infinitesimally. Stranger things happen when we allow the eigenvalues to vary.

One of the most direct ways of allowing eigenvalues to vary, in a physical context, is as follows.

Let $\hat{U}_H(t)$ specifically denote the unitary that effects time evolution:

$$\hat{U}_H(t) = \sum_n e^{-iE_n t} |E_n\rangle\langle E_n|. \tag{1.10}$$

Then, instead of considering the normal typicality of a distribution of observables $\hat{\Pi}_k$ for a given system with fixed $\hat{U}_H(t)$, as in Sec. 1.1.2, we can consider the likelihood of thermalization in a distribution of unitaries, say $\hat{U}_H(1)$, given a fixed set of observables.

In fact, as long as this distribution of $\hat{U}_H(1)$ is invariant under unitary transformations $\hat{U}_H(t) \to \hat{U}^\dagger \hat{U}_H(t)\hat{U}$, we will recover Eq. (1.4) because this transformation $|E_n\rangle \to \hat{U}|E_n\rangle$ can be transferred to the $\hat{\Pi}_k$ instead to recover $\hat{\Pi}_k \to \hat{U}^\dagger \hat{\Pi}_k \hat{U}$, obtaining the Haar distribution considered by von Neumann. This procedure is analogous to the passage between the Heisenberg and Schrödinger pictures for an arbitrary unitary transformation $\hat{U}$. This "transfers" normal typicality to typical time-evolution operators rather than observables. If the distribution of $\hat{U}_H(1)$ further favors eigenvalues E_n that are non-degenerate (with non-degenerate spacings) with probability 1, von Neumann's ergodicity result follows with probability 1 for every initial state (and a given orthogonal set of observables).

However, for historical and aesthetic reasons, it is customary to consider a specific uniform distribution of $\hat{U}_H(1)$ with respect to the actual Haar measure on the unitary group (formally, given exactly by the normalized Haar measure), which is additionally invariant under the group action (multiplication by unitaries) $\hat{U}_H(1) \to \hat{U}_H(1)\hat{V}$. This is the analogue of the microcanonical ensemble for the unitary group, and goes by the name of the "Circular Unitary Ensemble" (CUE) [28,36]. It also favors specific distributions of the eigenvalues e^{-iE_n}, which do satisfy von Neumann's non-degeneracy criteria with probability 1. Thus, from the above argument, Eqs. (1.4) and (1.7) hold with probability 1 for $\hat{U}_H(1)$ chosen from the CUE. This is a weaker typicality statement than canonical or normal typicality (because CUE is a more specific ensemble that restricts the eigenvalue distribution).

Mathematically, it may seem odd to enforce an eigenvalue distribution on E_n that doesn't particularly matter for the above typicality results. Here is where the historical reasons come into play: CUE is one of the cleanest examples of an eigenvalue phenomenon called "level repulsion", where it is not just degeneracies that are avoided, but the probability of finding eigenvalues *near* each other is also suppressed in specific, quantitative ways, as will be shortly derived in Sec. 1.2.2. Closely related examples include the Circular Orthogonal Ensemble (COE) and Circular Symplectic Ensemble (CSE) that are respectively uniform on the subsets of unitaries satisfying the additional constraints of symmetric matrix elements and self-duality [28, 36], which are associated with slight quantitative differences in the form of level repulsion. We will also sometimes work with the "Gaussian" counterparts of these Circular ensembles, i.e., GOE, GUE, GSE, which are distributions of Hermitian matrices $\hat{H}$ where each element is an independent Gaussian, invariant under the same set of transformations as the respective circular ensemble COE, CUE, CSE, and showing the same type of level repulsion [28, 36].

1.2.2 Level repulsion

For generality, consider a diagonalizable operator $\hat{A}$ with eigenvalues $\{a_k\}$ and respective eigenvectors $|a_k\rangle$. Under an infinitesimal similarity transformation $\hat{A} \to \hat{A}' = \hat{S}\hat{A}\hat{S}^{-1}$ where $\hat{S} = \hat{\mathbb{1}} + \epsilon\hat{X}$ for small $\epsilon > 0$, the matrix elements of the transformed operator in the original eigenbasis are given by:

$$\begin{aligned} A'_{jk} \equiv \langle a_j|\hat{A}'|a_k\rangle &= \langle a_j|\left(\hat{A} + \epsilon[\hat{X},\hat{A}] + O(\epsilon^2)\right)|a_k\rangle \\ &= a_k\delta_{jk} + \epsilon(a_k - a_j)X_{jk} + O(\epsilon^2). \end{aligned} \tag{1.11}$$

The key takeaway here is that the ability of a similarity transformation to change matrix elements is determined by the differences between eigenvalues: as an extreme case, if $a_j = a_k$, then $A'_{jk} = A_{jk}$ up to $O(\epsilon^2)$ corrections.

If we take $\mathrm{d}\hat{A}_{jk} = A'_{jk} - A_{jk}$ to be a Hermitian matrix (at least to leading order in ϵ), the independent (real) variables in its matrix elements are $\mathrm{d}A_{kk}$, $\mathrm{d}\,\mathrm{Re}A_{j>k}$, $\mathrm{d}\,\mathrm{Im}A_{j>k}$. A measure over Hermitian matrices that allows varying all these elements can then be written in terms of the eigenvalues a_k and the similarity transformation variables X_{jk} as follows:

$$\begin{aligned}\prod_k \mathrm{d}A_{kk} \prod_{j>k} \mathrm{d}\,\mathrm{Re}A_{jk} \mathrm{d}\,\mathrm{Im}A_{jk} &= \prod_k \mathrm{d}a_k \prod_{j>k} |a_j - a_k|^2 \epsilon^2 \mathrm{d}\,\mathrm{Re}X_{jk} \mathrm{d}\,\mathrm{Im}X_{jk} \\ &\propto \prod_k \prod_{j>k} |a_j - a_k|^2, \end{aligned} \tag{1.12}$$

where the last line is the square of the Vandermonde determinant (the determinant of the matrix whose entries are $\mathcal{V}_{kn}(a) = a_k^{n-1}$ where $n \geq 1$):

$$\mathcal{V}(a) \equiv \left|\det\left[a_k^{n-1}\right]_{kn}\right| = \prod_k \prod_{j>k} |a_j - a_k|. \tag{1.13}$$

This determinant representation will turn out to be useful later.

The main phenomenon of interest illustrated in Eq. (1.12) is that of "level repulsion" in random matrix theory [28, 36]: any distribution of the matrix A_{jk} that allows sufficiently many matrix elements $\mathrm{d}A_{jk}$ to vary suppresses the distribution of (several of) its eigenvalues by the differences $|a_j - a_k|$, favoring large separations between eigenvalues. In the above derivation, we have tried to keep our argument as general as possible. It is customary to specifically derive level repulsion in Haar distributions of Hermitian or unitary operators such as CUE, which are invariant under unitary transformations (or specific subgroups of the unitary group), in which case the similarity transformation $\hat{S} = \hat{U}$ is unitary, while $\hat{A}$ is also either Hermitian or unitary ($\mathrm{d}\hat{A}$ is Hermitian to $O(\epsilon)$ in both cases). In this case, the Haar probability distribution in terms of eigenvalues is precisely a power of the Vandermonde determinant (with $\beta_{\mathrm{RMT}} = 2$ when the similarity transformations are the full unitary

group) [28, 36],

$$P\left(\{a_k\}\right) \propto [\mathcal{V}(a)]^{\beta_{\rm RMT}} \tag{1.14}$$

But we see above that the crucial ingredient in level repulsion is that one has sufficient freedom in varying matrix elements (and, e.g., this generalizes [37] to beyond Hermitian dA_{jk}), and not so much the precise invariance properties of the Haar distribution. Eigenvalue distributions of this form are also called Wigner-Dyson distributions [28, 36], especially when $\beta_{\rm RMT} = 1$ (COE), 2 (CUE) and 4 (CSE).

There are elaborate formal techniques to derive the distribution of a small subset of eigenvalues, given such a Vandermonde distribution of the full set. One of these involves orthogonal polynomials [36], which will play a direct role in Chap. 5. There does not appear to be a quick intuitive path through this derivation, and we will eschew an overview of these techniques prior to Chap. 5, and merely state an important result here.

For the joint distribution function of pairs of energy levels, which is just a product of the distribution of levels:

$$R_2(E_1, E_2) = \left(\sum_m \delta(E_1 - E_m)\right)\left(\sum_n \delta(E_2 - E_n)\right), \tag{1.15}$$

averaging over $\beta_{\rm RMT} = 2$ ensembles such as CUE gives for small[4] $(E_1 - E_2)$,

$$\left\langle R_2(E_1, E_2)\right\rangle_{\beta_{\rm RMT}=2} = \mathcal{N}^{\#}(E_1)\mathcal{N}^{\#}(E_2)\left\{1 + \delta(E_1 - E_2) - \mathrm{sinc}^2\left[\pi(E_1 - E_2)\mathcal{N}^{\#}\left(\frac{E_1 + E_2}{2}\right)\right]\right\}. \tag{1.16}$$

Here, $\mathcal{N}^{\#}(E) = \langle\sum_n \delta(E - E_n)\rangle$ is the ensemble-averaged density of states near E, normalized to the total number of energy levels: $\int dE\ \mathcal{N}^{\#}(E) = D$. For individual elements of the ensemble, the ensemble-average may be replaced by averaging over a sufficiently small energy window.

[4] It is possible to state a more exact result for all $E_1 - E_2$ in CUE, see e.g. Refs. [28, 36], but here we want to anticipate the behavior of this quantity for generic systems.

1.2.3 Spectral form factors

One way to access the distribution of eigenvalues is through the traces of a matrix $\hat{A}$. For a $D \times D$ matrix with D eigenvalues, the set of traces $\mathrm{Tr}(\hat{A}^t)$ for $t = 1$ to D determines all the eigenvalues a_k up to permutations.

For our time evolution unitary operator $\hat{U}_H(t)$, this trace can be written as

$$\mathrm{Tr}\left[\hat{U}_H(t)\right] = \sum_n e^{-iE_n t}. \tag{1.17}$$

Here, we allow t to be continuous to accommodate Hamiltonian systems. At $t = 0$, this trace is D, and we expect it to gradually decay until the phases $E_n t$ all appear well distributed throughout the unit circle $[0, 2\pi)$. After this point, the sum in Eq. (1.17) is dominated by destructive interference between the phases, and its precise value will sensitively depend on the precise distribution of the E_n. As the quantization of energy levels is crucial for this behavior, we will call these contributions (which primarily track microscopic subtleties in the destructive interference of phases) the "quantum fluctuations" of the trace, which are expected to be $O(\sqrt{D})$ on average over the full range of times $t \in \mathbb{R}$ (e.g., if the phases add randomly, which is the generic behavior at most times).

To see how this connects to level repulsion, it is more convenient to focus on a closely related quantity, the spectral form factor (SFF) [28]:

$$K(t) \equiv \frac{1}{D^2}\left|\mathrm{Tr}\left[\hat{U}_H(t)\right]\right|^2 = \frac{1}{D^2}\sum_{n,m} e^{i(E_m - E_n)t} \tag{1.18}$$

$$= \frac{1}{D^2}\int \mathrm{d}E_1 \mathrm{d}E_2 R_2(E_1, E_2) e^{i(E_1 - E_2)t}. \tag{1.19}$$

Crucially, $K(t) \geq 0$ (and $K(t) \leq 1$), which also allows us to average it over ensembles (if desired) without entirely losing information about the quantum fluctuations. Our emphasis in this Dissertation will

be on the physics of individual systems rather than ensembles (except notably in Chap. 5, and partly Chap. 8); however, we will use ensembles here for convenience, which leads to simpler expressions.

Translating our comments on the trace to the SFF, we have $K(0) = 1$, which decays with time to generically $O(D^{-1})$ quantum fluctuations, whose precise behavior depends sensitively on the short-range energy level distribution via $R_2(E_1, E_2)$. We will also use t_f to (loosely) denote the "quantum fluctuations" time at which $K(t)$ first reaches an $O(D^{-1})$ value *and* begins to track the destructive interference of energy levels[5]. Averaging over an ensemble satisfying Eq. (1.16) and glossing over some subtleties [28,36] essentially gives the following behavior for the averaged quantum fluctuations (if $\mathcal{N}^{\#}(E)$ is constant in some interval and 0 outside, e.g., $\mathcal{N}^{\#}(E) = D/(2\pi)$ for CUE in $[0, 2\pi)$):

$$\langle K(t)\rangle_{\beta_{\text{RMT}}=2} = \begin{cases} \dfrac{1}{D}\dfrac{t}{t_{\text{H}}}, & \text{for } t_{\text{ramp}} < t \leq t_{\text{H}}. \\ \dfrac{1}{D}, & \text{for } t > t_{\text{H}}. \end{cases} \tag{1.20}$$

This $K(t > t_{\text{ramp}}) \propto t$ behavior is the celebrated "linear ramp" of the SFF, and it most directly measures a $\sqrt{t}$ growth in the size of quantum fluctuations in the trace in Eq. (1.17); $K(t > t_{\text{H}}) = \text{const.}$ is frequently called the "plateau" behavior [38]. Here, t_{ramp} represents the time scale (or inverse energy scale) after which Eq. (1.16) holds in a system (note that $t_f < t_{\text{ramp}}$ in $\beta_{\text{RMT}} > 0$ systems), and $t_{\text{H}} = 2\pi\mathcal{N}^{\#}(E)$. For ideal Poisson statistics, the level correlations due to the Vandermonde determinant do not exist, and it is convenient to set $\beta_{\text{RMT}} = 0$, which gives $R_2(E_1, E_2) = \mathcal{N}^{\#}(E_1)\mathcal{N}^{\#}(E_2)$. Here, we simply get the "plateau" throughout the time scale of quantum fluctuations:

$$\langle K(t > t_f)\rangle_{\beta_{\text{RMT}}=0} = \frac{1}{D}. \tag{1.21}$$

It is expected that this ensemble-averaging is a convenient proxy for describing the *average trend* of

[5] We can define this more formally in an ensemble, as the time where $D^{-2}\langle|\text{Tr}\,\hat{U}_H(t)|^2\rangle$ becomes appreciably larger than the disconnected average $D^{-2}\langle|\text{Tr}\,\hat{U}_H(t)|\rangle^2$.

$K(t)$ even in most individual systems of the ensemble (as may be obtained by a moving average over small time intervals to smoothen out the fluctuations). These quantum fluctuations will be a main focus of Chap. 2, as well as all subsequent chapters except Chaps. 3 and 4.

Significantly prior to t_f, the decay of $K(t)$ from 1 at $t = 0$ is well-described by the disconnected contribution $\mathscr{N}^{\#}(E_1)\mathscr{N}^{\#}(E_2)$ to $R_2(E_1, E_2)$ up to times where $K(t) \sim O(D^{-1})$:

$$\langle K(t \ll t_f)\rangle = \frac{1}{D^2}\left|\int \mathrm{d}E\ \langle \mathscr{N}^{\#}(E)\rangle e^{-iEt}\right|^2. \tag{1.22}$$

The quantitative form of this decay is specific to individual systems, and not as generic as the linear ramp; however, it is also largely agnostic to the quantum fluctuations that dominate $t > t_{\text{ramp}}$ (e.g. can be virtually identical for $\beta_{\text{RMT}} = 0$ and $\beta_{\text{RMT}} = 2$). This early-time behavior will play a crucial role in Chaps. 3 and 4. In the intermediate range of times from $\sim t_f$ up to t_{ramp}, spectral correlations come into play and can determine the precise value of t_{ramp} in system-specific ways [38]. Some of these generic features of the SFF are illustrated in Fig. 1.4.

The linear ramp is more intuitive to derive than the two level distribution in Eq. (1.16), for a sufficiently random matrix. The easiest case to illustrate this in is the CUE ensemble itself, for discrete time steps $t \in \mathbb{N}$, but the argument will be seen to be more general than CUE. All that is necessary is that the phase of each matrix element $\arg\langle j|\hat{U}_H(1)|k\rangle$ behaves as if it is sampled from a random distribution (see App. F.2 for more details). Taking $\hat{U}_H(1) = \hat{V}$, the trace at time t can be written as:

$$\mathrm{Tr}[\hat{V}^t] = \sum_{m_t,\ldots,m_1} V_{m_1 m_t}\ldots V_{m_3 m_2} V_{m_2 m_1}. \tag{1.23}$$

The sum is over D^t sets of indices $\{m_t, \ldots, m_1\}$, but not all of these are independent: all terms related by cyclic permutations of these t indices are *identical*, leaving only D^t/t independent choices of indices.

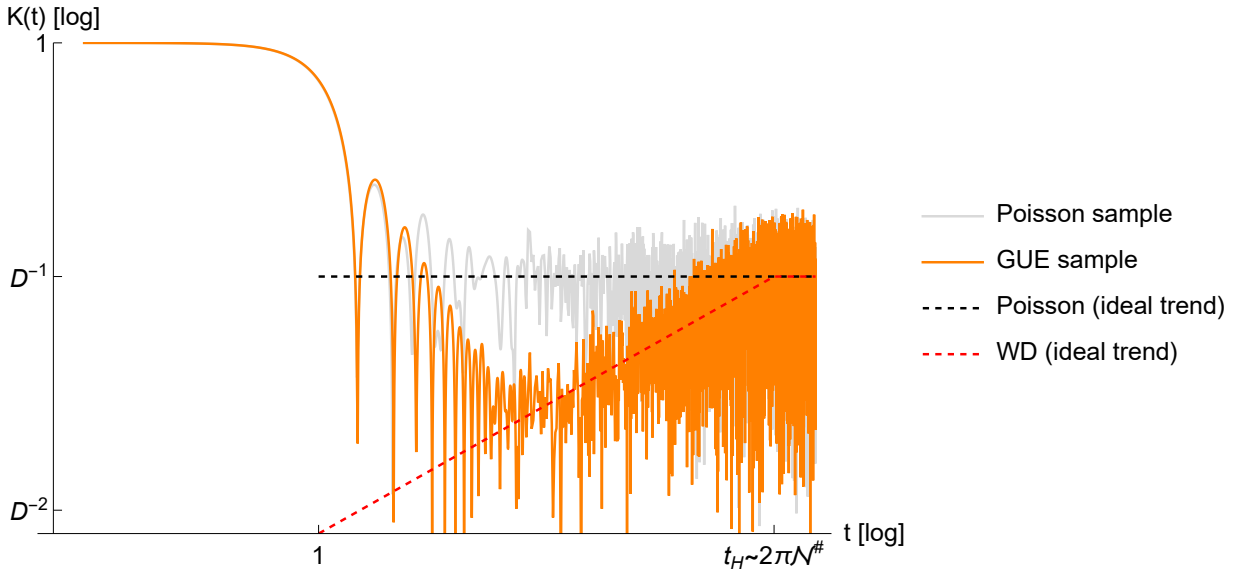


Figure 1.4: A log-log plot of $K(t)$ vs t showing examples of the early-time decay and the quantum fluctuations in the SFF, for a Hermitian Hamiltonian $\hat{H}$ sampled from GUE, as well as a random distribution of energy levels (Poisson) with the same density of states $\mathcal{N}^{\#}(E)$. The early-time decay is independent of level repulsion and similar for both cases, as both have the same overall density of states $\mathcal{N}^{\#}(E)$; the time $t = 1$ corresponds to the scale of the inverse width of the energy spectrum (specifically, we use units where the leading order ensemble-averaged GUE density of states [28, 36] is supported on $E \in [-2, 2]$). The ramp-plateau structure is reflective of level repulsion in GUE [the Heisenberg time $t_{\mathrm{H}} \sim 2\pi\mathcal{N}^{\#}$ depicted is from the center of the spectrum], while Poisson shows a pure plateau without a ramp in fluctuations. These features are compared with the "ideal" linear ramp and plateau for $\beta_{\mathrm{RMT}} = 2$ Wigner-Dyson distributions. One of our primary goals is to connect both the early-time decay and the late-time fluctuations to concrete quantum dynamical processes.

We have

$$\mathrm{Tr}[\hat{V}^t] = t\left(\sum_{[D^t/t \text{ independent terms}]} V_{m_1 m_t} \dots V_{m_3 m_2} V_{m_2 m_1}\right) \tag{1.24}$$

If each independent term has an independent phase that looks sufficiently random, the sum is $\sqrt{D^t/t}$ times the typical magnitude of each term:

$$\mathrm{Tr}[\hat{V}^t] \sim \sqrt{t}\left(\text{typical magnitude of } |D^{1/2}V_{jk}|^t\right). \tag{1.25}$$

The first $\sqrt{t}$ factor is the linear ramp contribution to $K(t)$, while the second factor oscillates around 1 for CUE unitaries (as $V_{jk} \sim D^{-1/2}$). Deviations from this behavior can be caused, for example, by

correlations between phases or larger matrix elements.

This argument has its origins in a semiclassical periodic orbit approach to energy levels that assumes randomly distributed and uncorrelated classical periodic orbits [28,39,40] (Eq. (1.23) can be thought of as a sum over periodic Feynman trajectories, with classical periodic orbits dominating in the classical limit). It also generalizes in various ways to sufficiently random ensembles of fully quantum systems such as kicked Ising chains or local random quantum circuits [35,41,42]. These ensembles are considerably more specialized to physically interesting situations than the mathematically idealized Haar ensembles. However, the generality of the linear ramp should not be too surprising from our viewpoint — the randomness conditions required for Eq. (1.25) are rather weak and do not require the full symmetry and other properties of the Haar measure itself. It is nevertheless a considerable challenge to *prove* such statements except in suitable ensembles (including CUE itself [43, 44]), and recovering these generic CUE-like (or COE/CSE-like) "periodic orbit" phase correlations from Feynman trajectories is often considered the closest we can get to *physical* explanations of the quantum chaos conjecture.

But we are still left with the following questions from an ergodic theory standpoint: is it mere statistical typicality that leads to the ramp in most "sufficiently random" systems? Does it directly determine any aspect of the quantum dynamics of an individual system? We are close to the point where we can describe an answer to such questions (see Sec. 1.4), but it is useful to first review one more property of spectra to build intuition.

1.2.4 Spectral fluctuations and rigidity

The level repulsion in Eqs. (1.12) and (1.13) affects not just neighboring energy levels, but also distant ones. The long-range effect of this level repulsion is to ensure that the energy levels end up

being fairly regularly spaced, with a perfect regular 1D lattice in the ideal case:

$$E_n - E_0 = \frac{n}{\langle \mathcal{N}^{\#}(E) \rangle}, \quad \text{for } \beta_{\text{RMT}} = \infty. \tag{1.26}$$

A convenient measure to quantitatively probe this effect [45], including in cases with finite β_{RMT}, is the number of energy levels in an energy interval of size ΔE (say, $[0, \Delta E)$), which is given by:

$$n(\Delta E) = \int_0^{\Delta E} \mathrm{d}E \sum_n \delta(E - E_n) = \int \frac{\mathrm{d}t}{2\pi} \operatorname{Tr}[\hat{U}_H(t)] \frac{1 - e^{i\Delta E t}}{it}. \tag{1.27}$$

The second equality follows from writing the sum of delta functions as a Fourier transform of Eq. (1.17). To retain information about the size of fluctuations after ensemble averaging (which we can take to represent averaging over different uncorrelated intervals ΔE), we square this quantity and use the fact that it is real:

$$\langle n^2(\Delta E) \rangle_{\beta_{\text{RMT}}} = \int \frac{\mathrm{d}t}{2\pi} \frac{\mathrm{d}t'}{2\pi} \frac{\langle \operatorname{Tr}[\hat{U}_H(t)] \operatorname{Tr}[\hat{U}_H^{\dagger}(t')] \rangle_{\beta_{\text{RMT}}}}{(t/2)(t'/2)} e^{i\Delta E(t-t')/2} \sin\left(\frac{\Delta E t}{2}\right) \sin\left(\frac{\Delta E t'}{2}\right) \tag{1.28}$$

To a good approximation (e.g., when $\mathcal{N}^{\#}(E) \approx$ const. in an interval much larger than ΔE, as in CUE), we have

$$\langle \operatorname{Tr}[\hat{U}_H(t)] \operatorname{Tr}[\hat{U}_H^{\dagger}(t')] \rangle_{\beta_{\text{RMT}}} \approx D^2 K(t) \delta(t - t') \tag{1.29}$$

which is essentially because the quantum fluctuations in $\operatorname{Tr}[\hat{U}_H(t)]$ behave as if they are uncorrelated for large $t - t'$. We get (as the integrand is even in t):

$$\langle n^2(\Delta E) \rangle_{\beta_{\text{RMT}}} \approx \frac{2D^2}{\pi^2} \int_0^{\infty} \mathrm{d}t \frac{\langle K(t) \rangle_{\beta_{\text{RMT}}}}{t^2} \sin^2\left(\frac{\Delta E t}{2}\right). \tag{1.30}$$

The $t < t_f$ contribution is from Eq. (1.22), and essentially just gives the disconnected $\langle n(\Delta E)\rangle^2_{\beta_{\text{RMT}}}$ contribution from the average density of states. The spectral fluctuations come from $t > t_f$, and we can write:

$$\langle n^2(\Delta E)\rangle_{\beta_{\text{RMT}}} \approx \langle n(\Delta E)\rangle^2_{\beta_{\text{RMT}}} + \frac{2D^2}{\pi^2}\int_{t_f}^{\infty} \mathrm{d}t\, \frac{\langle K(t)\rangle_{\beta_{\text{RMT}}}}{t^2} \sin^2\left(\frac{\Delta E t}{2}\right) \tag{1.31}$$

For $2\pi/t_{\text{H}} \ll \Delta E \ll 2\pi/t_f$, we can substitute Eqs. (1.20) and (1.21) for the behavior of the SFF for $\beta_{\text{RMT}} = 2$ (CUE) and $\beta_{\text{RMT}} = 0$ (Poisson) respectively, neglecting the $t \in (t_f, t_{\text{ramp}})$ contribution in the former (which is usually negligible compared to $t \in (t_{\text{ramp}}, t_{\text{H}})$. Due to the predominantly non-oscillatory behavior of the *ensemble-averaged* $\langle K(t)\rangle_{\beta_{\text{RMT}}}$ in both cases, we can also approximate $\sin^2(\Delta E t/2)$ by its average value $1/2$ over multiple oscillations, cutting off the lower limit of the integral at the period $t \sim O(\Delta E^{-1})$ of the first oscillation [as the $\sin^2(\Delta E t/2)$ factor behaves as $\Delta E^2 t^2$ for $t \ll O(\Delta E^{-1})$]. This gives

$$\langle n^2(\Delta E)\rangle_{\beta_{\text{RMT}}} - \langle n(\Delta E)\rangle^2_{\beta_{\text{RMT}}} \approx \frac{D^2}{\pi^2}\int_{O(\Delta E^{-1})}^{\infty} \mathrm{d}t\, \frac{\langle K(t)\rangle_{\beta_{\text{RMT}}}}{t^2} \sim \begin{cases} \dfrac{D}{\pi^2 t_{\text{H}}} \ln\left(t_{\text{H}}\Delta E\right), & \text{for } \beta_{\text{RMT}} = 2 \text{ (CUE)}, \\ \dfrac{1}{\pi^2} O(\Delta E), & \text{for } \beta_{\text{RMT}} = 0 \text{ (Poisson)}. \end{cases} \tag{1.32}$$

For less heuristic derivations, see, e.g., Refs. [36, 40, 45–47]. This means that in an energy window of d levels, i.e. $\Delta E t_{\text{H}} = 2\pi d$, we expect that (with $t_{\text{H}} \sim 2\pi\langle \mathscr{N}^{\#}(E)\rangle \sim O(D)$)

$$E_{d-1} - E_0 = \frac{d}{\mathscr{N}^{\#}(E)} \pm \frac{1}{\mathscr{N}^{\#}(E)} \begin{cases} O(\sqrt{\ln d}), & \text{for } \beta_{\text{RMT}} = 2 \text{ (CUE)}, \\ O(\sqrt{d}), & \text{for } \beta_{\text{RMT}} = 0 \text{ (Poisson)}. \end{cases} \tag{1.33}$$

The fluctuations around a regular $\beta_{\text{RMT}} = \infty$ spectrum, relative to the typical level spacings $1/\mathscr{N}^{\#}(E)$, are an enormously large $O(\sqrt{d})$ for Poisson statistics, compared to the logarithmic $O(\sqrt{\ln d})$ for Haar random matrices such as CUE (which generalizes to COE and CSE up to an overall β_{RMT}-dependent coefficient, see also Chap. 2). Consequently, it is often said that random matrix statistics is associated

with a "rigid" spectrum [40]. This difference in rigidity, illustrated in Fig. 1.3, is what will most directly enable us to make precise connections between the quantum chaos conjecture and quantum dynamics in Chap. 2 (see Sec. 1.4 for some intuition).

1.3 Aspects of quantum dynamics: Thermalization, scrambling, and entanglement

1.3.1 Quantum thermalization

Classically, it is possible for almost any (normalizable) probability distribution of states $\rho(x, t = 0)$ to *permanently* evolve, at long times, to the microcanonical distribution:

$$\lim_{t\to\infty} \int_{\mathscr{P}} \mathrm{d}\mu(x)\, \rho(x,t)\overline{\Theta}(x \in B) = \frac{\mu(B)}{\mu(\mathscr{P})}, \tag{1.34}$$

without requiring a time average. This is called *strong mixing* in ergodic theory [10–13, 48]. Thus, classical probability distributions *can* permanently thermalize without any obstacle from ergodicity or Poincaré recurrences, unlike individual trajectories (see the discussion following Eq. (1.3)).

However, strong mixing is impossible even for a density operator $\hat{\rho}(0)$ in quantum mechanics (under Hamiltonian evolution) due to the linearity of the theory: the linear flow of the phases $E_n t$ in the energy eigenbasis ensures that all of them can simultaneously return arbitrary close to 0 infinitely often. This is the phenomenon of quantum recurrences (connected to the classical Poincaré recurrences of linear flows on D-dimensional tori) [49], where the timescale of recurrence is expected to be exponential in the Hilbert space dimension, $T_{\mathrm{qr}} \sim \exp(\Theta(D))$. As $D \sim \exp(N)$ for an N-particle system, quantum mechanics allows the appearance of permanent thermalization over a much longer timescale $T_{\mathrm{qr}} \sim \exp(\exp(N))$ than the classical recurrences (which are merely $\exp(N)$). Intuitively, this is because the quantum Hilbert space is "larger" than the classical phase space due to superpositions.

In practice, we must account for these recurrences when considering quantum thermalization.

As we are working in a finite-dimensional Hilbert space, we must also allow for deviations from a thermal state that may be subleading in D (unlike the case of a continuum limit). A convenient starting point is to require that an initial state $\hat{\rho}$ evolves, after a time t_s, to remain within a certain "distance" of a thermal ensemble $\hat{\rho}_{\text{th}}$ up to a time $T_D \ll T_{\text{qr}}$:

$$\hat{\rho}(t : t_s < t < T_D) = \hat{\rho}_{\text{th}} + \delta\hat{\rho}, \tag{1.35}$$

where we can require $\text{Tr}(\delta\hat{\rho}^2) < \epsilon$ for some $\epsilon > 0$.

The thermal state $\hat{\rho}_{\text{th}}$ is usually a statistical ensemble of several pure states, and therefore a mixed state possessing a nonzero entanglement entropy (or equivalently, more than 1 nonzero eigenvalue) [50]. For example, the microcanonical distribution on the full Hilbert space $\mathscr{H}$ is the maximally mixed state $\hat{\rho}_{\text{th}} = \hat{\mathbb{1}}/D$. On the other hand, we will often want $\hat{\rho}(0)$ to be a pure state that can be prepared through projective measurements, having only one nonzero eigenvalue. Unitary evolution preserves the eigenvalues of $\hat{\rho}$, and cannot lead to Eq. (1.35).

One resolution is to follow the setting of von Neumann [16,17] (see Sec. 1.1.2), and only require that Eq. (1.35) appears to be satisfied with respect to the expectation values of a set of accessible projectors. As a special case of this requirement, the standard resolution [23,24,51] in a many-body context is to factorize the Hilbert space $\mathscr{H} = \mathscr{H}_S \otimes \mathscr{H}_E$ (say of N particles) into a subsystem $\mathscr{H}_S$ of interest (of N_S particles), and an external component $\mathscr{H}_E$ (of $N_E = N - N_S$ particles). In the language of Sec. 1.1.2, this corresponds to restricting to projectors that can only perform measurements within $\mathscr{H}_S$, acting trivially on $\mathscr{H}_E$:

$$\hat{\Pi}_k = \hat{\Pi}_{k,S} \otimes \hat{\mathbb{1}}_E. \tag{1.36}$$

For reduced density matrices $\hat{\rho}_S(t) = \text{Tr}_E[\hat{\rho}(t)]$ within $\mathscr{H}_S$, the effective evolution is no longer unitary, and Eq. (1.35) can be directly satisfied in $\mathscr{H}_S$. In this case, quantum dynamics within $\mathscr{H}_S$ must be capa-

ble of generating sufficient entanglement with $\mathscr{H}_E$ by $t = t_s$, so that even a pure state $\hat{\rho}_S(0) = |k\rangle_S\langle k|$ within $\mathscr{H}_S$ can evolve close to the mixed state $\hat{\rho}_{\text{th},S}$ within the subsystem. When the reduced initial state $\hat{\rho}_E(0)$ is a finite-temperature state in $\mathscr{H}_E$ with inverse temperature[6] β, this quantum thermalization problem also reduces to the classic thermodynamic problem of the thermalization (heating or cooling) of a system S in contact with a finite temperature bath E.

One significant consequence of Eq. (1.35) is that it becomes impossible to recover the initial state $\hat{\rho}_S(0)$ after t_s (and before T_D), as any such initial state must evolve to the same thermal state $\hat{\rho}_{\text{th},S}$. This is because the generation of entanglement between $\mathscr{H}_S$ and $\mathscr{H}_E$ distributes the information contained in $\mathscr{H}_S$ over the larger Hilbert space $\mathscr{H}$. This aspect of apparent information loss in quantum thermalization is also called "information scrambling" by the Hamiltonian for the full system. Thus, we see that thermalization, entanglement production, and information scrambling are inextricably linked in a quantum many-body system.

1.3.2 Fast scrambling

We recall from Sec. 1.1 that the question of *whether* thermalization occurs in a system may depend on the specific properties of the observables within $\mathscr{H}_S$, contained within Eq. (1.36). One nice simplification that occurs in this setting is that because we are concerned about *all* observables in the subsystem, we do not need to focus on the matrix elements of specific observables in the energy eigenbasis as in ETH (see Eq. (1.4) and the following discussion), but can directly look at fluctuations in the entanglement structure of the $|E_n\rangle$ within the subsystem itself. This "observable-independent" extension of ETH is called subsystem ETH [52], which we will return to in Chap. 8. However, given that it still depends on the choice of factorization into subsystems $\mathscr{H}_S \otimes \mathscr{H}_E$, we should not expect the energy eigenvalues E_n to play a direct role in the above question from the ergodic theory standpoint.

[6] Not to be confused with the exponent β_{RMT} occurring for the Vandermonde determinant in the Haar measure (see Eq. (1.14)); traditionally, the notation β is extremely standard for both quantities, and we use the subscript for the exponent to help differentiate them.

There is a closely related question in which the E_n will play a direct and crucial role, which is extremely relevant from the standpoint of an ergodic theory of quantum dynamics. There are several, essentially equivalent, ways of phrasing this question:

- If thermalization occurs in the subsystem, what is the earliest time t_s after which Eq. (1.35) can be satisfied in $\mathscr{H}_S$?
- What is the fastest rate at which entanglement can be generated between $\mathscr{H}_S$ and $\mathscr{H}_E$?
- What is the maximum allowed speed of information scrambling in $\mathscr{H}_S$?

The last phrasing, in particular, happens to be of interest in models of many-body quantum dynamics that are believed to capture aspects of quantum information loss in black holes [53,54]. The exact nature of these models with a precise correspondence to black holes is not rigorously established, and the specific details will not concern us here (although some aspects are reviewed in App. C.1 for completeness). What is of significant interest, however, is that these considerations suggest an N-dependent lower limit on t_s, namely that $t_s \geq O(\log N)$.

For a fully quantum mechanical argument that results in the same bound, Refs. [53, 54] also consider the problem of k-local quantum circuits (which are also believed to provide good models for scrambling dynamics in the black hole context). These are discrete-time systems, where each qubit can interact with at most $(k-1)$ other qubits over any discrete time step. Thus, in each successive time step, the information in one qubit can spread to at most k^t qubits. In that case, for the information in one qubit to spread to all of the N qubits in the system, one needs $k^t \geq N$, or $t \geq \log N / \log k$. As long as $k = O(1)$, we heuristically get the same $t_s \geq O(\log N)$ bound on information scrambling. However, for k that scales with N, this is clearly violated: a CUE unitary has $k = N$, and can scramble the entire system in a single time-step. Finally, for Hamiltonian systems (with $k = O(1)$-local terms) involving a finite-temperature bath, it is believed that the time t_s must be measured in units of the

inverse temperature β for this bound, i.e., $t_s/\beta \geq O(\log N)$.

From these considerations, Ref. [54] conjectured the following (adapted to our present context):

Conjecture 1.2 (Fast scrambling conjecture)**.** *(a) A "physical" (e.g., k-local) N-particle Hamiltonian system at an inverse temperature β can only scramble information after a time $t_s/\beta \gtrsim O(\log N)$. (b) Quantum many-body systems that model the scrambling dynamics of black holes saturate this bound.*

As with the quantum chaos conjecture (1.1), we should not take this to be a mathematical conjecture, but as a guiding principle for the kind of physical statements we should focus on. For example, the argument leading to this conjecture does not specify a precise notion of scrambling (but strongly suggests entanglement generation [53], as we consider for the quantum thermalization problem), nor does it rigorously specify which physical systems should qualify.

However, we can approach this question from a general viewpoint: irrespective of what kind of physical systems one is interested in, what kind of bounds on thermalization/scrambling/entanglement generation (whether trivial or non-trivial) are allowed by the mathematical structure of quantum mechanics itself? This does not require a specific choice of accessible observables, and is a natural question to attempt to tackle from a quantum "ergodic theory" standpoint.

As a hint towards how the spectrum influences early-time dynamics, there is already a well-known family of results that relates the width ΔE of the energy level distribution in an initial state $|\psi\rangle$, with the time Δt it takes for the state to evolve to a different state of low overlap with $|\psi\rangle$: the energy-time uncertainty principle (see Refs. [55–60] for more precise formulations):

$$\Delta E \, \Delta t \gtrsim \text{const.} \tag{1.37}$$

However, the speed limit $\Delta t \gtrsim \text{const.} \Delta E^{-1}$ does not adequately capture any potential dependence of this time on N (beyond what may be implicit in ΔE and the choice of state). This makes the standard

energy-time uncertainty principle almost entirely trivial for many-body systems, where due to the interaction of several particles, most time scales of relevance to quantum dynamics Δt grow with N relative to ΔE^{-1} (i.e., $\Delta E \Delta t \to \infty$ as $N \to \infty$, see e.g. Ref. [61]).

To get around this limitation, we leverage the detailed spectral information contained in the SFF to obtain tighter bounds than Eq. (1.37), which are useful for many-body processes (summarized in Sec. 1.4). Our approach to the problem is described in detail in Chaps. 3 and 4, which establish a precise version of statement (a) of Conjecture 1.1, as well as a set of counter-examples to statement (b) based on properties of the energy spectrum required by analogy with black holes.

1.4 Walkthrough and organization of this Dissertation

We have now set up the necessary background material for the rest of this Dissertation, such as Boltzmann's ergodic hypothesis (Sec. 1.1.1), von Neumann's approach to quantum ergodicity and (eigenstate) thermalization (Sec. 1.1.2), quantum chaos and random matrix theory (Sec. 1.2), as well as the connections between quantum thermalization, entanglement generation and information scrambling in a many-body system (Sec. 1.3). We placed a particular emphasis on the Quantum Chaos Conjecture (1.1) and the Fast Scrambling Conjecture (1.2) as guiding principles for our approach to quantum ergodicity.

We are now in a position to summarize the main results and organization of this Dissertation. The technical results presented in this Dissertation are primarily organized into three Parts and a set of Appendices. Part I: *Theory of Quantum Dynamics and the Energy Spectrum* develops the main theoretical ideas that tackle Conjectures 1.1 and 1.2 from an ergodic theory standpoint. As a prelude to the detailed analysis in the subsequent chapters, and due to the length of this Part, Sec. 1.4.1 provides an intuitive birds' eye view of the main conceptual ideas introduced in Part I, and explains how they fit together (see Fig. 1.5 for a pictorial depiction). The Chapter-wise organization of this part is then

summarized in Sec. 1.4.2.

The subsequent parts deal with examples, applications, and supporting calculations that (sometimes retrospectively) connect to the developments in this part. Part II: *Quantum Systems Beyond the Chaotic-Integrable Dichotomy* consists of several examples that illustrate the need for the precise notions developed in Part I as well as a few other subtleties (summarized in Sec. 1.4.3). Part III: *Experimental Probes of Many-Body Ergodicity* illustrates an experimental measurement protocol for the SFF sensitive to quantum ergodicity and thermalization, and also accounts for the effects of decoherence (Sec. 1.4.4). Part IV: *Appendices* consists of several additional calculations and numerics supporting the contents of the other Parts (Sec. 1.4.5).

1.4.1 A walkthrough of key ideas: Quantum ergodicity, scrambling, and the energy-time uncertainty principle

As stated in Sec. 1.1.3, our main focus is to address the following question: given just an energy spectrum, what can we say about the dynamics of a quantum system? It is convenient to separate out two aspects: (1) the spectral fluctuations within small energy windows, which are responsible for the late-time quantum fluctuations such as the ramp and plateau in the SFF, and (2) the overall profile of the energy spectrum, responsible for the early-time decay of the SFF from $K(0) = 1$. We will use (1) to address the quantum chaos conjecture 1.1, and we will tie (2) to the fast scrambling conjecture 1.2.

At a qualitative level, our entire approach can be summarized as different refinements of the energy-time uncertainty principle, according to which a more well-defined energy spectrum implies slower quantum dynamics, or longer persistence of an initial state. Specifically, we show that (see Fig. 1.5 for a schematic):

1. **Late times: Spectral rigidity $\leftrightarrow$ Quantum ergodicity**. A well-defined short-range structure of the energy spectrum, i.e., small spectral fluctuations, implies a longer persistence of a dynami-

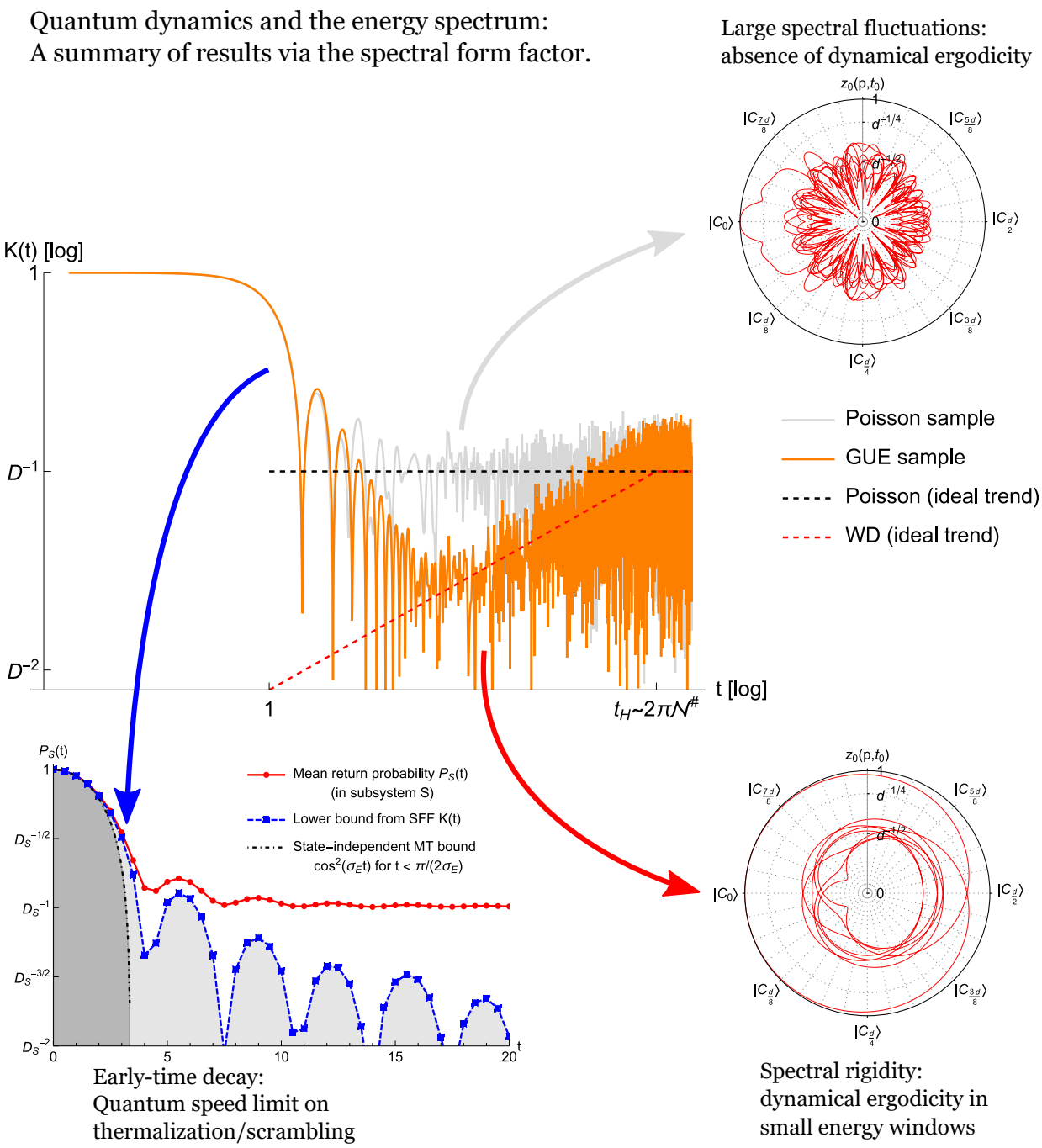


Figure 1.5: A summary of the primary results in Part I of this Dissertation, constituting parts of an ergodic theory of quantum information dynamics. For intuition, we refer to the different time scales of the SFF (see Fig. 1.4), corresponding to the structure of the spectrum at different energy scales, and find universal relations between this structure and forms of ergodic quantum dynamics. The late-time quantum fluctuations measure the presence/absence of quantum cyclic ergodicity in the Hilbert space (see Fig. 2.6 in Chap. 2), tying to the Quantum Chaos Conjecture 1.1, while the early-time decay of the SFF sets a universal quantum speed limit on scrambling (see Fig. 3.1 in Chap. 3) that ties to the Fast Scrambling Conjecture 1.2 (see Chap. 4).

.

cally ergodic structure in the Hilbert space.

2. **Early times: Narrower/sharper distribution of levels ⟷ Slower scrambling**. A well-defined long-range structure of the energy spectrum, i.e., a narrow distribution possibly with sharp features, implies a longer retention of memory of the initial state, and therefore slower scrambling.

We will provide more quantitative arguments for these below.

1.4.1.1 Quantum ergodicity

To motivate our strategy for (1), let us directly try to provide a quantum dynamical flavor to the level repulsion of random matrices, as sought in the quantum chaos conjecture. We return to the Vandermonde determinant $\mathcal{V}(e^{-iE_n})$ in Eq. (1.13) for a unitary $\hat{U}_H(1) = \hat{U}_H$, which weights the eigenvalue distribution of Haar random matrices. Given an unbiased superposition $|C_0\rangle = D^{-1/2}\sum_n |E_n\rangle$ of the energy eigenvectors, we consider D steps of the action of $\hat{U}_H$ on this state, for times $t = p$, $p \in \mathbb{Z}_D$:

$$\hat{U}_H^p |C_0\rangle = D^{-1/2} \sum_n e^{-iE_n p} |E_n\rangle, \quad p \in \{0, 1, \ldots, D-1\}. \tag{1.38}$$

The *volume* of the D-dimensional (complex) parallelepiped defined by these vectors in the Hilbert space is given by precisely by $\mathcal{V}(e^{-iE_n})$, up to proportionality:

$$\mathrm{Vol}\left(|C_0\rangle, \hat{U}_H|C_0\rangle, \ldots, \hat{U}_H^{D-1}|C_0\rangle\right) = D^{-D/2} \left| \det \begin{bmatrix} 1 & e^{-iE_0} & \ldots & e^{-iE_0(D-1)} \\ 1 & e^{-iE_1} & \ldots & e^{-iE_1(D-1)} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & e^{-iE_{D-1}} & \ldots & e^{-iE_{D-1}(D-1)} \end{bmatrix} \right| = \frac{\mathcal{V}(e^{-iE_n})}{D^{D/2}}. \tag{1.39}$$

Geometrically, our key takeaway from this is that random matrix ensembles such as CUE assign higher weights to unitaries $\hat{U}_H$ that lead to a larger volume of the parallelepiped defined by the discretized

trajectory of $|C_0\rangle$. A completely rigid ($\beta_{\text{RMT}} = \infty$) spectrum with the same range as CUE has, in contrast, $E_{n,C} = 2\pi n/D$; defining a unitary $\hat{U}_C$ with these eigenvalues (and the same eigenvectors $|E_n\rangle$), the set

$$\left\{|C_0\rangle, \hat{U}_C|C_0\rangle, \ldots, \hat{U}_C^{D-1}|C_0\rangle\right\} \tag{1.40}$$

defines a complete orthonormal basis for the Hilbert space, such that the corresponding parallelepiped volume is:

$$\text{Vol}\left(|C_0\rangle, \hat{U}_C|C_0\rangle, \ldots, \hat{U}_C^{D-1}|C_0\rangle\right) = D^{-D/2}\left|\det\begin{bmatrix} 1 & 1 & \ldots & 1 \\ 1 & \omega & \ldots & \omega^{D-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \omega^{D-1} & \ldots & \omega^{(D-1)(D-1)} \end{bmatrix}\right| = 1, \tag{1.41}$$

where $\omega = \exp(-2\pi i/D)$. The matrix in the above equation is in fact the unitary Discrete Fourier Transform (DFT) matrix [50], which attains the maximum possible volume of 1 for a parallelepiped of normalized vectors.

From Eqs. (1.39) and (1.41), we observe that by favoring larger values of $\mathcal{V}(a)$ random matrix distributions favor dynamical trajectories of $|C_0\rangle$ that are close to following a complete orthonormal basis. As the primary motivation for Chap. 2, we invert this observation and introduce a quantum dynamical notion of ergodicity as follows:

- The trajectory of a state $|C_0\rangle$ is "ergodic" if its trajectory remains "sufficiently close" to successive elements of a complete orthonormal basis.

As every observable in phase space should be associated with a complete orthonormal basis by the postulates of quantum mechanics [62], we consider this notion of ergodicity, which we call "cyclic ergodicity" (due to the cyclic permutation of the orthonormal basis implemented by $\hat{U}_C$), to be the

long-sought quantum analogue of Boltzmann's ergodic hypothesis.

However, directly realizing this connection in terms of Vandermonde determinants themselves seems challenging. Building on the above intuitive idea, but with separate quantitative techniques that are easier to connect to discrete forms of classical ergodicity, Chap. 2 identifies when a quantum system itself may be considered cyclic ergodic in a Hilbert space (namely, if the trajectory of all states $|C_k\rangle$ in a complete orthonormal basis follow successive elements in the same basis), and derives its quantitative relation to spectral fluctuations of the form considered in Sec. 1.2.4. We explicitly show that random matrix statistics is sufficient, but not necessary, for cyclic ergodicity. This establishes the first statement of the quantum chaos conjecture with a precise quantum notion of ergodicity, relevant for the late-time quantum fluctuations of the SFF.

1.4.1.2 Quantum speed limits and scrambling

Now, we wish to address the fast scrambling conjecture from properties of the energy spectrum corresponding to the early-time decay of the SFF. We recall from the discussion around Eq. (1.37) that the standard energy-time uncertainty principle does not have sufficient sensitivity for many-body processes. One way forward is to look for tighter bounds that take into account the full shape of the distribution of energy levels E_n, rather than just a single parameter such as ΔE. In Chap. 3, we show that the SFF can be used to obtain such a formulation of the energy-time uncertainty principle. We consider a complete set of observables $\hat{\Pi}_k$, such as the subsystem projectors $\hat{\Pi}_k = |k\rangle_S\langle k| \otimes \hat{\mathbb{1}}$ (see Eq. (1.36)) and initial states of the form $\hat{\rho}_k = \hat{\Pi}_k/D_E$, as in the quantum thermalization problem. We obtain the result that on average (in any basis $|k\rangle_S$ for the subsystem), the observable $\hat{\Pi}_k$ can not "forget" the initial state $\hat{\rho}_k$ at a speed faster than the decay of the SFF. In particular, the mean return probability $P_S(t)$ satisfies:

$$P_S(t) \equiv \frac{1}{D_S}\sum_k \mathrm{Tr}[\hat{\rho}_k(t)\hat{\Pi}_k(0)] \geq K(t). \tag{1.42}$$

This inequality generalizes a bound on a quantum dynamical feature called "cyclic aperiodicity" (connected to mixing, see Eq. (1.34)) in Chap. 2. It is also a more refined formulation of Eq. (1.37), as can be seen by noting that $K(t)$ initially decays over a time window of $\sim 1/\Delta E$, while the decay of $P_S(t)$ is a measure of Δt. Further, $K(t)$ in this expression is not restricted to unitary dynamics, but can be defined for any time-dependent completely positive quantum operation [50]. The crucial advantage of Eq. (1.42) is that $K(t)$ can continue to have a nontrivial asymptotic decay, determined by the structure of the energy spectrum. As the subsystem has Hilbert space dimension D_S, we want $P_S(t) = \mathrm{Tr}[\hat{\rho}_{\mathrm{th}}\hat{\Pi}_k] \sim D_S^{-1}$ for scrambling to the maximally mixed state. Thus, this can only occur if $K(t) \leq D_S^{-1}$, giving a bound on the scrambling time (assuming continuity):

$$t_s \geq \max K^{-1}[D_S^{-1}], \tag{1.43}$$

where $K^{-1}(x)$ denotes the set of times where $K(t) = x$ (with $t < T_D$ implicit).

In Chap. 4, we generalize the above result to an initial state $\hat{\rho} = |k\rangle_S\langle k| \otimes \hat{\rho}_{\beta E}$ where the external component $\mathscr{H}_E$ to the subsystem can be in any state, including one with a finite temperature β^{-1} (with a suitable generalization $K_\beta(t)$ of the SFF). Subsequently, analyticity properties of the SFF (which, at this time, do not seem to admit a short intuitive explanation) give the following (qualitative — we drop several quantities from the expression for intuitive accessibility) restriction:

$$K_\beta(t) \gtrsim \exp\left[-\exp\left(\frac{t}{c_1\beta}\right)\right] \tag{1.44}$$

(schematic; see Eqs. (4.17) and (4.21) for exact expressions).

Here, the parameter c_1 is determined by the region of analyticity in complex time. The key is the double-exponential structure: substituting this in Eq. (1.43) inverts the SFF to a double-logarithmic structure, giving $t_s/(c\beta) \gtrsim \log\log D_S = \log N_S$ as in statement (a) of the Fast Scrambling Conjec-

ture 1.2, as a simple mathematical property of scrambling independent of specific models.

However, we emphasize that the actual physical result is not in the N_S-dependence: parameters not included in the above schematic (as well as c_1) could have other forms of N-dependence that are not apparent in this semi-qualitative discussion. The key physical results instead directly connect the scrambling time t_s to the mean return probability of scrambled states $P_{S,\mathrm{scr}} = D_S^{-1} \sum_k \mathrm{Tr}[\hat{\rho}_{\mathrm{th}(k)} \hat{\Pi}_k]$, or their (2nd Rényi) entanglement entropy [63] $S_2 \sim -\ln \mathrm{Tr}[\hat{\rho}_{\mathrm{th}}^2]$, as follows:

$$t_s/\beta \gtrsim c_1 \ln\left(\ln P_{S,\mathrm{scr}}^{-1}\right) + \ldots, \tag{1.45}$$

$$t_s/\beta \gtrsim c_1 \ln S_2 + \ldots. \tag{1.46}$$

These dependencies hold irrespective of the scaling of other parameters; however, there can be thermodynamic regimes where the N-dependence is indeed given precisely by the conjectured $t_s/\beta \gtrsim O(\log N)$. Even in this regime, it should be considered quite surprising that we can prove a universal result, which happens to be similar to one originally motivated by entirely different arguments involving k-locality (mentioned in Sec. 1.3). See Chap. 4 for more precise statements.

For statement (b), we note in Chap. 3 that certain models such as the Sachdev-Ye-Kitaev models have a sharp edge in the spectrum, *due* to their correspondence with black holes (see also Chap. 4). This leads to $K(t) \sim 1/t^\alpha$, and therefore $t_s \sim \exp(N_S)$. Thus, scrambling in these systems is exponentially slow in the sense of entanglement generation (which carries over to the finite temperature setting of Chap. 4). This actually provides a set of counter-examples to statement (b) [as formulated for scrambling via entanglement generation].

For a Hamiltonian system, all these results make no assumptions other than the existence of the energy levels E_n. In this sense, our approach to the quantum chaos and fast scrambling conjectures fits naturally within the framework of an ergodic theory of quantum dynamics outlined in Sec. 1.1.3.

1.4.2 Part I: Theory of Quantum Dynamics and the Energy Spectrum

In Part I: Theory of Quantum Dynamics and the Energy Spectrum, we derive universal relations between the distribution of energy levels and the dynamics of quantum states that apply to every Hamiltonian (and sometimes, more general) quantum system. In doing so, we address the following questions:

1. *Which quantum systems are dynamically ergodic in the sense of Boltzmann?* In Chap. 2, we introduce quantum dynamical notions of ergodicity, and establish their quantitative connection to spectral fluctuations. We derive the spectral fluctuations of random matrices as a special case of this connection, and also show applicability to non-random matrix systems. We therefore address the long-standing question of what form of ergodicity is associated with the energy spectrum in the quantum chaos conjecture, while generalizing it beyond random matrices. (based on Ref. [1])

2. *How does the overall shape of the spectrum affect quantum dynamics?* In Chap. 3, we show that the early-time behavior of the (suitably generalized) SFF sets a universal quantum speed limit on arbitrary completely positive quantum operations, including Hamiltonian systems. This builds on one of the central concepts introduced in Chap. 2. We show that such a speed limit is significantly tighter than comparable versions of the energy-time uncertainty principle, and for the first time can be applied to many-body systems to constrain timescales that grow with the system size, without assuming specific types of interactions. This introduces a formalism to tackle the fast scrambling conjecture. (based on Ref. [2])

3. *What is the fastest possible rate of entanglement generation?* In Chap. 4, we apply the formalism developed in Chap. 3 to thermalization via arbitrary Hamiltonian interactions with a bath in an arbitrary initial state, including finite temperature states. From mathematical properties

of the SFF, we prove a precise quantitative statement of the fast scrambling conjecture, while generalizing it to systems with arbitrary nonlocal interactions. This also establishes a universal bound on the earliest possible time after which equilibrium statistical mechanics (where the system appears to permanently remain in a thermal macrostate except for quantum recurrences) can apply for the most general quantum mechanical Hamiltonian. (based on Ref. [3])

1.4.3 Part II: Quantum Systems Beyond the Chaotic-Integrable Dichotomy

In Part II: Quantum Systems Beyond the Chaotic-Integrable Dichotomy, we consider several examples of quantum systems where the quantum chaos conjecture (including the Berry-Tabor statement) and the fast scrambling conjecture (at least in the sense of which systems saturate the bound) do not hold in the most obvious and commonly expected ways. These examples illustrate the need for carefully paying attention to the intricate relationships between different measures of spectral statistics and quantum dynamics, and discourage the blanket classification of certain properties as "signatures of quantum chaotic systems" or "signatures of quantum integrable systems", particularly without reference to dynamical definitions of these notions. These examples are:

1. *Non-interacting fermions without (fully) Poisson statistics:* In Chap. 5, we show that formally fully integrable many-fermion systems do not satisfy all the assumptions that motivated the Berry-Tabor conjecture, essentially due to the Pauli exclusion principle. We show that they can display vestiges of spectral rigidity in their energy levels if the non-interacting single-particle spectra have rigidity, despite their integrability. This quantitatively explains an unexpected exponentially growing (rather than linear) ramp in the SFF of the non-interacting SYK-2 model. (based on parts of Ref. [4])

2. *Ergodicity, scrambling and chaos in interacting SYK:* In Chap. 6, we use the *interacting* SYK model, which is believed to be "maximally chaotic" and among the "fastest scramblers", to directly il-

lustrate different aspects of the theory developed in Part I. We show how the high-density sharp edges of its spectrum, crucial to its connection to black holes, lead to unusually slow thermalization even compared to certain other run-of-the-mill many-body spectra, via the bound of Chap. 3 (which generalizes to that of Chap. 4). We additionally provide a numerical illustration of ergodicity in the sense of Chap. 2 in this model, showing the selection of energy windows and how to account for discrete symmetries. (based on Ref. [2], with an expanded connection to Chap. 2)

3. *Spectral statistics in quantized Bernoulli shifts:* In Chap. 7 we numerically examine violations of the quantum chaos conjecture in different quantizations of maximally chaotic, ergodic classical maps. We illustrate that no set of familiar classical ergodic properties and symmetries is sufficient to determine the quantum spectrum, highlighting the necessity of fully quantum notions as developed in Part I. We also numerically show that the connection between the energy spectrum and fully quantum ergodic properties introduced in Chap. 2 seems to be respected by these maps, up to number-theoretic fluctuations inherent to these quantizations. (based on parts of Ref. [5])

1.4.4 Part III: Experimental Probes of Many-Body Quantum Ergodicity

In Part III: Experimental Probes of Many-Body Quantum Ergodicity, we examine the problem of probing these theoretical developments in experiments on many-body quantum systems, providing a concrete measurement protocol for the SFF that can access eigenstate properties in addition to the eigenvalue distribution.

1. *Simultaneous measurement protocol for eigenvalues and eigenstates:* In Chap. 8, we describe an efficient measurement protocol that can simultaneously access the SFF together with its generalizations to subsystems, Partial SFFs (PSFFs). Focusing on the properties of the PSFF, we show that it can probe the entanglement structure of the energy eigenstates in regard to ETH, without

any observable-dependence beyond a choice of subsystem. This establishes a simultaneous experimental probe of quantum dynamical ergodicity as encoded in the energy spectrum, as well as the (eigenstate) thermalization of all observables within accessible subsystems. (based on parts of Ref. [6])

2. *Ergodicity in an open system of trapped ions:* In Chap. 9, we describe an implementation of the above protocol in a highly decoherent trapped ion system. We review previously known concepts of ergodicity for open quantum systems that can be probed in this experiment, and explicitly illustrate how the generalized spectral form factor can diagnose these properties. Unlike the more elaborate constructions of Part I, these notions of ergodicity in open quantum systems only require certain non-degeneracy properties of eigenvalues, and tie directly to the thermalization of observables in von Neumann's sense (Sec. 1.1.2).

1.4.5 Part IV: Appendices

In Part IV: Appendices, we include several supporting and supplementary calculations and illustrations for the chapters included in the other Parts. These are as follows:

1. Appendix A: Supporting calculations and numerics for Chap. 2. (based on Ref. [1])

2. Appendix B: Supplementary calculations for Chap. 2, providing an alternate viewpoint with similar conclusions.

3. Appendix C: Supplementary discussion and calculations for Chaps. 3 and 4.

4. Appendix D: Supporting calculations and numerics for Chap. 7. (based on Ref. [5])

5. Appendix E: Supporting calculations and numerics for Chap. 8. (based on Ref. [6])

6. Appendix F: Supporting calculations for Chaps. 8 and 9.

Part I

Theory of Quantum Dynamics and the Energy Spectrum

Chapter 2: Dynamical quantum ergodicity from energy level statistics

2.1 Introduction: What ergodic phenomena are encoded in the quantum energy levels?

2.1.1 Background and motivation

Ergodic theory [10–12] concerns itself with a study of the statistical properties of classical dynamical systems, centered around a mathematically precise classification of classical dynamics into different levels of randomness called the ergodic hierarchy [13, 48] (see Fig. 2.1). These levels, such as ergodic, mixing, K-mixing and others [10–13, 48] (in order of increasing randomness, discussed in more detail in Sec. 2.2.1), can be used to motivate different elements of classical statistical mechanics [13, 48]: ergodicity justifies the use of the microcanonical ensemble, and mixing the approach to thermal equilibrium, while K-mixing is responsible for chaotic dynamics.

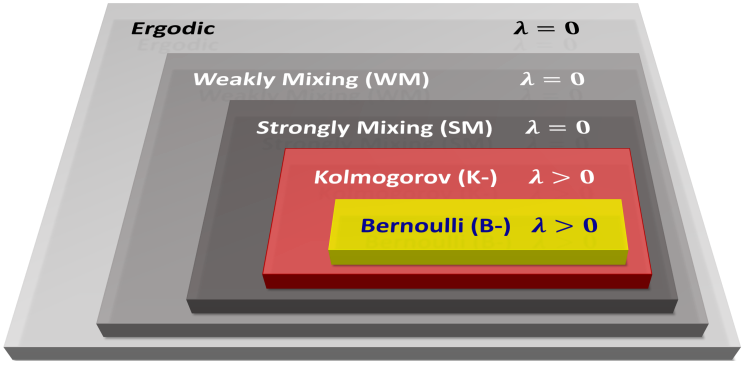


Figure 2.1: Venn diagram of the classical Ergodic Hierarchy [13, 48]. Each successive subset denotes a level of higher guaranteed randomness within the preceding level. λ indicates the maximal Lyapunov exponent, whose nonzero value is a defining signature of chaos [13, 48]. (Reused from Ref. [64], with permission)

In quantum mechanics, on the other hand, it has remained unclear if any pertinent "ergodic" dynamical properties can even be defined [28]. Instead, our present understanding of quantum statistical mechanics is founded on a much less precise, but empirically successful, connection to the statistical properties of random matrices [28, 36]. Direct contact with the thermalization of observables is made through a comparison of the energy eigenstates (or eigenvectors) in a "physical" basis of a system with random eigenvectors, via the Eigenstate Thermalization Hypothesis (ETH) [18–22, 52, 65, 66] and related approaches [17, 23, 24, 67–72]. While observables that satisfy ETH show thermalization behaviors resembling ergodicity and mixing [21, 22], every system also has observables that do not satisfy ETH, and it is not well understood from first principles which of these are to be regarded as the "physical" observables — except as determined empirically [21, 22, 69].

Such observable-dependent ambiguities are avoided in the comparison of the statistics of energy eigenvalues (i.e. level statistics) of a system with those of random matrices, at the apparent cost of direct dynamical relevance to thermalization. This approach is based on the observation that on quantization, typical classically non-ergodic systems show highly fluctuating energy spectra with Poisson (locally uncorrelated) level statistics [34], while typical classically chaotic systems show rigid spectra with the local level statistics of Wigner-Dyson random matrices [30–33] (within the eigenspaces of conserved quantities associated with all classical symmetries preserved on quantization; emergent quantum symmetries and conserved quantities must also be accounted for, e.g., in certain quantizations of billiards on arithmetic domains and cat maps [28]). A semiclassical "periodic orbit" argument for Wigner-Dyson level statistics soon followed [39, 40] (with further developments in, e.g., Refs. [73–77]), assuming semiclassical contributions from a "uniform" distribution of isolated periodic orbits in the classical phase space of a K-mixing system [28]. However, Wigner-Dyson statistics has also been seen numerically [78–80] on quantization of merely mixing [79, 80] and merely ergodic [78, 80] systems without K-mixing/chaos, and even without periodic orbits [78, 80, 81]. These observations

highlight the need for a theoretical understanding of spectral rigidity with regard to ergodicity, without relying on K-mixing or periodic orbits (as also anticipated in Ref. [81]).

Similar trends of spectral rigidity being associated with "chaotic" behavior have been observed analytically (often using basis-dependent analogues of periodic orbit theory, for ensembles of systems) and numerically in fully quantum many-body systems [4, 21, 35, 38, 41, 82–91], where judgements of the "chaoticity" of a system have been largely based on intuition. It remains unclear exactly what kind of ergodic dynamics is represented by spectral rigidity *especially* in such fully quantum systems, in addition to those with a classical limit. While correlation functions of *local* observables have been rigorously characterized in a manner similar to the ergodic hierarchy in the specific case of dual-unitary quantum circuits [92–94], no direct link to level statistics is known even in this case. Further, the only known dynamical signature of spectral rigidity — a (gradually weakening) *suppression* of extremely small (vanishing in the thermodynamic limit) quantum fluctuations representing recurrences at late times, called the "ramp" [82], appearing in autocorrelation functions [86, 95] and quantum simulation protocols designed specifically to measure spectral rigidity [6, 96] — is a subleading effect showing no transparent connection to the ergodic hierarchy. The fundamental open problem indicated here, of understanding the precise relationship between ergodicity and the basis-independent energy level statistics of a general quantum dynamical system, forms the central motivation for this work.

2.1.2 Summary of results

In this work, we develop an approach that provides a first general answer to the above problem within the fully quantum regime, by introducing suitable ergodic properties (independent of "physical" bases or observables) in the Hilbert space of a quantum dynamical system and deriving their formal connection to energy level statistics. Qualitatively, our approach is based on the following observations: 1. Cyclic permutations of a discrete set of states are the only invertible discrete processes (in

other words, permutations of states) whose repeated action is ergodic, i.e., every initial state evolves into every other state in the set; 2. A quantum cyclic permutation of an orthonormal basis of states is a unitary operator whose eigenvalues are regularly spaced (roots of unity); 3. Spectral rigidity, as usually studied in quantum chaos, is essentially a measure of how close a given energy spectrum is to a regularly spaced spectrum. Thus, by defining a quantum version of ergodicity in terms of "closeness" to cyclic permutation dynamics in a finite-dimensional Hilbert space (further justified by a similar approach to classical ergodicity in the mathematical literature [12, 97, 98]), we show that spectral rigidity is most directly a measure of this form of ergodicity in any quantum dynamical system. In physical terms, the suppression of recurrences of initial states due to spectral rigidity, indicated by the "ramp" in quantum fluctuations, is what allows certain initial states to "cyclically" *visit* other states under unitary dynamics (over the time scale of the ramp) due to the conservation of probability. This "cyclic" form of ergodicity encoded in the energy level statistics is, remarkably, quite distinct from the familiar quantum thermalization process of the *spreading* of an initial state over a "physical" basis (which occurs before most of the ramp [38]) associated with random eigenstates.

Formally, a rigorous treatment of classical ergodicity in terms of closeness to cyclic permutations of discretized cells in a continuum phase space (or "cyclic approximate periodic transformations") was developed in Refs. [97,98] (see also Refs. [11,12,99,100] for reviews and related results). Central to this treatment are a discrete counterpart to ergodicity and a discrete prerequisite for strong mixing (see Sec. 2.2.2 for the precise relationship), which we respectively call cyclic ergodicity and aperiodicity. An abstract "quantization" of these properties, where the discretized phase space cells correspond to pure states in an orthonormal basis (qualitatively representing the smallest Planck-sized phase space cells allowed by quantum mechanics), yields the desired quantum notions of ergodicity. These *extrapolate* the classical notions to fully quantum dynamics, rather than relying on any specific quantization techniques developed for various classical systems. Such an abstract approach is justified, and per-

haps even necessitated, by the established difficulty [71, 72] of relating energy levels to ergodicity in a "physical" basis (see Sec. 2.1.1). Correspondingly, our quantum definitions are in the context of a general autonomous (i.e. time-independent) unitary quantum dynamical system and do not rely on a classical limit.

Some key physical takeaways from our approach are:

1. Classical cyclic ergodicity and aperiodicity generalize to quantum mechanics in a surprisingly direct way (unlike the continuum ergodic/chaotic properties of Fig. 2.1) as fundamental quantum dynamical properties in the Hilbert space, including in systems without a classical limit. This allows a general observable-independent definition of quantum cyclic ergodicity and aperiodicity.

2. Quantum cyclic ergodicity requires the *existence* of an orthonormal basis where every time-evolving basis state "visits" every other (fixed) basis state successively in a cyclic sequence; quantum cyclic aperiodicity requires the *existence* of an orthonormal basis where no time-evolving basis state "visits" its original self; both requirements apply within a specified range of times. Here, a state "visits" another if their overlap exceeds that between two generic randomly chosen states.

3. Spectral rigidity directly characterizes the presence of quantum cyclic ergodicity and aperiodicity in a system, rather than conventional ergodicity, mixing or K-mixing as believed in the "quantum chaos conjecture" [31, 33]. This clarifies the precise dynamical role played by level statistics in relation to the ergodic hierarchy, beyond a generic "signature of quantum chaos" [28].

Thus, this work provides a system-independent framework that addresses the general connection between ergodic (quantum) dynamics and energy level statistics, in a way that captures the observable-independence of the latter. We further provide both analytical and numerical evidence for the applicability of this framework to different physically relevant types of energy level statistics: Wigner-Dyson

(seen near-universally in intuitively "quantum chaotic" systems) and Poisson (seen near-universally in "non-ergodic" systems) as associated with random matrix theory [28], and also the non-random-matrix-like spectral rigidity of quantized Kolmogorov-Arnold-Moser (KAM) tori (classically ergodic, non-mixing systems with no periodic orbits) and fluctuating spectra in rational tori (classically non-ergodic collections of periodic orbits) which occur as subsets of classically integrable systems [48]. A depiction of the resulting logical relationships is shown in Fig. 2.2.

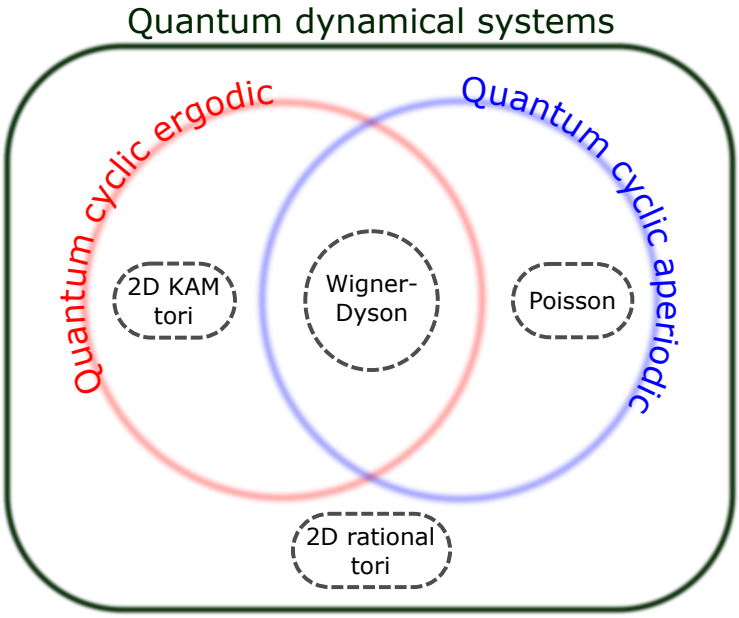


Figure 2.2: Depiction of the set of quantum dynamical systems in terms of the observable-independent ergodic properties proposed in this work, namely quantum cyclic ergodicity and aperiodicity. One distinct type of energy level statistics exemplifying each possibility for the presence or absence of these properties is shown (with dashed boundaries). The location of each such example in this diagram is analytically justified in the course of the Chapter.

2.1.3 Structure of this Chapter

The rest of this Chapter is organized as follows. The basic theory of observable-independent quantum ergodic properties and their connections to energy level statistics is set up in Secs. 2.2, 2.3, and 2.4. Sec. 2.2 reviews some necessary aspects of classical ergodic theory, including the use of cyclic permutations to "discretize" a dynamical system [11,12,97,98], and defines cyclic ergodicity and cyclic

aperiodicity (Definitions 2.1, 2.2) as discrete, primitive properties related to ergodicity and mixing that can be extended to quantum mechanics. Sec. 2.3 motivates and defines their analogues in quantum mechanics (Definitions 2.3, 2.4, 2.5). Sec. 2.4 shows that quantum cyclic ergodicity and aperiodicity in a system are quantitatively determined, through discrete Fourier transforms of the energy eigenstates[1] (Eq. (2.21)), by specific measures of energy level statistics: namely, mode fluctuations [79, 101, 102] (Eq. (2.37)) and the spectral form factor (SFF) [28] (Eq. (2.38)). Remarkably, Wigner-Dyson spectral rigidity emerges naturally at this stage, as a direct consequence of some simple dynamical conditions (Proposition 2.7); this allows an intuitive visualization of random matrix statistics in terms of ergodic Hilbert space dynamics (Fig. 2.6).

The subsequent sections are concerned with demonstrating applications to different situations of physical interest. Sec. 2.5 considers "typical" systems, relating quantum cyclic ergodicity to the "ramp" in the SFF via the conservation of probability (Sec. 2.5.3), and provides detailed analytical (Secs. 2.5.3, 2.5.4) and numerical (Fig. 2.7) evidence that quantum systems with Wigner-Dyson spectra are (quantum cyclic) ergodic and aperiodic, while those with Poisson spectra are aperiodic but not ergodic in the appropriate subspaces — covering the forms of level statistics associated with random matrices [28, 36]. Sec. 2.6 considers 2D linear flows on tori, for which we are able to explicitly construct cyclic permutations, and analytically prove both classical and quantum cyclic ergodicity and non-aperiodicity (Theorem 2.8) as well as higher-than-Wigner-Dyson spectral rigidity (Eq. (2.68)) for every 2D KAM torus (where the linear flow has an irrational frequency ratio), with an analytical argument that the non-ergodic and non-aperiodic rational tori (with a rational frequency ratio) have comparable, but not identical, spectral rigidity to Poisson. This covers some physically interesting systems with atypical level statistics, and suggests a wider applicability of cyclic ergodicity and

[1] We emphasize that the fact that the connection to level statistics is realized in specific bases, unambiguously specified relative to the energy eigenbasis, does not reduce the basis-independence of the definitions. This is in the same sense in which the energy eigenvalues are basis-independent, even though a unitary time evolution operator is only diagonal in the energy eigenbasis.

aperiodicity in characterizing quantum dynamics than random matrix theory, while establishing the possibility of analytical proofs of spectral rigidity for individual systems in this framework (which has been considered a mathematically and conceptually challenging problem [71, 72]). Finally, Sec. 2.7 discusses some insights about the connection between the above eigenvalue-based dynamical properties and thermalization as determined by eigenstates that may be gained from cyclic permutations, in a largely semi-qualitative manner that may motivate future rigorous work.

2.2 A short review of classical ergodic theory

In classical ergodic theory [10–13], one is concerned with dynamics on a phase space (or a smaller region of interest, such as an energy shell of a Hamiltonian system) $\mathscr{P}$, with an operator $\mathscr{T}^t : \mathscr{P} \to \mathscr{P}$ that evolves points in the space by the time t (which may be a continuous or discrete variable, corresponding to flows or maps). The main questions of interest are which regions of phase space are explored over time by an initial point, and how rapidly a typical point explores these regions. These questions are conveniently posed when there is a measure $\mu(A) \geq 0$ defined for subsets $A \subseteq \mathscr{P}$ that is preserved by time evolution, $\mu(\mathscr{T}^t A) = \mu(A)$ (in Hamiltonian dynamics, this measure is given by the phase space volume $\int_A \mathrm{d}^n q \mathrm{d}^n p$). An important feature of such systems is guaranteed by the *Poincaré recurrence theorem* [10–12]: for any $A \subseteq \mathscr{P}$ such that $\mu(A) > 0$, almost every point in A eventually returns to A, each within some (long) finite time (i.e. with the exceptions forming a set of measure zero). Given such a measure, how well an initial point explores the phase space is generally expressed through correlation functions of various sets, the behavior of which is classified into the ergodic hierarchy [13, 48]. In what follows, we normalize $\mu(\mathscr{P}) = 1$.

2.2.1 The classical ergodic hierarchy

We first ask whether almost all initial points explore every region of nonzero measure in $\mathscr{P}$. If so, the dynamics is said to be *ergodic* in $\mathscr{P}$. If not, $\mathscr{P}$ can be decomposed into (say) M subsets that are invariant under $\mathscr{T}$, i.e. $\mathscr{P} = \bigcup_{j=1}^{M} \mathscr{P}_j$ (each with a measure induced by μ), such that the dynamics is ergodic within each $\mathscr{P}_j$. Ergodicity in $\mathscr{P}$ can be shown to be equivalent [10–13, 48] to time-averaged correlations (e.g. presence of the system in a region $B \subseteq \mathscr{P}$) becoming independent of the initial condition (e.g. starting in the region $A \subseteq \mathscr{P}$) with only measure zero exceptions:

$$\lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \mathrm{d}t \; \mu\left[(\mathscr{T}^t A) \cap B\right] = \mu(A)\mu(B), \;\; \forall \; A, B \subseteq \mathscr{P}. \tag{2.1}$$

These measure zero exceptions could be isolated periodic orbits or other closed surfaces of lower dimension than $\mathscr{P}$. Here, we use $\mathrm{d}t$ either as a continuous integration measure or that corresponding to a discrete sum, depending on the domain of t.

Mixing is a property in which time evolution eventually becomes uncorrelated with initial conditions, and represents how rapidly typical points explore a phase space region $\mathscr{P}$ on which time evolution is ergodic. Generically, a system is mixing if any ensemble A with sufficiently many 'neighboring' initial states (i.e. of nonzero measure) in $\mathscr{P}$ eventually spreads out equally according to certain criteria over every region of $\mathscr{P}$ (with deviations from this behavior allowed for a vanishing fraction of times in the case of *weak* mixing). The simplest such criterion is expressed in terms of two element correlation functions eventually becoming uncorrelated [10–13, 48]:

$$\lim_{t\to\infty} \; \mu\left[(\mathscr{T}^t A) \cap B\right] = \mu(A)\mu(B), \;\; \forall \; A, B \subseteq \mathscr{P}, \tag{2.2}$$

and is conventionally merely called mixing (weak mixing corresponds to the limit converging except

at a vanishing fraction of times, instead of exactly for all $t \to \infty$ [10,13,48]). This can be extended to higher order correlation functions [12,103], and the dynamics is said to be *K-mixing* when all higher order correlation functions become uncorrelated in the above sense [12,13,48], which leads to chaotic behavior, e.g. nonzero Lyapunov exponents [13,48]. These criteria form a hierarchy in the sense that K-mixing implies mixing, which implies ergodicity [13,48]. Additional levels of randomness may also be considered [13,48]; see Fig. 2.1 for a depiction of the hierarchy of Ref. [13].

It is interesting to note that if one defines a unitary operator $U_{\mathscr{T}}$ induced by $\mathscr{T}$ on phase space functions $f(x \in \mathscr{P})$ via $U^t_{\mathscr{T}} f(x) \equiv f(\mathscr{T}^t x)$ (Koopman and von Neumann's Hilbert space representation of classical mechanics [10–12, 98–100]), some of these properties can be translated to those of the eigenvalues and eigenfunctions of $U_{\mathscr{T}}$, whose direct extensions to quantum mechanics have been previously considered [104]. For instance, ergodicity translates to non-degenerate eigenvalues with eigenfunctions of uniform magnitude, and weak mixing to a continuous spectrum with no non-constant eigenfunction, of $U_{\mathscr{T}}$ [10]. For a discrete and finite quantum spectrum that corresponds to phase spaces or energy shells of finite measure by Weyl's law [28], the eigenvalues are almost always non-degenerate (i.e. are non-degenerate or can be made so by infinitesimal perturbations) and the spectrum is necessarily discrete, prompting us to seek alternate avenues in which the above properties are at best emergent in the classical limit.

2.2.2 Discretizing ergodicity with cyclic permutations

We eventually want to understand how quantum mechanics with its discrete set of energy levels can lead to ergodic and mixing behaviors, defined classically for continuous systems. A useful bridge between the continuum and discrete descriptions is offered by the technique of discretizing an arbitrary dynamical system in terms of cyclic permutations, which have been studied as "cyclic approximate periodic transformations" in Refs. [11, 12, 97–100]. Here, we discuss and adapt the elements of this

framework that are most relevant for our purposes, following Ref. [98], leading to the definition of cyclic ergodicity and cyclic aperiodicity respectively as a discrete version of ergodicity and a discrete prerequisite for strong mixing. These ideas are then illustrated with simple examples, which are directly used later in the study of quantized KAM tori in Sec. 2.6.

A simple motivation for studying cyclic permutations is as follows. All invertible maps on a discrete set of states act as permutations of these states. Every permutation can be decomposed into a set of cyclic permutations, each acting on a separate subset of states. Among these, the only *ergodic* permutations — under whose repeated action every discrete state visits every other state — are cyclic permutations of *all* states with no further decomposition into subsets. This suggests the strategy of probing the ergodicity of an invertible autonomous dynamical system by comparing its dynamics to the repeated action of a cyclic permutation on some discretized states, in the continuum limit.

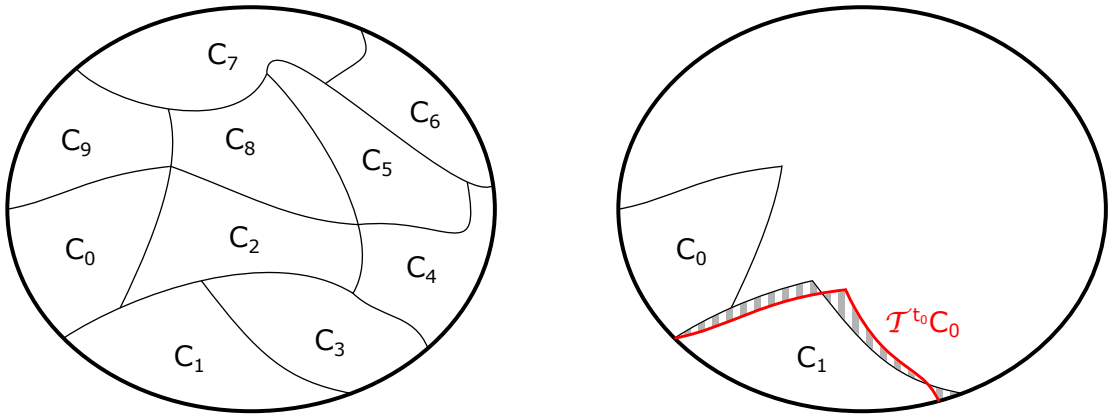
To this end, let $\mathtt{C} = \{C_k\}_{k=0}^{n-1}$ be a decomposition of the phase space $\mathscr{P}$ into a large number of μ-disjoint (i.e. with measure zero intersection) closed sets of identical measure, $\mu(C_k) = 1/n$; we will implicitly consider $\mathtt{C}$ to be part of a well-defined sequence of such decompositions, with $n \to \infty$ through a subset of the positive integers. Introduce a time evolution operator $\mathscr{T}_C$ on $\mathscr{P}$ that effects a cyclic permutation of the C_k, i.e. $\mathscr{T}_C C_k = C_{k+1}$ (with $(n-1)+1 \equiv 0$ i.e. the addition is modulo n). For some choice of discrete time step t_0, we would like to see if successive actions of $\mathscr{T}^{t_0}$ on the C_k closely resemble the cyclic permutation effected by $\mathscr{T}_C$. Thus, as a measure of how well $\mathscr{T}_C$ approximates $\mathscr{T}^{t_0}$ over a single time step, we define the single-step error of the permutation (differing from that in Ref. [98] by the factor of $1/2$):

$$\overline{\epsilon}_C(t_0) = \frac{1}{2}\sum_{k=0}^{n-1} \mu\big[(\mathscr{T}^{t_0} C_k) \triangle C_{k+1}\big], \tag{2.3}$$

where $A \triangle B = (A \cup B) - (A \cap B)$. The error measures the fraction of $\mathscr{T}^{t_0} C_k$ that lies outside C_{k+1},

averaged over all initial sets C_k.

We will often directly call $\mathtt{C}$ the cyclic permutation in place of $\mathscr{T}_C$ for brevity, as any $\mathscr{T}_C$ that cyclically permutes the elements of $\mathtt{C}$ has the same error. We also note that we can choose t_0 to depend on n, in particular with $t_0 \to 0$ as $n \to \infty$ for a flow with a continuous time variable, so that the discrete steps $\mathscr{T}^{pt_0}$ with $p \in \mathbb{Z}$ approach the continuous time flow $\mathscr{T}^t$ with $t \in \mathbb{R}$ as the decomposition becomes more refined with increasing n. As a simple example, the rotation $\mathscr{T}^t\theta_0 = \theta_0 + \omega t$ (modulo 2π, which will be left implicit in what follows) on a 1D circle with angular coordinate $\theta \in [0, 2\pi)$ is approximated by the n-element cyclic permutation $\mathtt{C}$ comprising of equal intervals of the circle $C_k = \{\theta \in [2\pi k/n, 2\pi(k+1)/n]\}$, with *zero* error if $t_0(n) = 2\pi/(n\omega)$. In contrast, higher dimensional ergodic systems typically take infinitely long to explore all of $\mathscr{P}$, and we expect $t_0 n \to \infty$ in addition to $t_0 \to 0$ in such cases. A schematic of a generic cyclic permutation is depicted in Fig. 2.3.



(a) Partitioning of phase space into $\{C_k\}_{k=0}^{10-1}$.

(b) Contribution to error from $(\mathscr{T}^{t_0}C_0)\triangle C_1$ (shaded region).

Figure 2.3: Schematic depiction of an $(n = 10)$-element cyclic permutation for some phase space $\mathscr{P}$ (interior of ellipse).

The single-step error $\overline{\epsilon}_C(t_0)$ can serve as a probe of ergodicity. Refs. [97, 98] showed that for a given dynamical system, the *existence* of any cyclic permutation with error $\overline{\epsilon}_C(t_0) < 2/n$ guarantees ergodicity as long as all nonzero measure sets can be constructed as unions of the C_k in the $n \to \infty$ limit, while $\overline{\epsilon}_C(t_0) < 1/n$ with $t_0 n \to \infty$ rules out (strong) mixing; a version of their proof is recounted

in Appendix A.1 for the interested reader. However, these results allow room for cyclic permutations with a larger single-step error to lead to ergodicity or prohibit mixing. Here, we define *cyclic ergodicity* (a notion implicit in their proofs) and *cyclic aperiodicity* (based on an observation in Ref. [98]) as more fundamental discrete dynamical properties based on cyclic permutations that can be used to partly characterize the ergodic and mixing nature of dynamical systems as described below.

2.2.2.1 Classical cyclic ergodicity and aperiodicity

Cyclic ergodicity is a discrete counterpart to ergodicity in a continuous system, that directly formalizes the notion of "(almost) all points visiting the neighborhood of every other point" in $\mathscr{P}$; the additional cyclic ordering of the "points" ensures that the discretized dynamical system is invertible (and first-order/"memoryless"). In terms of the decomposition $\mathbb{C}$, we want every C_k to visit (i.e. have nonzero intersection with) every other $C_{k'}$ at least once during its future and past time evolution in steps of t_0. The cyclic structure (in the ordering of the $C_{k'}$s for all C_ks) can also be motivated as follows: if a given initial set C_j intersects almost all of C_k after one step of $\mathscr{T}^{t_0}$ and then C_ℓ after two steps, C_k must intersect almost all of C_ℓ after one step. Formally, we have the following definition that considers a C_k to visit another if their *fractional* overlap (overlap divided by $1/n$, the measure of each C_k) is larger than a given precision $\eta(n)$ as $n \to \infty$, such as $\eta(n) = n^{-1}$ (depicted schematically in Fig. 2.4):

Definition 2.1 (**Classical cyclic ergodicity**). *A cyclic permutation $\mathbb{C}$ shows cyclic ergodicity with a given precision $\eta(n) > 0$ iff any element $C_j \in \mathbb{C}$ sequentially intersects a sufficiently large fraction of every other $C_k \in \mathbb{C}$ at least once under (future and past) time evolution:*

$$n\mu\left[(\mathscr{T}^{pt_0} C_j) \cap C_{j+p}\right] > \eta(n) \text{ as } n \to \infty, \ \forall \ j \text{ and } |p| \leq \frac{n}{2}, \tag{2.4}$$

where p represents the number of integer steps of time evolution in units of t_0.

For a given dynamical system, *if* there exists a sequence (implicitly labeled by n) of cyclic permutations $\mathtt{C}$ such that

1. $\mathtt{C}$ shows cyclic ergodicity for at least one choice of $\eta(n)$, and

2. $\mathtt{C}$ satisfies the *generating property*: every nonzero measure set in $\mathscr{P}$ contains at least one C_k in the $n \to \infty$ limit,

then it follows rigorously that every nonzero measure set visits every other, and $\mathscr{T}^{t_0}$ is therefore ergodic in $\mathscr{P}$ (see Refs. [12,98] and Appendix A.1; we also recall that generally $t_0 \to 0$ for a continuous flow). It is also convenient to call the system cyclic ergodic in $\mathscr{P}$ for a given $\eta(n)$, without reference to the generating property, if it admits a sequence of cyclic ergodic $\mathtt{C}$ with precision $\eta(n)$ (mainly to anticipate its quantum counterpart in Sec. 2.3).

For strong mixing (Eq. (2.2)), we require any initial region e.g. C_k to become spread out over all of $\mathscr{P}$ as $t \to \infty$. This requires that each C_k *on average* intersects no more than a vanishing fraction of itself in the $n \to \infty$ limit (so that $\mathscr{T}^t C_k$ is not preferentially distributed in C_k for almost all C_k), for any t with $t \to \infty$. Correspondingly, we call a system cyclic aperiodic if every sequence of cyclic permutations satisfies cyclic aperiodicity (a necessary condition for strong mixing):

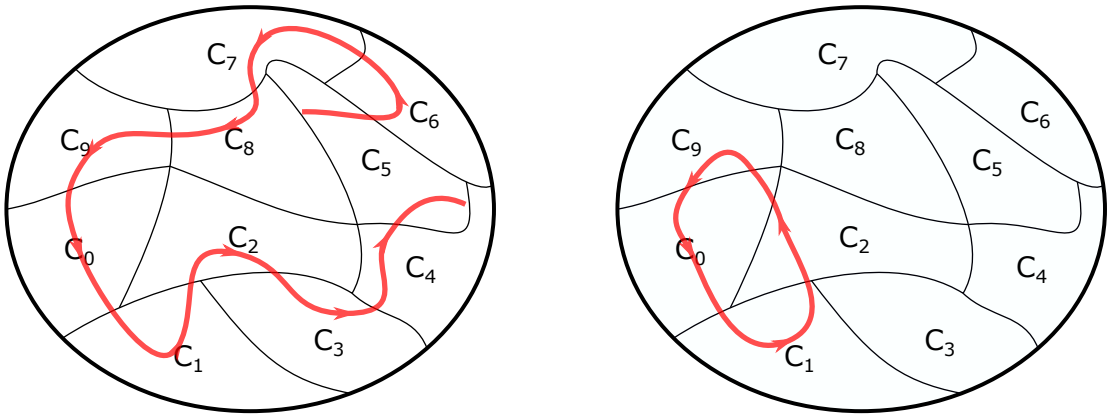
Definition 2.2 (**Classical cyclic aperiodicity**)**.** *A cyclic permutation $\mathtt{C}$ shows cyclic aperiodicity iff $\mathscr{T}^t C_k$ never returns to intersect a non-vanishing fraction of C_k on average, at all times later than $\bar{t}(n)$:*

$$\sum_{j=0}^{n-1} \mu\left[(\mathscr{T}^t C_j) \cap C_j\right] \to 0 \text{ as } n \to \infty, \ \forall\ t > \bar{t}(n), \tag{2.5}$$

for every $\bar{t}(n) \to \infty$ as $n \to \infty$.

In light of these definitions, Refs. [97, 98] effectively show that $\varepsilon_C(t_0) < 2/n$ for a cyclic per-

mutation implies cyclic ergodicity, while $\varepsilon_C(t_0) < 1/n$ with $t_0 n \to \infty$ implies a violation of cyclic aperiodicity (note that the reverse implication in both cases is not always true), as the error generated in each step is not sufficient to lead to zero overlap of $\mathcal{T}^{pt_0} C_j$ with C_{j+p} by $p = n/2$ (thereby maintaining cyclic ergodicity) or $p = n$ (thereby maintaining cyclic ergodicity and violating cyclic aperiodicity) respectively; see Appendix A.1. Thus, the existence of a sequence of cyclic permutations $\mathtt{C}$ with $\varepsilon_C(t_0) < 2/n$ or $\varepsilon_C(t_0) < 1/n$ for a dynamical system (satisfying the generating property, with $t_0 n \to \infty$) respectively implies that it is ergodic, or ergodic and not strongly mixing.



(a) Cyclic ergodicity.

(b) Non-ergodicity.

Figure 2.4: Schematic depiction of cyclic ergodicity and non-ergodicity for the cyclic permutation of Fig. 2.3. The trajectory may be thought of as the future and past history of the center of C_0, up to $\pm 5t_0$ (arrows indicate the forward flow of time).

2.2.2.2 Examples

Example 1: A simple illustration of these ideas is afforded by the example of rotations on a circle. A continuous rotation $\mathcal{T}^t \theta_0 = \theta_0 + \omega t$ is ergodic (any initial point covers the entire circle) and not-mixing (the angular width of any initial region is preserved); as discussed after Eq. (2.3), the n-element cyclic permutation $\mathtt{C}$ with

$$C_k = [2\pi k/n, 2\pi(k+1)/n] \text{ and } t_0 = 2\pi/(\omega n) \tag{2.6}$$

approximates this system with $\overline{\epsilon}_C(t_0) = 0$. $\mathtt{C}$ then shows cyclic ergodicity, and violates cyclic aperiodicity as $t \to \infty$ due to its periodicity, thereby implying ergodic and non-mixing behavior. It is also worth considering an additional static degree of freedom, e.g. a cylinder $\mathrm{Cyl} = \{(\theta, z)\}$ with $z \in [0,1]$ and $\mathscr{T}^t z = z$, in which case $\mathscr{T}$ is not ergodic on Cyl. A physical example of this type is the 1D harmonic oscillator of frequency ω, with conserved amplitude A (with, say, $z = \tanh A$) and phase $\theta(t)$. Here, the cyclic permutation $\overline{\mathtt{C}}$ comprised of "lengthwise strips" $\overline{C}_k = C_k \times [0,1]$ (where C_k is given by Eq. (2.6)) is cyclic ergodic for any given $0 < \eta(n) \leq 1$; however, arbitrary nonzero measure sets in Cyl (particularly those that do not span the entire range of z for all θ, e.g. sets with $z \in [0, 1/2]$) do not contain any $\overline{C}_k$. Thus, $\overline{\mathtt{C}}$ does not satisfy the generating property, remaining consistent with the non-ergodicity of $\mathscr{T}$ on Cyl in spite of its cyclic ergodicity.

Example 2: Somewhat more nontrivial is the discrete rotation on the circle,

$$\mathscr{T}^t \theta_0 = \theta_0 + 2\pi\alpha t, \ t \in \mathbb{Z}, \tag{2.7}$$

in steps of the angle $2\pi\alpha$ with $0 \leq \alpha < 1$, which is readily seen to be ergodic only for irrational α and decomposes into infinitely many periodic orbits for rational α (and is mixing in neither case). Here, the construction of cyclic permutations relies on the approximation of α by a rational number [105]. If $\alpha = (m + \delta_n)/n$ where m, n are coprime integers and $|\delta_n| < 1/2$, we can construct an n-element cyclic permutation given by the intervals

$$C_k = [2\pi km/n, \, 2\pi(km+1)/n] \text{ and } t_0 = 1,$$
$$\text{with error } \overline{\epsilon}_C(t_0 = 1) = |\delta_n|. \tag{2.8}$$

For all irrational α, we can find an infinite sequence of coprime m, n with $n \to \infty$ satisfying

$$|\delta_n| < \frac{c}{n}, \text{ for any constant } c \in \left[\frac{1}{\sqrt{5}}, 1\right), \tag{2.9}$$

by Dirichlet's and Hurwitz' theorems on Diophantine approximations [106]. The error then satisfies $\overline{\epsilon}_C(t_0) < 1/n$ in any such sequence, establishing ergodicity (as the C_k can be used to construct any finite interval as $n \to \infty$) as well as non-aperiodicity by the bounds [12, 98] discussed in Sec. 2.2.2.1 (see also Appendix A.1), with the latter implying the absence of mixing (as $t_0 n \to \infty$). This leaves the case of rational $\alpha = p/q$ with coprime p and q, for which it is useful to consider the regions $B_k = [2\pi kp/q, 2\pi(kp+1)/q]$; each point within any such B_k lies on a different periodic orbit and therefore cannot visit another point in the same B_k. Split any such B_k into two nonzero measure regions B_{k1} and B_{k2} which consequently never visit each other. Any cyclic permutation satisfying the generating property must possess some elements $C_{k1} \subseteq B_{k1}$ and $C_{k2} \subseteq B_{k2}$ as $n \to \infty$, which cannot visit each other by the preceding discussion; thus, no cyclic permutation satisfying the generating property can be cyclic ergodic for rational α.

In summary, this subsection discussed the connection between certain properties of discretized classical dynamics and levels of the ergodic hierarchy. In particular, the *existence* of at least one cyclic permutation satisfying cyclic ergodicity guarantees ergodicity (among regions of the phase space containing a discretized element), while mixing requires that every cyclic permutation satisfies aperiodicity (in the infinite time limit). We note that it is not generally known how to establish such properties for classical cyclic permutations except in cases with a sufficiently small single-step error [12, 97, 98, 105] (including the examples discussed above), but we will see that this difficulty is largely mitigated at a formal level after quantization.

2.3 Dynamical quantum ergodicity and cyclic permutations

This section aims to find parallels to the discretized classical ergodic properties of the previous section in quantum mechanics, which we propose as a starting point for a precise study of quantum ergodicity. In Sec. 2.3.1, we define quantum cyclic ergodicity and aperiodicity for cyclic permutations of orthonormal pure states in the Hilbert space, and subsequently the ergodicity and aperiodicity of a quantum system in any subspace of its Hilbert space — which is such that the energy levels of the system encode all the relevant properties. In Sec. 2.3.2, we briefly discuss the formal quantum analogues of the classical error bounds (see Sec. 2.2.2) which can be used to constrain the time evolution of a cyclic permutation based entirely on the single-step error in certain cases; a more general characterization of quantum cyclic permutations requires a quantitative study of their connection to energy level statistics, which is taken up in Sec. 2.4.

We consider a general autonomous quantum system with unitary time evolution. Let $\hat{U}_H(t) = \hat{U}_H^t(1)$ be the unitary time evolution operator, with D (possibly nonunique) eigenstates $\left\{|E_n\rangle\right\}_{n=0}^{D-1}$ and D (correspondingly, possibly degenerate) eigenvalues $e^{-iE_n t}$:

$$\hat{U}_H(t)|E_n\rangle = e^{-iE_n t}|E_n\rangle. \tag{2.10}$$

The time variable t can be chosen to be continuous or discrete, with E_n respectively corresponding to the eigenvalues of a Hamiltonian or eigenphases of a Floquet map. Without loss of generality, we will use terminology associated with Hamiltonians in what follows.

It is worth noting that such an autonomous unitary evolution is itself never ergodic in the Hilbert space (even after restricting to normalized states), but always has D conserved quantities $|\langle E_n|\psi(t)\rangle|^2$ for a time-evolving state $|\psi(t)\rangle$ [16]. Moreover, all systems with rationally incommensurate/generic energy levels (including the vast majority of those considered both "quantum chaotic"

and "integrable/non-ergodic") have no further conserved quantities, and consequently have the exact same number of ergodic subsets in the Hilbert space [16, 21, 22]. Thus, a different conceptual basis is necessary to define and understand quantum ergodicity in an observable-independent manner, while maintaining a connection to some meaningful notion of ergodicity. The main takeaway from this section is that suitably defined quantum cyclic permutations of pure states are a promising candidate for this purpose, allowing a natural quantization of cyclic ergodicity and aperiodicity.

Before embarking on a detailed discussion of quantum cyclic permutations (which may be defined in systems with or without a classical limit), we mention a semiclassical motivation for considering orthonormal pure state cyclic permutations. Weyl's law [28] (generally used for semiclassical calculations of the density of states) effectively assigns to each phase space region A a number of orthonormal pure states $P(A) \propto \mu(A)$ in the semiclassical regime; see also Refs. [26, 27] for a related "quantum measure algebra", and App. A.5 for a more formal justification. With $A = C_k$ in an n-element classical cyclic permutation $\mathbb{C}$, we have $\mu(C_k) = (1/n) \to 0$ classically as $n \to \infty$, suggesting that it is natural to associate the smallest number $P(C_k) = 1$ of pure states with each C_k, i.e. to consider pure state cyclic permutations in the fully quantum description to represent the classical $n \to \infty$ limit. The invertibility of the cyclic permutation in the discretized classical system translates to the unitarity of the associated quantum cyclic permutation of an orthonormal basis.

2.3.1 Pure state cyclic permutations for quantum dynamics

We work in an invariant subspace $\Sigma_d \subseteq \mathscr{H}$ (an 'energy subspace') spanned by any subset of suitably relabeled eigenstates $\{|E_n\rangle\}_{n=0}^{d-1}$; $\hat{U}_H(t)$ will henceforth refer to the restriction of the time evolution operator to Σ_d unless specified otherwise. In practice, Σ_d may be chosen depending on convenience to be e.g. in most cases, an energy shell of a physical system spanned by all levels with energies in a range $[E, E + \Delta E]$ (which is most likely to show "ergodicity" for any width ΔE less than

the energy scale to which spectral rigidity extends as discussed in Sec. 2.4.2), or the restriction of such a shell to a subspace with fixed values of conserved quantities showing spectral rigidity in systems with additional symmetries. The main question of physical interest is whether a physical system is ergodic within such a (restricted) energy shell. However, the formal considerations of this section apply quite generally to any energy subspace.

We seek pure state cyclic permutations that approximate $\hat{U}_H(t)$ within this energy subspace. To this end, let $\mathscr{C} = \{|C_k\rangle\}_{k=0}^{d-1}$ be an orthonormal basis spanning Σ_d with the unitary cycling operator $\hat{U}_C|C_k\rangle = |C_{k+1}\rangle$. The eigenvalues of $\hat{U}_C$ are necessarily distinct d-th roots of unity, $\{\exp(-2\pi i n/d)\}_{n=0}^{d-1}$. It is convenient to introduce the p-step persistence amplitudes of $U_H(pt_0)$ relative to the action of $\hat{U}_C^p$,

$$z_k(p,t_0) = \left|\langle C_{k+p}|\hat{U}_H(pt_0)|C_k\rangle\right|, \tag{2.11}$$

for some choice of t_0; these satisfy $z_k(p,t_0) \in [0,1]$, and represent the overlap amplitude between the time evolved $|C_k\rangle$ and the original $|C_{k+p}\rangle$. Then, we say that $\hat{U}_C$ approximates $\hat{U}_H(t_0)$ with p-step error

$$\varepsilon_C(p,t_0) = 1 - \left(\min_{k\in\mathbb{Z}_d} z_k(p,t_0)\right)^2, \tag{2.12}$$

where $\mathbb{Z}_n = \{0,\ldots,n-1\}$. A pure state approximation scheme for unitaries has been constructed in Ref. [100], in analogy with certain classical non-cyclic transformations (indirectly related to classical cyclic permutations [107]), to formalize results on e.g. the (non-)degeneracy of the classical unitary $U_{\mathscr{T}}$ in classical ergodic theory [10, 11, 98, 107]. As we will see in Sec. 2.4.1, the construction of pure state cyclic permutations as above allows us to go much further, and tackle non-trivial measures of the level statistics of $\hat{U}_H(t)$ that can e.g. distinguish between Wigner-Dyson and Poisson statistics, seen respectively in typical "quantum chaotic" and "non-ergodic" systems.

In analogy with the definitions for classical cyclic permutations (Eqs. (2.4) and (2.5)), we can

define cyclic ergodicity and cyclic aperiodicity for these pure state quantum cyclic permutations as below (see Fig. 2.5 for a schematic depiction, and Fig. 2.6 in Sec. 2.4.2 for examples with exact numerical data).

Definition 2.3 (**Quantum cyclic ergodicity**). *A pure state quantum cyclic permutation $\mathscr{C}$ shows cyclic ergodicity with precision $\eta(d)$ iff*

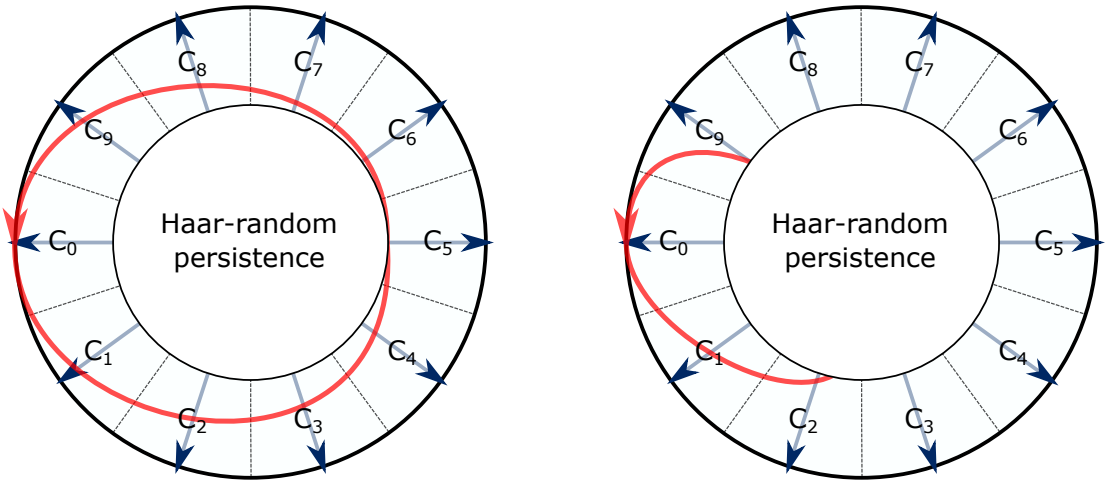
$$\left|\langle C_{k+p}|\hat{U}_H(pt_0)|C_k\rangle\right| > \eta(d),\ \forall\ k \text{ and } |p| \leq d/2. \tag{2.13}$$

Eq. (2.13) states that ($\mathscr{C}$ shows cyclic ergodicity iff) any initial state $|C_k\rangle \in \mathscr{C}$ "visits" all the other elements of $\mathscr{C}$ sequentially with sufficiently large overlap (i.e. $z_k(p, t_0) > \eta(d)$), at least once in its future and past evolution. Similarly, in place of a vanishing *average* self-intersecting fraction for classical cyclic aperiodicity in Eq. (2.5), we define quantum cyclic aperiodicity in terms of a sufficiently small mean overlap amplitude of a pure state in $\mathscr{C}$ with itself, with an additional restriction to a time interval (t_1, t_2) with $0 < t_1 < t_2$:

Definition 2.4 (**Quantum cyclic aperiodicity**). *A pure state quantum cyclic permutation $\mathscr{C}$ shows cyclic aperiodicity in a time interval (t_1, t_2) with precision $\eta(d)$ iff*

$$\frac{1}{d}\sum_k \left|\langle C_k|\hat{U}_H(t)|C_k\rangle\right| \leq \eta(d),\ \forall\ t:\ |t| \in (t_1, t_2). \tag{2.14}$$

The question of interest is now the choice of precision $\eta(d)$ and the time interval (t_1, t_2) that are most useful in physical situations. An important consideration for the former is suggested by canonical typicality [23, 24], which here refers to the observation that a typical "Haar random" [36]



(a) Cyclic ergodicity and aperiodicity.

(b) Non-ergodicity.

Figure 2.5: Schematic depiction of (a) cyclic ergodicity and aperiodicity, and (b) non-ergodicity, for a $(d = 10)$-element quantum cyclic permutation in a polar representation $(r, \theta) \in [0,1] \times [0, 2\pi)$ of the corresponding $(d = 10)$-dimensional Hilbert space Σ_d; the angular direction is parametrized as $\theta = 2\pi p/d$, and the radial coordinate is $r = f\left(\left|\langle C_0|(\hat{U}_C^\dagger)^p|\psi\rangle\right|\right)$ for any vector $|\psi\rangle \in \Sigma_d$, where $f : [0,1] \to [0,1]$ is some monotonic function and $\hat{U}_C^p$ is extrapolated to non-integer p in some convenient manner (e.g. connecting the $|C_k\rangle$ along some smooth path). The basis vectors $|C_k\rangle$ are depicted by arrows representing the corresponding axes. Each of (2.5a) and (2.5b) may be loosely regarded as a "quantization" of the respective classical versions in Fig. 2.4. The trajectory indicates the persistence amplitude $z_0(p, t_0)$ of the initial state $|C_0\rangle$ (1 at the outermost boundary, 0 at the center) in the radial direction up to $|p| = 5$ (visible trajectory) and beyond ("Haar-random persistence"). The region of "Haar-random persistence" refers to $z_0(p, t_0) \lesssim O(d^{-1/2})$, which includes all typical Haar random states by canonical typicality [23, 24]. Consequently, this region has by far the largest (Haar) volume in Σ_d, while the depicted outer regions of non-random overlap with the $|C_k\rangle$ together form a relatively tiny fraction of the space.

pure state $|\psi\rangle$ in Σ_d (by which we mean a normalized state $\langle\psi|\psi\rangle = 1$ that is randomly chosen with respect to the Haar/"uniform" distribution, invariant under unitary transformations, in the space of such states) has what we call a "random" overlap with every pure state in a given orthonormal basis $\{|\phi_k\rangle\}_{k=0}^{d-1} \in \Sigma_d$, namely $|\langle\phi_k|\psi\rangle| \in O(d^{-1/2})$, in a fairly strong sense (i.e. $\sum_{j=1}^{n} |\langle\phi_{k_j}|\psi\rangle|^2 \approx n/d$ for large n, appearing equally distributed [23, 24] over large collections of the $|\phi_k\rangle$). Thus, the action of $\hat{U}_H(t_0)$ on any generic choice of cyclic permutation $\mathscr{C} \subset \Sigma_d$, such that $\hat{U}_H(t_0)|C_k\rangle$ looks Haar random, is overwhelmingly likely to have p-step persistence amplitudes of $z_k(p; t_0) \in O(d^{-1/2})$ for almost all p in *any* system. Consequently, the simplest nontrivial definition of cyclic ergodicity for a quantum

cyclic permutation (i.e. one not automatically satisfied by a generic cyclic permutation in every system) would be one that requires a larger overlap than this "trivial" random value to consider a given state as having "visited" another state. Thus, the most physically useful choices of $\eta(d)$ should formally satisfy $\eta(d) \in O(d^{\varepsilon-1/2})$ for some $0 < \varepsilon < 1/2$ to account for all "non-trivial" overlaps and $\eta(d) \notin O(d^{-1/2})$ to exclude trivial overlaps as $d \to \infty$, while taking values physically regarded as "somewhat larger than" $d^{-1/2}$ for finite d. For any such choice of ε, we use the asymptotic notation:

$$f(d) > \eta(d) \implies f \gg O(d^{-1/2}), \tag{2.15}$$

$$f(d) \leq \eta(d) \implies f \lesssim O(d^{-1/2}). \tag{2.16}$$

Now we consider the physically appropriate time interval (t_1, t_2) (which stands in contrast to allowing/requiring infinitely long times in classical aperiodicity). The quantum recurrence theorem [49] (Poincaré recurrence for the flow $\varphi_n(t) = \varphi_n(0) - E_n t$ of phases $\varphi_n = \arg\langle E_n|\psi(t)\rangle$ in the energy eigenbasis) guarantees that aperiodicity will eventually be violated for every cyclic permutation after some extremely long time (possibly exponentially large in d; see e.g. [108] for a related discussion of recurrences), and we also do not want to rule out recurrences for small times $t \sim O(1)t_0$ when the state is still "near" the initial state. For this reason, $t_1/t_0 \notin O(1)$ and $d < t_2/t_0 \in O(d)$ are physically appropriate choices, in which case we will write

$$t_0 \ll |t| = O(t_0 d), \tag{2.17}$$

in place of $|t| \in (t_1, t_2)$ in Eq. (2.14). To summarize, Definitions 2.3 and 2.4 rigorously apply for any finite d, while our choice of precision $\eta(d)$ and aperiodicity time interval (t_1, t_2) suitable for physics are informed by (and mathematically rigorous in) the $d \to \infty$ limit.

We emphasize that it is cyclic ergodicity and aperiodicity — among classically relevant dynami-

cal properties — that have *direct* quantum dynamical counterparts as given by Definitions 2.3 and 2.4, allowing an observable-independent definition of quantum ergodicity in Definition 2.5 below. While "continuum" (or conventional) ergodicity and mixing are *classically* fundamental and observable-independent, any attempted quantization appears to acquire a strong dependence on the choice of "physical" observables [21, 22, 69] due to the necessary involvement of eigenstates (this is briefly discussed further in Sec. 2.7.1). Consequently, we will often drop the "cyclic" qualifier for quantum cyclic ergodicity and aperiodicity, and treat them as the fundamental quantum ergodic properties of interest in the remainder of this Chapter.

It is now convenient to define when a quantum system is ergodic or aperiodic in an energy subspace Σ_d. For ergodicity, we require the *existence* of an ergodic cyclic permutation in Σ_d, with the range of the time step t_0 restricted by $t_0 d < T$, for T small enough to avoid the time scale at which quantum recurrences may unavoidably occur in any system. For aperiodicity, we note that a cyclic permutation composed of the energy eigenstates, e.g. $|C_k\rangle = |E_k\rangle$, violates aperiodicity (and ergodicity) with the maximum possible value (1) for the mean overlap in Eq. (2.14); we are therefore able to construct non-aperiodic cyclic permutations (those that violate Eq. (2.14)) for any quantum system and cannot require all cyclic permutations to be aperiodic (as a prerequisite for mixing), a situation which has no classical analogue in the absence of superpositions. On the other hand, the aperiodicity of a cyclic permutation $\mathscr{C}$ is governed by the following inequality:

$$\frac{1}{d}\sum_k \left|\langle C_k|\hat{U}_H(t)|C_k\rangle\right| \geq \frac{1}{d}\left|\mathrm{Tr}\left[\hat{U}_H(t)\right]\right|, \tag{2.18}$$

which follows from writing the trace in the $\{|C_k\rangle\}$ basis and using the triangle inequality for complex numbers (we will see in Sec. 2.4.1 that this inequality is saturated by some cyclic permutation in every system). Thus, systems with sufficiently large ($\gg O(d^{1/2})$) trace of $\hat{U}_H(t)$ would possess no cyclic

permutation satisfying aperiodicity. Correspondingly, it is natural to call a quantum system aperiodic if it admits an aperiodic cyclic permutation. We then have the following definitions of ergodicity and aperiodicity, pertaining to a dynamical (i.e. time-domain) version of quantum ergodicity (which is distinct from the use of "quantum ergodicity" in the mathematical literature to refer to the delocalization of energy eigenstates in a given basis [71, 72]):

Definition 2.5 (**Ergodicity and aperiodicity of a quantum system**). *We call a quantum system (dynamically)* ***ergodic*** *with precision $\eta(d)$ in the energy subspace Σ_d within a time $T > 0$, if it admits at least one cyclic permutation that is ergodic with precision $\eta(d)$, satisfying $t_0 d < T$. Similarly, the system is* ***aperiodic*** *in (t_1, t_2) with precision $\eta(d)$ in Σ_d within the time T if it admits at least one cyclic permutation aperiodic in (t_1, t_2) with precision $\eta(d)$, satisfying $t_0 d < T$.*

While this definition is stated in general terms, we will assume the physically motivated precision and time interval specified by Eqs. (2.15), (2.16) and (2.17) in all physical applications considered later.

As a simple example, every system is always ergodic and not aperiodic in any subspace Σ_1 consisting of a single energy level. For typical quantum systems, we will implicitly assume a choice of T that is as large as possible while being much less than the quantum recurrence time scale for a generic quantum system, e.g. evolution by a (Haar) random unitary in Σ_d. In general, identifying which energy subspaces Σ_d of the system satisfy these properties would provide an observable-independent characterization of the ergodicity of a quantum system. In anticipation of Sec. 2.4.1, we emphasize that Definition 2.5 is insensitive to unitary transformations of Σ_d (which would simply map cyclic permutations in Σ_d to each other), and consequently independent of any observables or measurement bases that one may consider; whether a system is ergodic or aperiodic in the above sense is then determined entirely by the unitary invariants of $\hat{U}_H(t_0)$ within this subspace: the energy eigenvalues $\{E_n\}_{n=0}^{d-1}$.

2.3.2 Quantum bounds from the single-step error

Similar to the classical case, we can rigorously prove ergodicity or disprove aperiodicity for a cyclic permutation given just the single-step error $\varepsilon_C(1,t_0)$. In the quantum case, this is made possible by noting that the single-step deviation $\hat{U}_\Delta = \hat{U}_C^\dagger \hat{U}_H(t_0)$ of time evolution from a cyclic permutation is unitary, corresponding to a (complex) rotation in the Hilbert space Σ_d; it can thus be described by effective angles of deviation $\theta_k(p) = \arccos z_k(p,t_0)$ of each $\hat{U}_H(pt_0)|C_k\rangle$ from $\hat{U}_C^p|C_k\rangle$. A simple application of the triangle inequality in Σ_d (see Appendix A.2.1) shows that the p-step angle of deviation cannot exceed the sum of the corresponding single-step angles (with the bound being saturated for a 2D rotation of a vector in successive steps of the $\theta_k(p)$):

$$\min\left\{\theta_k(p), \frac{\pi}{2}\right\} \leq \sum_{q=0}^{p-1} \theta_{k+q}(1). \tag{2.19}$$

Using $\varepsilon_C(p,t_0) = \sin^2\theta_k(p)$, it follows from here that $\varepsilon_C(1,t_0) < \pi^2/d^2$ implies ergodicity and $\varepsilon_C(1,t_0) < \pi^2/(4d^2)$ implies non-aperiodicity for $\mathscr{C}$ (for large d); we emphasize that these are one-way implications. These results are of interest to the extent that they are the direct extension of the classical bounds of Refs. [97, 98] (see Sec. 2.2.2) to pure state cyclic permutations; however, they are similarly restricted in their applicability to special systems that admit sufficiently small single step errors as in the classical case. In the next section, we will show that an analysis in terms of the energy level statistics of Σ_d allows one to characterize ergodicity and aperiodicity in much more general terms in the quantum case.

2.4 Optimal cyclic permutations and energy level statistics

In this section, we explicitly identify how specific measures of energy level statistics determine the ergodicity and aperiodicity of a quantum system. Sec. 2.4.1 constructs certain cyclic permutations

directly related to the energy eigenstates (specifically, their discrete Fourier transforms) and establishes results concerning their optimality for determining ergodicity and aperiodicity. Building on these results, Sec. 2.4.2 shows that ergodicity is directly determined by the mode fluctuation distribution [79, 101,102] of the energy levels, while aperiodicity is directly determined by the spectral form factor [28].

2.4.1 Optimizing ergodicity and aperiodicity with Discrete Fourier Transforms of energy eigenstates

2.4.1.1 Discrete Fourier Transforms of energy eigenstates

To establish the ergodicity or aperiodicity of a general system in a subspace Σ_d, given the corresponding energy levels $\{E_n\}_{n=0}^{d-1}$, it is convenient to explicitly identify "optimal" cyclic permutations in Σ_d which are the "most likely" to be ergodic and/or aperiodic in a system. A special role in this regard is played by the set of cycling operators which are diagonal in the energy eigenbasis:

$$\hat{U}_C[q] = \sum_{n=0}^{d-1} e^{-2\pi i n/d} |E_{q(n)}\rangle\langle E_{q(n)}|, \tag{2.20}$$

where $q : \mathbb{Z}_d \leftrightarrow \mathbb{Z}_d$ ranges over all permutations of the indices $n \in \mathbb{Z}_d = \{0, \ldots, d-1\}$. This corresponds to cyclic permutations $\mathscr{C}$ which can be written as a discrete Fourier transform (DFT) of the energy eigenstates,

$$|C_k(q, \varphi_n)\rangle = \frac{1}{\sqrt{d}} \sum_{n=0}^{d-1} e^{-2\pi i n k/d} e^{-i\varphi_n} |E_{q(n)}\rangle, \tag{2.21}$$

for arbitrary phases φ_n that don't influence the persistence amplitudes.

In Sec. 2.4.1.2, we show that the minimum value $\varepsilon_{\min}(p, t_0)$ of $\varepsilon_C(p, t_0)$ for a given p, among all cyclic permutations $\mathscr{C}$ in Σ_d, occurs when the corresponding $\hat{U}_C$ has the form in Eq. (2.20) by 1. Theorem 2.6 when $\varepsilon_{\min}(p, t_0) < 2/d$ ("small" errors), and 2. a heuristic argument for "generic" energies $\{E_n\}_{n=0}^{d-1}$ when $\varepsilon_{\min}(p, t_0) > 2/d$ ("large" errors). This shows that an ergodic system is most

likely to possess an ergodic cyclic permutation among the set satisfying Eqs. (2.20) and (2.21). Further, in Sec. 2.4.1.3, we prove that the system is aperiodic in Σ_d *if and only if* any and all cyclic permutations of the form in Eq. (2.21) are aperiodic. Thus, the "most likely" aperiodic cyclic permutation is also given by the above set. The involvement of energy eigenstates in Eqs. (2.20) and (2.21) naturally connects ergodicity and aperiodicity to functions of the energy levels $\{E_n\}_{n=0}^{d-1}$. Later, in Sec. 2.4.2, we discuss how these properties are determined by concrete measures of energy level statistics.

We note that Eq. (2.21) allows an intuitive interpretation of why DFTs of energy eigenstates correspond to optimal cyclic permutations: states of this form (particularly when q sorts the $E_{q(n)}$ in ascending order with n) are the closest one can get to defining approximate "time eigenstates" $|C_k\rangle$ with quantized "time eigenvalue" k (of a fictitious "time operator" conjugate to the energy in Σ_d, keeping in mind the Fourier relation between e.g. canonically conjugate position and momentum variables in quantum mechanics [62]). Time evolution by $\hat{U}_H(t_0)$ should then ideally cause the time eigenvalue k to be incremented by 1 in such a state, essentially functioning as a cyclic permutation of the $|C_k\rangle$ (up to small errors caused by discretization). In other words, $\hat{U}_H(t_0) \approx \hat{U}_C$ with arguably minimal error for such "time" eigenstates.

2.4.1.2 Optimizing ergodicity via persistence amplitudes

The optimality of cyclic permutations satisfying Eq. (2.20) for small errors is formalized by the following theorem.

Theorem 2.6 (**Optimal cyclic permutations**). *If the system (in some energy subspace Σ_d) admits some cyclic permutation $\mathscr{C}'$ with p-step error $\varepsilon_{C'}(p, t_0) \leq (2/d)$ for a given p and t_0, then $\varepsilon_C(p, t_0)$ attains its minimum value among all cyclic permutations for a cyclic permutation $\mathscr{C}$, whose cycling operator $\hat{U}_C$ satisfies*

$$\lim_{\delta \to 0} \left[\hat{U}_H(t_0) e^{i\delta \hat{Y}}, \hat{U}_C \right] = 0. \tag{2.22}$$

Here, $\hat{Y}$ is any fixed Hermitian operator (which effectively selects a unique eigenbasis of $\hat{U}_H(t_0)$ if the latter is degenerate). In particular, the global minimum of the error is achieved by one such $\hat{U}_C$ for every choice of $\hat{Y}$.

Outline of proof. The proof of this statement is outlined below in the special case of $p = 1$ (with $\hat{U}_H \equiv \hat{U}_H(t_0)$); the full proof may be found in Appendix A.2.2.

We note that unitary transformations $\hat{V}$ on Σ_d, being the most general orthonormality preserving linear transformations, can generate all possible cyclic permutations from a given $\mathscr{C}$ via $|C_k\rangle \to \hat{V}|C_k\rangle$, which induces a transformation $\hat{U}_C \to \hat{V}\hat{U}_C\hat{V}^\dagger$. Our initial objective is to identify the global maximum of the 1-step mean persistence,

$$z(1,t_0) \equiv \frac{1}{d}\sum_{k=0}^{d-1} z_k(1,t_0), \tag{2.23}$$

over all cyclic permutations. We have $|\mathrm{Tr}(\hat{U}_C^\dagger \hat{U}_H)|/d \leq z(1,t_0)$ (as is evident from expressing the trace in the $\mathscr{C}$ basis); on the other hand, there always exists a unitary transformation $|C_k\rangle \to |C_k''\rangle = e^{i\varphi_k}|C_k\rangle$ to a basis $\mathscr{C}''$, leaving the $z_k(1,t_0)$ unchanged, such that $|\mathrm{Tr}(\hat{U}_{C''}^\dagger \hat{U}_H)|/d = z(1,t_0)$. Thus, a cyclic permutation that maximizes the trace overlap also necessarily maximizes $z(1,t_0)$, with the same maximum value:

$$\max_{\text{all } \hat{V}} \frac{1}{d}\left|\mathrm{Tr}\left[\hat{V}\hat{U}_C^\dagger \hat{V}^\dagger \hat{U}_H\right]\right| = \max_{\text{all } \mathscr{C}} z(1,t_0). \tag{2.24}$$

The maxima of the trace overlap can occur only when it is stationary with respect to small variations of $\hat{U}_C$ — effected by $\hat{V} = e^{i\hat{X}}$ for all small, Hermitian $\hat{X}$. Correspondingly, imposing stationarity via $|\mathrm{Tr}(e^{i\hat{X}}\hat{U}_C^\dagger e^{-i\hat{X}}\hat{U}_H)| = |\mathrm{Tr}(\hat{U}_C^\dagger \hat{U}_H)| + O(X^2)$ gives

$$\hat{U}_H\hat{U}_C^\dagger - \hat{U}_C^\dagger\hat{U}_H = \hat{F}e^{i\alpha} \text{ for some } \hat{F} = \hat{F}^\dagger, \tag{2.25}$$

where $\alpha = \arg[\mathrm{Tr}(\hat{U}_C^\dagger \hat{U}_H)]$, as a necessary condition for a given cyclic permutation $\mathscr{C}$ with cycling

operator $\hat{U}_C$ to maximize both $|\mathrm{Tr}(\hat{U}_C^\dagger \hat{U}_H)|$ and $z(1,t_0)$. We then have two distinct cases of interest.

1. $\hat{F} = 0$: Solutions to Eq. (2.25) with $\hat{F} = 0$ (equivalently, those satisfying Eq. (2.22)) exist in every system, corresponding precisely to the set of cyclic permutations satisfying Eq. (2.20). For such solutions, we have $z_k(1,t_0) = z(1,t_0)$ for all k as $\langle C_{k+\ell}|\hat{U}_H|C_{j+\ell}\rangle = \langle C_k|\hat{U}_H|C_j\rangle$. If the global maximum $z_{\max}(1,t_0)$ of the mean persistence occurs at such a solution, it must also have the smallest error by Eq. (2.12), as $\min_k z_{k,\max}(1,t_0) = z_{\max}(1,t_0)$ in this case, while

$$\min_k z_k(1,t_0) \leq z(1,t_0) \leq z_{\max}(1,t_0) \tag{2.26}$$

for any cyclic permutation.

2. $\hat{F} \neq 0$: The status of solutions with $\hat{F} \neq 0$ is much less clear, with respect to whether they exist in a given system, and if so, whether such solutions can include the global maximum of $z(1,t_0)$. To analyze this further, we write

$$\hat{U}_H \hat{U}_C^\dagger = e^{i\hat{A}} e^{i\alpha} \text{ and } \hat{U}_C^\dagger \hat{U}_H = e^{i\hat{B}} e^{i\alpha}, \tag{2.27}$$

for *Hermitian* $\hat{A}, \hat{B}$. We then have the following properties when $\hat{U}_C$ is a solution to Eq. (2.25):

$$\{\text{Eigenvalues of } e^{i\hat{A}}\} = \{\text{Eigenvalues of } e^{i\hat{B}}\}, \tag{2.28}$$

$$\mathrm{Tr}(\sin\hat{A}) = \mathrm{Tr}(\sin\hat{B}) = 0, \tag{2.29}$$

$$\sin\hat{A} = \sin\hat{B}, \tag{2.30}$$

with the first following from the definition of $e^{i\hat{A}}$ and $e^{i\hat{B}}$ as products of the same two operators ordered differently (which must have the same eigenvalues [109]), the second from the definition of α, and the

third from $\hat{F} = \hat{F}^\dagger$. From these, we can show that any $\hat{F} \neq 0$ solution must have $\varepsilon_C(1, t_0) > (2/d)$; this is done in Appendix A.2.2, but here we describe a more intuitive argument showing that a sufficiently small error rules out $\hat{F} \neq 0$. If $\hat{A}$, $\hat{B}$ and the error are "infinitesimal" (which is allowed by Eqs. (2.28), (2.29)), we have $\hat{F} \approx i(\hat{A} - \hat{B}) = 0$ to leading order by Eq. (2.30). In other words, the difference $\hat{F}$ between two unitaries $e^{i\hat{A}}$, $e^{i\hat{B}}$ sufficiently close to $\mathbb{1}$ must be (almost) *anti*-Hermitian, while Eq. (2.25) requires that $\hat{F}$ be Hermitian; the two requirements are only consistent for $\hat{F} = 0$.

On extending these considerations to general p (see Appendix A.2.2), we get Theorem 2.6.

□

Based on the above outline, we can also describe a heuristic argument for why we expect Eq. (2.20) to include the optimal cyclic permutation for "generic" $\hat{U}_H(t_0)$. $\hat{F} = 0$ solutions to Eq. (2.25) always exist without imposing any constraints on the eigenvalues of $\hat{U}_\Delta = \hat{U}_C^\dagger \hat{U}_H = e^{i\hat{B}+i\alpha}$ whatsoever; but for Eq. (2.30) to not reduce to $\hat{A} = \hat{B}$ (and therefore, $\hat{F} = 0$) requires $\sin\hat{B}$ to have at least one degeneracy (on account of Eq. (2.28)). We expect that this additional fine-tuning of the eigenvalues required for $\hat{F} \neq 0$ makes it unlikely for the optimal cyclic permutation to occur with this condition for a generic $\hat{U}_H(t_0)$. If this is the case, Eq. (2.20) generically includes the optimal cyclic permutation for small or large error[2].

2.4.1.3 Optimizing aperiodicity

Using $\langle C_k|\hat{U}_H(t)|C_k\rangle = \langle C_j|\hat{U}_C^{j-k}\hat{U}_H(t)\hat{U}_C^{k-j}|C_j\rangle = \langle C_j|\hat{U}_H(t)|C_j\rangle$ for DFT cyclic permutations from Eq. (2.21) (or equivalently, Eq. (2.20) or $[\hat{U}_H(t), \hat{U}_C^{k-j}] = 0$), we have

$$\frac{1}{d}\sum_k \left|\langle C_k|\hat{U}_H(t)|C_k\rangle\right| = \frac{1}{d}\left|\mathrm{Tr}\left[\hat{U}_H(t)\right]\right|. \tag{2.31}$$

[2]In numerical experiments with small d (up to $d = 5$) and arbitrarily chosen eigenvalues $E_n t_0$ of $\hat{U}_H(t_0)$, we see that the global minimum of the error over all unitary transformations $\hat{U}_C \to \hat{V}\hat{U}_C\hat{V}^\dagger$ does appear to occur for a cyclic permutation satisfying Eq. (2.20) for both small and large optimal error, providing some support for this argument.

This shows that the DFT cyclic permutations of Eq. (2.21) *all* saturate the lower bound of Eq. (2.18). Thus, the aperiodicity of any (and every) one of these DFT cyclic permutations is a necessary and sufficient condition for the system to be aperiodic in Σ_d.

2.4.2 Ergodicity, aperiodicity and energy level statistics

In a DFT basis, the p-step persistence amplitudes defined in Eq. (2.11) are equal, $z_k(p,t_0) = z(p,t_0)$, and the persistence amplitudes can be expressed directly in terms of the energy eigenvalues E_n:

$$\begin{aligned} z(p,t_0)[q] &= \left| \frac{1}{d} \operatorname{Tr}\left[\hat{U}_H(p t_0) \hat{U}_C^{-p}[q] \right] \right| \\ &= \left| \frac{1}{d} \sum_{n=0}^{d-1} \exp\left[ip \left(\frac{2\pi n}{d} - E_{q(n)} t_0 \right) \right] \right|. \end{aligned} \tag{2.32}$$

The corresponding p-step errors are given by $\varepsilon_C(p,t_0)[q] = 1 - z^2(p,t_0)[q]$, as per Eq. (2.12), among which the global minimum $\varepsilon_{\min}(p,t_0)$ is attained for some choice of q (if this minimum is less than $(2/d)$, as guaranteed by Theorem 2.6).

We will regard the $z(p,t_0)[q]$ as measures of the energy level statistics within Σ_d, in particular the deviation of the (permuted) energy levels from a regularly spaced spectrum. Namely, let

$$\Delta_n(t_0,q) = \left(\frac{t_0 d}{2\pi} E_{q(n)} \right) - n \tag{2.33}$$

represent the deviation of the $q(n)$-th level in a rescaled spectrum from the integer n. The persistence

as a function of time as given by

$$z^2(p,t_0)[q] = \left|\frac{1}{d}\sum_n e^{-i(2\pi p/d)\Delta_n(t_0,q)}\right|^2 \xrightarrow{d\to\infty} \left|\int \mathrm{d}\Delta\, f(\Delta;t_0,q)e^{-i(2\pi p/d)\Delta}\right|^2, \tag{2.34}$$

where $f(\Delta;t_0,q)$ is the probability density function of the $\Delta_n(t_0,q)$ in the $d\to\infty$ limit (or sufficiently large d).

Intuitively, the persistence at any time p would be maximized when the Δ_n are minimized. A practically reasonable choice of t_0 and q to estimate the global minimum of the 1-step error, for uniform density of states $\mathcal{N}^{\#}(\Sigma_d) = (d-1)/(E_{\max}-E_{\min})$ (i.e. appearing uniform over large energy windows [28] within Σ_d), is one in which the rescaled levels $t_0 dE_{q(n)}/(2\pi)$ are each close to the n-th integer. In other words, $t_0 \approx 2\pi\mathcal{N}^{\#}(\Sigma_d)/d$, with q chosen to be the sorting permutation $\overline{q}$ that sorts the energy levels in ascending order of the phase $(E_n t_0 \bmod 2\pi)$:

$$\{(E_{\overline{q}(n)}t_0 \bmod 2\pi) > (E_{\overline{q}(m)}t_0 \bmod 2\pi)\} \implies \overline{q}(n) > \overline{q}(m), \tag{2.35}$$

(essentially, $E_{\overline{q}(0)}$ is the lowest energy level in Σ_d, $E_{\overline{q}(1)}$ the next higher level and so on until the highest level $E_{\overline{q}(d-1)}$, for this choice of t_0), ensuring that $t_0 dE_{\overline{q}(n)}/(2\pi)$ remains reasonably close to $\overline{q}(n)$ for some choice of t_0. For a given t_0, it is shown in Appendix A.2.3 that Eq. (2.34) is indeed maximized at $p=1$ when $q=\overline{q}$, among a certain class of "small" permutations when $\Delta_n \ll d$. In other words, the sorting permutation is a (discrete version of a) local minimum for the error.

In this case, the Δ_n are essentially what have been called mode fluctuations [79, 101, 102] in the spectrum[3]; the Gaussianity of their distribution has been conjectured to be a "signature of

[3]The term "mode fluctuations" has been used with at least two different meanings in the literature [79, 101, 102]. In Refs. [101, 102] and related works cited there, it refers to the fluctuations of the *spectral staircase* around a straight line. Our usage is in the sense of Ref. [79], referring to deviations of the levels themselves from a straight line. The two are different

chaos" [101, 102]. A minor, but important, technical distinction between Δ_n and conventional mode fluctuations is that there is no unfolding [28, 36] — a convenient modification of the energy levels to make $\mathcal{N}^{\#}(\Sigma_d)$ appear uniform over large energy scales while preserving shorter range correlations — prior to calculating the Δ_n. Such a procedure, while useful for *combining* short or medium range level statistics from different parts of the spectrum for improved statistical quality in numerical studies, manually alters the long-range structure of the spectrum and does not preserve the dynamics of the system in the time domain. This makes unfolding unsuitable for any analytical approach aiming to relate the *unmodified* dynamics of a system to its energy levels, including the present study. Given this qualifier, Eq. (2.34) naturally states that the Fourier components of mode fluctuation distributions, obtained *without unfolding*, directly determine the optimal persistence of cyclic permutations. As we will see more quantitatively in the discussion around Eq. (2.41), this means that the persistence may be large enough to show cyclic ergodicity only when Σ_d is contained in a sufficiently narrow energy shell, within which larger scale variations in $\mathcal{N}(\Sigma_d)$ can readily be neglected.

Another relevant (and extensively studied) measure is the spectral form factor (SFF) [28] (here, of the energy levels within Σ_d), defined by

$$K(t) \equiv \left| \frac{1}{d} \operatorname{Tr}\left[\hat{U}_H(t)\right] \right|^2 = \frac{1}{d^2} \sum_{n,m} e^{i(E_n - E_m)t}. \tag{2.36}$$

The SFF is usually the central analytically tractable quantity in studies of "quantum chaotic" systems, as far as level statistics is concerned [28, 40]. Excluding $O(1)$ transients at early times (the "slope" [82]), a high degree of spectral rigidity in such systems is indicated by significantly suppressed late-time quantum fluctuations [110] in $K(t)$, over a length of time, to well below its "natural" yet small average value $\langle K(t) \rangle_t = d^{-1}$ seen over the longest time scales, e.g. to around $K(t) \sim [t/\mathcal{N}(\Sigma_d)]d^{-1}$ when

in general, but show close agreement in their statistical properties for Wigner-Dyson random matrix ensembles [46, 47] (see also Sec. 2.5.4.2).

$t \ll \mathcal{N}(\Sigma_d)$ for Wigner-Dyson statistics (these suppressed fluctuations form the "ramp" [82]). For low spectral rigidity, there is weaker or no suppression, e.g. for Poisson spectra, virtually all late-time fluctuations in $K(t)$ oscillate strongly around d^{-1}.

Calculations of $K(t)$ in several "quantum chaotic" systems (with various approximations) suggest that the low magnitude of the ramp is determined by generic randomness properties and low levels of recurrence/periodicity of certain physical processes, rather than any direct notion of ergodicity — e.g. an appropriate "uniform"[4] and minimally correlated distribution of strictly isolated periodic orbits in systems with a chaotic classical limit [28, 39, 40, 73], analogous properties of closed Feynman paths in random Floquet systems [35, 41, 42], or small return probabilities in diffusive processes [73, 111, 112]. Indeed, from Eqs. (2.31) and (2.36), it is clear that the SFF is most directly associated with aperiodicity. Nevertheless, in Sec. 2.5, we will show how the behavior of the SFF ramp influences *cyclic ergodicity* within Σ_d — connecting an observable-independent notion of quantum ergodicity to these somewhat better understood recurrence properties of specific physical processes in some systems, while remaining applicable to more general systems.

In terms of the measures $z(p, t_0)[q]$ and $K(t)$, we have the following *direct* requirements on the energy level statistics within Σ_d, for ergodicity (Eq. (2.13)) and aperiodicity (Eq. (2.14)) as per Definition 2.5:

1. *Ergodicity in Σ_d:*

$$z^2(p, t_0)[q] \gg O(d^{-1}), \ \forall \ |p| \leq \frac{d}{2} \tag{2.37}$$

is a sufficient condition for ergodicity if satisfied at all $|p| \leq d/2$ for at least one choice of q (usually $q = \overline{q}$) by Eq. (2.32). It is also necessary that Eq. (2.37) be satisfied by at least one

[4] This distribution of periodic orbits, called the Hannay-Ozorio de Almeida sum rule [39], is often motivated in terms of its similarity to ergodicity [28]. It is worth emphasizing that in spite of the mathematical similarity, the two are logically and conceptually distinct, partly due to the fact that isolated periodic orbits are of measure zero in Eq. (2.1); as noted in Ref. [39], KAM tori are ergodic systems that do not satisfy this sum rule.

$q = q_p$ for each $|p| \leq d/2$ in a cyclic ergodic system, if the heuristic argument for the large error version of Theorem 2.6 holds.

2. *Aperiodicity in* Σ_d*:*

$$K(t) \lesssim O(d^{-1}),\ \forall\ t_0 \ll |t| = O(t_0 d) \tag{2.38}$$

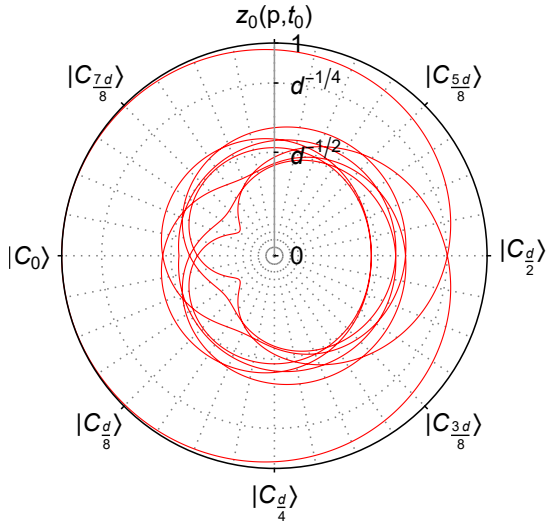
is a necessary and sufficient condition for aperiodicity, by Eqs. (2.31) and (2.36).

A further question is if the simplest i.e. $p = 1$ measures of level statistics can be used to study these dynamical properties for all times p, say for a given q.To enable this, we first have the formal bound of Eq. (2.19), which can be expressed in terms of $\varepsilon_C(1, t_0)$ as a lower bound on the decay of the persistence (via $z = \cos\theta$):

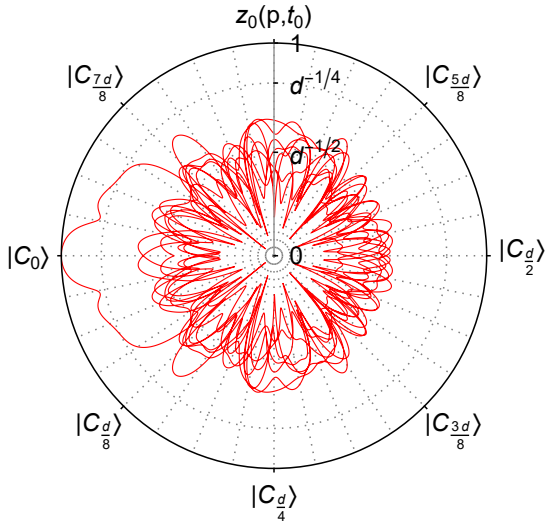
$$z(p, t_0) \geq \begin{cases} \cos\left(p\sqrt{\varepsilon_C(1,t_0)}\right), & |p| < \pi/\sqrt{4\varepsilon_C(1,t_0)}, \\ 0, & |p| \geq \pi/\sqrt{4\varepsilon_C(1,t_0)}, \end{cases} \tag{2.39}$$

neglecting $O[p\varepsilon_C^{3/2}(1, t_0)]$ contributions.

In general, by the reciprocal relation of Fourier variables, a slow decay of the persistence corresponds to a narrow distribution $f(\Delta; t_0, q)$ e.g. as measured by the variance $\sigma_\Delta^2 \equiv \langle f^2\rangle_\Delta - \langle f\rangle_\Delta^2$, which is related to $z(1, t_0)$ via Eq. (2.34). When q is a sorting permutation, a low σ_Δ^2 essentially implies a high rigidity of the spectrum . While the precise connection between ergodicity, aperiodicity and σ_Δ^2 depends on the functional form of $f(\Delta; t_0, q)$, it is natural to pay special attention to the case of an ideal Gaussian distribution of mode fluctuations, as has been seen in typical "quantum chaotic" systems [79, 101, 102]. In this case, we get $z(p, t_0) \approx e^{-\gamma p^2}$ (up to random fluctuations) and $\varepsilon_C(1, t_0) = 2\gamma = 4\pi^2\sigma_\Delta^2/d^2$ (for $\varepsilon_C \ll 1$) from Eq (2.34). More significant is the following proposition, which follows from substituting the Gaussian form of $z(p, t_0)$ in Eq. (2.37), as well as in Eq. (2.38) after identifying $z^2(d, t_0) = K(t_0 d)$ (detailed in Sec. 2.5.4.1):



(a) Cyclic ergodicity and aperiodicity (staying closer to the boundary than $O(d^{-1/2})$ for more than one and less than two full rotations); data for a single Circular Unitary Ensemble (CUE [28, 36]) random matrix with E_n spanning $[0, 2\pi)$.



(b) Non-ergodicity (staying closer to the boundary than $O(d^{-1/2})$ for less than one full rotation); data for a single realization of Poisson/uncorrelated energy levels spanning $[0, 2\pi)$.

Figure 2.6: Exact numerical data for a polar representation of (a) cyclic ergodicity and aperiodicity, as well as (b) non-ergodicity, via the trajectory of $|C_0\rangle$ in a Hilbert space Σ_d with $d = 2048$, and $t_0 = 2\pi/\Delta E = 2\pi\mathcal{N}(\Sigma_d)/d = 1$. Essentially, the cyclic permutation basis elements $|C_k\rangle$ are points on the boundary (with $p = k$) of the polar representation, and the permutation is ergodic if the actual trajectory of any such point remains close to the boundary for a full rotation of the angular coordinate (including future and past evolution). The angular coordinate is $\theta = 2\pi p/d$ (depicted here for $p \in \mathbb{Z}$), and the radial coordinate represents $z_0(p, t_0)$ via the map $r = g(z_0(p, t_0))/g(1)$ with $g(x) = \{1 + \tanh[\ln(x^2 d/2)/6]\}$. The trajectories extend up to $|p| = 4d$. The chosen cyclic permutation in both cases is the sorted DFT cyclic permutation. This figure anticipates the ergodicity of Wigner-Dyson level statistics and non-ergodicity of Poisson level statistics (Sec. 2.5). The central region $z_0(p, t_0) \in O(d^{-1/2})$ of Haar-random persistence again corresponds to nearly all of the Haar volume of Σ_d by canonical typicality [23, 24], and is where the trajectory typically remains for long times with the exception of occasional recurrences near the boundary. See Fig. 2.7 for a different depiction of similar data.

Proposition 2.7 (Ergodicity, aperiodicity and Wigner-Dyson spectral rigidity). *If* $z(p, t_0) = e^{-\gamma p^2} + O(d^{-1/2})$ *with* $\gamma > 0$ *when* $|p| \in O(d)$ *for some DFT cyclic permutation* $\mathscr{C}$*, then as* $d \to \infty$*,*

$$\textit{Cyclic ergodicity and aperiodicity of } \mathscr{C} \iff \left[\sigma_\Delta^2 = \frac{\alpha^2}{4\pi^2}\ln d, \textit{ with } \alpha \in [1, 2), \textit{ and Eq. (2.38) holds}\right]. \tag{2.40}$$

This form of σ_Δ^2 is precisely that of the Wigner-Dyson circular random matrix ensembles — COE, CUE, CSE respectively corresponding to maximal ensembles of symmetric, unconstrained, or self-dual unitaries [28, 36] — for $\alpha = 2, \sqrt{2}, 1$, which also satisfy Eq. (2.38). Moreover, in typical "quantum chaotic" systems, the energy levels show Wigner-Dyson statistics (in particular, possessing the spectral rigidity of the circular Wigner-Dyson ensembles and Gaussian mode fluctuations) [28,36,79,101,102], if Σ_d is chosen to be a sufficiently narrow energy shell spanning $[E, E+\Delta E]$ with $\Delta E \lesssim 2\pi/t_{\text{ramp}}(E)$; here, $t_{\text{ramp}}(E)$ can be directly identified [38] as the time beyond which the ramp appears in the SFF around the energy E after eliminating the effect of early-time transients as described below[5]. This ramp time is system specific, e.g. ranging from $t_{\text{ramp}} = 1$ in some quantized chaotic maps [113], $O(1)$ to growing logarithmically with system size in some Floquet many-body systems [35, 38, 41, 84], or the time scale of diffusion in some disordered systems [38]. In general, the definite width ΔE of Σ_d introduces a large $\text{sinc}^2(t\Delta E/2)$ transient (see also Ref. [114]) owing to the Fourier sum in Eq. (2.36), due to which the SFF effectively takes the following form when $\Delta E \lesssim 2\pi/t_{\text{ramp}}$:

$$K(t) = \text{sinc}^2\left(t\frac{\Delta E}{2}\right) + \text{ramp}(t). \tag{2.41}$$

Thus, by choosing $t_0 = 2\pi/\Delta E$ $(= 2\pi\mathcal{N}^{\#}(\Sigma_d)/d$, in agreement with our earlier estimate after Eq. (2.34); we also note that this is the minimum time scale for nontrivial dynamics in an energy window of width ΔE, suggested by the energy-time uncertainty principle) so that integer steps pt_0 coincide with the zeros of the transient, we can eliminate the influence of this transient and obtain cyclic permutations that are directly determined by the intrinsic spectral rigidity of the system repre-

[5] As an illustration, a Hamiltonian CUE-class Wigner-Dyson system with an energy spectrum of width $\mathscr{E}$ has $|\text{Tr}\,\hat{U}_H(t)|^2 = \mathscr{E}t/(2\pi)$ (up to fluctuations [110]) for sufficiently small $t \gtrsim t_{\text{ramp}}$ [38]. Each (disjoint) energy shell $\Sigma(k)$ (with unitary $\hat{U}_{H(k)}(t)$) of width $\Delta E_t = 2\pi/t$ contributes phases $E_n(k)t$ that span $[0, 2\pi)$ with $|\text{Tr}\,\hat{U}_{H(k)}(t)|^2 = 1$ (*irrespective* of the number of levels d_k in the shell), effectively behaving like a CUE random unitary [28, 36] at the time t. The direct sum of all $M = \mathscr{E}/\Delta E = \mathscr{E}t/(2\pi)$ such CUE-like unitaries has the individual traces combine with random phases, giving $|\text{Tr}\,\hat{U}_H(t)|^2 = M$ (see also Refs. [38, 112]). In other words, the quantitative form of the Wigner-Dyson ramp in Hamiltonian systems is due to all energy shells of some width $\Delta E \lesssim 2\pi/t_{\text{ramp}}$ behaving like circular ensemble random unitaries at the time $t = 2\pi/\Delta E$.

sented by the ramp (this will be justified further in Sec. 2.5). In addition, this means that the period $t_0 d$ of such cyclic permutations is given precisely by the Heisenberg time $t_{\mathrm{H}} \equiv 2\pi \mathscr{N}^{\#}(\Sigma_d)$ of Σ_d, at which the individual energy levels typically dephase completely marking the end of the ramp [28].

Overall, this proposition suggests that all energy shells of width $\Delta E \lesssim 2\pi/t_{\mathrm{ramp}}$ in systems with Wigner-Dyson level statistics are ergodic and aperiodic with $t_0 = 2\pi/\Delta E$ (by Definition 2.5), as the sorted DFT cyclic permutation is both ergodic and aperiodic, which is anticipated by the numerical data in Fig. 2.6 for CUE and Poisson level statistics. From a dynamical standpoint, that the ergodicity of such systems only holds within thin energy shells is not surprising — this just reflects the fact that Hamiltonian systems with a classical limit are only ergodic in thin energy shells, and not over phase space volumes covering a wide range of energies.

2.5 Cyclic permutations for typical systems

In this section, we study the behavior of DFT cyclic permutations for a "typical" system with sufficiently random fluctuations in the energy levels. Based on a general decomposition of $\hat{U}_H(pt_0)$ into a periodic part that follows the cyclic permutation and orthogonal random fluctuations in Sec. 2.5.1, we motivate a Gaussian estimate for the time dependence of persistence amplitudes for sufficiently early p in typical systems in Sec. 2.5.2. In Sec. 2.5.3, we show how the Gaussian estimate can be used to derive a lower bound for the error from the SFF for the system sampled at discrete times pt_0 given by Eq. (2.52), directly connecting cyclic ergodicity to the size of quantum fluctuations in the SFF ramp (which is often analytically tractable). In Sec. 2.5.4, we discuss detailed analytical and numerical evidence for Proposition 2.7, showing that in the ideal case where the Gaussian estimate remains valid for longer times ($pt_0 \lesssim t_{\mathrm{H}}$) in an individual system, cyclic ergodicity and aperiodicity are equivalent to requiring the spectral rigidity of energy levels to be in the range spanned by the Wigner-Dyson circular ensembles: COE, CUE and CSE.

2.5.1 Periodic and random parts of time evolution

For a cyclic permutation $\mathscr{C}$ with cycling operator $\hat{U}_C$ that commutes with $\hat{U}_H$ (i.e. a cyclic permutation of a DFT basis), the p-step persistence probability $z^2(p,t_0)$ is given by the analogue of the SFF for the error unitary $\hat{U}_\Delta = \hat{U}_C^{-1}\hat{U}_H(t_0)$:

$$z^2(p,t_0) = \left|\frac{1}{d}\,\mathrm{Tr}\left[\hat{U}_\Delta^p\right]\right|^2. \tag{2.42}$$

To study the development of the persistence over time, it is convenient to write a general expression for $\hat{U}_\Delta^p$ in terms of the p-step errors $\varepsilon_C(p,t_0)$. On account of $[\hat{U}_\Delta,\hat{U}_C]=0$, we have $\langle C_{k+\ell}|\hat{U}_H|C_{j+\ell}\rangle = \langle C_k|\hat{U}_H|C_j\rangle$, due to which $\hat{U}_\Delta^p$ can be expressed simply in terms of powers of $\hat{U}_C$:

$$\hat{U}_\Delta^p = \left[\left(\sqrt{1-\varepsilon_C(p,t_0)}\right)\hat{\mathbb{1}} + \left(\sqrt{\varepsilon_C(p,t_0)}\right)\sum_{m=1}^{d-1}\nu_m(p)\hat{U}_C^m\right]e^{i\phi_\Delta(p)}, \tag{2.43}$$

for some phases $\phi_\Delta(p)$ and complex error coefficients $\nu_m(p)$. Unitarity $\hat{U}_\Delta^\dagger\hat{U}_\Delta = \hat{U}_\Delta\hat{U}_\Delta^\dagger = \hat{\mathbb{1}}$ translates to nonlinear constraints on the $\nu_m(p)$:

$$\sum_{m=1}^{d-1}|\nu_m(p)|^2 = 1, \tag{2.44}$$

$$\nu_m(p) + \nu_{-m}^*(p) = -g_p\sum_{k=1}^{d-1}\nu_k^*(p)\,\nu_{k+m}(p), \text{ for } m\neq 0, \tag{2.45}$$

where $\nu_0(p)\equiv 0$, and $g_p \equiv \sqrt{\varepsilon_C(p,t_0)/[1-\varepsilon_C(p,t_0)]}$.

As a matter of nomenclature, we call the first term proportional to $\hat{\mathbb{1}}$ in Eq. (2.43) the "periodic part", and the remaining terms involving $\hat{U}_C^m$ (orthogonal to the periodic part) the "random part", of time evolution. This is because the former becomes a term proportional to $\hat{U}_C^p$ in $\hat{U}_H(pt_0)$ which is periodic in p, while we expect the $\nu_m(p)$ to generally (but not necessarily) look "random". In fact, (a

subset of) the $\nu_m(p)$ are directly related to the SFF of $\hat{U}_H(pt_0)$ within the subspace Σ_d, via:

$$K(pt_0) = \left|\frac{1}{d}\,\mathrm{Tr}\left[\hat{U}_H(pt_0)\right]\right|^2 = \varepsilon_C(p,t_0)|\nu_{-p}(p)|^2, \tag{2.46}$$

and the expectation of randomness in the $\nu_m(p)$ reflects the randomness in the SFF ramp [35,86,110] (more precisely, particularly in the phases of $\mathrm{Tr}[\hat{U}_H(t)]$). Additionally, $K(pt_0)$ serves as a (rather weak) lower bound for the p step errors. In particular, $\varepsilon_C(p,t_0) = O(1)$ if $K(pt_0) = O(1)$, establishing the impossibility of finding cyclic permutations that are reasonably close to $\hat{U}_H(t_0)$, when $t_0 \ll t_{\text{ramp}}$ i.e. in the early-time "slope" regime of the SFF. To refine this bound, we will need a generic expression for the p-dependence of the right hand side, derived in the following subsection.

2.5.2 Gaussian estimate for persistence amplitudes

Using Eq. (2.43), one can readily express the persistence at arbitrary time p in terms of the 1-step parameters $\varepsilon_C(1,t_0)$ and $\nu_m(1)$. The resulting expression involves a complicated multinomial expansion in the $\nu_m(1)$ (with $\binom{p}{s}$ representing binomial coefficients),

$$\hat{U}_\Delta^p e^{-ip\phi_\Delta(1)} = \left[\left(1-\varepsilon_C(1,t_0)\right)^{p/2}\sum_{s=0}^{p}\binom{p}{s}g_1^s\sum_{m_1,\ldots,m_s}\nu_{m_1}(1)\ldots\nu_{m_s}(1)\hat{U}_C^{m_1+\ldots+m_s}\right], \tag{2.47}$$

which is hard to extract general predictions out of. To simplify the expression, we invoke a heuristic argument that relies on the expected randomness of the ν_m.

Specifically, we assume that each $\nu_m(1)$ is well described by an ensemble of complex numbers with a fixed magnitude and random phases, subject to the constraints Eq. (2.44) and Eq. (2.45). Further, if one neglects $O(\sqrt{\varepsilon_C(p,t_0)})$ corrections to the ν_m, Eq. (2.45) essentially becomes

$$\nu_m(p) \approx -\nu_{-m}^*(p). \tag{2.48}$$

Thus, pairings of $\nu_m(1)$ and $\nu_{-m}(1)$ in Eq. (2.47) have a definite phase and generate contributions that potentially interfere constructively, while the remaining random terms add out of phase. This suggests following a strategy similar to methods based on the pairing of closed Feynman paths in studies of generic semiclassical [28, 40, 45, 75, 76] and quantum [35, 42] "chaotic" systems: we evaluate the contribution from terms dominated by pairings of $\nu_m(1)$ and $\nu_{-m}(1)$ with at most one free $\nu_{m_k}(1)$, assuming (with no proof beyond the above argument) that the remaining terms are negligible. As is common with these methods, other contributions would eventually dominate at large enough times, when $\varepsilon_C(p,t_0)$ is sufficiently large and $\nu_m(p)$ is sufficiently random, invalidating Eq. (2.48) for such p.

The assumed dominance of paired error coefficients can be used to derive a general form of $\hat{U}_\Delta^p$ for small p, and from there an estimate for $z(p,t_0)$ using a recurrence relation; this is detailed in Appendix A.3, with numerical evidence for error coefficient pairing. For $\varepsilon_C(1,t_0) \ll 1$ and $p \ll 1/\sqrt{\varepsilon_C(1,t_0)}$, the general form is

$$\hat{U}_\Delta^p e^{-ip\phi_\Delta(1)} \approx \frac{z(p,t_0)}{\sqrt{1-\varepsilon_C(1,t_0)}}\left[\sqrt{1-\varepsilon_C(1,t_0)}\hat{\mathbb{1}} + p\sqrt{\varepsilon_C(1,t_0)}\sum_{r=1}^{d-1}\nu_r(1)\hat{U}_C^r\right]. \tag{2.49}$$

In other words, time evolution for small p simply manifests as a relative growth of the random part in comparison to the periodic part, up to an overall phase. This gives a simple Gaussian expression for the persistence amplitude (in the same regime of small error and time):

$$z(p,t_0) \approx \exp\left[-\frac{1}{2}\frac{\varepsilon_C(1,t_0)}{1-\varepsilon_C(1,t_0)}p^2 - \frac{1}{2}\varepsilon_C(1,t_0)|p|\right]. \tag{2.50}$$

The second (linear) term in the exponent is negligible until $|p| \sim 1/\varepsilon_C(1,t_0)$, and we will simply drop it in further calculations. The Gaussian follows the sinusoidal lower bound in Eq. (2.39) rather closely, suggesting that typical cyclic permutations are surprisingly close to saturating the lower bound. In

other words, $\hat{U}_\Delta^p$ remains close to a 2D rotation in Hilbert space, until a time $p \sim 1/\sqrt{\varepsilon_C(1,t_0)}$ when the cyclic permutation develops a large (~ 1) error.

2.5.3 Lower bound on the error from the SFF

Now we are in a position to quantitatively analyze the connection between the SFF ramp and the persistence of cyclic permutations. The 1-step error coefficients $\nu_m(1)$ can be related to the SFF $K(pt_0)$ in the $p \ll 1/\sqrt{\varepsilon_C(1,t_0)}$ regime, using Eqs. (2.43), (2.46) and (2.49):

$$|\nu_{-p}(1)|^2 \approx \frac{1-\varepsilon_C(1,t_0)}{z^2(p,t_0)\varepsilon_C(1,t_0)}\frac{K(pt_0)}{p^2}. \tag{2.51}$$

Summing over $p=-\overline{p}$ to $\overline{p}$ excluding 0, the left hand side can be at most 1 on account of the normalization constraint, Eq. (2.44). Expanding $z^2(p,t_0) = 1 - O(\varepsilon_C(1,t_0)p^2)$ and using $K(t)=K(-t)$, we get

$$\sum_{p=1}^{\overline{p}} K(pt_0)\left\{\frac{1}{p^2} + O[\varepsilon_C(1,t_0)]\right\} \lessapprox \frac{\varepsilon_C(1,t_0)}{2(1-\varepsilon_C(1,t_0))}. \tag{2.52}$$

Every term on the left hand side is positive. Considering only the first term and choosing the largest possible $\overline{p}$ for which the second term is negligible then gives a reasonably restrictive lower bound on $\varepsilon_C(1,t_0)$. Correspondingly, we take $\overline{p} = 1/(M\sqrt{\varepsilon_C(1,t_0)})$ where M is some large number satisfying $M = O(1) \geq 1$.

Eq. (2.52) is the main relation of interest connecting the recurrence properties represented by the SFF to cyclic ergodicity via $\varepsilon_C(1,t_0)$ (assuming typical $\nu_m(p)$). It demonstrates that the suppression of recurrences indicated by the small magnitude of the SFF ramp is essential for the dynamics to be able to closely follow a cyclic permutation, essentially due to the conservation of probability (Eq. (2.44)). As it involves only integer steps $p \in \mathbb{Z}$ of time pt_0, we see directly that choosing $t_0 = 2\pi/\Delta E$ for an energy shell prevents the $\mathrm{sinc}^2(t\Delta E/2)$ transient in Eq. (2.41) from influencing ergodicity.

We can also derive explicit bounds for specific cases. As such sums of the SFF over time are generally self-averaging (i.e. fluctuations of the ramp average out to give a more steady sum) [35], we replace $K(pt_0)$ with a smooth power law expression for the ramp: $K(t) = \lambda t^\gamma$ for $t_0 \leq t \ll t_0 d$, $\gamma \geq 0$, and with $\lambda \ll 1$, which accounts for the behavior of a wide variety of systems[6]; all such systems have $K(t_0 \leq t = O(t_0 d)) \lesssim O(d^{-1})$ and are consequently aperiodic. Evaluating the sum in Eq. (2.52) for this power law (Appendix A.3.3) gives the following constraints on the error:

$$\varepsilon_C(1, t_0) \gtrapprox \begin{cases} 2\lambda t_0^\gamma \zeta(2-\gamma), & \text{for } 0 \leq \gamma < 1, \\ \lambda t_0^\gamma \ln \dfrac{1}{\lambda}, & \text{for } \gamma = 1, \\ \left[2\lambda t_0^\gamma \dfrac{\gamma-1}{M^{\gamma-1}}\right]^{\frac{2}{\gamma+1}}, & \text{for } \gamma > 1, \end{cases} \tag{2.53}$$

where $\zeta(z)$ is the Riemann zeta function. Now we consider the most important (i.e. typical) cases of practical interest. Poisson statistics [28] corresponds to $\lambda = d^{-1}$ and $\gamma = 0$, for which we obtain

$$\varepsilon_C(1, t_0)|_{\text{Poisson}} \gtrapprox \frac{\pi^2}{3d}. \tag{2.54}$$

Together with the conditions for Eq. (2.22), this implies that every (DFT and non-DFT) cyclic permutation for a system with Poisson level statistics has $\varepsilon_C(1, t_0) > (2/d)$. As long as $z(p, t_0)$ is not drastically different from a Gaussian in p, as expected from the typicality considerations of Sec. 2.5.2, it follows that the persistence decays to the Haar random value by $p \lesssim O(d^{1/2}) \ll d/2$, and no DFT cyclic permutation is even remotely close to being ergodic for a typical system with Poisson statistics. On the other hand, the circular Wigner-Dyson ensembles [28, 36] have $\gamma = 1$ and $\lambda = 2/(\beta_{\text{RMT}} d^2)$

[6] In the following sense: $\gamma = 0$ and $\lambda = d^{-1}$ corresponds to generic integrable systems with Poisson statistics [28, 34]; $\gamma = 1$ corresponds to generic "chaotic" systems when $\lambda \in O(d^{-2})$ [28, 36], and those with macroscopic conserved quantities for larger magnitudes of λ [85, 112]; integer $\gamma > 1$ with $\lambda \in O(d^{-2})$ corresponds to tensor products of γ independent chaotic systems, as well as the γ-particle sectors of single-particle chaotic systems with $\lambda \in O(\gamma! d^{-2})$ (for large d), in which the many-particle SFF shows an exponential ramp [4, 87, 88].

with $\beta_{\text{RMT}} = 1, 2, 4$ for COE, CUE, CSE respectively. With $t_0 = 1$ $(= 2\pi\mathscr{N}^{\#}/d)$, the error satisfies

$$\varepsilon_C(1,1)|_{\text{Wigner-Dyson}} \gtrapprox \frac{4}{\beta_{\text{RMT}} d^2} \ln d, \text{ with } \beta_{\text{RMT}} \in \{1,2,4\}. \tag{2.55}$$

These relations encode the following property: any system admits cyclic permutations with large error, but only sufficiently rigid spectra can admit cyclic permutations with small error, quantifying the discussion in Sec. 2.4.2. For instance, if a system is known to have a cyclic permutation with error smaller than $(2/d)$, we can rule out Poisson statistics for that system.

2.5.4 Cyclic permutations and Wigner-Dyson level statistics

2.5.4.1 Spectral rigidity for ergodic, aperiodic systems with almost exactly Gaussian persistence amplitudes

From the viewpoint of the Gaussian estimate, an idealized situation is when the persistence amplitude $z(p, t_0)$ remains exactly Gaussian as it decays all the way through to the random state (order of magnitude) value $z(p, t_0) \in O(d^{-1/2})$. Writing $g_1^2 = \varepsilon_C(1, t_0)/[1 - \varepsilon_C(1, t_0)]$, we can solve for g_1 corresponding to ergodic or non-aperiodic evolution by imposing:

$$\exp\left[-\frac{1}{2\alpha^2} g_1^2 d^2\right] \geq c d^{-1/2}, \tag{2.56}$$

where $\alpha = 2$ for ergodicity and $\alpha = 1$ for non-aperiodicity (from Eqs. (2.13), (2.14)), while c is some $O(1)$ positive constant. From Eq. (2.34), we also obtain a Gaussian distribution for mode fluctuations given some g_1 (assuming that the DFT cyclic permutation under discussion corresponds to a level permutation function q),

$$f(\Delta; t_0, q) = \frac{1}{\sqrt{2\pi\sigma_\Delta^2}} \exp\left[-\frac{1}{2\sigma_\Delta^2}\Delta^2\right], \tag{2.57}$$

with variance $\sigma_\Delta^2 = g_1^2 d^2/(4\pi^2)$. Requiring ergodicity and aperiodicity therefore gives:

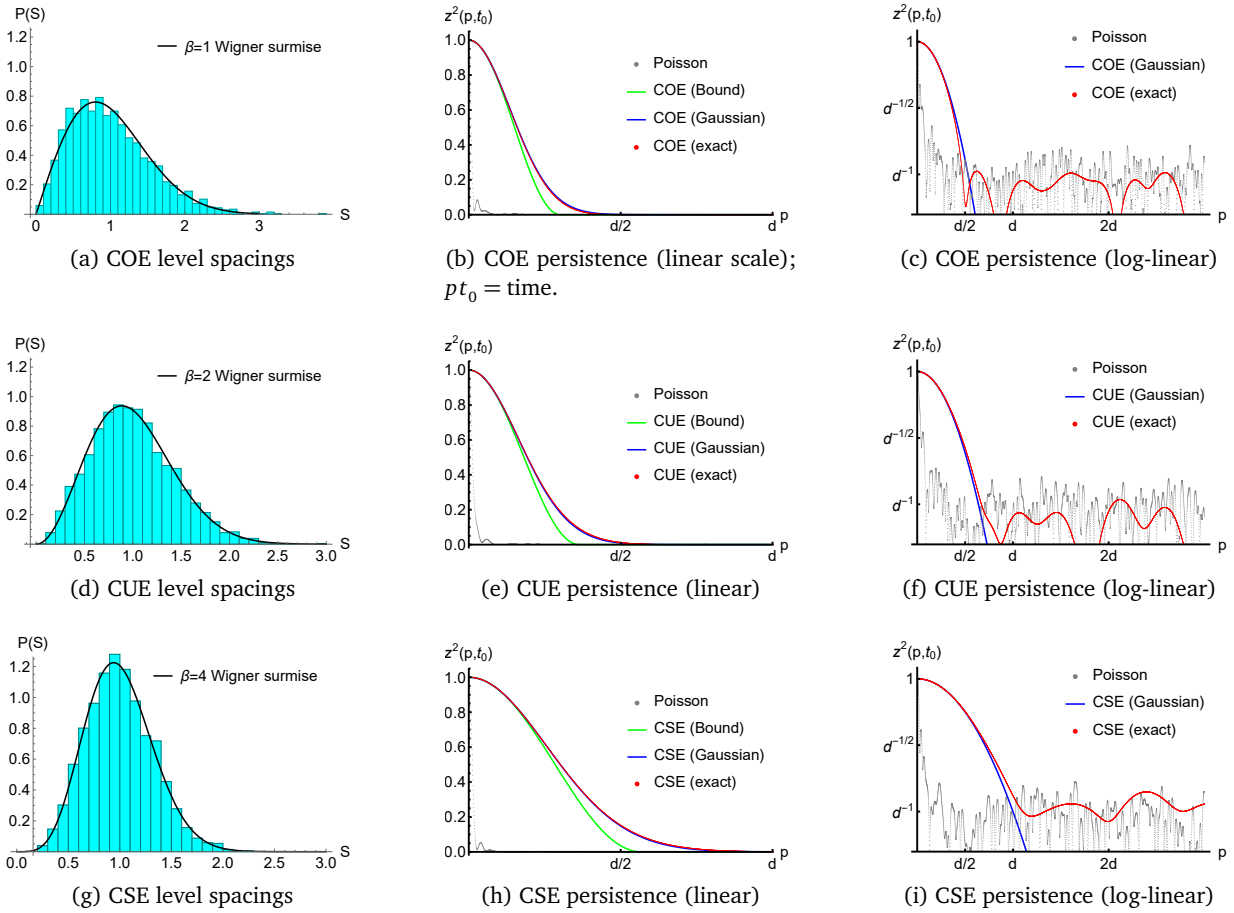
$$\sigma_\Delta^2 = \frac{\alpha^2}{4\pi^2} \ln d + O(1), \text{ with } \alpha \in [1,2). \tag{2.58}$$

This amounts to a derivation of Eq. (2.40).

The logarithmic growth of the variance of mode fluctuations with the dimension d of the energy subspace is a direct consequence of the Gaussianity of the persistence. In less idealized situations, it is possible to have a non-Gaussian tail in Eq. (2.50), for $p \gtrsim 1/\sqrt{\varepsilon_C(1,t_0)}$, even if the Gaussian estimate holds for smaller times. It is worth noting that non-Gaussian tails at long times would show up as non-Gaussianities near $\Delta \approx 0$ in the mode fluctuation distribution; such deviations from Gaussianity are largely determined by the complicated correlations between the errors ν_m, partly encoded in the fluctuations of the SFF $K(pt_0)$. The main takeaway here is instead the extremely specific numerical range $\alpha \in [1,2)$, of the coefficient multiplying the logarithm, demanded by ergodicity and aperiodicity. For non-Gaussian tails, one would have a similarly specific range of some other parameter.

2.5.4.2 Ergodicity and aperiodicity for Wigner-Dyson spectral rigidity

Now we consider Wigner-Dyson random matrix ensembles as well as individual systems with Wigner-Dyson level statistics within energy subspaces Σ_d with $\Delta E \lesssim 2\pi/t_{\text{ramp}}$. There is numerical evidence that the mode fluctuation distribution is exactly Gaussian [79,101,102] (as well as an analytical proof of Gaussianity for a closely related measure, number fluctuations [115]) especially near $\Delta \approx 0$, suggestive of an almost Gaussian persistence even at late times. In Refs. [46,47], the leading behavior of the variance σ_Δ^2 (there called Δ^*) for Wigner-Dyson ensembles has been shown to be equal to that of the (spectrum or ensemble averaged) spectral rigidity parameter $\Delta_3(d)$ [36,40,116] — measuring the variance of the "spectral staircase" around a best fit straight line — when $t_0 = 2\pi/\Delta E$ is determined by the slope of the straight line and $q = \overline{q}$ is the sorting permutation. Moreover, $\Delta_3(d)$ can be



(a) COE level spacings

(b) COE persistence (linear scale); $pt_0 = \text{time}$.

(c) COE persistence (log-linear)

(d) CUE level spacings

(e) CUE persistence (linear)

(f) CUE persistence (log-linear)

(g) CSE level spacings

(h) CSE persistence (linear)

(i) CSE persistence (log-linear)

Figure 2.7: Numerical support for ergodicity and aperiodicity of realizations of Wigner-Dyson random matrix ensembles, for $d = 2048$, $t_0 = 1$ and q being the sorting permutation. Level spacing data, showing the distribution of nearest neighbor level spacings $S \in \{E_{\overline{q}(k+1)} - E_{\overline{q}(k)}\}$, depicts the closeness of the realization to an ideal Wigner-Dyson distribution as given by the appropriate Wigner surmise [28, 36]. Persistence is plotted (in red) in terms of persistence probabilities $z^2(p, t_0)$. The lower bound ("Bound", green) of Eq. (2.39) is satisfied, and good agreement is seen with the Gaussian estimate ("Gaussian", blue) of Eq. (2.50) including the tail at long times; both are calculated based on the numerical value of $\varepsilon_C(1, t_0)$ for the realization. The Poisson persistence probability fluctuations (for a sorted, uncorrelated distribution of points in the same range of energies/eigenphases; in gray) are included to provide a visual reference for the range of persistence probabilities that should be considered $O(d^{-1})$ for random states, while simultaneously substantiating the non-ergodicity of Poisson statistics. The time scales when the random matrix persistence amplitudes reach $O(d^{-1})$ are consistent with $p = d/2$ for COE, $p = d/\sqrt{2}$ for CUE and $p = d$ for CSE as predicted by Eqs. (2.56), (2.58) and (2.59).

calculated exactly [36,40,45] by using the appropriate Wigner-Dyson ensemble averaged SFF $\overline{K_{\beta_{\text{RMT}}}(t)}$ for the energy subspace.

In fact, the leading contribution for large d comes only from the ramp at $t \ll t_{\text{H}}$, given by

$\overline{K_{\beta_{\mathrm{RMT}}}(t)} \approx t/(\beta_{\mathrm{RMT}}\pi\mathcal{N}^{\#}d)$ with $\beta_{\mathrm{RMT}} = 1, 2, 4$ respectively for COE, CUE and CSE (much like in the derivation of Eq. (2.55)). The result is a logarithmic dependence of σ^2_Δ on d to leading order (for $t_0 = 1$ and q being the sorting permutation),

$$\sigma^2_\Delta\big|_{\text{Wigner-Dyson}} = \frac{1}{\beta_{\mathrm{RMT}}\pi^2}\ln d + O(1), \text{ with } \beta_{\mathrm{RMT}} \in \{1, 2, 4\}. \tag{2.59}$$

This precisely corresponds (via $\sigma^2_\Delta = g_1^2 d^2/(4\pi^2)$) to an error that saturates the lower bound in Eq. (2.55), providing an important sanity check. Comparing this with Eq. (2.58) (or Eq. (2.40)), we see that the Wigner-Dyson ensembles span exactly the range of allowed coefficients for ergodic, aperiodic systems with a Gaussian persistence. CUE is well within this range, whereas COE is at the upper bound and barely ergodic while CSE is at the lower bound and just about aperiodic (here, it is worth noting that for CSE, we have considered only one non-degenerate half of the doubly-degenerate spectrum as is conventional [28, 36]).

It remains to be verified that $z(p, t_0)$ for Wigner-Dyson random matrix ensembles is indeed well approximated by a Gaussian all the way until $p \in [d/2, d]$ as suggested by the ideal Gaussian distribution of mode fluctuations, so that the identification between Eq. (2.59) and Eq. (2.58) can be made with some confidence. We provide numerical support for this statement in Fig. 2.7 for $d = 2048$.

2.6 Cyclic ergodicity and spectral rigidity in 2D KAM tori

In this section, we study the quantum ergodic properties of linear flows on 2D tori, a class of non-mixing dynamical systems for which the framework of quantum cyclic permutations allows an analytical proof of spectral rigidity when the system is ergodic — thereby rigorously demonstrating the link between cyclic ergodicity and spectral rigidity at the level of an individual system, without relying on ensembles or heuristic arguments. It is convenient to classify these tori into the ergodic

KAM tori (with an irrational frequency ratio) and non-ergodic rational tori (with a rational frequency ratio). For some physical context, we note that both occur as invariant subsets of classically integrable systems, with almost all KAM tori remaining stable under small perturbations [48, 117]. It is particularly interesting to note that KAM tori, though ergodic, possess neither periodic orbits nor random matrix (-like) level statistics as required in semiclassical explanations of spectral rigidity [28, 40, 73]; the results of this section (together with Sec. 2.5 concerning random matrix level statistics) therefore suggest a wider applicability of cyclic permutations as a way to understand spectral rigidity in terms of ergodic properties.

In Sec. 2.6.1, we discuss a simple quantization procedure for these tori, motivated by the standard quantization of integrable systems. In Sec. 2.6.2, building on the example of irrational rotations discussed in Sec. 2.2.2.2, we prove that all 2D KAM tori are both classical and quantum cyclic ergodic (if suitably quantized), while all rational tori are non-ergodic. Finally in Sec. 2.6.3, we show using the corresponding cyclic permutations and Theorem 2.6 that all 2D KAM tori possess much higher spectral rigidity than Poisson and even Wigner-Dyson statistics (as indicated by a smaller $\varepsilon_C(1, t_0)$ or σ^2_Δ), and verify this result numerically. We note that this remains consistent with the Berry-Tabor conjecture of Poisson statistics in typical integrable systems [34], as the latter typically contain several uncorrelated KAM tori overlapping with each other in an energy shell, leading to enhanced spectral fluctuations (similar to Refs. [91, 112]) and low spectral rigidity of the integrable system.

2.6.1 Quantization of linear flows on a torus

The Hamiltonian of a linear flow with frequencies $\boldsymbol{\omega} = (\omega_x, \omega_y)$ on a 2D torus is given by

$$H = \mathbf{J} \cdot \boldsymbol{\omega} = J_x \omega_x + J_y \omega_y, \tag{2.60}$$

with angle variables $\boldsymbol{\theta} = (\theta_x, \theta_y) \in [0, 2\pi)^2$ conjugate to the action variables $\mathbf{J} = (J_x, J_y)$. The equation of motion of the linear flow is $\mathscr{T}^t \boldsymbol{\theta} = \boldsymbol{\theta} + \boldsymbol{\omega} t$.

The ergodicity of this system on the 2D phase space $\mathscr{P}_\mathbf{J} = \{(\theta_x, \theta_y)\}$ with fixed $\mathbf{J}$ is characterized [11, 12] by the ratio $\alpha = \omega_y/\omega_x$. When α is irrational, the dynamics is ergodic on this phase space, and the system may be called a KAM torus[7] [48, 118]; but when α is rational, $\mathscr{P}_\mathbf{J}$ decomposes into an infinite number of invariant ergodic and periodic subsets, which share the same period. In both cases, there is no mixing. We emphasize that the system of Eq. (2.60) is different from a free particle moving on a torus. The latter is never ergodic in its *phase space*, but possibly (depending on initial conditions) merely visits all points among its position coordinates with conserved momentum. The Hamiltonian of the free particle is quadratic in $\mathbf{J}$ rather than linear, and its level statistics has been found to be Poissonian [119, 120] (see also Ref. [102] for its mode fluctuations) in accordance with the Berry-Tabor conjecture [34] for fully integrable systems.

The quantization of Eq. (2.60) can be motivated by considering how such tori occur as invariant subsets of integrable systems [34, 48, 117]. The Hamiltonian of a 2D integrable system with constants of motion $\mathbf{I} = (I_x, I_y)$ is given by $H = H(\mathbf{I})$, with quantization restricting each $\hat{I}_k \in \mathbf{N}_0$ to the non-negative integers due to the periodicity of the conjugate angle variables. Classically, expanding $\mathbf{I} = \mathbf{I}_0 + \mathbf{J}$ in the neighborhood of $\mathbf{I}_0$ (say, with $J_x, J_y \geq 0$) gives Eq. (2.60) to leading order (up to an $\mathbf{I}_0$-dependent constant) with $\boldsymbol{\omega}(\mathbf{I}_0) = [\partial H/\partial \mathbf{I}](\mathbf{I}_0)$. Thus, a natural way to quantize Eq. (2.60) in terms of operators $\hat{\mathbf{J}}$, $\hat{\boldsymbol{\theta}}$ is to restrict the eigenvalues J_x, J_y of $\hat{J}_x, \hat{J}_y$ each to a finite (but large) set of consecutive integers; up to additive constants in H, this is equivalent to working in the Hilbert space Σ_d spanned by

$$\left\{ |\mathbf{J}\rangle : J_x \in \mathbb{Z}_{d_x}, J_y \in \mathbb{Z}_{d_y} \right\}, \tag{2.61}$$

[7] There is an additional formal requirement that $|\alpha - m/n| \geq c(|n| + |m|)^{-\gamma}$ for some constant c, with $\gamma \geq 2$, and all $m, n \in \mathbb{Z} - \{0\}$ in the strict mathematical definition of a KAM torus, required for stability under perturbations [118]. All but a negligible fraction of irrationals satisfy this property (especially as γ can be arbitrarily large), and we will ignore this restriction when using the terminology.

with $d = d_x d_y$. Of particular interest is the following manifold of $\boldsymbol{\theta}$-states corresponding to points on the torus:

$$|\boldsymbol{\theta}\rangle = \frac{1}{\sqrt{d_x d_y}} \sum_{J_x \in \mathbb{Z}_{d_x}} \sum_{J_y \in \mathbb{Z}_{d_y}} e^{-i\mathbf{J}\cdot\boldsymbol{\theta}} |\mathbf{J}\rangle \ \in \Sigma_d. \tag{2.62}$$

We can construct infinitely many complete orthonormal bases for Σ_d in this manifold, each comprised of d such states, such as $\{(\theta_x, \theta_y) = (2\pi n_x/d_x, 2\pi n_y/d_y) : n_x \in \mathbb{Z}_{d_x}, n_y \in \mathbb{Z}_{d_y}\}$. It is straightforward to see that time evolution reproduces the classical equation of motion in this manifold:

$$\hat{U}_H(t)|\boldsymbol{\theta}\rangle \equiv e^{-i\hat{H}t}|\boldsymbol{\theta}\rangle = |\boldsymbol{\theta} + \boldsymbol{\omega}t\rangle, \tag{2.63}$$

which provides a natural way to map classical cyclic permutations onto Σ_d.

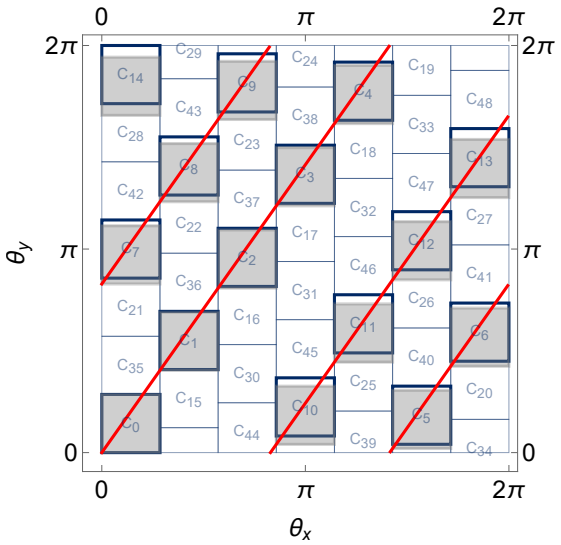
2.6.2 Classical and quantum cyclic ergodicity

In this and the next subsection, we will see how cyclic permutations can be used to rigorously prove spectral rigidity for KAM tori. First, we show that all 2D KAM tori admit a classical cyclic permutation (sequence) that is cyclic ergodic and not aperiodic. This is made possible by identifying the discretized flow on the torus as a sequence of irrational rotations, allowing us to use the cyclic permutations constructed for the latter in Sec. 2.2.2.2. Using Eqs. (2.62) and (2.63) to adapt this construction to a quantum cyclic permutation of an orthonormal basis in Σ_d, we will show the following:

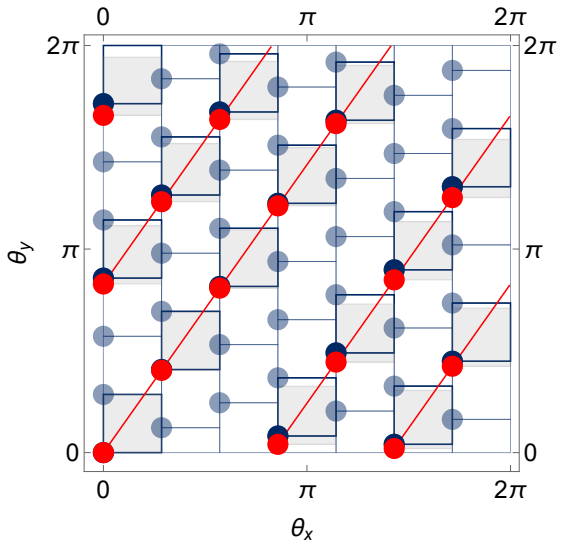
Theorem 2.8 (**Quantum cyclic ergodicity and non-aperiodicity of 2D KAM tori**)**.** *For every linear flow on a 2D torus with irrational frequency ratio $\alpha = \omega_y/\omega_x$, there exists an infinite sequence of energy subspaces Σ_d with $d \to \infty$ (such that $d_x, d_y \to \infty$) in which the system is ergodic and not aperiodic, within any time $T > \min(2\pi d_y/\omega_x, 2\pi d_x/\omega_y)$.*

Thus, the quantum dynamical properties of 2D KAM tori reflect their classically cyclic ergodic and non-aperiodic nature, which is further consistent with their ergodicity and lack of mixing.

2.6.2.1 Classical cyclic permutation



(a) The classical cyclic permutation of Eq. (2.64) with $n = 49$, depicting the C_k (blue-edged squares; thicker edges for $k = 0$ to 14). The evolution $\mathscr{T}^{pt_0}C_0$ of the set C_0 (gray,filled) and the trajectory $\mathscr{T}^t\boldsymbol{\theta}$ of the point $\boldsymbol{\theta} = (0,0)$ (red) are depicted respectively for $p = 0$ to 14 and $t = 0$ to $14t_0$. The corresponding p-step errors are determined by the portion of each highlighted square that does not intersect a gray region.



(b) The quantum cyclic permutation of Eq. (2.65) with $d = 49$, overlaid on a depiction of the corresponding classical cyclic permutation as per Fig. 2.8a. Filled disks represent pure states in $|\boldsymbol{\theta}\rangle$-space. The depicted pure states are the $|C_k\rangle$ (blue; highlighted/opaque for $k = 0$ to 14, transparent otherwise) and $\hat{U}_H(pt_0)|C_0\rangle$ for $p = 0$ to 14 (red, lying on classical trajectory).

Figure 2.8: The (a) classical cyclic permutation of Eq. (2.64) with $n_x = n_y = 7$, $m_y = 10$ and $t_0 = 2\pi/(\omega_x n_x)$, and (b) the corresponding quantum cyclic permutation of Eq. (2.65) with $d_x = d_y = 7$, $m_y = 10$ and $t_0 = 2\pi/(\omega_x d_x)$, for an irrational flow on the 2D torus with $\alpha = \sqrt{2}$. For these parameters, we have $\delta_{n_y} \approx 0.704/n_y < 1/n_y$ (and likewise with d_y), making both the classical and quantum cyclic permutation ergodic and non-aperiodic. Trajectories starting from $\boldsymbol{\theta} = (0,0)$ are depicted up to $p = 2d_x = 14$, which corresponds to $r = 2$ steps of the irrational rotation $\mathscr{T}^{rT_x}\theta_y$ at $\theta_x = 0$.

For the classical construction, we first observe that for $T_x \equiv 2\pi/\omega_x$ (the period along the θ_x direction), the map $\mathscr{T}^{rT_x}(\theta_x,\theta_y) = (\theta_x,\theta_y + 2\pi\alpha r)$ with $r \in \mathbb{Z}$ is a discrete rotation of any $\theta_x = \text{const.}$ circle in the θ_y direction, with $\alpha = \omega_y/\omega_x$. Writing $\alpha = (m_y + \delta_{n_y})/n_y$ as in Sec. 2.2.2.2 with coprime integers m_y and n_y, we know that every rotation admits an n_y-element cyclic permutation $\mathtt{C}_y$ with error $|\delta_{n_y}|$ by Eq. (2.8). To adapt this to the torus, we divide the θ_x direction into n_x elements and consider time steps in units of $t_0 = T_x/n_x$, such that each step captures a $1/n_x$ fraction of the error generated in the full rotation by $T_x = n_x t_0$ (see Fig. 2.8a). Formally, our cyclic permutation $\mathtt{C}$ for the

2D torus consists of:

$$C_k = \left[2\pi\frac{k}{n_x}, 2\pi\frac{k+1}{n_x}\right]_{\theta_x} \times \left[2\pi\frac{\frac{k}{n_x}m_y}{n_y}, 2\pi\frac{\frac{k}{n_x}m_y+1}{n_y}\right]_{\theta_y},$$
$$\text{with error } \overline{\epsilon}_C\left(t_0 = \frac{2\pi}{\omega_x n_x}\right) = \frac{|\delta_{n_y}|}{n_x}. \tag{2.64}$$

It is easily verified that the intersection of the C_{qn_x} (for $q \in \mathbb{Z}$) with $\theta_x = 0$ gives precisely the cyclic permutation of Eq. (2.8) for the rotation $\mathscr{T}^{rT_x}$. From the discussion of irrational rotations in Sec. 2.2.2.2, we can find a sequence of coprime $m_y, n_y \to \infty$ such that $|\delta_{n_y}| < 1/n_y$ for all irrational α. Thus, we have $\overline{\epsilon}_C(t_0) < 1/n$ with $n \to \infty$ for all n_x, establishing the existence of a *classical* cyclic permutation showing cyclic ergodicity and non-aperiodicity (by the bounds discussed in Sec. 2.2.2) for all 2D KAM tori.

There are two straightforward extensions of this construction that are worth mentioning: 1. choosing $t_0 = 2\pi/(\omega_y n_y)$ and considering irrational rotations in the θ_y direction gives essentially similar results, and 2. for rational α, the non-ergodicity of the flow implies that no cyclic permutation with the C_k reducing to (neighborhoods of) points in the $n \to \infty$ limit is ergodic or aperiodic, as can be shown by suitably adapting the corresponding discussion for rational rotations in Sec. 2.2.2.2.

2.6.2.2 Quantum cyclic permutation

We note that among the $\boldsymbol{\theta}$-states of Eq. (2.62), any two states $|\boldsymbol{\theta}\rangle$, $|\boldsymbol{\theta}'\rangle$ are orthogonal if $(\theta_x - \theta_x') = 2\pi\ell/d_x$ or $(\theta_y - \theta_y') = 2\pi\ell/d_y$ for some $\ell \in \mathbb{Z}$. A pure state cyclic permutation corresponding to Eq. (2.64) can then be constructed using these states, by formally setting $d_x = n_x$, $d_y = n_y$ and considering the set of points given by the d "lower" corners of the classical C_k in Eq. (2.64) (see also

Fig. 2.8b), namely:

$$|C_k\rangle = \left|\boldsymbol{\theta} = \left(2\pi\frac{k}{d_x}, 2\pi\frac{\frac{k}{d_x}m_y}{d_y}\right)\right\rangle, \text{with } \hat{U}_C|\boldsymbol{\theta}\rangle = \left|\boldsymbol{\theta} + \left(\frac{2\pi}{d_x}, \frac{2\pi m_y}{d_x d_y}\right)\right\rangle, \tag{2.65}$$

and $t_0 = 2\pi/(\omega_x d_x)$. From Eqs. (2.63) and (2.65), it is clear that $[\hat{U}_H(t), \hat{U}_C] = 0$ as both $\hat{U}_H(t)$ and $\hat{U}_C$ are additive maps on the torus; consequently, the above cyclic permutation corresponds to a DFT of the energy eigenstates.

Using Eqs. (2.62) and (2.63), it is straightforward to evaluate the p-step persistence probabilities $z^2(p, t_0) = |\langle C_{k+1}|\hat{U}_H(pt_0)|C_k\rangle|^2$ in terms of the frequency ratio $\alpha = (m_y + \delta_{d_y})/d_y$. We obtain the following exact expression for $p \neq 0$:

$$z^2\left(p, t_0 = \frac{2\pi}{\omega_x d_x}\right) = \frac{\sin^2\left[p\pi\frac{\delta_{d_y}}{d_x}\right]}{d_y^2 \sin^2\left[p\pi\frac{\delta_{d_y}}{d_x d_y}\right]}. \tag{2.66}$$

To discuss ergodicity and aperiodicity in the quantum case, we need a slightly more refined implication of Eq. (2.9) concerning Diophantine approximations than in the classical case, namely that every irrational α admits an infinite sequence of m_y, d_y with $|\delta_{d_y}| < c/d_y$ for some constant $c < 1$ (this constant ensures that $|\delta_{d_y}|$ does not even approach $1/d_y$ asymptotically, as $d_y \to \infty$). Using $(\sin x)^2 \leq x^2$ in the denominator of Eq. (2.66), we get that $z^2(pt_0) \geq \text{sinc}^2[p\pi\delta_{d_y}/d_x]$; as the smallest zeros of the right hand side occur at $|p| = d_x/\delta_{d_y}$, $z^2(p, t_0)$ is guaranteed to remain $\Theta(1)$, i.e., no less than $O(1)$ for any d_y belonging to the above sequence, until some $|p| > d_x d_y/c > d$ (as $c < 1$). It follows that $z^2(p, t_0) \gg O(d^{-1})$ for $|p| \leq d$, making the constructed cyclic permutation ergodic and non-aperiodic (the latter via $z^2(\pm d, t_0) = K(\pm t_0 d)$). As this is a DFT cyclic permutation, it follows from the considerations of Sec. 2.4.1.3 that no cyclic permutation in Σ_d is aperiodic. On using these results in Definition 2.5 with $t_0 d = 2\pi d_y/\omega_x$, while noting that analogous results would hold if we

swap the θ_x and θ_y directions, we directly get Theorem 2.8.

When $\alpha = m/\ell$ is rational with coprime m and ℓ, the motion in $|\boldsymbol{\theta}\rangle$-space is periodic with period $T_\alpha = 2\pi\ell/\omega_x$. In particular, we have

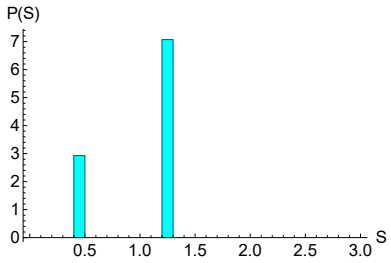
$$K(rT_\alpha) = 1, \ \forall \ r \in \mathbb{Z}, \tag{2.67}$$

immediately ruling out aperiodicity in any Σ_d over any time $T > T_\alpha$. Further, from Eq. (2.46), we get $z_k(p, t_0) = 0$ for every cyclic permutation in some Σ_d with $pt_0 = rT_\alpha$, which rules out ergodicity for any cyclic permutation with $t_0 = rT_\alpha/p$ with $(p \in \mathbb{Z}) \leq d/2$. Ruling out ergodicity for more general t_0 would require a more detailed study of spectral properties.

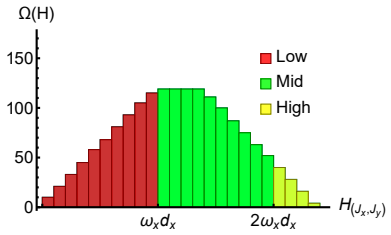
2.6.3 Spectral rigidity

Having established the classical and quantum cyclic ergodicity of 2D KAM tori, as well as non-aperiodicity in both the irrational and rational cases, we now turn to an analytical and numerical study of the spectral rigidity of the quantized systems, with the main results depicted in Fig. 2.9. The distribution of the energy levels depends sensitively on α (via $H = (J_x + J_y\alpha)\omega_x$ for $J_x \in \mathbb{Z}_{d_x}$, $J_y \in \mathbb{Z}_{d_y}$), and this dependence has been studied numerically and analytically in the context of the 2D harmonic oscillator in e.g. Refs. [34, 121, 122]. Their main analytical result is that the nearest-neighbor level spacings $S \in \{E_{\bar{q}(k+1)} - E_{\bar{q}(k)}\}$ within sufficiently small energy windows takes the form of at most 3 delta functions, $P(S) = \sum_{j=1}^{3} c_j \delta(S - S_j)$ with $c_j, S_j \geq 0$; this behavior is shown in our context in Fig. 2.9a. This distribution is quite unlike the smooth level spacing distributions of Wigner-Dyson (compare with Fig. 2.7) or Poisson level statistics [28], and these results leave open the question of how the spectrum may be quantitatively characterized in direct relation to some dynamical property of the system.

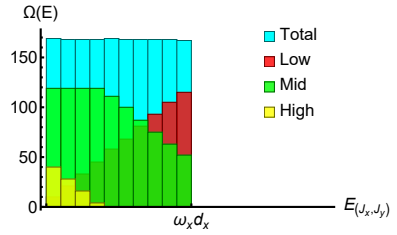
The cyclic permutation of Eq. (2.65) provides precisely such a characterization, via the persis-



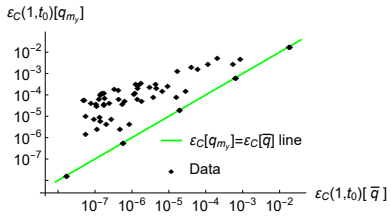
(a) Level spacings.



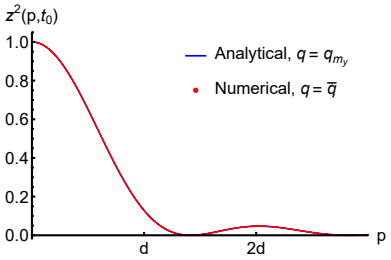
(b) Non-uniform density of states in H.



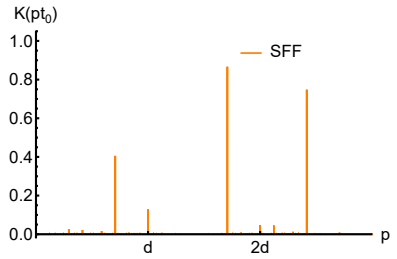
(c) Uniform density of states in E.



(d) Analytically derived error ($q = q_{m_y}$) vs numerical sorted DFT error ($q = \overline{q}$); $1 \le d_x = d_y \le 100$ (with d ranging up to 10^4), excluding points where m_y and d_y are not coprime.



(e) Cyclic ergodicity in terms of the analytically derived persistence and the numerical sorted DFT persistence (linear), showing extremely close agreement.



(f) Numerical SFF (linear), showing non-aperiodicity.

Figure 2.9: Spectral properties of 2D KAM tori with $\alpha = \sqrt{2}$, compared with analytical predictions corresponding to the cyclic permutation with $t_0 = 2\pi/(\omega_x d_x)$ constructed in Sec. 2.6.2; $d_x = d_y = 41$ and $d = 1681$, except in (d) which considers a range of d_y. (a) Level spacings, required to be at most 3 delta functions. (b) and (c): comparison of unwrapped and wrapped density of states, showing how the choice of $t_0 = 2\pi/(\omega_x d_x)$ leads to a uniform "wrapped" density of states, when "low", "mid" and "high" energy portions of the spectrum demarcated by multiples of $\omega_x d_x$ are considered separately. (d) Comparison of the analytical prediction of Eq. (2.68), where m_y is numerically chosen to be the nearest integer to $d_y\alpha$, with the sorted DFT error. (e) The persistence amplitude of Eq. (2.66) for $m_y = 58$ (with $\delta_{d_y} \approx 0.707/d_y$) compared with the sorted DFT persistence obtained numerically; this also shows that $d_x = d_y = 41$ belongs to the infinite sequence of ergodic, non-aperiodic Σ_d in Theorem 2.8. (f) Numerically computed SFF, showing a peak at $p = d$ (as expected analytically) and elsewhere, confirming non-aperiodicity.

tence probabilities in Eq. (2.66). As discussed in Sec. 2.4.2, these quantities are themselves measures of energy level statistics when the cyclic permutation is a DFT of energy eigenstates, due to Eq. (2.32). The associated map q that permutes the energy levels before comparing them to regularly spaced levels is determined by the choice of m_y; for convenience, we denote this map q_{m_y} for the cyclic permutation of Eq. (2.65). Then, setting $d_x = d_y$ for convenience and using Eq. (2.66) with $\delta_{d_y} < 1/d_y$ to obtain

the single-step DFT error $\varepsilon_C(1,t_0)[q_{m_y}] = 1 - z^2(1,t_0)[q_{m_y}]$, we get

$$\varepsilon_C\left(1, t_0 = \frac{2\pi}{\omega_x d_x}\right)\left[q_{m_y}\right] < \frac{\pi^2}{3d^2} + O(d^{-3}), \tag{2.68}$$

where d belongs to the infinite sequence occurring in Theorem 2.8 (with $d_x = d_y$), for every irrational α. By the discussion in Sec. 2.5.3, we expect $\varepsilon_C(1,t_0)[q] \gtrsim O(d^{-1})$ for Poisson statistics, and $\varepsilon_C(1,t_0)[q] \gtrsim O(d^{-2}\ln d)$ for Wigner-Dyson statistics, for all q; Eq. (2.68) thus constitutes an analytical proof that all 2D KAM tori have *significantly* "more rigid" spectra than both Wigner-Dyson and Poisson statistics in the appropriate energy subspaces. In Appendix A.4, we argue that rational tori are non-ergodic over a range of t_0 and show spectral rigidity of the same order of magnitude as Poisson statistics.

Another measure of spectral rigidity is the sorted DFT single-step error $\varepsilon_C(1,t_0)[\overline{q}]$, which is most directly related to the conventionally studied mode fluctuations according to Sec. 2.4.2. Given our choice of t_0, the permutation $q = \overline{q}$ corresponds to a direct sorting of the "wrapped" energy levels:

$$\begin{aligned} E_{(J_x,J_y)} t_0 &= H(J_x,J_y)\frac{2\pi}{\omega_x d_x} \bmod 2\pi \\ &= 2\pi\left[\left(\frac{J_x}{d_x} + \frac{J_y\alpha}{d_x}\right) \bmod 1\right], \end{aligned} \tag{2.69}$$

which are essentially the eigenphases of the time evolution operator $\hat{U}_H(t_0 = 2\pi/[\omega_x d_x])$. Indeed, while the density of states is highly nonuniform with linearly growing and decaying parts when expressed in terms of $H(J_x,J_y)$, it naturally appears uniform in terms of the wrapped energies $E_{(J_x,J_y)}$ at t_0 (or the eigenphases of $\hat{U}_H(t_0)$); see Figs. 2.9b, 2.9c. It is this "wrapping" that allows the spectrum to have energy levels close to a regularly spaced spectrum (and correspondingly high spectral rigidity) as per the discussion in Sec. 2.4.2.

We should also expect the analytically calculated DFT error of Eq. (2.68) to not be less than the

sorted DFT error (see also the corresponding discussion in Sec. 2.4.2):

$$\varepsilon_C(1,t_0)[q_{m_y}] \geq \varepsilon_C(1,t_0)[\overline{q}]. \tag{2.70}$$

We numerically verify this to be the case for Σ_d (with $d_x = d_y$) both belonging to and outside the infinite sequence to which Theorem 2.8 applies; however, we specifically appear to have $\varepsilon_C(1,t_0)[q_{m_y}] \approx \varepsilon_C(1,t_0)[\overline{q}]$ for several Σ_d including all those belonging to the infinite sequence (i.e. with $|\delta_{d_y}| < 1/d_y$). This is shown in Fig. 2.9d. Further, we find that the persistence amplitudes $z^2(p,t_0)[\overline{q}]$ are also in close numerical agreement with Eq. (2.66) for Σ_d in this sequence, as seen in Fig. 2.9e. Finally, the SFF is shown in Fig. 2.9f, with a large $O(1)$ peak expected at $p = \pm d$ from Eq. (2.66).

In conclusion, we have analytically proven and numerically verified that all quantized 2D KAM tori admit an infinite sequence of energy subspaces that are cyclic ergodic and non-aperiodic, with an exact expression for the spectral rigidity that is of a higher order of magnitude than even Wigner-Dyson level statistics (i.e. has lower error). On the other hand, quantized rational tori were shown to be non-aperiodic and argued (in Appendix A.4) to have spectral rigidity comparable to Poisson statistics, making them unlikely to be ergodic in any subspace. These quantum properties directly reflect the discretized classical ergodic properties of these systems. Some additional numerical evidence (outside the scope of this Chapter) further indicates that it is possible, and perhaps even typical, to have cyclic ergodic and non-aperiodic subspaces for KAM tori even when Σ_d is not in the infinite sequence for which we are able to prove these properties analytically. These subspaces often show an almost ideal Gaussian persistence and an apparent $O(d^{-2}\ln d)$ error (e.g., for $\alpha = \sqrt{2}$), but we do not yet have a quantitative understanding of this behavior.

2.7 Discussion: Dynamical ergodicity and thermalization

In classical dynamical systems, it is usually unnecessary to differentiate between two equivalent notions of ergodicity. The first is "dynamical ergodicity", by which we mean the dynamical behavior where almost all initial states in phase space visit the neighborhood of every other state. The second is "thermal ergodicity", referring to the initial-state-independence of time-averaged expectation values of any observable, as quantified by Eq. (2.1). The direct equivalence [10, 11] between the dynamical and thermal property is at the foundation of classical statistical mechanics [13].

The analogous picture for quantum dynamical systems is the main subject of this section, which we will approach in a semi-qualitative manner. As noted in Sec. 2.1, ETH has been proposed as a quantum analogue (or at least, a sufficient condition) for the thermal ergodicity of physical observables, and is largely a statement about energy eigenstates in some "physical" basis (with minimal conditions on the energy eigenvalues themselves, namely non-degeneracy and rational incommensurability) [21, 22, 52]. The detailed study of quantum cyclic permutations in the previous sections strongly suggests that quantum cyclic ergodicity, which directly decodes spectral rigidity into a dynamical property in Hilbert space, is the appropriate quantum counterpart to dynamical ergodicity. The open problem of the connection between dynamical and thermal properties in quantum statistical mechanics is therefore essentially equivalent (through the framework of cyclic permutations) to that of the connection between the observable-independent energy eigenvalues and observable- (or basis-)dependent energy eigenstates of a quantum system.

The close correspondence between quantum and classical cyclic ergodicity suggests that one may explore these connections in systems with a classical limit, and apply it to a more general class of quantum systems that show spectral rigidity and ETH. In particular, we may treat a quantum cyclic permutation $\mathscr{C}$ in some Σ_d as representing a fictitious phase space (or energy shell) $\mathscr{P}[\mathscr{C}]$ in systems

with or without a classical limit. By Weyl's law [28], an ensemble of $\mu(A)d$ pure states in $\mathscr{C}$ represents a region $\mathscr{P}[\mathscr{C}]$ of measure $\mu(A)$ (see Sec. 2.3, and Refs. [26, 27]). Such an ensemble corresponds to a mixed state of the form

$$\hat{\rho}(A) = \frac{1}{\mu(A)d} \sum_{k \in C(A)} |C_k\rangle\langle C_k|, \tag{2.71}$$

where $C(A) \subseteq \mathbb{Z}_d$ is a set of indices of size $|C(A)| = \mu(A)d$ (with $|\cdot|$ denoting the cardinality of a finite set). Interestingly, while such a connection between mixed states and phase space regions could in principle be used to quantitatively connect the quantum error $\varepsilon_C(1, t_0)$ to that of the classical cyclic permutations, it is not clear how to do this in practice. This is because the dominant contribution to the classical error is in the lack of overlap of these mixed states, which is determined more strongly by their support in Hilbert space [123, 124] than the much smaller contribution from the error in the overlap of the constituent pure states, except in the trivial $\mu(C_k) \to 1/d$ pure state limit.

It is also worth noting that not all $\mathscr{P}[\mathscr{C}]$ correspond to the actual phase space $\mathscr{P}$ in systems with a classical limit, but likely only those that satisfy the *generating property* with respect to classical phase space regions (see Sec. 2.2.2). At a strictly formal level, any n-element partition $B_k(n)$ of $\mathscr{P}$ can be used to define non-degenerate observables $b[B_k(n)]$ that take distinct values on each $B_k(n)$ as $n \to \infty$ (e.g. b is essentially $\boldsymbol{\theta}$ for the tori of Sec. 2.6); such observables admit an orthonormal eigenbasis $\mathscr{B} = \{|B_k\rangle\}$ by the standard procedure of quantization, such as the postulates in Ref. [62]. The quantum analogue of the generating property then appears to be

$$(\mathscr{C} = \mathscr{B}) \implies (\mathscr{P}[\mathscr{C}] = \mathscr{P}), \tag{2.72}$$

for some such $\mathscr{B}$, and is therefore an observable-dependent eigenstate property, unlike cyclic ergodicity.

2.7.1 Quantum analogues of classical ergodicity and mixing

If such a generating $\mathscr{C}$ has error $\varepsilon_C(1,t_0)$, we can approximate $\hat{U}_H(pt_0) \approx \hat{U}_C^p$ whenever $|p| \ll 1/\sqrt{\varepsilon_C(1,t_0)}$ (from the bound of Eqs. (2.19), (2.39)). As we will see below, this approximation implies that late-time quantum dynamics in the fictitious phase space basis $\mathscr{P}[\mathscr{C}]$ can be described in nearly classical terms up to extremely long times for sufficiently rigid spectra (e.g. $t_{\mathrm{H}}/\sqrt{\ln(t_{\mathrm{H}}/t_{\mathrm{ramp}})}$ with $t_0 = t_{\mathrm{ramp}}$, $t_0 d = t_{\mathrm{H}}$ for Wigner-Dyson spectra), standing in contrast to rapid "quantum" thermalization processes such as scrambling that cause an initial state to rapidly spread in some "local" basis over much shorter times (closer to t_{ramp}) without any significant influence of spectral rigidity beyond $t > t_{\mathrm{ramp}}$ [38, 41, 92, 95, 112, 114, 125].

Using the above approximation with Eq. (2.71) for $A, B \subseteq \mathscr{P}$, we have

$$\mu(\mathscr{T}^{pt_0}A \cap B) \approx \frac{1}{d}\left|C(\mathscr{T}^{pt_0}A) \cap C(B)\right|. \tag{2.73}$$

In this way, thermal properties of fictitious phase space regions A,B can be represented over this time scale as simple correlation functions of subsets $C(A), C(B)$ of $\mathbb{Z}_d$ under relative shifts, which may obey e.g. Eq. (2.1) or Eq. (2.2). We further expect that not all such regions A, B are "physically relevant" in the classical limit: some may correspond to e.g. scattered points or regions that do not contain the neighborhood of any point in the appropriate physical variables (such as the highlighted set $\{C_k\}_{k=0}^{14}$ in Fig. 2.8). The definition of such regions requires considerably higher resolution (of measure $\sim 1/d$) in the physical variables than classically accessible, and arguably become unphysical in the classical limit. Thus, if all *physical* regions $A, B, \ldots$ satisfy Eq. (2.1) we have the emergence of effectively classical ergodic behavior, and if they further satisfy Eq. (2.2) we have mixing behavior, in the fictitious phase space $\mathscr{P}[\mathscr{C}]$.

We note that such statements, about the physicality of fictitious phase space regions with respect

to observables, is again an observable-dependent eigenstate property and not determined by spectral rigidity. This suggests that the classical ergodic hierarchy is significantly encoded in the energy eigenstates (with spectral rigidity determining the time scale of applicability), while it is their discretized counterparts (cyclic ergodicity and aperiodicity) that are fully encoded in the energy eigenvalues. Such a conclusion is consistent with the mathematical observation [99] that classical cyclic permutation errors cannot fully identify classical mixing behaviors, which instead requires an "incongruity" of the C_k when expressed in terms of the physical variables in the $n \to \infty$ continuum limit (essentially, the C_k must be arranged erratically in phase space such that non-constant eigenfunctions $f_{m\neq 0}(x \in \mathscr{P})$ of the classical unitary $\mathscr{U}_{\mathscr{T}} \approx \mathscr{U}_{\mathscr{T}_{\mathscr{C}}}$, which satisfy $f_m(x \in C_k) \approx \exp(-2\pi i k m/n)$ for $0 < |m| \ll n$, do not exist as smooth, well-behaved functions of the phase space coordinates x in the $n \to \infty$ limit; see also Sec. 2.2.1). An identical conclusion is suggested by the numerical observation [79, 80] that similar spectral properties may be seen in mixing and merely ergodic systems, and one may require explicitly identified physical variables to distinguish them (see also a related discussion of physical bases and the role of eigenstates in Ref. [69]). Nevertheless, in Sec. 2.7.2, we argue that a connection between spectral rigidity and eigenstate properties can be made if we introduce an appropriate dynamical criterion to identify physical observables.

2.7.2 Poincaré recurrences and eigenstate thermalization

It is natural to ask if we can characterize a physical phase space $\mathscr{P}[\mathscr{C}]$ directly based on some physical property, without relying on observables such as $\mathscr{B}$ (which may only be possible to identify in quantum systems with an explicit phase space description [126], and even then, may be complicated by the non-classical behavior of e.g. Wigner quasiprobability distributions [127–129], especially for small $\mu(C_k)$). In this subsection, we consider a quantum analogue of Poincaré recurrences as a candidate for this role, noting that it is satisfied by the physical phase space in all classical systems — ergodic

and non-ergodic[8].

We introduce the following *ad hoc* definition based on the classical statement of the Poincaré recurrence theorem [10–12] (see also Sec. 2.2): any subspace $\mathscr{H}(A) \subseteq \Sigma_d$ with projector $\hat{\Pi}(A)$ (and density matrix $\hat{\rho}(A) = \hat{\Pi}(A)/\operatorname{Tr}[\hat{\Pi}(A)]$) is Poincaré recurrent if there exists an orthonormal basis $\mathscr{A}$ for $\mathscr{H}(A)$, such that any pure state $|\psi\rangle \in \mathscr{A}$ returns to have a larger-than-random overlap with the subspace at some time t, with $t_0 \ll |t| \lesssim O(t_0 d)$:

$$\operatorname{Tr}\left[\hat{\Pi}(A)\hat{U}_H(t)|\psi\rangle\langle\psi|\hat{U}_H^\dagger(t)\right] \gg O(d^{-1}). \tag{2.74}$$

We note the similarity of the restriction on the range of t to that in the definition of cyclic aperiodicity (Eq. (2.14)).

For simplicity, we assume that it is sufficient to consider DFT cyclic permutations as representing any (sub-)regions in phase space. We also assume that the persistence amplitude is the only greater-than-random component of any state with respect to a DFT basis of interest (in other words, none of the terms $\varepsilon_C^{1/2}(p, t_0)\nu_m(p)$ exceed $O(d^{-1/2})$ in magnitude), as is typically the case for e.g. Wigner-Dyson or Poisson statistics (partly due to Eq. (2.46)). Let $t_R(\mathscr{C})$ then represent the randomization time of a DFT cyclic permutation $\mathscr{C}$ — the smallest time for which we have $z(t_R/t_0, t_0) \in O(d^{-1/2})$, which generically decreases with increasing $\varepsilon_C(1, t_0)$.

The time t_R determines the minimum dimension of Poincaré recurrent subspaces $\mathscr{H}(A)$ that have the diagonal form in Eq. (2.71). Specifically, only subspaces with

$$\dim \mathscr{H}(A) = \mu(A)d \geq \frac{t_0 d}{t_R(\mathscr{C})}, \tag{2.75}$$

are guaranteed to be Poincaré recurrent. Non-aperiodic cyclic permutations have $t_R(\mathscr{C}) > t_0 d$ and

[8] We recall that quantum recurrences [49] of phases in the energy eigenbasis occur over times possibly exponentially large in d [108], and are not directly relevant at earlier times.

every subspace is recurrent; for ergodic ones, $t_R(\mathscr{C}) > t_0 d/2$, and any subspace with $\mathscr{H}(A) \geq 2$ (in other words, every subspace that is not a pure state) is recurrent.

If we impose Poincaré recurrence for all regions of the form of Eq. (2.71) containing more than one pure state ($\mu(A) > 1/d$)[9], as a criterion to identify a physical phase space $\mathscr{P}[\mathscr{C}] = \mathscr{P}$, it follows that the only DFT cyclic permutations that can satisfy this requirement are ergodic ones. In other DFT bases, Poincaré recurrence fails for certain regions A with $2/d \leq \mu(A) < t_0/t_R(\mathscr{C})$. Thus, if we want to construct a fictitious phase space for some energy subspace Σ_d that satisfies Poincaré recurrence for arbitrarily small regions (of greater than the smallest volume $1/d$), there are two possibilities given our assumptions:

1. Σ_d is itself ergodic, and a corresponding ergodic (DFT) cyclic permutation $\mathscr{C}$ allows the definition of mixed states corresponding to Poincaré recurrent phase space regions according to Eq. (2.71).

2. Σ_d must be decomposed into M ergodic subspaces $\Sigma_{d_1}(1), \ldots, \Sigma_{d_M}(M)$, each spanned by d_k energy levels (respectively) that add up to d. Poincaré recurrent phase space regions can then be defined according to Eq. (2.71) on the combination of their respective ergodic (DFT) cyclic permutations $\mathscr{C}(1), \ldots, \mathscr{C}(M)$. The previous case corresponds to $M = 1$.

This is the key component of our argument: imposing an *ad hoc* Poincaré recurrence requirement to identify a physical phase space basis $\mathscr{P}[\mathscr{C}] = \mathscr{P}$ forces this basis to be a collection of ergodic cyclic permutations. Thus, we obtain a quantum analogue of the classically more trivial statement that the phase space $\mathscr{P}$ is either itself ergodic or decomposes into ergodic subsets [10–12].

[9] If we imagine pure states as corresponding to regions with negligible ($= 1/d$) phase space measure, they are effectively measure zero sets which are not required to satisfy Poincaré recurrence even classically, unless they are on e.g. periodic orbits. More realistically, we should expect recurrence times to increase with decreasing measure.

In either case, the projectors $\hat{\Pi}(A)$ can be written as

$$\hat{\Pi}(A) = \sum_{m=1}^{M} \sum_{k \in C_m(A)} |C_k(m)\rangle\langle C_k(m)| \tag{2.76}$$

where each $C_m(A)$ is a set of $n_m(A) = |C_m(A)|$ indices of elements of $\mathscr{C}(m)$, and $\sum_m n_m(A) = n(A) \approx \mu(A)d$.

The connection to ETH emerges if one asks for the matrix elements of these projectors in the energy eigenbasis. Let $\{|E_k(m)\rangle\}_{k=0}^{d_m-1}$ be the energy eigenstates contained in $\Sigma_{d_m}(m)$, whose form is explicitly known as DFTs of the $|C_k(m)\rangle$. Assuming that (in the generic case) each $C_m(A)$ is randomly distributed on $\mathscr{C}(m)$ (so the phases of the DFT can be taken to be random for $1 \ll n_m \ll d_m$), we have

$$\langle E_k(m)|\hat{\Pi}(A)|E_j(m)\rangle = \frac{n_m(A)}{d_m}\delta_{kj} + O\left(\frac{\sqrt{n_m(A)}}{d_m}\right)R_{kj}(m), \tag{2.77}$$

for some random $d_m \times d_m$ Hermitian matrix $R_{kj}(m)$ with $O(1)$ matrix elements (with weak correlations ensuring $\hat{\Pi}^2 = \hat{\Pi}$). On the other hand, the statement of ETH for an n-dimensional projector $\hat{\Pi}$ may be motivated by random matrix arguments (e.g. similar to Refs. [6, 21, 52]; see also Ref. [130] for a discussion of the role of highly degenerate projectors), giving

$$\langle E_k|\hat{\Pi}|E_j\rangle = \frac{n}{d}\delta_{kj} + O\left(\frac{\sqrt{n}}{d}\right)R_{kj}, \tag{2.78}$$

in the full energy subspace Σ_d. We see that Eq. (2.77) and Eq. (2.78) are guaranteed to agree for $M = 1$. For $M > 1$, the first (diagonal) terms of Eq. (2.77) and (2.78) can only agree through a statistically unlikely coincidence $n_m/d_m = n/d$; even in the rare instance that this holds, the second (fluctuation) term of the former typically has matrix elements of size $O[(\sqrt{Mn})/d]$ (or zero when k, j correspond to different $\Sigma_{d_m}(m)$), and ETH can be satisfied only for $M \in O(1)$.

We have essentially argued (under some simplifications, e.g. focusing on DFT bases with typical random parts), that a quantum correspondence between dynamical ergodicity (equivalently, spectral rigidity) and thermal ergodicity (essentially, observable-dependent ETH or eigenstate properties) can be established within the framework of cyclic permutations, given a single additional principle to identify physical observables. In particular, given Poincaré recurrence, ETH is satisfied by generic projectors onto "physical" phase space regions (which then show quantum analogues of thermal ergodicity and mixing, albeit over long times $t \gg t_{\text{ramp}}$) only if the system is cyclic ergodic in Σ_d. It would be interesting to see if such an argument can be made more rigorous, and suitably generalized to systems without a classical limit where a similar correspondence between eigenvalue and eigenstate statistics is observed [21] but not yet understood. To do so, it would be necessary to determine if there's any link between a Poincaré recurrence requirement for the fictitious phase space observables considered here (relevant for $t \gtrsim t_{\text{ramp}}$ dynamics), and the *local* or few-body observables (relevant for $t \lesssim t_{\text{ramp}}$) that are the subject of conventional ETH [18–22, 51, 52].

2.8 Conclusions

We have identified fully quantum dynamical properties in the Hilbert space that resemble discretized classical ergodic properties, namely cyclic ergodicity and aperiodicity, and shown how these are determined by precise measures of energy level statistics, thereby introducing genuine dynamical elements into the study of quantum ergodicity and "chaos". These properties were illustrated with four physically relevant types of level statistics: Wigner-Dyson, Poisson, and that of 2D KAM and rational tori. A key takeaway is that we can isolate the precise aspects of random matrix behavior — an exact Gaussian distribution of mode fluctuations with variance σ_Δ^2 in the range of Wigner-Dyson spectral rigidity — that are sufficient conditions for cyclic ergodicity, identifying which of the numerous measures of level statistics [36] are physically important (from the present viewpoint). However,

random matrix behavior is not necessary for cyclic ergodicity, as typified by 2D KAM tori. Our study of the latter included an analytical proof of spectral rigidity in an infinite sequence of subspaces, demonstrating that cyclic permutations may offer an intuitive path towards the open problem of rigorously connecting spectral and ergodic properties in individual systems [71, 72].

Our dynamical construction also clarifies that spectral rigidity should be most directly associated with (quantum) cyclic ergodicity and aperiodicity, rather than stronger notions of chaos as is the norm in the literature [28]. This provides a general theoretical framework for justifying numerical observations [78–80] in systems with a non-chaotic, ergodic classical limit. At the same time, the precise relationship between these eigenvalue-based "cyclic" dynamical properties and more familiar (semi-)classical properties (generally involving both eigenvalues and eigenstates in the classical limit) merits further investigation, perhaps based on a classical limit of Eqs. (2.46) and (2.52) that relate quantum cyclic ergodicity to mean recurrence properties (such as, possibly, the semiclassical effect of periodic orbits [28, 39, 40]) in a DFT cyclic permutation. As argued in Sec. 2.7, the classical limit may also provide inspiration for identifying fully quantum connections between dynamical (or eigenvalue) and thermal (or eigenstate) properties in arbitrary quantum systems.

The approach developed here may generalize to "quantizing" other dynamical properties of physical interest. Connections between quantum dynamical properties and a suitably defined Kolmogorov-Sinai (KS) entropy (a measure of chaos closely related to Lyapunov exponents [48], see e.g. Ref. [131] for candidate definitions) may be accessible through non-cyclic permutations, which are related to the classical KS entropy in Refs. [12, 98]. Generalizing further to non-unitary dynamical structures could help associate precise dynamical properties with the recently observed analogues of spectral rigidity in autonomous dissipative quantum systems, whose evolution is generated by non-unitary Liouvillian superoperators [132–134]. Finally, the precise understanding of the spectral rigidity of 2D KAM tori in terms of their dynamical properties obtained in Sec. 2.6 may serve as a starting point to adapt elements

of KAM theory [48, 118] — the classical study of the development of ergodicity under perturbations to integrable systems — to fully quantum systems. The rapid development of ergodicity in many-particle systems even under small perturbations, while not well understood in quantum mechanics, is believed to be essential for the applicability of statistical mechanics [22, 135].

Chapter 3: Exact universal bounds on quantum dynamics and fast scrambling

3.1 Synopsis: An energy-time uncertainty principle for many-body systems

Introduction— The energy-time uncertainty principle sets fundamental limits on the speed of quantum dynamical processes. Specific formulations of this principle [55–60,136], such as the Mandelstam-Tamm (MT) [55,56] and Margolus-Levitin (ML) [57,136] bounds (see also Refs. [59,60] for reviews and extensions), are expressed in terms of a *single* parameter Δ_E characterizing e.g. the standard deviation of energy (MT) or the mean energy relative to the ground state (ML), in a given initial state. In general, these allow a given decay in the return probability of the state only after a sharp time (proportional to Δ_E^{-1}). However, such sharp bounds do not provide useful constraints on many-body processes that typically occur over timescales much larger than Δ_E^{-1}, such as the thermalization of interacting many-body systems.

Thermalization has been at the focus of several developments in nonequilibrium statistical mechanics [60, 114, 137, 138], many-body quantum chaos [18, 19, 21, 22, 41, 66, 69, 92–94, 139] and quantum information [41, 92–94, 125, 139]. In our current understanding, many-body thermalization is driven by the generation of a high degree of entanglement between the interacting particles [17,23,24,52,140]. The question of how fast this entanglement can be generated, irrespective of any restrictions on the nature of interactions, has come to be of fundamental interest, partly motivated by a conjectured correspondence between the black hole information problem and a form of thermalization called the *scrambling* of information in Hamiltonian many-body systems [53, 54, 141–143].

On the other hand, useful many-body speed limits known so far require highly restrictive assumptions such as spatially local interactions [60, 144–146] or limited external control parameters [61], preventing an exact solution of this problem in a general setting.

In this work, we derive a universal bound on quantum dynamics by considering the evolution of a complete set of coarse-grained initial states, e.g., a complete set of states for a subsystem of particles. This bound is directly given in terms of the spectral form factor [28] (SFF; see Eq. (3.3)) that characterizes the full profile of the energy spectrum in Hamiltonian systems, and also generalizes to non-Hamiltonian systems. We use this bound to constrain the speed of scrambling of information within subsystems of a many-body system. For Hamiltonian systems, we argue that any subsystem can typically *remain* scrambled for a sustained length of time only after an asymptotically long scrambling time in the subsystem size. Our bound constrains the "sustained" scrambling time of a particular Hamiltonian system in terms of its density of states. Finally, we map the problem of bounding the fastest scrambling time to a purely mathematical statement, which in turn is related to the as yet unresolved mathematical problem of the necessary asymptotic conditions for a sufficiently well-behaved function (related to the SFF) to have a non-negative Fourier transform (related to a physical constraint on the density of states to be non-negative).

Correspondingly, our main results are: 1. Eq. (3.5), providing a universal bound on quantum dynamics with relevance to long times, 2. Eq. (3.9), which supplies a rigorous necessary condition on the SFF to allow scrambling at a given time, and 3. Eq. (3.12) that relates the fastest allowed sustained scrambling time of Hamiltonian many-body systems to the mathematical properties of its density of states, leading to Statements 3.1 and 3.2 describing universal constraints on fast scrambling. These results are illustrated analytically (Eq. (3.14)) and numerically (Fig. 3.1) for the Sachdev-Ye-Kitaev (SYK) model. We use the formal asymptotic notation [50, 147] in this Chapter, with the symbols ω, Ω, Θ, O, o respectively representing the order-of-magnitude versions of $>, \geq, =, \leq, <$.

Setup, and quantities of interest— We consider a general quantum mechanical system with a Hilbert space $\mathscr{H}$ of dimension D (e.g, $D = 2^N$ for N qubits), whose state at time t is specified by a density operator $\hat{\rho}(t)$. Given an initial state $\hat{\rho}_0$, its time evolution may be generated by a Hamiltonian $\hat{H}$,

$$\hat{\rho}(t) = e^{-i\hat{H}t}\hat{\rho}_0 e^{i\hat{H}t}, \tag{3.1}$$

or more generally, any linear dynamics of the form:

$$\hat{\rho}(t) = \hat{\mathscr{T}}^t[\hat{\rho}_0] \equiv \sum_{r=1}^{M} \hat{\mathscr{K}}_r(t)\hat{\rho}_0\hat{\mathscr{K}}_r^\dagger(t). \tag{3.2}$$

Eq. (3.2) represents the most general completely positive linear quantum operation [50] with time-dependent Kraus operators $\hat{\mathscr{K}}_r(t)$, and accounts for both unitary and nonunitary e.g. dissipative dynamics. A restriction to $M = 1$ and $\hat{\mathscr{K}}_1(t) = \hat{U}(t)$ corresponds to unitary evolution generated by $\hat{U}(t)$,

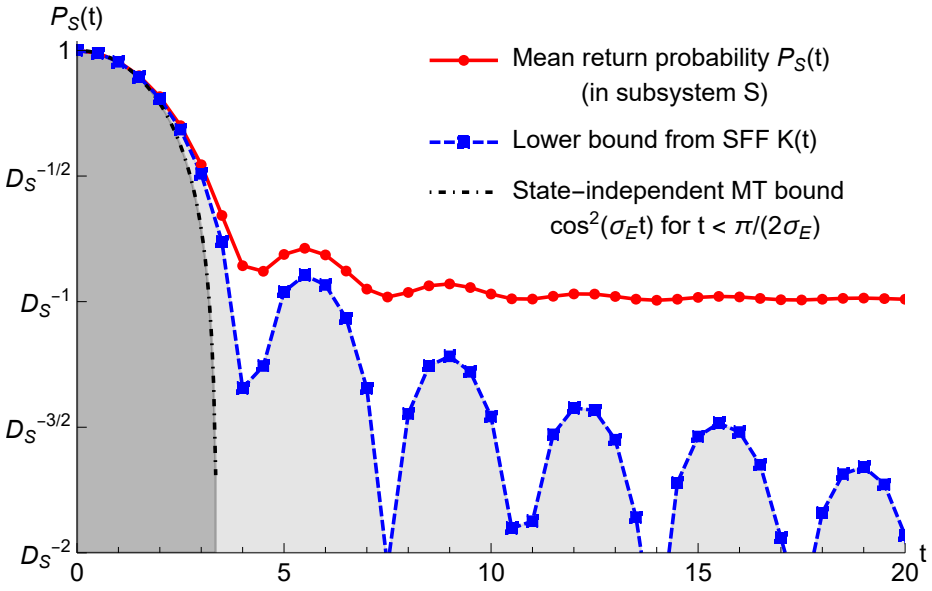


Figure 3.1: Illustration (log-linear) of scrambling speed limits via Eqs. (3.5), (3.6), (3.8) and (3.9) in a realization of the SYK-4 model of randomly interacting fermions, for initial fermion Fock states given by Eq. (3.13) in a subsystem of $N_S = 7$ sites ($D_S = 128$) among $N = 10$ total sites ($D = 1024$). The power-law decay of $K(t)$ (whose nontrivial effect is evident here up to the first "peak" at $t \approx 6$) leads to exponentially slow sustained scrambling via Eq. (3.14) as $N, N_S \to \infty$.

reducing further to Hamiltonian dynamics when $\hat{U}(t) = e^{-i\hat{H}t}$. We note that the $M = 1$ case also allows for nonunitary "regularized" or "filtered" Hamiltonian evolution of the form $\hat{\mathscr{K}}_1(t) = g(\hat{H})e^{-i\hat{H}t}$, which has been of interest in some applications [38, 114].

A state- and observable-independent characteristic of such a generic quantum dynamical system is given by the generalized SFF:

$$K(t) \equiv \frac{1}{D^2}\sum_{r=1}^{M}\left|\mathrm{Tr}\left[\hat{\mathscr{K}}_r(t)\right]\right|^2. \tag{3.3}$$

For Hamiltonian dynamics, $K(t)$ is the Fourier transform of the 2-point energy level correlation function satisfying $0 \leq K(t) \leq K(0) = 1$, and was introduced in the study of quantum chaos [28]. Its late-time quantum fluctuations (usually of $O(D^{-1})$ magnitude) play a central role in characterizing energy level statistics [4, 35, 38, 79, 82–84, 86–89, 148] and quantum dynamical ergodicity [1]. Time-dependent and non-unitary analogues have additionally been considered in various contexts [149–153], and Eq. (3.3) also accounts for these cases. A key feature of the SFF is that in spite of its state- and observable-independence, it lends itself to direct experimental measurement in many-body systems through recently developed measurement protocols [6, 96]; see Chaps. 8 and 9 for a description and experimental implementation of one of these protocols. Our primary interest in this Chapter is to obtain universal constraints on the dynamics of quantum thermalization in terms of this quantity.

To study thermalization, we should consider states and physical observables. In our initial abstract setting, this role will be fulfilled by a complete set $S = \{\hat{\Pi}_k\}_{k=1}^{D_S}$ of D_S orthonormal projection operators $\hat{\Pi}_k$, satisfying Hermiticity $\hat{\Pi}_k^\dagger = \hat{\Pi}_k$, orthonormality $\hat{\Pi}_k\hat{\Pi}_\ell = \delta_{k\ell}\hat{\Pi}_k$ and completeness $\sum_{k=1}^{D_S}\hat{\Pi}_k = \hat{\mathbb{1}}$. These may be interpreted as D_S different coarse-grained initial states $\hat{\rho}_{0,k} = \hat{\Pi}_k/\mathrm{Tr}[\hat{\Pi}_k]$ of the system within $\mathscr{H}$, as well as observables $\hat{\Pi}_k$ measuring the probability of overlap of a given state with these coarse-grained states (called a projection valued measure [50]). Such projectors have a no-

table merit for our present purposes: the eigenvalues of $\hat{\Pi}_k$ are uniquely fixed to be 1 (with $\mathrm{Tr}\,\hat{\Pi}_k$-fold degeneracy) and 0 (with $(D - \mathrm{Tr}\,\hat{\Pi}_k)$-fold degeneracy). This avoids complications arising from a non-universal set of eigenvalues of more conventional observables such as particle positions or spin/qubit states – which serve to *label* specific measurement outcomes if $\hat{\mathscr{T}}^t$ is already specified – and allows one to focus on the intrinsic aspects of quantum dynamics $\hat{\mathscr{T}}^t$ itself.

We focus on the *mean return probability* $P_S(t)$ of the projectors at time t, noting that return probabilities are the main quantities of interest in the MT and ML bounds [55–59]. This is given by:

$$P_S(t) \equiv \frac{1}{D}\sum_{k=1}^{D_S} \mathrm{Tr}\left[\hat{\Pi}_k(t)\hat{\Pi}_k\right], \tag{3.4}$$

where $\hat{\Pi}_k(t) = \mathrm{Tr}[\hat{\Pi}_k]\hat{\rho}_{0,k}(t)$ is the time-evolving initial state weighted by its degeneracy. As shown later, $P_S(t)$ partly measures the failure of an initial state to thermalize (see also Eq. (3.8)), but the above is sufficient setup to state our first key result.

A universal dynamical inequality— In this general setting of a quantum system whose dynamics is given by Eq. (3.2) with the SFF defined by Eq. (3.3), our primary result is the following inequality on the dynamics of the mean return probability defined in Eq. (3.4) for any set of projectors S (see Fig. 3.1):

$$P_S(t) \geq K(t). \tag{3.5}$$

This is derived in Sec. 3.2 as a consequence of two simple inequalities: the contribution to $P_S(t)$ from each $\hat{\mathscr{K}}_r(t)$ cannot be less than the *squared* mean return amplitude of any set of orthonormal pure states constituting the projectors under the action of $\hat{\mathscr{K}}_r(t)$, while this mean amplitude is further bounded from below by the magnitude of $\mathrm{Tr}[\hat{\mathscr{K}}_r(t)]$ due to the triangle inequality. The latter bound has been previously used in Ref. [1], in the restricted context of Hamiltonian (and Floquet) unitary dynamics, to characterize the "aperiodicity" of quantum dynamical systems in terms of their SFF (see

Chap. 2); in that case, the bound is saturated in any basis of discrete Fourier transformed [1] (more generally, complex Hadamard transformed, see Sec. 3.2) energy eigenstates. While other relations between suitable return probabilities and $K(t)$ can be derived by averaging over all operators in the entire Hilbert space [86,114,154,155] and losing basis-specific information (see Eq. (3.10)), or in the late-time regime of quantum fluctuations by typicality arguments or an ensemble average over $\hat{\mathscr{T}}^t$ for certain "physical" operators [73,111,112], Eq. (3.5) is an *exact* relation that holds for all time in *any* given basis.

Eq. (3.5) provides an alternative formulation of the energy-time uncertainty principle for the dynamics of the most general completely positive quantum operation, introducing a nontrivial sensitivity — even in the Hamiltonian case — to long times and microscopic values of $P_S(t)$ through the asymptotics of $K(t)$. Its relation to the more familiar MT and ML bounds [55–59,136] in a Hamiltonian system is obtained through corresponding bounds on the SFF [156]. For instance, the MT bound states that the return probability of a single initial state with energy variance Δ_E^2 cannot decay faster than $\cos^2(\Delta_E t)$, which is related to Eq. (3.5) via the following inequality [156]:

$$K(t) \geq \cos^2(\sigma_E t), \text{ for } |t| < \frac{\pi}{2\sigma_E}, \tag{3.6}$$

where $\sigma_E^2 = -K''(0)/2$ is the variance of the energy spectrum itself, rather than any particular state (see Fig. 3.1). This is generalized in Sec. 3.2, from which it follows that the bound of Eq. (3.5) is tighter than *state-independent* bounds on $P_S(t)$ stemming from combining known speed limits with the properties of the energy spectrum itself (or Kraus operators), such as σ_E for a state-independent MT bound $P_S[|t| < \pi/(2\sigma_E)] \geq \cos^2(\sigma_E t)$. These form a class of bounds applying universally to *any* mean return probability $P_S(t)$ in the system, among which Eq. (3.5) is consequently the tightest. It is worth emphasizing a trade-off: while the MT and ML bounds apply to the return probability of

individual states instead of an average as in $P_S(t)$, working with $P_S(t)$ in Eq. (3.5) has the advantage of capturing the full spectral and dynamical information encoded in $K(t)$ over all timescales. This advantage is crucial for the application we consider next.

Application to scrambling — Now, we specialize to many-body systems, with an implicit $D \to \infty$ thermodynamic limit. Here, we consider subsystems consisting of a subset of the degrees of freedom e.g. spins/qubits, which corresponds to a factorization of $\mathscr{H} = \mathscr{H}_S \otimes \mathscr{H}_E$ into the D_S-dimensional subsystem Hilbert space $\mathscr{H}_S$ (e.g. $D_S = 2^{N_S}$ for a subsystem of N_S qubits), and the effective $D_E = D/D_S$ dimensional component $\mathscr{H}_E$ of the system "external" to $\mathscr{H}_S$ formed by the remaining degrees of freedom, which can act as a thermalizing bath [17,23,24,52]. In this context, each projector in S can be chosen to reduce to a given pure state in an orthonormal basis $\mathscr{B}_S = \{|k\rangle_S\}_{k=0}^{D_S-1}$ for the subsystem, with $D_E^{-1}\,\mathrm{Tr}_E[\hat{\Pi}_k] = |k\rangle_S\langle k|$ (e.g., computational basis states of qubits within $\mathscr{H}_S$).

Following Ref. [141], our criterion for the subsystem $\mathscr{H}_S$ to be scrambled at time t, in a given basis $\mathscr{B}_S$, requires *all* basis states to evolve to have indistinguishable overlap with the original basis states to leading order within $\mathscr{H}_S$:

$$\frac{1}{D_E}\,\mathrm{Tr}[\hat{\Pi}_k(t)\hat{\Pi}_\ell] = \frac{1}{D_S} + o(D_S^{-1}). \tag{3.7}$$

This is necessary (but not sufficient) for "thermalization to infinite temperature" [51, 141] and, typically, maximal entanglement [23, 24] at time t, i.e., the overlap of the time-evolved states $\hat{\rho}(t) = \hat{\Pi}_k(t)/D_E$ with the observables $\hat{\Pi}_\ell$ look like those of the "infinite temperature" maximally mixed state $\hat{\rho}_{(\infty)} \equiv \hat{\mathbb{1}}/D_S$ in $\mathscr{H}_S$, to leading order. A necessary condition for scrambling in the sense of Eq. (3.7) is:

$$P_S(t) = \frac{1}{D_S} + o(D_S^{-1}). \tag{3.8}$$

This is further readily shown to be a necessary condition for scrambling within the subsystem $\mathscr{H}_S$, of

a basis of *pure product states* for $\mathscr{H}$ (see also Sec. 3.3 for a brief discussion of this [141] and related notions of scrambling). From Eqs. (3.5) and (3.8), we obtain the following *necessary* condition for scrambling at the time t (usually satisfied within the early time decay of the SFF):

$$K(t) \leq \frac{1}{D_S} + o(D_S^{-1}). \tag{3.9}$$

Eq. (3.9) is, however, not a *sufficient* condition for scrambling in most bases. For instance, in a basis where $\hat{\mathscr{T}}^t$ has a simple structure e.g. locality, $K(t)$ can even decay to nearly $O(D^{-2}) \ll D_S^{-1}$ by the scrambling timescale [38, 41, 139]. However, a typicality result in random matrix theory [86, 114, 154, 155] states that in *almost all* (with respect to the Haar measure [24, 36]) subsystems $\mathscr{H}_S$ and bases $\mathscr{B}_S$ in the Hilbert space of a given system,

$$\frac{1}{D_E}\mathrm{Tr}[\hat{\Pi}_k(t)\hat{\Pi}_j] = \frac{1+O(D_E^{-1/2})}{D_S} + \left(\delta_{kj} - \frac{1}{D_S}\right)K(t). \tag{3.10}$$

This means that $K(t) = o(D_S^{-1})$ is a sufficient condition for almost all subsystems of dimension $D_S = o(D)$ in the Hilbert space (i.e., $N_S/N < 1-\epsilon$) to scramble at time t, which may or may not include a given subsystem of interest. Further, such random bases "almost" saturate Eq. (3.5) up to $O(D_S^{-1})$ terms.

Fast scrambling in Hamiltonian systems—In Hamiltonian many-body systems, Eq. (3.9) translates to a constraint on the energy spectrum as follows: if $\mathscr{N}_D(E) \geq 0$ is the density of states *per level* of the D energy levels of the system (given, at this stage, by a set of delta function spikes that integrates to 1), then $K(t) = |\widetilde{\mathscr{N}_D}(t)|^2$, where $\widetilde{\mathscr{N}_D}(t) = \mathrm{Tr}[\hat{U}(t)]$ is the Fourier transform of $\mathscr{N}_D(E)$:

$$\widetilde{\mathscr{N}_D}(t) \equiv \int \mathrm{d}E\ \mathscr{N}_D(E)e^{-iEt}. \tag{3.11}$$

This relation yields nontrivial constraints on the fastest allowed scrambling time in Hamiltonian systems based entirely on the density of states, where we take $t = \Theta(1)$ to be the timescale of a $\Theta(1)$ decrease in $K(t)$ from $K(0) = 1$. An important consideration is suggested by a global Haar random (say, CUE [28, 36]) Floquet unitary $\hat{U}_{\text{CUE}}$, which *instantly* scrambles [154] an initial state if acting on the system in discrete time steps. However, any *continuous-time* Hamiltonian dynamics that reproduces the action of $\hat{U}_{\text{CUE}}$ at a time $\tau = \Theta(1)$ [with $e^{-i\hat{H}\tau} = \hat{U}_{\text{CUE}}$] typically has long-time transients in the SFF, e.g. $K(t) = \text{sinc}^2(\pi t/\tau) \sim t^{-2}$ in the simplest case [1, 114] (see Eq. (2.41)), preventing *sustained* scrambling for a *continuous* length of time until $t_{s,\text{CUE}} = \Omega(\sqrt{D_S})$, even with global interactions.

We therefore seek to constrain the time t_s, after which a given system can *remain* scrambled up to a long time $T \gg t_s$ beyond all timescales of physical interest (potentially the long $T \sim \exp[\Theta(D)]$ quantum recurrence time [49, 108]). From Eq. (3.9) and $K(t) = |\widetilde{\mathcal{N}}(t)|^2$ (where we drop the D dependence for convenience) we see that this time is determined by the asymptotic requirement on the density of states:

$$\left|\widetilde{\mathcal{N}}(t > t_s)\right|^2 \leq \frac{1}{D_S} + o(D_S^{-1}), \tag{3.12}$$

with $t < T$ implicit. From Eq. (3.12), it is clear that the fastest t_s is constrained by what kind of asymptotic behaviors $\widetilde{\mathcal{N}}(t \to \infty)$ may exist in a "physical" system. For instance, power law asymptotics $|\widetilde{\mathcal{N}}(t \to \infty)|^2 \sim t^{-n}$ implies $t_s \gtrsim D_S^{1/n} = 2^{N_S/n}$, giving a scrambling time that is at least exponentially long in the subsystem size N_S. This asymptotic behavior is generally restricted by the non-negativity of a physical density of states $\mathcal{N}(E) \geq 0$ (as discussed in Sec. 3.4), but the quantitative effect of this restriction on the Fourier transform $\widetilde{\mathcal{N}}(t)$ remains an open mathematical problem.

This is because it is not fully understood mathematically [157, 158] how the non-negativity of $\mathcal{N}(E)$ is reflected in the local or asymptotic behavior of $\widetilde{\mathcal{N}}(t)$. What is known are some necessary [158] and sufficient conditions [157] on $\widetilde{\mathcal{N}}(t)$ that do not directly address asymptotics, see Sec. 3.4. It is also customary to consider many-body systems with additional "physicality conditions" like analytic

dynamics (in the thermodynamic limit) in many cases [159, 160], which enforces $t_s = \omega(1) \gg \Theta(1)$ due to the smooth asymptotic decay of $\widetilde{\mathcal{N}}(t)$. Given the subjectivity of imposing "physicality" conditions on the density of states, we instead formulate two exact system-independent statements on fast scrambling for specific cases of wide relevance:

Statement 3.1. *The Gaussian function $\widetilde{\mathcal{N}}(t) = \exp(-at^2/2)$ is analytic with a positive Fourier transform (also a Gaussian), and consequently quantum systems with fully analytic dynamics that scramble subsystems by any $t_s = \omega(\sqrt{\log D_S})$ do exist, by Eq.* (3.10).

Statement 3.2. *Due to a Theorem [156, 161, 162] stating that the Fourier transform of a function $\mathcal{N}(E)$ with one-sided bounded support decays at a slower-than-exponential rate $|\widetilde{\mathcal{N}}(t \to \infty)| \neq O(e^{-\Theta(t)})$, every quantum system with a finite spectral edge [156, 161] as $D \to \infty$ [e.g. $\mathcal{N}(E < E_0) \to 0$, where $E_0 = O(1)$ over the scale of $\Theta(1)$ variations in $\mathcal{N}(E)$], required for a nontrivial (state-independent) ML bound [57, 136], can only scramble subsystems to infinite temperature after a time $t_s = \omega(\log D_S)$, by Eq.* (3.9).

Illustration: Slow scrambling in the SYK model— Consider any typical individual realization of the Majorana SYK-q model [163, 164] of fermions on N sites ($2N$ Majoranas, $D = 2^N$) subject to random ($q \geq 4$)-body interactions without fermion number conservation (see Chap. 6 for details), which is regarded as "maximally chaotic" and a "fast" scrambler due to saturating a finite temperature bound on a local-operator measure of "chaos" [159]. However, complete scrambling via *entanglement generation* in large subsystems of N_S sites ($D_S = 2^{N_S}$), prepared in Fock states with a maximally mixed (completely uncertain) E-component (where $n_{j(k)} \in \{0, 1\}$ is the fermion occupancy number in the j-th site):

$$\hat{\Pi}_k = |n_{1(k)}, \ldots, n_{N_S(k)}\rangle_S \langle n_{1(k)}, \ldots, n_{N_S(k)}| \otimes \hat{\mathbb{1}}_E, \tag{3.13}$$

can proceed "exponentially" slowly on account of Eq. (3.5).

This is because $\mathcal{N}(E)$ has sharp edges for SYK (similar to Statement 3.2, albeit at asymptotically large energies) leading [38,82] to $K(t \to \infty) \sim \Theta(D^{-\alpha})t^{-3} + O(D^{-1})$, where $\alpha < 1$ ($\alpha \approx 0.66$ for $q = 4$, $\alpha \propto q^{-2}$ for large q) is determined [38] by the zero-temperature entropy [163,165] $\mathcal{S}_0 = N(1-\frac{\alpha}{2})\ln 2$. This effect of the edges cannot be "filtered" [38] out of the dynamics due to the completeness of the $\hat{\Pi}_k$ (see Chap. 6). Thus, from the discussion following Eq. (3.12), a subsystem in SYK can remain fully scrambled to any $P_S(t) = O(D_S^{-1})$ (including D_S^{-1}) only after an exponentially long time in N_S:

$$t_{s,\mathrm{SYK}} = \Omega\left[2^{(N_S-\alpha N)/3}\right], \text{ nontrivial for } N_S > \alpha N. \tag{3.14}$$

This $t_{s,\mathrm{SYK}} \gtrsim e^{\Theta(N_S)}$ behavior should be contrasted with, say, systems with a fully Gaussian density of states without sharp edges (e.g., [38]), in which t_s approaches $N_S^{1/2}$ even for finite N_S/N by Statement 3.1, provided that Eq. (3.10) holds as is typical for systems with sufficiently random energy eigenstates [114].

Conclusions— Eq. (3.5) is a general speed limit on the quantum dynamics of a complete set of states, which remains nontrivial for longer times than basis-independent versions of the MT and ML bounds and typically even for asymptotically long times. We showed that it can be used to constrain [Eqs. (3.8) and (3.9)] characteristically slow many-body processes, such as the generation of entanglement associated with scrambling or thermalization to infinite temperature. In particular, it enables the problem of the fastest allowed scrambling timescale of a Hamiltonian many-body system to be mapped to a mathematical property of the density of states, irrespective of any interaction structure in the Hamiltonian. As an illustration, we showed that the SYK model, though "maximally chaotic" by certain measures, cannot allow the sustained scrambling (via entanglement generation) of large fermion subsystems up to *exponentially long times in the subsystem size*, due to the sharp edges in its density of states. Another application of Eq. (3.5) is to constrain entanglement generation even in

finite size systems (e.g., Fig 3.1) in terms of the decay of $K(t)$.

Outline of this Chapter— In the rest of this Chapter, we discuss the derivation of the speed limit in Eq. (3.5) as well as its saturation and relation to other speed limits (Sec. 3.2), some of the notions of scrambling and their relation to Eq. (3.7) (Sec. 3.3), some formal aspects of relating scrambling and the density of states (Sec. 3.4), and the details of the SYK model as well as how our speed limit leads to a slow sustained scrambling time in this model, in the sense of entanglement generation in large subsystems (Sec. 6.2).

3.2 Basis-independent quantum speed limit

3.2.1 Derivation of the inequality $P_S(t) \geq K(t)$

We have a complete set of orthonormal projectors $\{\hat{\Pi}_k\}$, with

$$P_S(t) \equiv \frac{1}{D}\sum_{k=1}^{D_S} \mathrm{Tr}\left[\hat{\Pi}_k(t)\hat{\Pi}_k\right]. \tag{3.15}$$

where $\hat{\Pi}_k(t) = \sum_{r=1}^{M} \hat{\mathscr{K}}_r(t)\hat{\Pi}_k\hat{\mathscr{K}}_r^\dagger(t)$ Let $D_k = \mathrm{Tr}[\hat{\Pi}_k]$ represent the dimensionality of each projector. As $\hat{\Pi}_k$ has D_k eigenvalues equal to 1, and $(D-D_k)$ eigenvalues equal to 0, there is an orthonormal set of vectors $\{|k;\ell\rangle\}_{\ell=1}^{D_k}$ for each $\hat{\Pi}_k$ such that:

$$\hat{\Pi}_k = \sum_{\ell=1}^{D_k} |k;\ell\rangle \otimes \langle k;\ell|. \tag{3.16}$$

We note that the full set $\bigcup_{k=1}^{D_S}\{|k;\ell\rangle\}_{\ell=1}^{D_k}$ containing all these vectors forms an orthonormal basis for the Hilbert space $\mathscr{H}$. Substituting this expression in Eq. (3.15) gives

$$\begin{aligned} P_S(t) &= \frac{1}{D}\sum_{r=1}^{M}\sum_{k=1}^{D_S}\sum_{\ell,\ell'=1}^{D_k}\left|\langle k;\ell'|\hat{\mathscr{K}}_r(t)|k;\ell\rangle\right|^2 \\ &\geq \frac{1}{D}\sum_{r=1}^{M}\sum_{k=1}^{D_S}\sum_{\ell=1}^{D_k}\left|\langle k;\ell|\hat{\mathscr{K}}_r(t)|k;\ell\rangle\right|^2. \end{aligned} \tag{3.17}$$

In the second line, we have dropped the $\ell \neq \ell'$ terms and retained only the diagonal $\ell = \ell'$ terms (incidentally, such a simple step is also a key element in the proof of the Shnirelman wavefunction ergodicity theorem [71]). Now, we consider the contribution to Eq. (3.18) from each $\hat{\mathscr{K}}_r(t)$, for which we get

$$\frac{1}{D}\sum_{k=1}^{D_S}\sum_{\ell=1}^{D_k}\left|\langle k;\ell|\hat{\mathscr{K}}_r(t)|k;\ell\rangle\right|^2 \geq \left[\frac{1}{D}\sum_{k=1}^{D_S}\sum_{\ell=1}^{D_k}\left|\langle k;\ell|\hat{\mathscr{K}}_r(t)|k;\ell\rangle\right|\right]^2. \tag{3.18}$$

This is just the inequality $(1/n)\sum_{j=1}^{n} x_j^2 \geq [(1/n)\sum_{j=1}^{n} x_j]^2$ with $x_{k;\ell} = \left|\langle k;\ell|\hat{\mathscr{K}}_r(t)|k;\ell\rangle\right|$ (essentially $\langle x^2\rangle \geq \langle x\rangle^2$, familiar from statistics). The sum on the right hand side can be further constrained (as in the context of "cyclic aperiodicity" in Ref. [1], see Chap. 2) using the triangle inequality $\sum_j |y_j| \geq |\sum_j y_j|$ with complex $y_{k;\ell} = \langle k;\ell|\hat{\mathscr{K}}_r(t)|k;\ell\rangle \in \mathbb{C}$, giving

$$\frac{1}{D}\sum_{k=1}^{D_S}\sum_{\ell=1}^{D_k}\left|\langle k;\ell|\hat{\mathscr{K}}_r(t)|k;\ell\rangle\right|^2 \geq \frac{1}{D^2}\left|\sum_{k=1}^{D_S}\sum_{\ell=1}^{D_k}\langle k;\ell|\hat{\mathscr{K}}_r(t)|k;\ell\rangle\right|^2 = \frac{1}{D^2}\left|\mathrm{Tr}[\hat{\mathscr{K}}_r(t)]\right|^2. \tag{3.19}$$

Combining Eq. (3.19) with Eq. (3.17), we get the desired inequality, which is Eq. (3.5):

$$P_S(t) \geq \frac{1}{D^2}\sum_{r=1}^{M}\left|\mathrm{Tr}\left[\hat{\mathscr{K}}_r(t)\right]\right|^2 = K(t). \tag{3.20}$$

3.2.2 Condition for saturation

Eq. (3.20) is saturated when $P_S(t) = K(t)$. Here, we will identify a sufficient set of conditions for this to happen, which is exhaustive for pure initial states ($D_S = D$) in the case of time-independent unitary evolution (Hamiltonian or Floquet systems). Mixed initial states and other kinds of time evolution present significant additional complications and are outside the scope of this work.

In Hamiltonian and Floquet systems ($M = 1$, $\hat{\mathscr{K}}_1(t) = \hat{U}(t)$), we can expand any pure initial state $|\psi\rangle$ in the (quasi-)energy eigenbasis $\{|E_n\rangle\}_{n=0}^{D-1}$ as $|\psi\rangle = \sum_n c_n |E_n\rangle$. In the special case that the state is equally distributed over all energy eigenstates, $|c_n|^2 = 1/D$, we have

$$\left|\langle\psi|\hat{U}(t)|\psi\rangle\right|^2 = \frac{1}{D^2}\sum_{n,m} e^{i(E_n - E_m)t} = K(t). \tag{3.21}$$

Thus, it follows that any orthonormal basis of states $\{|k\rangle\}_{k=0}^{D-1}$ which are *all* equally distributed over the energy eigenbasis satisfies $P_S(t) = K(t)$ and saturates Eq. (3.20). A simple example of this is a basis of discrete Fourier transforms (DFTs) of the energy eigenstates:

$$|k\rangle = \frac{1}{\sqrt{D}}\sum_{n=0}^{D-1} e^{-2\pi i k n/D}|E_n\rangle. \tag{3.22}$$

The fact that the above speed limit is saturated in this basis of energy DFTs plays a key role in an ergodic classification of quantum dynamical systems by their level statistics [1], as noted in Chap. 2.

A more general basis of pure states that saturate the bound, by the above consideration, can be expressed by transforming the energy eigenbasis using a unitary complex Hadamard matrix X_{kn} (which, by definition [50], are unitary matrices such that $|X_{kn}|^2 = 1/D$, ensuring equal distribution in

the energy eigenbasis and orthonormality of the transformed basis by virtue of unitarity):

$$|k\rangle = \sum_{n=0}^{D-1} X_{kn} |E_n\rangle. \tag{3.23}$$

The additional complications for other kinds of states and dynamics are as follows: 1. *(for mixed states)* When one has mixed initial states comprised of many pure states, the return probability of the mixed state can involve overlaps of the pure states with each other, and may not saturate the bound even if diagonal in a Hadamard transformation of the energy eigenbasis (where the contribution from the self-overlap of each constituent pure state would already saturate the bound), and 2. *(for non-Hamiltonian/Floquet dynamics)* without a clearly defined basis of energy eigenstates, it is not clear how to identify pure states *within* the Hilbert space whose individual return probability is $K(t)$, though this may be possible for the special case of e.g. time-independent non-unitary (i.e. Lindbladian) evolution which admits its own eigenoperators. However, we note that Eq. (3.10) represents an "approximate" saturation of the bound (in an order of magnitude sense) for large subsystems in Haar random bases, as its right hand side is $K(t) + O(D_S^{-1})$, which only differs from $K(t)$ by a small quantity.

3.2.3 Tightness compared to other basis-independent speed limits

Now we will show quite generally, based on the techniques of Ref. [156], that the speed limit supplied by Eq. (3.20) is tighter than other basis-independent speed limit for a complete set of states that may be derived from known speed limits (such as the MT or ML bounds) on individual states, by replacing state-specific properties with those of the spectrum or Kraus operators. Eq. (3.32) is our main result in this regard, while we also state it in the special case of Hamiltonian systems in Eq. (3.34) to aid intuition.

Let $|\psi\rangle$ be a given pure state such that its return probability,

$$P(|\psi\rangle;t) \equiv \sum_{r=1}^{M} \left|\langle\psi|\hat{\mathscr{K}}_r(t)|\psi\rangle\right|^2, \tag{3.24}$$

satisfies a speed limit set by a function L that depends on the expectation value of powers of the Kraus operators in $|\psi\rangle$:

$$P(|\psi\rangle;t) \geq L\left(\{\langle\psi|[\hat{\mathscr{K}}_r(t)]^m|\psi\rangle\};\ t\right). \tag{3.25}$$

For instance, if we are considering the MT bound with a Hamiltonian $\hat{H}$,

$$L\left(\{\langle\psi|e^{-im\hat{H}t}|\psi\rangle\};\ t\right) = \cos^2(\Delta_E t), \tag{3.26}$$

where $\Delta_E^2 = \langle\psi|\hat{H}^2|\psi\rangle - \langle\psi|\hat{H}|\psi\rangle^2$ can be extracted from the $\{\langle\psi|e^{-im\hat{H}t}|\psi\rangle\}_m$, e.g., through a power series expansion. Of course, one may consider other (e.g., ML) speed limits, in which case L would depend on the expectation values of other functions of $\hat{H}$ in $|\psi\rangle$ in place of the variance Δ_E^2.

With $\mathscr{H}$ being the D-dimensional Hilbert space of a system of interest with an orthonormal basis $\{|k\rangle\}_{k=1}^D$, we consider a maximally entangled state $|\Psi_{\text{TFD}}\rangle$ in the doubled Hilbert space $\mathscr{H}\otimes\mathscr{H}$, given by:

$$|\Psi_{\text{TFD}}\rangle \equiv \frac{1}{\sqrt{D}}\sum_{k=1}^{D}|k\rangle\otimes|k\rangle. \tag{3.27}$$

This is equivalent to the infinite temperature thermofield double (TFD) state when $\{|k\rangle\}$ is chosen to be the energy eigenbasis in a Hamiltonian system [156], but we stress that this does not restrict our results to infinite temperatures alone: finite temperature TFD states can always be accounted for through the Kraus operators, e.g., $M=1$, $\hat{\mathscr{K}}_1(t) = g(\hat{H})e^{-i\hat{H}t}$ with $g(\hat{H}) \propto e^{-\beta\hat{H}}$ (which has the same effect as choosing a finite temperature TFD state [156] of temperature β^{-1} for Hamiltonian evolution). Using the above TFD state, however, has the advantage of allowing us to include non-Hamiltonian e.g.

dissipative or time-dependent systems (which don't have obvious energy eigenstates or candidates for finite temperature TFD states) in our considerations.

We assume that time evolution in this doubled Hilbert space acts only on the "first" copy with no dynamics in the "second", i.e., for density operators $\hat{\rho}$ and $\hat{\sigma}$ in the respective Hilbert spaces (using the notation of Eq. (3.2)),

$$\hat{\mathscr{T}}^t_{\mathscr{H}\otimes\mathscr{H}}[\hat{\rho}\otimes\hat{\sigma}] = \hat{\mathscr{T}}^t[\hat{\rho}]\otimes\hat{\sigma}. \tag{3.28}$$

Then, the observation in Ref. [156] that the SFF $K(t)$ for Hamiltonian systems in $\mathscr{H}$ is the return probability of the TFD state also applies to the generalized SFF in Eq. (3.3):

$$K(t) = P(|\Psi_{\text{TFD}}\rangle; t). \tag{3.29}$$

Here, it is important to re-emphasize that $K(t)$ is a *state-independent* quantity (being just a double trace of the time evolution operator) of the original system in $\mathscr{H}$, while $P(|\Psi_{\text{TFD}}\rangle; t)$ is a naturally state-dependent quantity in the doubled Hilbert space $\mathscr{H}\otimes\mathscr{H}$. Equating the two does not affect the state-independence of the former; instead, the equality states that the dynamics of the specific doubled-Hilbert-space state $|\Psi_{\text{TFD}}\rangle \in \mathscr{H}\otimes\mathscr{H}$ captures *state-independent* aspects of the dynamics within one copy of the Hilbert space $\mathscr{H}$.

Using Eq. (3.29) in (3.25), and noting that the TFD state captures only the *state-independent* traces of the Kraus operators within the Hilbert space $\mathscr{H}$,

$$\langle\Psi_{\text{TFD}}|[\hat{\mathscr{K}}_r(t)]^m|\Psi_{\text{TFD}}\rangle = D^{-1}\,\text{Tr}\left[[\hat{\mathscr{K}}_r(t)]^m\right], \tag{3.30}$$

immediately gives our desired relation:

$$K(t) \geq L\left(\left\{D^{-1}\,\mathrm{Tr}\left[\left[\hat{\mathscr{K}}_r(t)\right]^m\right]\right\};\ t\right). \tag{3.31}$$

We note that it is not essential to deal with a doubled Hilbert space to derive this inequality in certain cases. For Hamiltonian (or Floquet) systems, this can also be derived from $P(|k\rangle; t) = K(t)$ for the states of Eq. (3.23) or as a special case of a geometric bound on the growth of errors in unitary approximations of quantum dynamical systems [1] (see Sec. 2.3.2). But the thermofield double appears to provide a more useful general technique for non-Hamiltonian and non-Floquet evolution where one has no further information about the structure of the Kraus operators.

From $P_S(t) \geq K(t)$, it further follows that

$$P_S(t) \geq K(t) \geq L\left(\left\{D^{-1}\,\mathrm{Tr}\left[\left[\hat{\mathscr{K}}_r(t)\right]^m\right]\right\};\ t\right). \tag{3.32}$$

Let us review the physical content of Eq. (3.32). It states that provided one has a formula for a *pure state* quantum speed limit L in terms of the expectation value of powers of the Kraus operators in a specific state (which may be generalizations of MT, ML or tighter formulations), one can obtain a basis-independent speed limit on $P_S(t)$ (which we recall is the *mean return probability* of a complete set of states, rather than an individual state) from L, but now depending on the state-independent traces of powers of the Kraus operators. However, $K(t)$ also satisfies this speed limit [Eq. (3.31)], and Eq. (3.20) therefore sets a tighter basis-independent bound on $P_S(t)$ than L.

Let us see how this works for the Hamiltonian case to build intuition. Here, the traces $\{\mathrm{Tr}[e^{-im\hat{H}t}]\}_m$ can be transformed a power series expansion (for convenience) directly to the state-independent mo-

ments of the eigenvalues of the Hamiltonian:

$$\langle E^m \rangle \equiv D^{-1} \operatorname{Tr}[\hat{H}^m]. \tag{3.33}$$

These are the statistical properties of the energy spectrum itself, rather than any particular state within $\mathscr{H}$. In this case, Eq. (3.32) becomes:

$$P_S(t) \geq K(t) \geq L\left(\left\{\langle E^m \rangle\right\};\ t\right). \tag{3.34}$$

Thus, L yields a state-independent speed limit on $P_S(t)$ as determined by properties of the energy spectrum itself, but the speed limit set by $K(t)$ is always tighter. When $L = \cos^2(\Delta_E t)$ corresponding to the MT bound, we replace Δ_E^2 (the variance of a state) with σ_E^2 (the variance of the energy spectrum itself) to get Eq. (3.6).

Finally, as already stated in Sec. 3.1, we emphasize that $K(t)$ does not set any direct speed limit on *individual* states, and is not always tighter than state or basis-*dependent* speed limits for specific states. For instance, any state in the energy eigenbasis of a Hamiltonian system has $\Delta_E = 0$ and $P(|E_n\rangle; t) = 1 = \cos(0t)$, saturating the MT bound but not Eq. (3.20) for the basis states; however, almost all other states in the Hilbert space $\mathscr{H}$ with nontrivial dynamics violate this state-specific MT speed limit set by $\cos(0t) = 1$, as $P(t) \leq 1$ in general.

Thus, what Eq. (3.32) shows is that Eq. (3.20) essentially remains the tightest possible state-independent speed limit that applies to *all complete sets of states* in the Hilbert space, as any other speed limit L for return probabilities also sets a state-independent bound [Eq. (3.31)] on the SFF. We note that this makes it especially suited to address questions such as the fastest *allowed* scrambling time for any basis in the system, which is the key application we focus on in this work.

3.3 Types of scrambling and thermalization

Here, we discuss some of the different physically interesting types of scrambling and their relation to thermalization.

3.3.1 The scrambling of added degrees of freedom to infinite temperature

We begin with the type of scrambling directly defined in Sec. 3.1, in the setting of a Hilbert space $\mathscr{H} = \mathscr{H}_S \otimes \mathscr{H}_E$ where the subsystem $\mathscr{H}_S$ is D_S-dimensional, and the component $\mathscr{H}_E$ of the system external to the subsystem is D_E-dimensional. According to Eq. (3.7), scrambling at time t corresponds to:

$$\frac{1}{D_E}\mathrm{Tr}[\hat{\Pi}_k(t)\hat{\Pi}_\ell] = \frac{1}{D_S} + o(D_S^{-1}), \tag{3.35}$$

where $\hat{\Pi}_k/D_E$ are a complete set of (D_E-dimensional) orthonormal initial states, that have the form $\hat{\Pi}_k = |k\rangle_S\langle k| \otimes \hat{\mathbb{1}}_E$, where the $|k\rangle_S$ form an orthonormal basis in $\mathscr{H}_S$. We noted in Sec. 3.1 that this corresponds to thermalization to infinite temperature, which essentially means that $\hat{\Pi}_k(t)/D_E$ "looks" like the maximally mixed (or infinite temperature) thermal state $\hat{\rho}_S(\infty) = \hat{\mathbb{1}}_S/D_S$ within $\mathscr{H}_S$ (see also the discussion in Ref. [51]); correspondingly, Eq. (3.35) also follows by requiring the overlaps of $\hat{\Pi}_k(t)/D_E$ with all $\hat{\Pi}_\ell$ to look like those of $\hat{\rho}_S(\infty)$ to leading order. We note, however, that Eq. (3.35) is slightly more general, and may be satisfied if $\hat{\Pi}_k(t)/D_E$ is not close to a maximally mixed state within $\mathscr{H}_S$, e.g. even if it reduces to a pure state that is completely delocalized over the original basis. In this sense, the above notion of scrambling is only a *necessary* condition for the generation of maximal entanglement within $\mathscr{H}_S$.

Let us consider a physical procedure where the above form of scrambling is relevant. As the initial state $\hat{\rho}_k(0) = \hat{\Pi}_k/D_E = |k\rangle_S\langle k| \otimes (\hat{\mathbb{1}}_E/D_E)$ is pure in the subsystem $\mathscr{H}_S$ and maximally mixed in the external component $\mathscr{H}_E$, this corresponds to the following situation: say we start with a many-

body system with Hilbert space $\mathscr{H}_E$, which has already thermalized to infinite temperature via maximal entanglement with some larger space $\mathscr{H}_R$. At time $t = 0$, we couple it to the subsystem $\mathscr{H}_S$ comprised of a set of qubits in a product state $|k\rangle$, and allow the combined system $\mathscr{H} = \mathscr{H}_S \otimes \mathscr{H}_E$ to evolve (without any interaction with $\mathscr{H}_R$); Eq. (3.35) then corresponds directly to the scrambling of this $\hat{\rho}_k(t)$ for all initial product states of the *added* subsystem.

This type of scrambling of "added degrees of freedom" $\mathscr{H}_S$ to an already scrambled system $\mathscr{H}_E$, the latter being entangled with an additional system $\mathscr{H}_R$, is of the most direct relevance to the Hayden-Preskill protocol and the black hole information problem [53, 54, 141]. In the restricted case of $D_S = \Theta(1)$, the bounds offered by Eq. (3.20) only constrain scrambling to occur slower than some $\Theta(1)$ time (over which $K(t)$ changes by a $\Theta(1)$ amount), which is much shorter than the associated scrambling time of e.g. systems with $[k = O(1)]$-local interactions that is expected [53, 54, 141–143] to be $t = \Omega(\log \log D)$ (especially in large subsystems). This is largely because Eq. (3.20) allows us to consider only the return probabilities *within* the added $\mathscr{H}_S$ (i.e. overlaps of $\hat{\Pi}_k(t)$ with the original $\hat{\Pi}_k$), as opposed to a fuller consideration of the scrambling of these initial states within $\mathscr{H}_S$ over arbitrary subsystems much larger than $\mathscr{H}_S$. We note that bounds corresponding to asymptotically large times $t = \omega(1) \gg \Theta(1)$ are still directly obtained using $P_S(t)$ if one considers adding much larger subsystems with $D_S = \omega(1)$ in this procedure. Additionally, our bound allows us to nontrivially consider arbitrarily large subsystems in a related case of more immediate physical interest as discussed below: the scrambling of initial states that are product states in the *full* system.

3.3.2 The scrambling of pure product states to infinite temperature

Now, given a set of $\{\hat{\Pi}_k\}$, consider initial states corresponding to any basis of their constituent pure states, e.g., $\hat{\Pi}_{k;\ell} \equiv |k;\ell\rangle\langle k;\ell|$ (see also Eq. (3.16)). Then, we have from Eqs. (3.15) and (3.16):

$$P_S(t) = \frac{1}{D}\sum_{k=1}^{D_S}\sum_{\ell=1}^{D_E} \mathrm{Tr}[\hat{\Pi}_{k;\ell}(t)\hat{\Pi}_k]. \tag{3.36}$$

Each term in Eq. (3.36) corresponds to the following situation: prepare the pure state $|k;\ell\rangle \in \mathscr{H}$, and consider its overlap with its restriction $|k\rangle_S \in \mathscr{H}_S$. For instance, the initial states $\hat{\Pi}_{k;\ell}$ could be product states like $|0_1 1_2 0_3 1_4\rangle$ in a computational basis of 4 qubits, while $P_S(t)$ measures their mean overlap with e.g. the corresponding restricted states $|0_1 1_2\rangle_S \in \mathscr{H}_S$ in a 2-qubit subsystem (of, say, the first 2 qubits). Here, we are free to choose $\mathscr{H}_S$ to be whichever subsystem of qubits we want (in particular, of any size), and the scrambling of all such product states in a computational basis within any chosen subsystem of dimension D_S at time t requires that $P_S(t) = D_S^{-1} + o(D_S^{-1})$. Thus, Eq. (3.8) is again a necessary condition for this more accessible form of scrambling.

Such bounds would also apply to a version of the Hayden-Preskill [53] protocol in place of the discussion in the preceding subsection, if it can be extended to a basis of pure initial states in the full $\mathscr{H}_S \otimes \mathscr{H}_E \otimes \mathscr{H}_R$ Hilbert space to eventually appear scrambled within arbitrarily large subsystems of the smaller $\mathscr{H}_S \otimes \mathscr{H}_E$ interacting subspace (we recall that $\mathscr{H}_R$ does not participate in dynamics).

3.3.3 Thermalization to finite temperature

Similar to the infinite temperature case, we may consider a state to have thermalized to a finite temperature β^{-1} if it appears indistinguishable from a given finite temperature density matrix $\hat{\rho}_{(\beta^{-1})}$ within a subsystem. Such density matrices are usually obtained as the reduction of a microcanonical

state,

$$\hat{\rho}_{E_\beta,\Delta E} \propto \sum_{E_n \in [E_\beta, E_\beta + \Delta E]} |E_n\rangle\langle E_n| \tag{3.37}$$

supported on a narrow range of energies $[E_\beta, E_\beta + \Delta E]$, to the subsystem [21, 22, 52]. The simplest way to apply the bound of Eq. (3.20) to this case is to consider a complete set of states within the smaller Hilbert space spanned by the energy eigenstates within $[E_\beta, E_\beta + \Delta E]$.

A complication arises when one considers subsystems in place of an arbitrary basis of states within the microcanonical window. A complete set of states within the subsystem is necessarily supported on the *entire* Hilbert space (by virtue of its completeness) rather than a microcanonical window, and its dynamics consequently involves the full range of energies available in the system — thereby corresponding to infinite temperature thermalization, if one were to require any form of scrambling with respect to such states. It therefore appears that, in the case of finite temperature thermalization, one cannot presume to have complete control over the state of subsystems to the extent of preparing any pure state in a given basis; subsystem states must instead be "coarse grained" in some way to reflect only those states of the subsystem supported within $[E_\beta, E_\beta + \Delta E]$.

Alternatively, it is customary to consider "regularized" time evolution operators such as $g(\hat{H})e^{-i\hat{H}t}$ as accounting for such finite temperature effects, e.g. when $g(\hat{H})$ is significantly supported [114] only within $[E_\beta, E_\beta + \Delta E]$ or sometimes [38, 82] with $g(\hat{H}) \propto e^{-\beta\hat{H}}$. For such a prescription to be useful in our context, in a way that allows applying Eq. (3.20) to a physically meaningful process of thermalization to finite temperature, the choice of $g(\hat{H})$ will have to be justified by an appropriate restriction on the kind of initial states that one may prepare within the subsystem. In either case, our inequality continues to constrain the thermalization of such coarse-grained states states or under regularized time evolution operators to the appropriate finite temperature value of $P_S(t)$.

3.3.4 Entanglement vs operator spreading

In addition to the form of scrambling represented by Eq. (3.35), which we noted to be a necessary condition for the generation of a maximal degree of entanglement, another commonly used notion of scrambling is that of the saturation of out-of-time order correlators (OTOCs) to $\Theta(1)$ values, which measures the extent of "operator spreading" in a many-body system. While this can often take asymptotically long times when the growth of OTOCs is slow, the criterion of saturation to $\Theta(1)$ values does not appear to have sufficient resolution to accurately track the generation of entanglement over a large number of degrees of freedom (e.g. it seems analogous to requiring $P_S(t) \ll 1$ in our case, as opposed to the more precise $P_S(t) = D_S^{-1} + o(D_S^{-1})$). It has also been noted that the saturation of OTOCs can occur in $\Theta(1)$ times even in k-local systems, much faster than the time it takes for entanglement to be generated [142, 143]. For this reason, Eq. (3.35) appears to be the most relevant definition of scrambling for the present study, even independent of its natural connection with Eq. (3.20). However, in App. C.2, we show that $P_S(t) \geq K(t)$ also implies a *formal* bound on the growth of certain (nonlocal) OTOCs.

3.4 Fast Scrambling and Fourier transforms: Formal statement

This section formally states the connection between fast scrambling and properties of Fourier transforms noted in Sec. 2.1 in a certain time regime.

The question of *fast scrambling*, in a form that takes the discussion of "sustained" scrambling in Sec. 2.1 into account, considers whether it is possible for thermodynamically large subsystems of size D_S (implicitly given in terms of D, e.g. $D_S = D^{1/4}$) to become and remain scrambled after a given time $t_s = \Theta[\lambda^{-1}(D_S)]$ up to a long time $T > t_s$ (potentially the long $T \sim \exp[\Theta(D)]$ quantum recurrence time [49, 108]). Here, $\lambda(x) > 0$ is a given monotonically increasing function of $x > 0$, with inverse

$\lambda^{-1}(\lambda(x)) = x$. By Eq. (3.9) of Sec. 3.1, this requires (with $t < T$ implicit):

$$K\left(t > t_s = \Theta[\lambda^{-1}(D_S)]\right) \leq \frac{1}{D_S} + o(D_S^{-1}). \tag{3.38}$$

Further, Eq. (3.10) of Sec. 3.1 implies that

$$K\left(t > t_s = \Theta[\lambda^{-1}(D_S)]\right) = o(D_S^{-1}) \tag{3.39}$$

is a sufficient condition for fast scrambling by $t_s = \Theta[\lambda^{-1}(D_S)]$ in *some* subsystem of dimension D_S in the Hilbert space. We note that the $\Theta(1)$ time scale is set here by that of significant $\Theta(1)$ variations in the SFF (i.e. the time scale over which $K(t)$ decreases from $K(0) = 1$ by a finite amount). The fastest scrambling system is one that satisfies the above criterion with the fastest growing $\lambda(x)$, or equivalently, the slowest growing t_s (say, among a set of systems subject to certain "physicality" conditions on the SFF). While it has been conjectured and argued that $t_s = \Omega(\log\log D)$ — with a suitable prefactor (e.g. inverse temperature) that sets an $O(1)$ time scale — in systems with $[k = O(1)]$-local interactions [54, 141, 142], we would like to explore this question in a general setting. In our approach, on setting $K(t) = |\widetilde{\mathcal{N}_D}(t)|^2$ in Eqs. (3.38) and (3.39) (where $\mathcal{N}_D(t)$ is the Fourier transform of the density of states $\mathcal{N}(E)$, as in Eq. (3.11) of Sec. 3.1), the problem of fast scrambling is determined by the $D_S, D \to \infty$ asymptotic behavior of $|\widetilde{\mathcal{N}_D}(\Omega[\lambda^{-1}(D_S)])|^2$.

To simplify the present analysis, we will focus on an early-time regime $|t| < t_D$ where we assume negligible *explicit* D-dependence (to avoid complications from having 2 independent asymptotic parameters D_S and D in this initial study):

$$\widetilde{\mathcal{N}_D}\left(t : |t| < t_D\right) = \widetilde{\mathcal{N}}(t) + o\left(\widetilde{\mathcal{N}}(t)\right), \tag{3.40}$$

for a D-independent function $\widetilde{\mathcal{N}}(t)$. By Eq. (3.38), $\widetilde{\mathcal{N}_D}(t)$ can be replaced with $\widetilde{\mathcal{N}}(t)$ in the fast scrambling problem for subsystems of dimension $D_S = o(D_{\max})$, where $D_{\max} \equiv 1/K(t_D)$; we will correspondingly call $\widetilde{\mathcal{N}}(t)$ the "smoothened" form of $\widetilde{\mathcal{N}_D}(t)$. It is further convenient to smoothen out the finer structure of $\mathcal{N}_D(E)$ (by e.g. convolving it with a Gaussian of width $\sigma \sim 2\pi/t_D$), so that it essentially equals the D-independent Fourier transform $\mathcal{N}(E) \geq 0$ of $\widetilde{\mathcal{N}}(t)$ (as such a convolution suppresses the $|t| > t_D$ oscillations of $\widetilde{\mathcal{N}_D}(t)$). Several physical systems can even have $D_{\max} \sim D$ (corresponding to effective D-independence at early times when $K(t) \gg D^{-1}$), such as typical realizations of the random matrix ensembles [28, 36] and the (exact or numerically apparent) behavior of some of the many-body systems considered in Refs. [4, 38, 87, 148], where this smoothening can be made rigorous through ensemble averaging (noting that "each" energy level typically follows a smooth distribution, though not necessarily Gaussian, over infinitely many ensemble realizations). As another example, the SYK-($q \geq 4$) model [82, 163, 164] considered in Sec. 3.1 and Sec. 6.2 appears to have $D_m \sim D^\alpha$ with $\alpha < 1$ (see Eq. (6.8)), due to which an extension of these formal considerations that accounts for D-dependence is necessary to directly tackle Eq. (6.9) applying to larger subsystems.

With this simplification (which is essentially an argument to use a smooth D-independent density of states), we can formally state the combination of Eqs. (3.38) and (3.39), using (3.40) and the fact that $|\widetilde{\mathcal{N}}(0)| = 1$ from the properties of $K(t)$, as a proposition relating the question of whether the scrambling of subsystems can occur before a given time scale $\lambda^{-1}(D_S)$ to mathematical necessary conditions on the asymptotic decay of a (sufficiently well-behaved) function to ensure a nonnegative Fourier transform:

Proposition 3.3 (**Fast scrambling and Fourier transform nonnegativity**). *Let $\mathscr{F}$ be the set of functions $\widetilde{\mathcal{N}} : \mathbb{R} \to \mathbb{C}$ normalizable to $|\widetilde{\mathcal{N}}(0)| = 1$ with a non-negative real-valued Fourier transform $\mathcal{N}(E) \geq 0$, and $\mathscr{F}_p \subseteq \mathscr{F}$ be any subset of these functions satisfying as yet unspecified "physicality"*

conditions. ***If*** *every such "physical"* $\widetilde{\mathcal{N}}(t) \in \mathscr{F}_p$ *necessarily has a slow asymptotic decay satisfying:*

$$|\widetilde{\mathcal{N}}(t \to \infty)|^2 \neq o\left[\frac{1}{\lambda(\Theta(t))}\right], \tag{3.41}$$

where $\lambda(x > 0) > 0$ *is a given monotonically increasing function,* ***then*** *any quantum system whose dynamics is subject to the same physicality conditions [i.e. with smoothened* $\widetilde{\mathcal{N}}(t) \in \mathscr{F}_p$ *and the corresponding smoothened density of states* $\mathcal{N}(E)$*] can scramble subsystems of dimension* $D_S = o(D_{\max})$ *only after a time* $t_s = \Omega(\lambda^{-1}(D_S))$ *[by Eq. (3.38)].* ***Further****, a quantum system with* $\widetilde{\mathcal{N}}(t) \in \mathscr{F}_p$ *that scrambles subsystems within* $t_s = O(\lambda^{-1}(D_S))$ *is guaranteed to exist if there is at least one function* $\widetilde{\mathcal{N}_1}(t) \in \mathscr{F}_p$ *that decays faster than in Eq. (3.41), i.e.* $|\widetilde{\mathcal{N}_1}(t \to \infty)|^2 = o\left[1/\lambda(\Theta(x))\right]$ *[from Eq. (3.39)].*

This proposition is our main formal statement as far as Hamiltonian fast scrambling is concerned, setting lower limits on the scrambling time within the regime of effective D-independence in $\widetilde{\mathcal{N}_D}(t)$. The subset of functions $\mathscr{F}_p$ left unspecified above may be chosen according to certain "physicality" conditions on the spectrum, e.g. the analyticity of $K(t)$ in the $|t| < t_D$ regime. What makes the resulting problem nontrivial is that one has to simultaneously satisfy $\mathcal{N}(E) \geq 0$ *and* the physicality conditions.

As noted in Sec. 3.1, it is not known [157, 158] how such conditions determine the asymptotic behavior of Fourier transforms. What is known are certain necessary conditions [158] such as maximality $\widetilde{\mathcal{N}}(0) \geq \widetilde{\mathcal{N}}(t)$ (automatically satisfied in our case due to the properties of $K(t)$) and concavity $\widetilde{\mathcal{N}}''(t = 0) < 0$ at the origin, and sufficient conditions [157] such as convexity $\widetilde{\mathcal{N}}''(t > 0) > 0$ for real-valued and symmetric $\widetilde{\mathcal{N}}(t) = \widetilde{\mathcal{N}}(-t)$. Such convex functions can be made to decay as fast as desired (corresponding to t_s as small as desired in physical systems): an extreme limiting case with $t_s = 1$ is $\widetilde{\mathcal{N}}(t) = \max\{1-|t|, 0\}$, with $\mathcal{N}(E) = \mathcal{N}(E) = \mathrm{sinc}^2(\pi E) \geq 0$. However, they are non-analytic at $t = 0$, possessing an infinite energy variance $\sigma_E^2 = -K''(0)/2$; consequently, the scrambling time *relative* to

the time scale set by σ_E is infinite even in this case, i.e., $\sigma_E t \to \infty$. We also note that the (state-independent) ML bound for $P_S(t)$ in this case would be trivial, $P_S(t \neq 0) \geq 0$, as $\mathcal{N}(E) = \mathrm{sinc}^2(\pi E)$ admits a spectrum of infinite width, without a finite ground or ceiling state in the $D \to \infty$ limit (alternatively, as noted in Statement 2 of Sec. 3.1, a nontrivial ML bound would automatically imply a higher-than-logarithmic scrambling time in D_S, and cannot lead to $t_s = 1$).

In the next Chapter, we will derive a precise statement that goes beyond the formal statement above and simultaneously tackles the difficulties posed by subjective physicality conditions and arguments required for a thermodynamic limit. This is made possible by noting that analyticity in complex time is a universal property of every finite dimensional quantum system, and consequently, we do not need to take the thermodynamic limit prior to obtaining a rigorous bound on the scrambling time. Remarkably, we will find that a mathematical relation for complex analytic functions establishes a general bound of the form $t_s \gtrsim \Omega(\log\log D)$ for sufficiently large D_S, suggested after Eq. (3.39) above based on the fast scrambling conjecture for k-local systems [54, 141, 142], as a universal property of all Hamiltonian quantum systems including arbitrarily nonlocal ones.

Chapter 4: Proof of a universal speed limit on fast scrambling in quantum systems

4.1 Synopsis: The maximum speed of Hamiltonian dynamics

Introduction— Consider a quantum mechanical system with Hilbert space $\mathscr{H}_S$, e.g., a collection of N_S qubits with dimension $D_S = 2^{N_S}$, that is initially in a pure basis state $|k\rangle_S$ (which we call the computational basis). At the time $t = 0$, we couple $\mathscr{H}_S$ to an external system (say, of N_E qubits and dimension $D_E = 2^{N_E}$) in a state $\hat{\rho}_{\beta E}$ (e.g., a thermal state with inverse temperature β), and allow their mutual interactions to drive time evolution in the combined system. The initial state of the overall system $\mathscr{H}_S \otimes \mathscr{H}_E$ of dimension $D = D_S D_E$ is given by the density operator:

$$\hat{\rho}_k(0) = |k\rangle_S\langle k| \otimes \hat{\rho}_{\beta E}. \tag{4.1}$$

Generically, information about the initial state of $\mathscr{H}_S$ is "scrambled" by interactions, being lost to $\mathscr{H}_E$ through the generation of entanglement, and not entirely recoverable from $\mathscr{H}_S$ alone. To examine to what extent any basis $|\alpha\rangle_S$ of the subsystem $\mathscr{H}_S$ retains this information, we probe the system with measurements represented by the orthonormal projectors

$$\hat{\Pi}_\alpha = |\alpha\rangle_S\langle\alpha| \otimes \hat{\mathbb{1}}_E. \tag{4.2}$$

The above is a minimal description of quantum information scrambling that has far-reaching connections to statistical mechanics. Most directly, it is the prototypical setting for the nonequilibrium

process of thermalization [16,17,21,23,24,51,60,69,138,140], of the system $\mathscr{H}_S$ via interactions with a thermal reservoir $\mathscr{H}_E$, which (if it occurs) is ultimately responsible for the validity of equilibrium statistical mechanics in $\mathscr{H}_S$. In a separate but related context, the same model is also believed to capture some essential aspects of information loss and recovery in black holes, even without a precise microscopic model of black holes as may emerge from an eventual theory of quantum gravity [53]. Here, one thinks of the thermal reservoir $\mathscr{H}_E$ as representing the information already contained within a black hole, which is in the thermal state $\hat{\rho}_{\beta E}$ due to entanglement with the Hawking radiation emitted before $t = 0$ (a separate quantum system $\mathscr{H}_R$ that does not interact with $\mathscr{H}_E$). The system $\mathscr{H}_S$ then represents infalling matter into the black hole, and the above setup models how infalling information spreads among a black hole's degrees of freedom, and whether it can be recovered at any given time from the previously emitted Hawking radiation [53].

In both cases, a problem of fundamental interest is to determine the fastest allowed scrambling time t_s, after which $\mathscr{H}_S$ retains no memory of its initial state up to thermodynamic parameters like temperature. From the perspective of quantum statistical mechanics, this amounts to constraining the earliest time after which equilibrium statistical mechanics can apply to any subsystem of interest. Separately, a scrambling time of $t_s \gtrsim O(\beta \ln N)$ has been conjectured in sufficiently k-local systems of $N = N_S + N_E$ particles [53, 54], based on arguments attempting to reconcile models of quantum information in black holes with ideas in quantum gravity. A number of studies of the scrambling time with different notions of scrambling appear to be consistent with this conjecture [141, 142, 159, 163] in certain classes of systems, but it is not expected that a nontrivial bound on t_s exists for a general quantum mechanical system without any restrictions on the nature of interactions [141, 154].

Qualitative summary— In this Chapter, we tackle the challenge of proving a bound on fast scrambling assuming "just the validity of quantum mechanics" (as posed in Ref. [141]), going beyond the specific models of local interactions that were previously considered necessary. Our approach is en-

abled by direct connections between quantum dynamics and the structure of the energy spectrum developed in Refs. [1,2], assuming little beyond the existence of quantum energy levels. In particular, due to a refinement of the energy-time uncertainty principle resulting from these connections [1, 2] (see Sec. 1.4.1 and Chaps. 2 and 3), the problem of formulating speed limits on *infinite temperature* scrambling within quantum mechanics can be mapped to the asymptotic properties of the Fourier transform of the density of states (distribution of energy levels) of a Hamiltonian system [2] (as discussed in Chap. 3). Here, we extend this result in two ways: (1) We generalize the previously obtained speed limit to a reservoir in an arbitrary initial state $\hat{\rho}_{\beta E}$, including finite temperature states [Eqs. (4.13), (4.14)]; (2) We identify (and derive a quantitative version of) a rigorous theorem on the asymptotics of Fourier transforms that provides an exact speed limit for any Hamiltonian system. Surprisingly, our speed limit holds for arbitrarily nonlocal systems (including random matrix Hamiltonians [36]). It establishes a quantitative limit on fast scrambling [Eq. (4.8)] that is logarithmic in the scrambled entanglement entropy [Eq. (4.10)], as a property of quantum mechanics itself rather than any particular class of systems.

Quantitative overview of results— To formulate our results in a manner similar to Ref. [2] (Chap. 3), we work with the mean return probability of the computational basis states in $\mathscr{H}_S$ (i.e., selecting $\hat{\Pi}_k$ with $\alpha = k$ in Eq. (4.2)) under evolution by a Hamiltonian $\hat{H}$, given the corresponding initial state $\hat{\rho}_k(0)$ in Eq. (4.1):

$$P_S(t) \equiv \frac{1}{D_S} \sum_{k=1}^{D_S} \mathrm{Tr}[e^{-i\hat{H}t} \hat{\rho}_k(0) e^{i\hat{H}t} \hat{\Pi}_k]. \tag{4.3}$$

This quantity partially measures the extent to which the *computational basis*, among all the bases in $\mathscr{H}_S$, retains information about the initial state, and constrains the generation of entanglement between the two systems which is at the root of quantum thermalization [17,23,24,52]. For example, the average

purity of the reduced states $\hat{\rho}_{k,S}(t) \equiv \mathrm{Tr}_E\, \hat{\rho}_k(t)$ is constrained by (see Sec. 4.2),

$$\frac{1}{D_S}\sum_{k=1}^{D_S} \mathrm{Tr}_S\left[\hat{\rho}_{k,S}^2(t)\right] \geq [P_S(t)]^2. \tag{4.4}$$

Initially, $P_S(0) = 1$, signifying complete retention of this information and pure basis states in $\mathscr{H}_S$. At subsequent times, we expect the scrambling process to cause $P_S(t)$ to decay to a "scrambled" value P_{scr}. For example, when information about the initial state is completely and uniformly delocalized (to leading order) over the D_S basis states $|k\rangle_S$, we have $P_{\mathrm{scr}} \sim D_S^{-1}$, which is typically necessary for maximal entanglement [2] (as also noted in Chap. 3). Here, we use formal asymptotic notation [147], with $\omega, \Omega, \Theta, O, o$ representing $>, \geq, =, \leq, <$ up to any multiplicative constant, and $\gtrsim, \sim, \lesssim$ for leading-order (in)equalities, when a parameter such as D diverges. To maintain generality, we do not assume a specific scrambled value at this stage, and we allow D to remain finite without an explicit limit, except where asymptotic notation is used.

Our main concern is the *sustained* [2] scrambling time t_s, introduced in Chap. 3, after which $P_S(t)$ remains no larger than the corresponding scrambled value P_{scr}. We must also implicitly only consider times $|t| < T_D$, where T_D is a cutoff beyond all time scales of physical interest, chosen to avoid quantum recurrences [49] to large $P_S(t)$. Thus, t_s is set by the condition

$$P_S(t > t_s) \leq P_{\mathrm{scr}}. \tag{4.5}$$

Given this setup, we can summarize our main result. We find that t_s is bounded in any quantum mechanical system in terms of the following three quantities: (1) the scrambled return probability P_{scr}, (2) the information-theoretic *fidelity* [50] $0 \leq f_\beta \leq 1$ between the thermal state $\hat{\rho}_{\beta E}$ and the

infinite temperature maximally mixed state $\hat{\rho}_{0E} \equiv \hat{\mathbb{1}}_E / D_E$ in $\mathscr{H}_E$,

$$f_\beta \equiv \mathrm{Tr}_E \sqrt{\sqrt{\hat{\rho}_{0E}}\, \hat{\rho}_{\beta E} \sqrt{\hat{\rho}_{0E}}} = \frac{1}{\sqrt{D_E}} \mathrm{Tr}_E \left[\hat{\rho}_{\beta E}^{1/2} \right], \tag{4.6}$$

and (3) an effective partition function $Z_\beta(\tau) = \widetilde{\mathscr{N}_\beta}(-i\tau)$ [see Eq. (4.15)] in a Euclidean time interval containing $\tau = 0$:

$$\tau \in [\tau_1, \tau_2],\ \tau_1 \leq 0 \leq \tau_2,\ \text{of width}\ (\tau_2 - \tau_1) \equiv c\beta. \tag{4.7}$$

Specifically, we show that for *every* interval $[\tau_1, \tau_2]$ as in Eq (4.7), the sustained scrambling time satisfies:

$$t_s \geq \frac{c\beta}{\pi} \ln \left[\frac{1}{\Lambda_{\text{eff}}} \ln \left(\frac{f_\beta^2 Z_{\max}^2}{P_{\text{scr}}} \right) \right]. \tag{4.8}$$

The parameters $Z_{\max} = \max_{\tau \in [\tau_1, \tau_2]} Z_\beta(\tau)$ (attained at $\tau = \tau_1$ or τ_2, as discussed in Sec. 4.4) and Λ_{eff} [Eqs. (4.23), (4.24)] depend on $Z_\beta(\tau)$ within the interval. One can further optimize the choice of interval in Eq. (4.7), including $c\beta$, to yield the tightest bound on t_s, but this depends on the behavior of $Z_\beta(\tau)$ and must be done in system-specific ways. Eq. (4.8) provides a direct, exact bound on the generation of entanglement [via Eq. (4.4)], and applies to arbitrary Hamiltonians $\hat{H}$, in contrast to the (conjectured) bound on chaos in out-of-time-ordered correlators (OTOCs) of local operators that is shown to hold only for special classes of Hamiltonians [142, 143, 159, 166]. Further, the two bounds do not necessarily imply each other and refer to distinct physical processes associated with scrambling [2, 142, 143, 167].

We can alternatively directly constrain t_s in terms of a minimum 2nd Rényi entanglement entropy [63] $S_{2,S}$ required for a state to be considered scrambled in $\mathscr{H}_S$:

$$S_2[\hat{\rho}_{k,S}^2(t > t_s)] \equiv -\ln \mathrm{Tr}[\hat{\rho}_{k,S}^2(t)] \geq S_{2,S}. \tag{4.9}$$

By Eq. (4.4), P_{scr} can then be replaced by $\exp(-S_{2,S}/2)$ in Eqs. (4.5) and (4.8). This is discussed in more detail in Sec. 4.2. To simplify the resulting bound on t_s, we consider a thermodynamic limit $N_S, N_E \to \infty$, such that $Z_\beta(\tau)$ remains $O(1)$ in $[\tau_1, \tau_2]$ and $f_\beta \geq \exp[-o(S_{2,S})]$ is not too small. In this regime, we obtain to leading order:

$$t_s \gtrsim \frac{c\beta}{\pi} \ln S_{2,S}. \tag{4.10}$$

Further, a canonical thermal state typically admits $c = (1-\epsilon)/2$ with a small $\epsilon > 0$ for $Z_\beta(\tau) = O(1)$ [as motivated near Eq. (4.20)]. For volume-law (including maximal) entanglement [51, 140], $S_{2,S} \sim \mu N_S$ for $N_S < N_E$. Consequently, the conjectured [53, 54] fast scrambling bound $t_s = \Omega(\beta \ln N)$ holds for subsystems with $N_S \sim N^\kappa$ (where $0 < \kappa \lesssim 1$).

Effective nonunitary dynamics at finite temperatures— We begin our technical derivation with a brief summary of a result of Ref. [2] (Chap. 3). There, we derived a universal speed limit on $P_S(t)$ under arbitrary time-dependent quantum operations [50] $\hat{\rho}_k(t) = \sum_{r=1}^{M} \hat{\mathscr{K}}_r(t) \hat{\rho}_{k,0} \hat{\mathscr{K}}_r^\dagger(t)$, but only for a special choice of initial states $\hat{\rho}_{k,0} = \hat{\Pi}_k / D_E$ corresponding to the infinite temperature state $\hat{\rho}_{0E}$ in Eq. (4.1). To adapt this bound for the more general family of initial states in Eqs. (4.1) and (4.3), we can apply a simple trick that trades off unitary dynamics with an arbitrary initial state for nonunitary dynamics with infinite temperature initial states.

Specifically, in Eq. (4.3), we write

$$\hat{\rho}_k(t) = e^{-i\hat{H}t} \hat{\rho}_k(0) e^{i\hat{H}t} = \frac{1}{D_E} \hat{\mathscr{K}}_1(t) \hat{\Pi}_k \hat{\mathscr{K}}_1^\dagger(t), \tag{4.11}$$

$$\text{with } \hat{\mathscr{K}}_1(t) \equiv e^{-i\hat{H}t} \hat{\rho}_{\beta E}^{1/2} D_E^{1/2}. \tag{4.12}$$

Thus, $P_S(t)$ can also be regarded as the mean return probability of the infinite temperature initial states $\hat{\rho}_{k,0} = \hat{\Pi}_k / D_E$ with effective time-dependent nonunitary dynamics given by $\hat{\mathscr{K}}_1(t)$. In this case, the

speed limit of Ref. [2] [see Eq. (3.5)] gives the following bound, derived in Sec. 4.3, which underlies the main results of this Chapter:

$$P_S(t) \geq K_\beta(t). \tag{4.13}$$

In the above inequality,

$$K_\beta(t) = \frac{D_E}{D^2} \left| \mathrm{Tr}\left(e^{-i\hat{H}t} \hat{\rho}_{\beta E}^{1/2} \right) \right|^2 \tag{4.14}$$

is the *regularized* spectral form factor (SFF) corresponding to the density operator $\hat{\rho}_{\beta E}$, occurring here as a special case of the generalized SFF [2] for time-dependent quantum operations $K(t) \equiv D^{-2} \sum_{r=1}^{M} \left| \mathrm{Tr}\, \hat{\mathscr{K}}_r(t) \right|^2$ [see Eq. (3.3)], which extends and unifies time-dependent [149] and dissipative [150–153] SFFs.

For context, we note that SFFs are a key staple of quantum chaos that characterize energy level correlations [28] and the presence of observable-independent ergodic properties in quantum dynamics [1,5]; see Chap. 2. Further, they can be directly measured in experiments [6,96,168], as described in Chaps. 8 and 9. Eq. (4.14) specifically belongs to a family of regularized SFFs [82,89,156,160,169–171] and correlators [159,163], where the splitting of roots of a density operator between different factors (usually for analytical convenience) has been regarded as somewhat artificial to engineer in experiments [96,172]. That the regularized SFF $K_\beta(t)$ universally constrains pure Hamiltonian dynamics via Eq. (4.13) now provides an exact operational physical justification for regularization, at least in one context.

Properties of regularized density of states— From Eqs. (4.13), (4.14), it is clear that we can obtain universal bounds on the decay of $P_S(t)$ if we can constrain the decay of $K_\beta(t)$ for any Hamiltonian system. To do so, it is useful to focus on its "square root" up to constant factors, the trace of $\hat{\mathscr{K}}_1(t)$,

$$\widetilde{\mathscr{N}}_\beta(t) \equiv \frac{\mathrm{Tr}\, \hat{\mathscr{K}}_1(t)}{\mathrm{Tr}\, \hat{\mathscr{K}}_1(0)} = \frac{D_E^{1/2}}{f_\beta D} \sum_n \langle E_n | \hat{\rho}_{\beta E}^{1/2} | E_n \rangle e^{-iE_n t}, \tag{4.15}$$

normalized to $\widetilde{\mathscr{N}_\beta}(0) = 1$, so that $|\widetilde{\mathscr{N}_\beta}(t \geq 0)| \leq 1$. $\widetilde{\mathscr{N}_\beta}(t)$ is the Fourier transform of the *regularized* density of states (a probability distribution that integrates to $\widetilde{\mathscr{N}_\beta}(0) = 1$),

$$\mathscr{N}_\beta(E) = \frac{D_E^{1/2}}{f_\beta D} \sum_n \langle E_n | \hat{\rho}_{\beta E}^{1/2} | E_n \rangle \delta(E - E_n) \geq 0. \tag{4.16}$$

In terms of this quantity, we have

$$K_\beta(t) = f_\beta^2 \left| \widetilde{\mathscr{N}_\beta}(t) \right|^2. \tag{4.17}$$

Now, we can use this in conjunction with Eqs. (4.5) and (4.13) to directly relate the scrambling time t_s to the regularized density of states,

$$\left| \widetilde{\mathscr{N}_\beta}(t > t_s) \right|^2 \leq \frac{P_{\text{scr}}}{f_\beta^2}, \tag{4.18}$$

generalizing the corresponding infinite temperature result of Ref. [2] (Chap. 3), i.e., Eq. (3.5). For this to give a nontrivial $t_s > 0$, we must have $f_\beta^2 > P_{\text{scr}}$ (as $|\widetilde{\mathscr{N}_\beta}(t)| \leq 1$). This condition can always be satisfied for any given P_{scr} by any $\hat{\rho}_{\beta E}$ with sufficient entanglement entropy in $\mathscr{H}_E$ (see Sec. 4.3), and for any given initial state $\hat{\rho}_{\beta E}$ by choosing N_S large enough to have sufficiently small $P_{\text{scr}} < f_\beta^2$ [e.g., if $P_{\text{scr}} \sim D_S^{-1}$].

To constrain the decay of $\widetilde{\mathscr{N}_\beta}(t)$, we extend the time variable to the complex plane, $\mathbf{t} = t - i\tau$, with $t, \tau \in \mathbb{R}$. In this case, $\widetilde{\mathscr{N}_\beta}(\mathbf{t})$ is a weighted sum of D (analytic) exponentials $e^{-iE_n \mathbf{t}}$, and is therefore analytic everywhere on $\mathbb{C}$. It is also real-valued on the imaginary axis, with $\widetilde{\mathscr{N}_\beta}(-i\tau) = Z_\beta(\tau)$ being the partition function of the regularized density of states at inverse temperature (or Euclidean time) τ. From Eq. (4.15), this partition function sets an upper bound on $|\widetilde{\mathscr{N}_\beta}(\mathbf{t})|$ at each τ:

$$|\widetilde{\mathscr{N}_\beta}(t - i\tau)| \leq Z_\beta(\tau), \ \forall \ t \in \mathbb{R}, \tag{4.19}$$

which is saturated (at least) at $t = 0$.

While $Z_\beta(0) = 1$, it can diverge for other values of τ, making $\widetilde{\mathcal{N}_\beta}(\mathbf{t})$ unbounded in a thermodynamic limit. For example, for a canonical thermal state $\hat{\rho}_{\beta E}$ with $N_S \ll N_E$ whose energy eigenbasis matrix elements behave as $e^{-\beta E_n}$,

$$Z_\beta(\tau) \approx \frac{\sum_n e^{-(\beta+2\tau)E_n/2}}{\sum_n e^{-\beta E_n/2}}. \tag{4.20}$$

For most $-\beta/2 < \tau \leq 0$, the numerator remains exponentially suppressed, and we expect $Z_\beta(\tau) = \Theta(1)$ if the density of states doesn't grow exponentially with E_n. However, for $\tau \approx -\beta/2$, the exponential suppression in the numerator is lost and $Z_\beta(-\beta/2) \approx D_E^{1/2}/f_\beta$ diverges as $N_E \to \infty$. This motivates considering horizontal strips containing $\tau = 0$, restricted to a specific width $c\beta$ as in Eq. (4.7), in which quantities remain bounded. While this argument suggests that a strip in $(-\beta/2, 0]$ with $c = (1-\epsilon)/2$ is appropriate for thermal states, we formulate our results for an arbitrary width $c\beta \geq 0$ to retain exactness and universal applicability to arbitrary initial states.

Here, we also note previous physical bounds involving similar properties of analyticity and boundedness in a strip, including that expected for regularized OTOCs [159], as well as at a point of inflection of the regularized SFF [160], which however differ qualitatively and quantitatively from the results of interest [e.g., Eq. (4.21)] in this Chapter.

Universal bound on fast scrambling— In our case, given that $\widetilde{\mathcal{N}_\beta}(\mathbf{t})$ is analytic and bounded in the strip given by Eq. (4.7), results from the theory of analytic functions [173–177] constrain $\widetilde{\mathcal{N}_\beta}(t \geq 0)$ to decay no faster than a double exponential in t for *nearly all* $t > 0$, *except* in an "exceptional set" $W_{\text{ex}} \subset \mathbb{R}_{>0}$ of finite length in t. In particular, recalling that $Z_{\max} = \max_{\tau\in[\tau_1,\tau_2]} Z_\beta(\tau)$, we have (see Sec. 4.4)

$$|\widetilde{\mathcal{N}_\beta}(t \notin W_{\text{ex}})|^2 \geq Z_{\max}^2 \exp\left[-\Lambda \exp\left(\frac{\pi t}{c\beta}\right)\right], \tag{4.21}$$

where Λ is a finite constant [173,174]. For intuition behind the exceptional set W_{ex}, consider [1,2,114] the function $\widetilde{\mathcal{N}_\beta}(t) = \text{sinc}^2(\pi t/\beta)$ (see also related discussions in Chap. 2 near Eq. (2.41), and Chap. 3 near Eq. (3.11)), which is oscillating and vanishes at the nodes $t = n\beta$ for positive integers n. While its asymptotic behavior is generally t^{-2}, satisfying Eq. (4.21) for most t, the oscillations bring it below the bound of Eq. (4.21) in a (shrinking with n) neighborhood of each node, which comprise the finite-length set W_{ex}. The presence of the exceptional set excludes all discrete-time systems from a nontrivial application of Eq. (4.21), such as Haar random unitaries that maximally scramble $\mathscr{H}_S$ after just a single time-step $t = 1$, but *not* their continuous-time Hamiltonian extrapolations.

However, in this form, Eq. (4.21) is not sufficiently predictive for our purposes. This is because the "finite" constant Λ and the "finite" duration of W_{ex} are undetermined, and could still be larger or smaller than any scale of interest for a system with finite D. Fortunately, we can adapt the methods of Ref. [176] to explicitly obtain their values. The details of this calculation are given in Sec. 4.4, and here we summarize the resulting finite-D expressions. We find that a rescaled length of W_{ex} can be made smaller than any chosen $\ell > 0$,

$$\int_{W_{\text{ex}}} \mathrm{d}t \leq \frac{c\beta}{\pi}\ell, \tag{4.22}$$

if one chooses Λ according to

$$\Lambda_\ell = \frac{(2+\ell)}{\ell}\left\{\min_{\tau\in[\tau_1,\tau_2]} \frac{2(\pi^2+8)\ln\left[\frac{Z_{\max}}{Z_\beta(\tau)}\right]}{\cos\left[\frac{\pi}{c\beta}\left(\tau - \frac{\tau_1+\tau_2}{2}\right)\right]}\right\}. \tag{4.23}$$

We can now obtain the scrambling time from Eq. (4.18), with the caveat that the exceptional set introduces an uncertainty in the scrambling time with range $[t_s, t_s + (c\beta\ell/\pi)]$ from Eq. (4.22). We get, using Eq. (4.18), (4.21) and (4.23),

$$t_s \geq \frac{c\beta}{\pi}\left\{\ln\left[\frac{1}{\Lambda_\ell}\ln\frac{f_\beta^2 Z_{\max}^2}{P_{\text{scr}}}\right] - \ell\right\}, \tag{4.24}$$

where the second term accounts for the uncertainty by subtracting the length of W_{ex} from the bound obtained. To optimize tightness, we choose ℓ to maximize the right hand side of Eq. (4.24). We get $\ell = \sqrt{3}-1$, from which we recover our main universal result, Eq. (4.8), with $\Lambda_{\text{eff}} \equiv e^{\sqrt{3}-1}\Lambda_{(\sqrt{3}-1)}$.

Further, an important question is whether a near-optimally decaying double exponential function as in Eq. (4.21) can emerge in a system with a nonnegative (regularized) density of states $\mathcal{N}_\beta(E) \geq 0$ (e.g., [2]; see Chap. 3 for context). This is indeed the case: a class of examples was constructed in Refs. [178], [175, p.35], and we construct other (simpler) examples in Sec. 4.5 by taking the thermodynamic limit of $\widetilde{\mathcal{N}_\beta}(t)$ to be the convolution $g * g$, where g is the double-exponentially decaying function $g(t) = \exp(-e^{\frac{\pi}{c\beta}t})\exp(-e^{-\frac{\pi}{c\beta}t})$. Given sufficiently (including Haar [28]) random energy eigenstates, these examples scramble by $t_s \lesssim (2c\beta/\pi)\ln N_S$, nearly saturating Eq. (4.10) [2, 114, 154, 155] [see the discussion following Eq. (3.10)].

Discussion—We consider the implications of our results in light of the fast scrambling conjecture [54], which states that (a) $t_s \gtrsim O(\beta \log N)$ for some definition of scrambling time and some class of systems (neither rigorously specified [141–143, 166]), and (b) quantum systems believed to describe black holes (again not rigorously specified) saturate this bound. For a successful analysis, we must specify at least one of the three unspecified features and examine its implications for the others.

In this Chapter, we have considered the notion of scrambling in Eq. (4.5) that directly constrains the time required for sustained entanglement generation and the applicability of equilibrium statistical mechanics in subsystems. In this case, Eq. (4.8) provides a universal quantum mechanical bound on information scrambling by *any* Hamiltonian system; Eq. (4.10) further shows that even the conjectured $t_s = \Omega(\beta \ln N)$ form holds *universally* for volume-law entanglement in a certain thermodynamic regime. This completely specifies and establishes statement (a) of the fast scrambling conjecture for this notion of scrambling, generalizing it beyond any specific assumptions of interaction structure.

For statement (b), we first consider the example of the "maximally chaotic" Sachdev-Ye-Kitaev

models [163], which saturate the OTOC bound for local correlators [159]. A key feature of these models is a zero-temperature entropy (extensive clustering of states near the ground state) [163, 165, 179] that is directly believed to capture some aspects of black hole physics [180, 181]. However, this feature leads to a subleading slow decay [82] of $K_\beta(t)$, giving t_s exponential in N_S (for large N_S) as shown in Ref. [2] (see Chaps. 3 and 6; see also Ref. [182] for an interesting parallel to glassy metastable states). Thus, for scrambling via entanglement generation, we can limit the applicability of statement (b) by noting that systems with a nonvanishing zero temperature entropy are not fast scramblers in large subsystems. At the same time, the discussion after Eq. (4.24) illustrates what the energy spectrum of a "nearly fast scrambler" may look like, showing that systems nearly saturating Eq. (4.10) do *formally* exist. An intriguing open challenge is to address the following question: What is the natural physical setting in which we can expect to find such fast scramblers in the sense of sustained entanglement generation?

Outline of the Chapter— In the remaining sections of this Chapter, we derive various results presented in this synopsis. In Sec. 4.2, we derive quantitative relations between scrambling in terms of return probabilities and in terms of entanglement entropies [supporting Eqs. (4.4) and (4.10) of Sec. 4.1]. In Sec. 4.3, we derive the inequality $P_S(t) \geq K_\beta(t)$ that formulates the energy-time uncertainty principle with sensitivity to microscopic values of $P_S(t)$ and nontriviality at arbitrarily long times (as in Ref. [2], Chap. 3), generalized to any initial state of the external system including finite temperature states [Eq. (4.13) of Sec. 4.1]. In Sec. 4.4, we derive our main result, the universal bound on the scrambling time t_s [Eq. (4.8) of Sec. 4.1] by proving the double exponential bound on analytic functions with quantitatively determined parameters [Eqs. (4.21)-(4.24) of Sec. 4.1]. Finally, in Sec. 4.5, we show that quantum mechanical systems that nearly saturate our bound on the scrambling time formally exist, by discussing explicit examples with a suitable (regularized) density of states.

4.2 Sustained scrambling and entanglement entropies

4.2.1 The mean return probability bounds the average purity of the evolving basis states

In this subsection, we show that the mean return probability $P_S(t)$ constrains the generation of entanglement, as stated in Eq. (4.4) of Sec. 4.1. We consider the average purity on the left hand side of this equation, and split it into contributions from matrix elements in the computational basis:

$$\begin{aligned}\frac{1}{D_S}\sum_{k=1}^{D_S}\mathrm{Tr}[\hat{\rho}_{k,S}^2(t)] &= \frac{1}{D_S}\sum_{k,k_1,k_2=1}^{D_S}\left|{}_S\langle k_1|\,\mathrm{Tr}_E\left[\hat{\rho}_k(t)\right]|k_2\rangle_S\right|^2 \\ &\geq \frac{1}{D_S}\sum_{k,k_1=1}^{D_S}\left|{}_S\langle k_1|\,\mathrm{Tr}_E\left[\hat{\rho}_k(t)\right]|k_1\rangle_S\right|^2 \\ &\geq \frac{1}{D_S}\sum_{k=1}^{D_S}\left|{}_S\langle k|\,\mathrm{Tr}_E\left[\hat{\rho}_k(t)\right]|k\rangle_S\right|^2 . \end{aligned} \tag{4.25}$$

We have dropped the off-diagonal matrix elements $k_1 \neq k_2$ and used the non-negativity of the corresponding terms in the second line. In the third line, we have further dropped the $k \neq k_1$ contributions, which are also non-negative, to obtain an inequality focusing on the 3D diagonal $k = k_1 = k_2$ contributions. Though a simple technique, dropping off-diagonal terms in such inequalities has been considerably useful in other physical contexts as well [2, 71] (for instance, Sec. 3.2).

We can rewrite the last expression in Eq. (4.25) in terms of the projectors $\hat{\Pi}_k = |k\rangle_S\langle k| \otimes \hat{\mathbb{1}}_E$, which gives

$$\frac{1}{D_S}\sum_{k=1}^{D_S}\mathrm{Tr}[\hat{\rho}_{k,S}^2(t)] \geq \frac{1}{D_S}\sum_{k=1}^{D_S}\mathrm{Tr}\left[\hat{\rho}_k(t)\hat{\Pi}_k\right]^2 . \tag{4.26}$$

The right hand side is the mean (over all values of k) of the squares of the return probabilities $P_k(t) \equiv \mathrm{Tr}\left[\hat{\rho}_k(t)\hat{\Pi}_k\right] \geq 0$, which must be at least the square of their mean. Thus, we obtain:

$$\frac{1}{D_S}\sum_{k=1}^{D_S}\mathrm{Tr}[\hat{\rho}_{k,S}^2(t)] \geq \frac{1}{D_S}\sum_{k=1}^{D_S}P_k^2(t) \geq \left(\frac{1}{D_S}\sum_{k=1}^{D_S}P_k(t)\right)^2 = \left[P_S(t)\right]^2 , \tag{4.27}$$

which is Eq. (4.4) in Sec. 4.1.

4.2.2 The scrambling time t_s in terms of the scrambled Rényi entanglement entropy

Here, we express the scrambling bound of Eq. (4.8) in terms of the entanglement entropy of a scrambled state, deriving Eq. (4.10) of Sec. 4.1.

The second Rényi entanglement entropy of the state $\hat{\rho}_k(t)$ in the subsystem $\mathscr{H}_S$ is defined as

$$S_2[\hat{\rho}^2_{k,S}(t)] \equiv -\ln \mathrm{Tr}\left[\hat{\rho}^2_{k,S}(t)\right]. \tag{4.28}$$

This is a direct measure of entanglement [50]. For example, volume-law entanglement corresponds to $S_2 \sim \min\{N_S \ln 2, N_E \ln 2\}$ to leading order, usually associated with maximal scrambling as we also have the general bound:

$$S_2 \leq \{N_S \ln 2, N_E \ln 2\} \tag{4.29}$$

We now introduce the scrambled entropy cutoff $S_{2,S}$ as the minimum necessary value of the entanglement entropy at which a state is considered scrambled in $\mathscr{H}_S$. In particular, we require that every scrambled state satisfies (as in Eq. (4.9) of Sec. 4.1):

$$S_2[\hat{\rho}^2_{k,S}(t > t_s)] \geq S_{2,S} \tag{4.30}$$

(where $t < T_D$ is implicit to avoid quantum recurrences [49]). It follows that the mean purity on the left hand side of Eq. (4.27) is at most $\exp[-S_{2,S}]$ for $t > t_s$. Then Eq. (4.27) [or Eq. (4.4) in Sec. 4.1] implies

$$P_S(t > t_s) \leq \exp\left(-\frac{1}{2} S_{2,S}\right), \tag{4.31}$$

corresponding to replacing $P_{\mathrm{scr}} \to \exp(-S_{2,S}/2)$ in Eq. (4.5) of Sec. 4.1. We can substitute this in the

bound on the sustained scrambling time t_s, given by Eq. (4.8) of Sec. 4.1, which yields

$$t_s \geq \frac{c\beta}{\pi} \ln\left[\frac{S_{2,S} + 4\ln(f_\beta Z_{\max})}{2\Lambda_{\text{eff}}}\right]. \tag{4.32}$$

In a thermodynamic limit with $|\ln(f_\beta Z_{\max})| = o[S_{2,S}]$ and $\Lambda_{\text{eff}} = \Theta(1)$ (requiring $Z_\beta(\tau) = O(1)$ in the interval $\tau \in [\tau_1, \tau_2]$), we get the leading order inequality

$$t_s \gtrsim \frac{c\beta}{\pi} \ln S_{2,S}, \tag{4.33}$$

which is Eq. (4.10) of Sec. 4.1. In this thermodynamic regime, we can qualitatively phrase our fast scrambling result in simple terms: the scrambling time is at least logarithmic in the scrambled (2nd Rényi) entanglement entropy. While Eqs. (4.32) and (4.33) directly capture the time required to generate a given degree of entanglement in our setup, we note two reasons for preferring the expression in terms of the scrambled return probability P_{scr} in Eq. (4.8) of Sec. 4.1.

1. Most significantly, the bound in terms of P_{scr} can be made tighter. Intuitively, this is because $P_S(t)$ can continue to track aspects of quantum dynamics and scrambling that the entanglement entropy is insensitive to; for example, $S_{2,S}$ is subject to the bound of Eq. (4.29), while P_{scr} can be taken to be as small as desired. For concreteness, consider maximal scrambling (e.g., to Haar random behavior), for which $P_{\text{scr}} \sim D_S^{-1}$ and $S_{2,S} \sim \min\{N_S, N_E\} \ln 2$. When $N_S \leq N_E$, both give identical bounds to leading order, $t_s \gtrsim (c\beta/\pi) \ln N_S$. But for $N_S > N_E$, while $-\ln P_{\text{scr}} \sim N_S$ continues to hold and constrains $t_s \gtrsim (c\beta/\pi) \ln N_S$ from Eq. (4.8) of Sec. 4.1, the leading contribution to the entropy comes from $\mathscr{H}_E$ with $S_{2,S} \sim N_E \ln 2$, due to which Eq. (4.32) sets a much weaker bound $t_s \gtrsim (c\beta/\pi) \ln N_E$. For a physical example where this can be relevant, see Ref. [2] (Chaps. 3 and 6), where scrambling is considered for subsystem sizes including $N_S > N_E$ (e.g., in the SYK-4 model).

2. $P_{\rm scr}$ requires only projective measurements in a computational basis with the D_S projectors $\hat{\Pi}_k$, and directly determines the dynamics of observables diagonal in this basis. In contrast, $S^{(2)}_{\rm scr}$ in principle probes more of the full structure of the reduced density matrices $\hat{\rho}_{k,S}$ in the Hilbert space H_S. In particular, it may be dominated by contributions (such as from the off-diagonal matrix elements of $\hat{\rho}_{k,S}$) not relevant to a specific observable of interest.

On the other hand, $S^{(2)}_{\rm scr}$ can be efficiently measured in experiments, for instance, using local randomized measurements [183]; this means that Eqs. (4.32) and (4.33) are also testable experimentally, in addition to the bound in terms of $P_S(t)$.

4.3 A quantum speed limit for arbitrary initial states $\hat{\rho}_{\beta E}$

In this section, we describe the passage from the infinite temperature speed limit of Ref. [2] [i.e., Eq. (3.5)] to the inequality $P_S(t) \geq K_\beta(t)$ for arbitrary initial states under Hamiltonian evolution with more details than in Sec. 4.1, beginning with a brief review of the former. This is essentially an expanded version of the derivation containing Eq. (4.11) through Eq. (4.14) in Sec. 4.1.

4.3.1 Derivation of $P_S(t) \geq K_\beta(t)$

Consider a general time-dependent completely positive quantum operation [50] acting on an initial reference state $\hat{\rho}_0$ in $\mathscr{H}$, with any set of M time-independent Kraus operators $\{\hat{\mathscr{K}}_r(t)\}_{r=1}^M$,

$$\hat{\rho}(t) = \sum_{r=1}^{M} \hat{\mathscr{K}}_r(t)\hat{\rho}_0\hat{\mathscr{K}}_r^\dagger(t). \tag{4.34}$$

Note that the initial state $\hat{\rho}(0)$ is not necessarily equal to the reference state $\hat{\rho}_0$, and further that the quantum operation is not required to be trace preserving [50]. Quantum dynamics of this type may

be characterized by a generalized SFF,

$$K(t) \equiv \frac{1}{D^2} \sum_{r=1}^{M} \left| \mathrm{Tr}\, \hat{\mathscr{K}}_r(t) \right|^2 \tag{4.35}$$

Given this setup, the mean return probability for specific initial reference states $\hat{\rho}_{0,k} = \hat{\Pi}_k / D_E$ (or equivalently, the states of Eq. (4.1) in Sec. 4.1 with $\hat{\rho}_{\beta E} = \hat{\rho}_{0E} \equiv \hat{\mathbb{1}}_E / D_E$),

$$P_S(t)[\hat{\Pi}_k] \equiv \frac{1}{D} \sum_{k=1}^{D_S} \mathrm{Tr}[\hat{\Pi}_k(t) \hat{\Pi}_k(0)] = \frac{1}{D} \sum_{k=1}^{D_S} \sum_{r=1}^{M} \mathrm{Tr}[\hat{\mathscr{K}}_r(t) \hat{\Pi}_k(0) \hat{\mathscr{K}}_r^\dagger(t) \hat{\Pi}_k(0)]. \tag{4.36}$$

is constrained by [2] (from Eq. (3.5)):

$$P_S(t)[\hat{\Pi}_k] \geq K(t), \tag{4.37}$$

for *any* complete, orthonormal choice of projectors $\hat{\Pi}_k$.

For the above speed limit, it is crucial that the initial reference states $\hat{\Pi}_k / D_E$ form a complete, orthogonal set for the full Hilbert space $\mathscr{H}$. It is this completeness that allows constraining the basis-dependent $P_S(t)$ with the basis-independent $K(t)$ that involves a trace over the entire space $\mathscr{H}$. However, the initial states $\hat{\rho}_k(0) = |k\rangle_S \langle k| \otimes \hat{\rho}_{\beta E}$ of Eq. (4.1) in Sec. 4.1 are complete only in $\mathscr{H}_S$, and generally not in $\mathscr{H}_E$ except in the specific case of the infinite temperature state $\hat{\rho}_{0E}$.

The resolution to this difficulty comes from writing the initial state $\hat{\rho}_k(0)$ in terms of nontrivial Kraus operators $\hat{\mathscr{K}}_r(0)$ acting on the initial reference state $\hat{\rho}_{0,k} = \hat{\Pi}_k / D_E$. This is enabled by noting that $\hat{\rho}_{\beta E}$ is a positive operator (has non-negative eigenvalues) [50], and therefore admits a positive Hermitian square root $\hat{\rho}_{\beta E}^{1/2}$ (with non-negative eigenvalues). In particular, we have

$$\hat{\rho}_k(0) = \hat{\rho}_{\beta E}^{1/2} \left(|k\rangle_S \langle k| \otimes \hat{\mathbb{1}}_E \right) \hat{\rho}_{\beta E}^{1/2} = \left[D_E^{1/2} \hat{\rho}_{\beta E}^{1/2} \right] \hat{\rho}_{0,k} \left[D_E^{1/2} \hat{\rho}_{\beta E}^{1/2} \right]^\dagger. \tag{4.38}$$

Now, we can set $M = 1$ with $\hat{\mathscr{K}}_1(0) = D_E^{1/2}\hat{\rho}_{\beta E}^{1/2}$, i.e., a single nonvanishing Kraus operator. As subsequent time evolution is generated by a Hamiltonian $\hat{H}$, we have the time-dependent Kraus operators:

$$\hat{\mathscr{K}}_r(t) = e^{-i\hat{H}t}\hat{\mathscr{K}}_r(0) = D_E^{1/2}e^{-i\hat{H}t}\hat{\rho}_{\beta E}^{1/2}\delta_{r,1}. \tag{4.39}$$

Substituting these Kraus operators in Eq. (4.36), we obtain precisely the mean return probability in Eq. (4.3) of Sec. 4.1:

$$P_S(t)[\hat{\Pi}_k] = \frac{1}{D}\sum_{k=1}^{D_S} D_E \,\mathrm{Tr}[e^{-i\hat{H}t}\hat{\rho}_{\beta E}^{1/2}\hat{\Pi}_k\hat{\rho}_{\beta E}^{1/2}e^{i\hat{H}t}\hat{\Pi}_k] = P_S(t). \tag{4.40}$$

Further, the SFF for these Kraus operators is

$$K(t) = \frac{1}{D^2}\left|\mathrm{Tr}\,\hat{\mathscr{K}}_1(t)\right|^2 = \frac{D_E}{D^2}\left|\mathrm{Tr}\left(e^{-i\hat{H}t}\hat{\rho}_{\beta E}^{1/2}\right)\right|^2 \equiv K_\beta(t). \tag{4.41}$$

Now, Eq. (4.37) for these Kraus operators gives $P_S(t) \geq K_\beta(t)$, which is Eq. (4.13) in Sec. 4.1.

4.3.2 Criteria for nontriviality

We should emphasize that, unlike the $\beta = 0$ case where $P_S(0) = K(0) = 1$ under Hamiltonian dynamics, for $\beta \neq 0$ their values at $t = 0$ can be different. In particular, $P_S(0) = 1$, while $K_\beta(0) = f_\beta^2 \leq 1$ [see Eqs. (4.6) and (4.18) of Sec. 4.1]. Further, as $K_\beta(t) \leq K_\beta(0)$, we can only obtain a nontrivial bound on the scrambling time from $P_S(t) \geq K_\beta(t)$ if

$$f_\beta^2 > P_{\mathrm{scr}}. \tag{4.42}$$

In this subsection, we ask under what conditions a system may satisfy Eq. (4.42). These considerations are equivalent to, and expand on, the discussion following Eq. (4.18) of Sec. 4.1. The need to

consider such criteria is in contrast to the $\beta = 0$ case of Ref. [2] (Chap. 3), where as long as $P_{\text{scr}} < 1$, some nontrivial > 0 bound on t_s always exists.

However, the tradeoff in our case still admits two kinds of universality:

1. For a given initial state with a certain f_β^2, we get nontrivial values when $P_{\text{scr}} < f_\beta^2$. This can translate to a restriction on the system size of $\mathscr{H}_S$. For example, if we are interested in maximal scrambling with $P_{\text{scr}} \sim D_S^{-1}$, the number of qubits N_S in the subsystem $\mathscr{H}_S$ must be as large as

$$N_S > 2 \log_2 \frac{1}{f_\beta}, \tag{4.43}$$

for a nontrivial scrambling time. An interesting special case is when the initial state is a pure state, $\rho_{\beta E} = |\ell\rangle_E\langle\ell|$, which has $f_\beta = D_E^{-1/2}$. Then, we get $N_S > N_E$ for a nontrivial bound for pure states. For any state that is not entirely pure, it follows that we get nontrivial scrambling time bounds even for some values of $N_S \leq N_E$.

2. For a given scrambling value $P_{\text{scr}} < 1$, the condition $f_\beta^2 > P_{\text{scr}}$ restricts the entanglement of the initial state $\hat{\rho}_{\beta E}$. For example, this is because the 1/2-order Rényi entanglement entropy [63] of this state is determined by the fidelity:

$$S_{1/2,E} \equiv 2 \ln \left[D_E^{1/2} f_\beta \right]. \tag{4.44}$$

Thus, for a nontrivial bound, $S_{1/2,E} > \ln \left[D_E P_{\text{scr}} \right]$, implying that a smaller value of P_{scr} requires less entanglement in the initial state for our bound to be useful.

4.4 Derivation of a universal bound on the scrambling time

In this section, we derive a quantitative refinement of a known asymptotic double exponential bound on bounded analytic functions which is originally given in terms of finite but undetermined constants [173–176]. This refinement provides explicit values for the undetermined constants in the previous versions. As mentioned in Sec. 4.1, this allows us to constrain the scrambling time even for finite dimensional systems and obtain concrete numerical bounds for t_s that can potentially be tested in experiments, instead of just an asymptotic estimate in a thermodynamic limit. This fills in the technical details for the discussion around Eqs. (4.21) through (4.24) of Sec. 4.1.

Sec. 4.4.1 states our quantitative version of the double-exponential bound in Theorem 4.1, and derives the scrambling time as per Eqs. (4.8) and (4.24) of Sec. 4.1. Sec. 4.4.2 proves the double exponential bound by conformally mapping the strip to a half-plane, based on two lemmas on analytic functions on the half-plane, closely following the methods of Ref. [176].

As a prelude to our technical discussion, we provide some local power series intuition for why analyticity should constrain the decay of a bounded function. For simplicity, let an analytic $F(t)$ have a saddle point at $F(t=0)=1$. On account of analyticity, we can write a power series expansion in a neighborhood of this point:

$$F(t-i\tau) = 1 - \frac{1}{2}c_2(t-i\tau)^2 + O((t-i\tau)^3), \tag{4.45}$$

where $c_2 = -F''(0)$. We take $0 < c_2 \in \mathbb{R}$, so that $F(i\tau) \in \mathbb{R}$ [to $O(\tau^2)$] grows as $1 + c_2\tau^2/2$ away from $t = 0$ while $F(t) \in \mathbb{R}$ [to $O(t^2)$] decays as $1 - c_2 t^2/2$; in addition, $F(t)$ then locally satisfies $|F(t-i\tau)| < F_R(-i\tau)$ to $O((t-i\tau)^2)$ [like in Eq. (4.19) of Sec. 4.1] so it is sufficient to consider $F(-i\tau)$ for boundedness. The key restriction from analyticity here is that the same coefficient c_2

determines both the growth and decay. Thus, if $F(-i\tau) < F_{\max}$ in $\tau \in (0, \tau_2)$, we have

$$F(t) \geq 1 - (F_{\max} - 1)\left(\frac{t}{\tau_2}\right)^2 + O(t^3). \tag{4.46}$$

In other words, the decay rate is bounded by $F_{\max}$. This qualitatively connects analyticity and boundedness to (local) decay rates in an intuitive example, but we emphasize that the mathematical details of our (global) double exponential bound [Eq. (4.48), or Eq. (4.21) in Sec. 4.1] and method of proof are quite different. Mathematically inclined readers may find closer intuition to our result through Jensen's formula from complex analysis or the Poisson kernel formula for harmonic functions [177].

4.4.1 The scrambling time from a double-exponential bound on analytic functions

4.4.1.1 A quantitative theorem: analytic functions are bounded by a double-exponential in time

The key to deriving our quantitative bound on the scrambling time is the following theorem bounding the decay of analytic functions, which quantifies the undetermined parameters in the previously known double-exponential bounds of Refs. [173–176] (the latter being stated in Eq. (4.21) of Sec. 4.1):

Theorem 4.1 (**Quantitative decay rate of an analytic function on a strip**). *Let the function $F(\mathbf{t})$ be analytic with $|F(\mathbf{t})| < 1$ in the open strip $\{\mathbf{t} = t - i\tau \in \mathbb{C}:\ \tau \in (\tau_1, \tau_2)\}$, and continuous in the corresponding closed strip $\tau \in [\tau_1, \tau_2]$. We also require that $F(\mathbf{t})$ is not identically 0 in the strip. Then for any $0 < \ell < \infty$, there is an "exceptional set" $W_{ex} \subset [0, \infty)$ of times t whose length is at most*

$$\int_{W_{ex}} \mathrm{d}t \leq \frac{(\tau_2 - \tau_1)}{\pi}\ell, \tag{4.47}$$

such that $|F(\mathbf{t})|$ *is at least a double exponential in* t *everywhere in the strip except when* t *is in the exceptional set, i.e.,*

$$|F(t-i\tau)| \geq \exp\left[-\frac{\Lambda_\ell}{2}\exp\left(\frac{\pi t}{\tau_2-\tau_1}\right)\right],$$
$$\text{for all } t \in [0,\infty)\setminus W_{ex}, \text{ and } \tau \in [\tau_1,\tau_2]. \tag{4.48}$$

Here, the parameter $\Lambda_\ell \geq 0$ *is given in terms of* ℓ *and* $F(-i\tau)$ *by* $\Lambda_\ell = 2(\pi^2+8)(2+\ell)\Gamma/\ell$*, in which*

$$\Gamma = \inf_{\tau\in(\tau_1,\tau_2)}\left\{\sec\left[\frac{\pi}{\tau_2-\tau_1}\left(\tau-\frac{\tau_2+\tau_1}{2}\right)\right]\ln\frac{1}{|F(-i\tau)|}\right\}. \tag{4.49}$$

Proof. See Sec. 4.4.2. The proof closely follows the techniques used to prove Theorem 7.32 in Ref. [176], which is related to the present theorem, but unlike Eq. (4.48) does not quantitatively determine the constants in Eq. (4.21) of Sec. 4.1. □

Note that if we set $\ell = 0$, the expression in the exponent of Eq. (4.48) diverges, so we can never guarantee that the exceptional set vanishes for the double exponential form in this equation. However, we can make it as small as desired.

4.4.1.2 Applying the bound to $\widetilde{\mathcal{N}_\beta}(t)$

Here, we derive Eqs. (4.21) through (4.23) of Sec. 4.1 from Theorem 4.1. As noted in Eq. (4.15) of Sec. 4.1, $\widetilde{\mathcal{N}_\beta}(\mathbf{t})$ is a weighted sum of exponentials with non-negative weights,

$$\widetilde{\mathcal{N}_\beta}(t-i\tau) = \frac{D_E^{1/2}}{f_\beta D}\sum_{n=1}^{D}\langle E_n|\hat{\rho}_{\beta E}^{1/2}|E_n\rangle e^{-iE_n t-E_n\tau}, \tag{4.50}$$

and is consequently analytic and bounded on any strip $t \in \mathbb{R}$, $\tau \in (\tau_1,\tau_2)$, as well as continuous up to the boundary of the strip. As we want to focus on evolution in real time, we take $\tau = 0$ to lie in the

strip, $\tau_1 \leq 0 \leq \tau_2$. Further, $\widetilde{\mathcal{N}_\beta}(0) = 1$ making $\widetilde{\mathcal{N}_\beta}(\mathbf{t})$ not identically 0. Thus, $\widetilde{\mathcal{N}_\beta}(\mathbf{t})$ already satisfies most of the properties required by Theorem 4.1 for $F(\mathbf{t})$, except not necessarily that $|F(\mathbf{t})| < 1$ in the interior of the strip.

To normalize $\widetilde{\mathcal{N}_\beta}(\mathbf{t})$ to satisfy this last property, we consider the maximum of $\widetilde{\mathcal{N}_\beta}(t)$ in the closed strip $\tau \in [\tau_1, \tau_2]$. From Eq. (4.50) we have $Z_\beta(\tau) \equiv \widetilde{\mathcal{N}_\beta}(-i\tau) \in [0, \infty)$ and $|\widetilde{\mathcal{N}_\beta}(t - i\tau)| \leq Z_\beta(\tau)$ [Eq. (4.19) of Sec. 4.1], which amount to the "ridge property" [175] of Fourier transforms of probability distributions — implying that the maximum of $\widetilde{\mathcal{N}_\beta}(\mathbf{t})$ occurs on the imaginary axis $t = 0$.

Further, $Z_\beta(\tau)$ is also a convex function of τ (see, e.g., Ref. [175]); in our case, this is because

$$\frac{\mathrm{d}^2 Z_\beta}{\mathrm{d}\tau^2}(\tau) = \frac{D_E^{1/2}}{f_\beta D} \sum_n E_n^2 \langle E_n | \hat{\rho}_{\beta E}^{1/2} | E_n \rangle e^{-E_n \tau} \geq 0. \tag{4.51}$$

Here, there are two distinct cases of interest:

1. All energy levels with $\langle E_m | \hat{\rho}_{\beta E}^{1/2} | E_m \rangle \neq 0$ have $E_m = 0$. In this (trivial) case, $\widetilde{\mathcal{N}_\beta}(\mathbf{t}) = 1$ everywhere, automatically satisfying the double exponential bound [Eq. (4.21) of Sec. 4.1] for any choice of decay rate $\Lambda \geq 0$.

2. Otherwise, the partition function is strictly convex, $\mathrm{d}^2 Z_\beta(\tau)/\mathrm{d}\tau^2 > 0$ in $[\tau_1, \tau_2]$ (as all terms in Eq. (4.51) are nonnegative, and not all of them vanish). In this case, the maximum

$$Z_{\max} \equiv \max_{\tau \in [\tau_1, \tau_2]} Z_\beta(\tau) \tag{4.52}$$

is attained exclusively on either $\tau = \tau_1$ or $\tau = \tau_2$, with $Z_\beta(\tau) < Z_{\max}$ in $\tau \in (\tau_1, \tau_2)$. Then, we can define

$$F(\mathbf{t}) \equiv \frac{\widetilde{\mathcal{N}_\beta}(\mathbf{t})}{Z_{\max}}, \text{ satisfying } |F(t - i\tau)| < 1 \text{ in } \tau \in (\tau_1, \tau_2). \tag{4.53}$$

Theorem 4.1 now applies to this $F(\mathbf{t})$, and we obtain the quantitative double exponential bound,

$$|\widetilde{\mathscr{N}_\beta}(t)| \geq Z_{\max} \exp\left[-\frac{\Lambda_\ell}{2} \exp\left(\frac{\pi t}{\tau_2 - \tau_1}\right)\right], \text{ for } t \geq 0 : t \notin W_{\text{ex}}. \tag{4.54}$$

corresponding to Eqs. (4.21), (4.22) and (4.23) of Sec. 4.1 (respectively from Eqs. (4.48), (4.47) and (4.49) of the Theorem, with Eq. (4.23) obtained by extending the right hand side of Eq. (4.49) to be continuous at τ_1, τ_2 to replace the infimum with a minimum over $[\tau_1, \tau_2]$).

A key point to emphasize is that these results specify the previously undetermined constants such as Λ_ℓ entirely in terms of the behavior of $\widetilde{\mathscr{N}_\beta}(t - i\tau)$ at the initial (real) time $t = 0$, or equivalently, the partition function $Z_\beta(\tau) = \widetilde{\mathscr{N}_\beta}(-i\tau)$. This is due to taking the physical viewpoint that we should not expect to know aspects of the late-time dynamics of the system [e.g., $\widetilde{\mathscr{N}_\beta}(t > 0)$] when constraining this very dynamics. From a purely mathematical standpoint, we can potentially make Γ smaller (i.e., obtain a tighter bound) in some cases if we already know some aspects of the $t > 0$ behavior (by extending Lemma 4.3 in Sec. 4.4.2 to $|z| \geq 1$; see also [8]).

Finally, to provide some physical intuition for the decay rate Γ, we can carry out the minimization in Eq. (4.49) by finding a stationary point τ_0 (if it exists). Differentiation yields the implicit equation:

$$\frac{\pi}{\tau_2 - \tau_1} \tan\left[\frac{\pi}{\tau_2 - \tau_1}\left(\tau_0 - \frac{\tau_2 + \tau_1}{2}\right)\right] \ln \frac{Z_{\max}}{Z_\beta(\tau_0)} = \left.\frac{\mathrm{d} \ln Z_\beta(\tau)}{\mathrm{d}\tau}\right|_{\tau = \tau_0}. \tag{4.55}$$

If the stationary point does not exist, the minimum occurs at $\tau_{\max} \in \{\tau_1, \tau_2\}$ where $Z_\beta(\tau_{\max}) = Z_{\max}$. Whether τ_0 exists or not, the value of Γ is in all cases conveniently expressed, using familiar thermodynamic relations [65, 184], in terms of the expectation value of the energy $\mathscr{E}_\beta(\tau) \equiv -\mathrm{d} \ln Z_\beta(\tau)/\mathrm{d}\tau$ and the excess free energy $\Delta\mathscr{F}_\beta(\tau) \equiv -\tau^{-1}[\ln Z_\beta(\tau) - \ln Z_{\max}]$ at Euclidean time τ [with $\Delta\mathscr{F}_\beta(\tau_{\max}) = 0$].

We get:

$$\Gamma = \sqrt{\tau_0^2[\Delta\mathscr{F}_\beta(\tau_0)]^2 + \frac{(\tau_2-\tau_1)^2}{\pi^2}[\mathscr{E}_\beta(\tau_0)]^2}. \tag{4.56}$$

This expresses the decay rate in terms of thermodynamic quantities. The location of the point τ_0 and the significance of the (free) energy at this point are system-dependent, and it appears that we cannot obtain further universal insights from this expression.

4.4.1.3 Derivation of the scrambling time

We can obtain a bound on the scrambling time t_s by using Eqs. (4.48) and (4.53) in Eq. (4.18) of Sec. 4.1, which states that (implicitly with $t < T_D$, the time scale of quantum recurrences [49])

$$\left|\widetilde{\mathscr{N}_\beta}(t > t_s)\right|^2 \leq \frac{P_{\text{scr}}}{f_\beta^2}, \tag{4.57}$$

However, we should carefully account for the exceptional set W_{ex}. Any part of the exceptional set in the region $t \geq t_s$ does not affect the scrambling time, as it refers to when $|\widetilde{\mathscr{N}_\beta}(t)|$ itself already satisfies our scrambling criterion and whether it violates Eq. (4.48) becomes immaterial. However, the length of the portion of the exceptional set before t_s is uncertain: it can be 0 at minimum, or as much as the maximum allowed length $(\tau_2-\tau_1)\ell/\pi$ of the exceptional set. So we can only constrain the scrambling time to within this maximum possible length. This corresponds to requiring $|\widetilde{\mathscr{N}_\beta}(t_s+(\tau_2-\tau_1)\ell/\pi)|^2$ to be less than the scrambling value. Assuming $0 \in [\tau_1, \tau_2]$ and using Eq. (4.48) with (4.53), we get the inequality

$$Z_{\max}\exp\left[-\frac{(2+\ell)}{\ell}(\pi^2+8)\Gamma\exp\left(\frac{\pi t_s}{\tau_2-\tau_1}+\ell\right)\right] \leq \frac{P_{\text{scr}}^{1/2}}{f_\beta}. \tag{4.58}$$

Rearranging this equation, we obtain

$$t_s \geq \frac{\tau_2-\tau_1}{\pi}\left\{\ln\left[\frac{1}{2}\ln\left(\frac{f_\beta^2 Z_{\max}^2}{P_{\text{scr}}}\right)\right] - \ln\left[\frac{(2+\ell)}{\ell}(\pi^2+8)\Gamma\right] - \ell\right\}. \tag{4.59}$$

This leads to Eq. (4.24) of Sec. 4.1. Since the choice of ℓ is arbitrary, it is best to choose a value that maximizes the bound for t_s (though this is not crucial in a thermodynamic limit, where this term is usually subleading). The corresponding maximization of $\ln[\ell/(2+\ell)]-\ell$ over $\ell > 0$ gives

$$\ell = \sqrt{3}-1. \tag{4.60}$$

Using this value of ℓ in Eq. (4.59) yields (4.8) of Sec. 4.1.

4.4.2 Proof of Theorem 4.1 on the quantitative decay of analytic functions

4.4.2.1 Conformal map from a strip to a half-plane

In order to prove Theorem 4.1, it will be convenient to work on the right *half-plane* instead of the strip, since several formulas from complex analysis are simpler on the half-plane. There is a conformal map between the two regions shown in Fig. 4.1 and defined below, which can translate results on the right half-plane to results on the strip $\tau_1 < -\mathrm{Im}\,\mathbf{t} < \tau_2$.

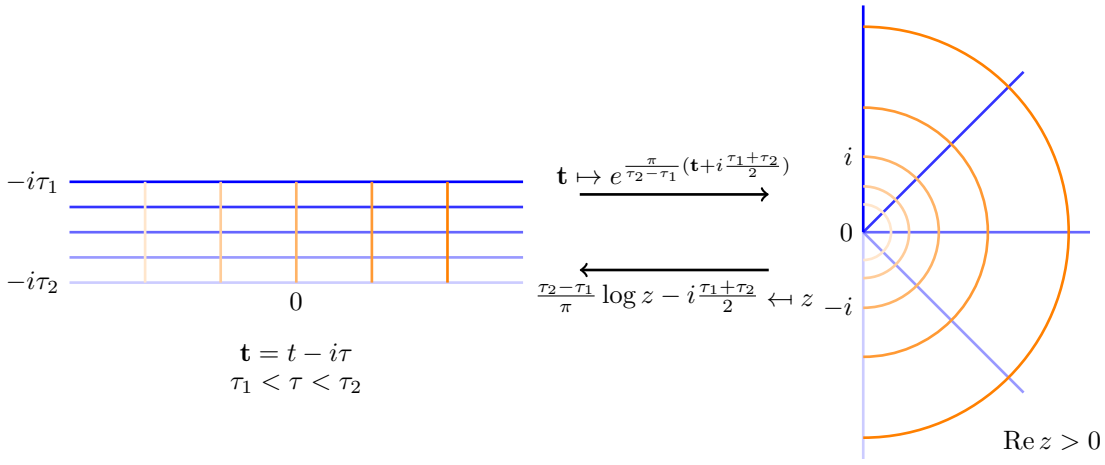


Figure 4.1: Conformal map between the strip $\{\mathbf{t} = t - i\tau : \tau_1 < \tau < \tau_2\}$ and right half-plane $\mathrm{Re}\, z > 0$, with the boundaries included. Horizontal lines in the strip are mapped to radial rays $re^{i\theta}$, $r > 0$, in the half-plane.

Specifically, we define the variables ($z \in \mathbb{C}$; $r, \theta, x, y \in \mathbb{R}$) via the transformation

$$z = re^{i\theta} = x + iy = \exp\left[\pi \frac{\mathbf{t} + i\frac{\tau_2+\tau_1}{2}}{\tau_2 - \tau_1}\right]. \tag{4.61}$$

Then for $(-\mathrm{Im}\,\mathbf{t}) \in [\tau_1, \tau_2]$,

$$r = \exp\left[\frac{\pi\,\mathrm{Re}\,\mathbf{t}}{\tau_2 - \tau_1}\right] \in (0, \infty), \qquad \theta = \pi \frac{\mathrm{Im}\,\mathbf{t} + \frac{\tau_2+\tau_1}{2}}{\tau_2 - \tau_1} \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \tag{4.62}$$

which is seen to correspond to the closed half-plane $\{z = x + iy : x \geq 0, y \in \mathbb{R}\}$ (with $z = 0$ corresponding to values $-\infty - i\tau$ in the strip). We will freely switch between these different coordinates for the right half-plane in what follows.

4.4.2.2 Constraint on regions of double-exponential or faster decay in t

On account of Eq. (4.62), a double exponential decay in t corresponds to an exponential decay in r. Our goal is now to constrain the size of regions $r \in V_\lambda$ in which an analytic function $F(z)$ may decay faster than the exponential rate $\exp(-\lambda r)$. This is accomplished using the following Lemma, which is an immediate consequence of Theorem 7.27 in Ref. [176]:

Lemma 4.2 (**Constraining regions of exponential or faster decay**). *Let $F(z)$ satisfying $|F(z)| < 1$ be an analytic function that is not identically 0 in the open half-plane $x > 0$. For any $\lambda > 0$, let $V_\lambda \subset (0, \infty)$ be a measurable set such that*

$$\inf_{\theta \in (-\frac{\pi}{2}, \frac{\pi}{2})} |F(re^{i\theta})| \leq e^{-\lambda r}, \quad \textit{for all } r \in V_\lambda. \tag{4.63}$$

The set V_λ consists of radii r in which $|F(re^{i\theta})|$ is at most $e^{-\lambda r}$ for some value of θ in the right-half-plane. Then, if we define the logarithmic length of V_λ in r (corresponding in the strip to its actual length in t, up

to constants) by

$$L[V_\lambda] \equiv \int_{V_\lambda} \frac{\mathrm{d}r}{r}, \tag{4.64}$$

the following bound is satisfied everywhere on the half-plane $x > 0$:

$$\ln \frac{1}{|F(z)|} \geq \frac{4}{(2 + L[V_\lambda])(\pi^2 + 8)} \int_{V_\lambda} \mathrm{d}r' \frac{x}{x^2 + (r' + |y|)^2} \lambda r'. \tag{4.65}$$

Proof. This follows from setting $u(z) = \ln|F(z)|$ and $f(r) = \lambda r$ in Theorem 7.27 of Ref. [176] (see also Theorem 7.32 of Ref. [176]). □

To provide some relation to more familiar expressions, we note that the integral on the right hand side is related to electrostatics (i.e., solutions of the Laplace equation) on the half-plane [173, 176]. In particular, the Poisson kernel formula [177, 185] gives the following solution to Laplace's equation for the electrostatic potential:

$$\phi(x, y) = \frac{1}{\pi} \int_{-\infty}^{\infty} \mathrm{d}y' \frac{x}{x^2 + (y' - y)^2} \phi(0, y'), \tag{4.66}$$

which we will use with the boundary condition $\phi(0, y) = \lambda y$ for $y \in V_\lambda$ and 0 otherwise. With this context, Eq. (4.65) can be expressed for $y \leq 0$ as follows [176]:

$$\ln \frac{1}{|F(x + iy)|} \geq \frac{4\pi}{(2 + L[V_\lambda])(\pi^2 + 8)} \phi(x, y). \tag{4.67}$$

The right hand side should be considered [173] an indirect measure of the size of the set V_λ.

We make two quick remarks concerning the mathematical context of the above lemma. First, in Ref. [176], the result is stated more generally for subharmonic functions u, which includes $u(z) = \ln|F(z)|$ for analytic $F(z)$ as a special case. The subharmonicity suggests why one might expect to

see a Poisson kernel-like formula in Eq. (4.65). Second, the actual Poisson kernel formula (see e.g. Ref. [177, §IIIG.2]) for $\ln|F(z)|$ can be used in place of Lemma 4.2 to obtain quantitative decay rate estimates like in Theorem 4.1, but only on the *boundary* of the strip. In [8], we explore the utility of these Poisson kernel boundary estimates in other situations. However, in the present setting, since we will often work in the *interior* of a strip, using Lemma 4.2, which provides decay rate estimates on the entirety of the strip, will produce better t_s bounds whenever $\tau_1, \tau_2 \neq 0$. This is because larger strips, which are not restricted to have a boundary at $\tau = 0$, will increase the prefactor $c\beta$ measuring the strip width in our bound on t_s; in the appropriate thermodynamic limit [Eq. (4.10) of Sec. 4.1 or (4.33) here], this is the *only* parameter available to improve the tightness of the bound [provided that $Z_\beta(\tau) = O(1)$ continues to hold in the larger strip].

4.4.2.3 Direct constraint on the decay parameter λ

Lemma 4.2 can be used to derive the following Lemma (closely related to Theorem 7.32 in Ref. [176]):

Lemma 4.3 (**Constraining the decay parameter**). *Let $F(z)$ be as in Lemma 4.2. For any $\lambda > 0$ and $R > 1$, let $V_\lambda \subseteq [1, R]$ be a set of exponential-or-faster decay satisfying Eq. (4.63), and therefore subject to Lemma 4.2. Then the logarithmic length $L[V_\lambda] \leq \ln R < \infty$ of V_λ is related to the decay rate λ by*

$$\lambda \leq \frac{(2 + L[V_\lambda])}{L[V_\lambda]}(\pi^2 + 8)\left\{\frac{1}{\cos\theta}\ln\frac{1}{|F(e^{i\theta})|}\right\}, \tag{4.68}$$

where $z = e^{i\theta} : \theta \in (-\pi/2, \pi/2)$ is any point on the unit semi-circle in the right half-plane (which corresponds to the vertical segment $t = 0$ in the original strip).

Proof. Our proof will closely follow that of Theorem 7.32 in Ref. [176]. In Eq. (4.65), we set $z = e^{i\theta}$,

and specifically consider only these points in the integral:

$$\int_{V_\lambda} \mathrm{d}r' \frac{x}{x^2 + (r' + |y|)^2} \lambda r'. \tag{4.69}$$

In this integral, $|x|, |y| \leq 1$ due to the restriction to $|z| = 1$, and further $r' \geq 1$ as $V_\lambda \subseteq [1, \infty)$ by assumption; this implies $|z|, |y| \leq r'$. The denominator thus satisfies the inequality

$$x^2 + (r' + |y|)^2 = |z|^2 + (r')^2 + 2(r')|y| \leq 4(r')^2. \tag{4.70}$$

Using this inequality in Eq. (4.65) with $x = \cos\theta$ gives

$$\ln \frac{1}{|F(e^{i\theta})|} \geq \frac{\lambda \cos\theta}{(2 + L[V_\lambda])(\pi^2 + 8)} \int_{V_\lambda} \frac{\mathrm{d}r'}{r'}. \tag{4.71}$$

Rearranging factors and noting that the integral on the right hand side equals $L[V_\lambda]$ by Eq. (4.64), we get Eq. (4.68). □

The above result differs from Theorem 7.32 of Ref. [176], which determines the limit of $r^{-1} \ln|F(re^{i\theta})|$ as $r \to \infty$, outside an exceptional set of some (unspecified) finite logarithmic length. However, Eq. (4.68) provides an explicit quantitative constraint on the logarithmic length of the exceptional set V_λ pertaining to a specific λ. As an aside, we note that Theorem 7.32 of Ref. [176] has connections with the Phragmén-Lindelöf principle [173, 174, 176] (which generalizes the maximum modulus principle to e.g. unbounded domains). This principle also plays a role in the conjectured bound on chaos in out-of-time-ordered correlators (OTOCs) [159], albeit in a quantitatively very different form.

4.4.2.4 Completing the proof

Only a few quick observations remain to obtain Theorem 4.1 from Eq. (4.68). We note that the function $(2+\ell)/\ell$ is monotonically decreasing with ℓ. Further, as $F(\mathbf{t})$ is analytic and not identically zero in the open strip by assumption, the identity theorem [186, 187] implies that $F(e^{i\theta})$ is not identically zero in $\theta \in (-\pi/2, \pi/2)$. Thus, for any given $0 < \ell < \infty$, if we choose

$$\lambda = \frac{(2+\ell)}{\ell}(\pi^2+8)\left\{\frac{1}{\cos\theta}\ln\frac{1}{|F(e^{i\theta})|}\right\} \tag{4.72}$$

for any $\theta \in (-\pi/2, \pi/2)$ such that $F(e^{i\theta}) \neq 0$, then it follows from Eq. (4.68) that

$$\frac{2+\ell}{\ell} \leq \frac{(2+L[V_\lambda])}{L[V_\lambda]}, \quad \text{which implies } L[V_\lambda] \leq \ell. \tag{4.73}$$

Eq. (4.73) holds for any value of R, so we can take $R \to \infty$ in Lemma 4.3. This implies that given any $\ell > 0$, by choosing λ according to Eq. (4.72), we can guarantee that any set $V_\lambda \in [1, \infty)$ in which $|F(re^{i\theta})| \leq \exp(-\lambda r)$ for some θ must have a logarithmic length no greater than ℓ.

Given that the choice of θ in Eq. (4.72) is arbitrary, we will get a smaller value of λ, and therefore a tighter exponential in Eq. (4.63), by minimizing the right hand side of Eq. (4.72) with respect to θ, while Eq. (4.73) continues to hold. At this stage, transforming back to the strip $\mathbf{t} = t - i\tau$ gives (as the unit semi-circle corresponds to the vertical $t = 0$ segment in the strip)

$$\lambda = \frac{(2+\ell)}{\ell}(\pi^2+8)\inf_{\tau\in(\tau_1,\tau_2)}\left\{\sec\left[\frac{\pi}{\tau_2-\tau_1}\left(\tau-\frac{\tau_2+\tau_1}{2}\right)\right]\ln\frac{1}{|F(-i\tau)|}\right\}, \tag{4.74}$$

as an appropriate choice of λ so that any set $W_{\text{ex}} \in [0, \infty)$ (the transformed version of V_λ) such that

$$\inf_{\tau \in (\tau_1, \tau_2)} |F(t - i\tau)| \leq \exp\left[-\lambda \exp\left(\frac{\pi t}{\tau_2 - \tau_1}\right)\right], \quad \text{for all } t \in W_{\text{ex}}, \tag{4.75}$$

must have a length of at most

$$\int_{W_{\text{ex}}} \mathrm{d}t \leq \frac{(\tau_2 - \tau_1)}{\pi} \ell. \tag{4.76}$$

Finally, as $F(\mathbf{t})$ is continuous up to the boundary of the strip, Eq. (4.75) can also be extended to $\tau = \tau_1, \tau_2$. This completes the proof of Theorem 4.1. □

4.5 Formal examples of nearly fast scramblers

In this section, we construct formal examples of systems that nearly saturate our bound $t_s \gtrsim (c\beta/\pi) \ln S_{2,S}$ on fast scrambling in the thermodynamic limit. Specifically, we construct the *regularized* density of states of these systems at an assumed inverse temperature β, assuming that a suitable initial state $\hat{\rho}_{\beta E}$ exists, and show that with sufficiently random eigenstates $t_s \lesssim (2c\beta/\pi) \ln S_{2,S}$ for these systems. This corresponds to the discussion of "nearly fast scramblers" after Eq. (4.24) in Sec. 4.1.

4.5.1 Regularized density of states and its Fourier transform

In this section, we construct simple examples of continuum limits of $\widetilde{\mathcal{N}_\beta}(t)$ that attain a double exponential decay rate for $t \in \mathbb{R}$, subject to the requirement that its inverse Fourier transform (the regularized density of states) satisfies $\mathcal{N}_\beta(E) \geq 0$ and $\int_{-\infty}^{\infty} \mathcal{N}_\beta(E)\, dE = 1$. (See also [178], [175, p.35, Appendix II] for a different class of examples.) These examples demonstrate that the double exponential decay rate in Eq. (4.21) of Sec. 4.1 is essentially optimal (up to a constant in the exponent), even with the constraint that $\mathcal{N}_\beta(E) \geq 0$. In [8], we also formally verify that this leads to similar scrambling time bounds for corresponding finite dimensional systems.

First consider the analytic function

$$g(\mathbf{t}) = \exp(-e^{\frac{\pi}{c\beta}\mathbf{t}})\exp(-e^{-\frac{\pi}{c\beta}\mathbf{t}}), \tag{4.77}$$

for $\mathbf{t} = t - i\tau \in \mathbb{C}$ and a fixed $c\beta > 0$, which has absolute value

$$|g(\mathbf{t})| = \exp\left(-e^{\frac{\pi t}{c\beta}}\cos\left(\frac{\pi\tau}{c\beta}\right)\right)\exp\left(-e^{-\frac{\pi t}{c\beta}}\cos\left(\frac{\pi\tau}{c\beta}\right)\right). \tag{4.78}$$

The function g is thus bounded in the strip $|\operatorname{Im}\mathbf{t}| \le c\beta/2$, but does not meet the requirement that its inverse Fourier transform is nonnegative. (One can either check this numerically, or note that g is everywhere analytic, but not maximized at $t = 0$ when $\tau = c\beta$, and so even its normalized version cannot be the Fourier transform of a nonnegative density [175, §II.3].)

However, recall the convolution of two functions, $g_1 * g_2(t) \equiv \int_{-\infty}^{\infty} g_1(y)g_2(t-y)\,dy$, behaves with the inverse Fourier transform $g^{\vee}(E) \equiv \frac{1}{2\pi}\int_{-\infty}^{\infty} g(t)e^{itE}\,dt$ as

$$(g_1 * g_2)^{\vee}(E) = (2\pi)g_1^{\vee}(E)g_2^{\vee}(E). \tag{4.79}$$

Therefore the normalized self-convolution

$$\widetilde{\mathscr{N}_\beta}(t) = \frac{g * g(t)}{\|g\|_2^2}, \text{ where } \|g\|_2^2 = \int_{-\infty}^{\infty} g(t)^2\,dt, \tag{4.80}$$

has nonnegative inverse Fourier transform

$$\mathscr{N}_\beta(E) = \frac{2\pi}{\|g\|_2^2}(g^{\vee}(E))^2 \ge 0, \tag{4.81}$$

which is also properly normalized since $\int_{-\infty}^{\infty}\mathscr{N}_\beta(x)\,dx = \widetilde{\mathscr{N}_\beta}(0) = 1$. Additionally, $\widetilde{\mathscr{N}_\beta}(t)$ has similar

decay rate on the real line as g:

$$\widetilde{\mathcal{N}_\beta}(t) = \int_{-\infty}^{\infty} g(y)g(t-y)\,dy \le \left(\int_{|y|\ge t/2} + \int_{|y-t|\ge t/2}\right)[g(y)g(t-y)]\,dy \quad \le 2g(t/2)\int_{-\infty}^{\infty} g(y)\,dy, \tag{4.82}$$

so that $|\widetilde{\mathcal{N}_\beta}(t)| \le Cg(t/2)$, where $C = 2\int_{-\infty}^{\infty} g(t)\,dt$ is independent of t. The rapid decay of g for $|\operatorname{Im}\mathbf{t}| < c\beta/2$ also ensures that $\widetilde{\mathcal{N}_\beta}(\mathbf{t})$ is defined, bounded, and integrable along horizontal lines in any smaller strip $|\operatorname{Im}\mathbf{t}| \le c\beta/2 - \epsilon$.

The double exponential decay rate $\propto \exp(-e^{\frac{\pi}{2c\beta}|t|})$ can thus be achieved by $\widetilde{\mathcal{N}_\beta}(t)$ analytic and bounded in a strip $|\operatorname{Im}\mathbf{t}| < c\beta/2 - \epsilon$ for any $\varepsilon > 0$, with $\mathcal{N}_\beta(E) \ge 0$ and normalized. Verifying the conditions in e.g. [188, Theorem IX.14] shows the density of states $\mathcal{N}_\beta(E)$ satisfies an exponential decay bound. Plots demonstrating the decay of these $\widetilde{\mathcal{N}_\beta}(t)$ and $\mathcal{N}_\beta(E)$ are shown in Fig. 4.2.

Surprisingly, $\mathcal{N}_\beta(E)$ and $\widetilde{\mathcal{N}_\beta}(t)$ look very similar to a Gaussian distribution expected for generic many-body systems [38, 47]; for $c\beta/\pi = 1$ as in Fig. 4.2, the deviation of the latter from a Gaussian becomes appreciable (more than 10% of the Gaussian) only around $t \approx 1.9$, at which $|\widetilde{\mathcal{N}_\beta}(1.9)|^2 \approx 0.014 \in (2^{-7}, 2^{-6})$. This suggests that the difference in quantum dynamics between the examples constructed here and generic many-body systems (in terms of the contribution from the energy eigenvalues) may become significant only for systems of $N \gg 6$ qubits [so that the $O(D^{-1})$ fluctuations in $|\widetilde{\mathcal{N}_\beta}(t)|^2$, which is the square of a sum of D phase factors, are negligible compared to its value at this time].

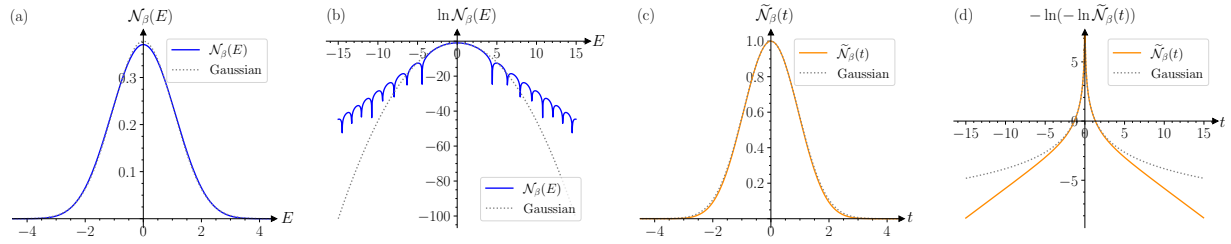


Figure 4.2: (a)-(b) Plots of $\mathcal{N}_\beta(E) = \frac{2\pi}{\|g\|_2^2}(g^\vee(E))^2$, for $g(t) = \exp(-e^t)\exp(-e^{-t})$ which corresponds to $c\beta = \pi$ in Eq. (4.77), compared against a Gaussian probability density with the same variance. The spacing between energy samples is $\Delta E = 0.025$. While the plot of $\mathcal{N}_\beta(E)$ in (a) looks similar to a Gaussian, higher precision integration using the mpmath library [189] and a logarithmic scale in (b) demonstrate the exponential (linear in $\ln \mathcal{N}_\beta(E)$) decay, rather than Gaussian decay. (c)-(d) Plots of $\widetilde{\mathcal{N}_\beta}(t) = \frac{g*g(t)}{\|g\|_2^2}$ compared to a Gaussian $e^{-t^2\sigma^2/2}$, where σ^2 is the variance from (a). The spacing between time samples is $\Delta t = 0.05$. While the plot of $\widetilde{\mathcal{N}_\beta}(t)$ in (c) again looks similar to a Gaussian, higher precision integration and a doubly-logarithmic scale in (d) demonstrate the much faster double-exponential (linear in $-\ln(-\ln \widetilde{\mathcal{N}_\beta}(t))$) decay of $\widetilde{\mathcal{N}_\beta}(t)$.

4.5.2 Upper bound on the scrambling time for these examples

The constraint $|\widetilde{\mathcal{N}_\beta}(t)| \leq C g(t/2)$ leads to a double exponential decay of the regularized SFF for large t:

$$K_\beta(t \to \infty) \lesssim f_\beta^2 C^2 \exp\left[-2\exp\left(\frac{\pi t}{2c\beta}\right)\right] \tag{4.83}$$

It is also known that for systems with energy eigenstates that are "sufficiently random" [i.e., behave as if sampled from the Haar distribution, including but not restricted to Haar random energy eigenstates in the computational basis], the expectation values of observables $\hat{\Pi}_\alpha$ in initial states $\hat{\rho}_k(0)$ [Eqs. (4.1) and (4.2) of Sec. 4.1] essentially track the SFF to decay to their (maximally) scrambled values [2, 114, 154, 155]:

$$\mathrm{Tr}[\hat{\rho}_k(t)\hat{\Pi}_\alpha] = \frac{1 + O(D_E^{-1/2})}{D_S} + \left\{\mathrm{Tr}[\hat{\rho}_k(0)\hat{\Pi}_\alpha] - \frac{1}{D_S}\right\} K_\beta(t). \tag{4.84}$$

This corresponds to Eq. (3.10) in Chap. 3. Thus, all these expectation values are guaranteed to maximally scramble after a time t_s, i.e.,

$$\mathrm{Tr}[\hat{\rho}_k(t)\hat{\Pi}_\alpha] \sim \frac{1}{D_S}, \quad \text{for } t : t_s < t < T_D, \tag{4.85}$$

if $K_\beta(t > t_s) = o(D_S^{-1})$ [2], e.g., $K_\beta(t > t_s) \leq D_S^{-1-\varepsilon}$ for any $\varepsilon > 0$. Using Eq. (4.83) with this condition gives for the scrambling time:

$$t_s \lesssim \frac{2c\beta}{\pi} \ln\Big[\ln\Big(f_\beta^2 C^2 D_S^{1+\varepsilon}\Big)\Big]. \tag{4.86}$$

Recalling that $S_{2,S} \sim N_S \ln 2$ for maximal entanglement (for $N_S < N_E$), and further assuming an initial state such that $|\ln(C f_\beta)| = o(N_S)$, we get a leading order scrambling time of $t_s \lesssim (2c\beta/\pi) \ln S_{2,S}$. We conclude that the systems constructed in this section nearly saturate the logarithmic-in-entanglement-entropy bound of Eq. (4.33) [Eq. (4.10) in Sec. 4.1], if the energy eigenbasis is sufficiently random with respect to the computational basis.

Part II

Quantum Systems Beyond the Chaotic-Integrable Dichotomy

Chapter 5: Spectral rigidity imprints in integrable many-body systems: Complex SYK-2

5.1 Introduction: Does the Berry-Tabor conjecture apply to many-fermion systems?

Recall that the Berry-Tabor conjecture [34] expects the standard action-angle quantizations of classically integrable systems to show Poisson statistics, as long as the integrable Hamiltonian is "generic" (e.g., the tori of Sec. 2.6.1 are notable exceptions [34]). The key ingredient for their argument is the idea that the Hamiltonian for quantized integrable systems can be expressed in terms of N action variables $\hat{J}_k$ with eigenvalues $J_k \in \mathbb{N}_0$:

$$\hat{H} = \hat{H}(\{\hat{J}_k\}). \tag{5.1}$$

Then, for sufficiently "generic" Hamiltonians, Ref. [34] pictures surfaces of constant energy $\hat{H}(\{J_k\}) = E$ in the $\hat{J}_k$-eigenbasis as generic surfaces in the space $\{J_k\} \in \mathbb{R}^N$. This surface evolves in this space as energy increases and randomly intersects the energy eigenstates represented by different points on an N-dimensional integer lattice of the $\{J_k\}$, with E giving the corresponding eigenvalue at each intersection. These random encounters of the surface with the lattice sites leads to Poisson statistics for the energy eigenvalues.

In a non-interacting many-fermion system, there are two aspects that differ significantly from this picture. The first is that the integrals of motion are the fermion numbers $n_k \in \{0, 1\}$ in each of the N single-particle levels. These do not lead to an integer lattice as above, but an N-dimensional cube

in the space of integrals of motion. The second is that it is less clear if a "generic" surface of constant energy would intersect points on this cube in such a way as to reproduce Poisson statistics.

In this chapter, we derive the spectral statistics corresponding to this problem, averaging over ensembles of Haar-random single particle Hamiltonians to model generic single-particle behavior. The single-particle energy levels therefore have Wigner-Dyson statistics for almost all systems in this ensemble. We find that the many-body levels show correlations at energy separations corresponding to the single-particle Wigner-Dyson correlations, but nevertheless resemble Poisson statistics by and large at smaller energy scales. This reformulates and refines the Berry-Tabor conjecture for a class of integrable systems outside its original premise. In the context of the dynamical notions of ergodicity in Chapter 2, we argue that an energy spectrum exhibiting this separation between correlations at the single-particle scale and Poisson statistics at the many-body energy scale is consistent with non-ergodicity in the full many-body Hilbert space, and cyclic ergodicity in the single-particle subspace.

The quantitative derivation of these single-particle imprints in the many-body energy spectrum also addresses another significant question in the context of SYK-q models with q-body interactions. For $q \geq 4$, these models are interacting, non-integrable, and expected to show Wigner-Dyson statistics [82, 86] (see Chapter 6 for a more detailed discussion). However, while the $q = 2$ model is non-interacting and formally integrable, numerical studies of the non-interacting $q = 2$ case revealed an unexpected ramp (of previously unknown quantitative form) representing nontrivial spectral correlations [190]. The model considered in this Chapter is equivalent to SYK-2 with complex fermions, and provides an analytical derivation of these correlations. In this context, this Chapter (and the paper [4] it is based on) is a close counterpart of the parallel work Ref. [87] focusing on SYK-2 with Majorana fermions.

5.2 Gaussian random matrices and complex SYK-2

Model— We first consider a Gaussian Unitary ensemble (GUE) [28, 36] of $N \times N$ Hermitian single-particle Hamiltonians, h, following the distribution function:

$$P(\{h_{ij}\}) = 2^{N(N-1)/2}\left(\frac{N}{2\pi}\right)^{N^2/2}\exp\left[-\frac{N}{2}\mathrm{Tr}\left(h^2\right)\right], \tag{5.2}$$

with the local level statistics of $\hat{h}$ falling into the unitary Wigner-Dyson (WD) class [28,36]. Populating these single-particle sites [corresponding to the basis in which h looks like a random matrix as in Eq. (5.2)] with fermions $(\hat{c}_i, \hat{c}_i^\dagger)$ with a chemical potential μ then defines the many-body Hamiltonian,

$$\hat{H} = \sum_{i,j}\hat{c}_i^\dagger(h_{ij} - \mu\delta_{ij})\hat{c}_j. \tag{5.3}$$

This is an integrable model, and the particle number at each single-particle level is a constant of the motion. Specifically, it is the complex SYK-2 model: a version of the SYK-($q = 2$) model with complex fermions [163, 179]. Each site essentially has two basis states, and the Hilbert space dimension is therefore given by $D = 2^N$.

In general, for such a statistical ensemble of Hamiltonians, we can define an *ensemble-averaged* 2-point spectral form factor (SFF) [28, 40],

$$K(t) \equiv \langle|Z(it)|^2\rangle = \frac{1}{D^2}\left\langle\sum_{n,m}e^{i(E_m - E_n)t}\right\rangle, \tag{5.4}$$

where $Z(it) \equiv D^{-1}\,\mathrm{Tr}(e^{-i\hat{H}t}) = D^{-1}\sum_n e^{-iE_n t}$ is the normalized partition function with $\beta = 0$ as in Chap. 4 (here, without regularization), with E_n being the eigenvalues of $\hat{H}$, and the angular bracket represents ensemble averaging. Even after ensemble averaging, we continue to have that $0 \leq |K(t)| \leq$

1 with $K(0) = 1$, and further that $K(\infty) = D^{-1}$ if degeneracies are statistically insignificant in the ensemble.

The (ensemble-averaged) SFF is essentially a Fourier transform of the (ensemble-averaged) joint two-level distribution function (see also Eq. (5.15)) [28, 36]. For an ensemble with Poisson statistics (effectively independently distributed energy levels in small energy windows), $K(t)$ decays from 1 at $t = 0$, gradually approaching D^{-1} at a time scale much smaller than the inverse mean level separation. However, Hamiltonians obeying WD statistics are characterized by level repulsion at a scale Δ corresponding to the typical level spacing. This results in an SFF that "slopes" down below D^{-1} up to around a dip time $t \sim t_d$, where it reaches a minimum, then grows in an approximately linear "ramp", generally reaching a $K(t) = D$ "plateau" for $t \geq t_*$. We will call t_* the plateau time. This should be differentiated from the Heisenberg time $t_{\mathrm{H}} \sim 2\pi\mathscr{N}^{\#}$, determined by the typical density of states $\mathscr{N}^{\#} \sim \exp(\Theta(N))$. The ramp and plateau have their origins in the Fourier transform of the level repulsion component of the distribution, which implies that $t_* \sim 1/\Delta$. We note that due to ensemble averaging, the SFF precisely follows these trends without fluctuations [110].

Results— For the system given by Eq. (5.3), which is our primary concern in this Chapter, we find three approximate expressions that closely describe the SFF in different regions, in the large N limit, i.e.

$$K(t) \approx \begin{cases} K_1(t), & 0 < t \ll O(1), \\ K_2(t), & O(1) \ll t \ll O\left(N/\log_2 N\right), \\ K_3(t), & \sqrt{2}N < t < \infty, \end{cases} \tag{5.5}$$

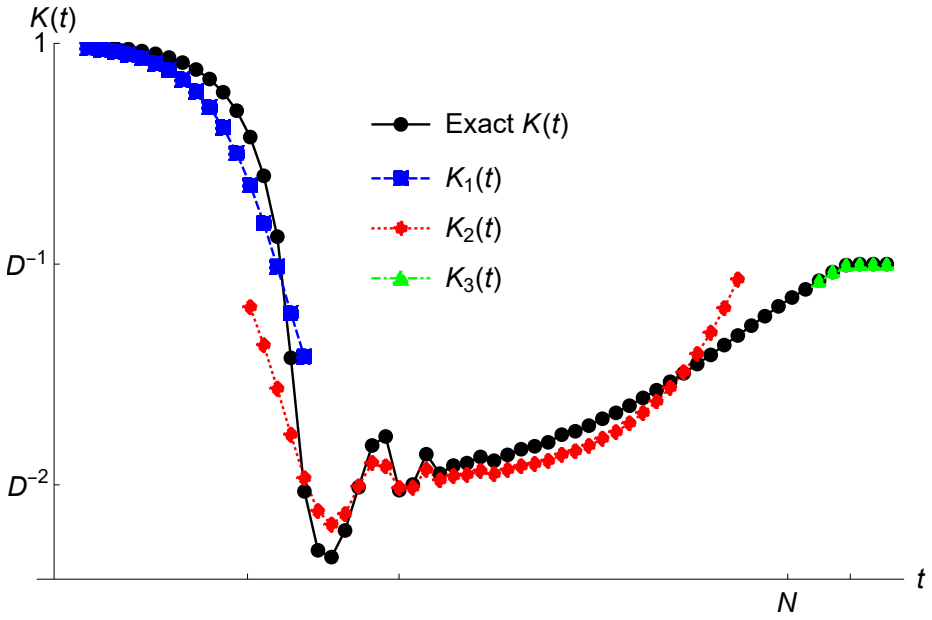


Figure 5.1: A log-log plot of the many-body SFF of the non-interacting GUE model (Eqs. (5.2) and (5.3)) computed for $N = 400$ levels with chemical potential $\mu = 0$: numerically from Eqs. (5.12), (5.13) ("exact"), compared with the three approximate expressions $K_{1,2,3}(t)$ given by Eqs. (5.6a), (5.6b) and (5.6c), respectively.

where

$$K_1(t) = \cos^{2N}\left(\frac{\mu t}{2}\right)\exp\left[N\left(\frac{J_1(2t)}{t}-1\right)\cos(\mu t)\right], \tag{5.6a}$$

$$K_2(t) = D^{-2}\left(\frac{N}{8}e^{\gamma_E}\right)^{t/4}\exp\left[N\frac{J_1(2t)}{t}\cos(\mu t)\right], \tag{5.6b}$$

$$K_3(t) = D^{-1}\exp\left[-\frac{(4N^2-t^2)^{\frac{3}{2}}}{12\pi N t}\Theta(2N-t)\right], \tag{5.6c}$$

with $J_1(z)$ being the Bessel function of the first kind, $\Theta(x)$ the unit step function, and $\gamma_E = 0.577...$ the Euler-Mascheroni constant. $K_1(t)$ describes the initial downward slope region; $K_2(t)$ is related to the transition from an oscillatory region up to $t \sim O((N/\ln N)^{\frac{2}{5}})$ to an exponential beginning of the ramp; and $K_3(t)$ gives the late-time ramp approaching the plateau. These approximations are illustrated in Fig. (5.1), where they are compared with a numerical calculation based on Eqs. (5.12) and (5.13) discussed later in this Chapter. We will now sketch some important steps in the derivation of these

expressions, relegating the full details to the subsequent sections of this Chapter.

Details of calculation — From the definition of the SFF (Eq. (5.4)), it is straightforward to show that for an ensemble of Hamiltonians described by Eqs. (5.2) and (5.3), $K(t)$ is given by

$$K(t) = D^{-1} \int \mathrm{d}\varepsilon_1 ... \mathrm{d}\varepsilon_N P(\varepsilon_1, ..., \varepsilon_N) \prod_{k=1}^{N} \left[1 + \cos\left((\varepsilon_k - \mu)t\right)\right] \tag{5.7}$$

Here $P(\varepsilon_1, ..., \varepsilon_N)$ is the joint probability density function of the N single-particle energy levels $\{\varepsilon_i\}$ and is given by

$$P(\varepsilon_1, \varepsilon_2, ..., \varepsilon_N) = C_N \exp(-\frac{N}{2}\varepsilon_i^2) \prod_{1 \le i < j \le N} |\varepsilon_i - \varepsilon_j|^2, \tag{5.8}$$

where C_N is a normalization constant.

By expanding the product of cosines and integrating out the ε_k that do not occur in the cosine terms, Eq. (5.7) can be expressed in terms of the n-point single-particle energy level correlation functions $\mathrm{R}_n(\varepsilon_1, ..., \varepsilon_n)$, defined by

$$\mathrm{R}_n(\varepsilon_1, ..., \varepsilon_n) = \frac{N!}{(N-n)!} \int d\varepsilon_{n+1} ... d\varepsilon_N P(\varepsilon_1, ..., \varepsilon_N). \tag{5.9}$$

This gives the probability density of finding an energy level around each ε_i, for $i = 1, ..n$, irrespective of the remaining levels and independent of the labeling. It has been found that, for energy levels described by the GUE probability distribution (Eq. (5.8)), $\mathrm{R}_n(\varepsilon_1, \varepsilon_2 .... \varepsilon_n)$ is given by the determinant of the kernel $\mathrm{K}(\varepsilon_i, \varepsilon_j)$ [36],

$$\mathrm{R}_n(\varepsilon_1, \varepsilon_2 ... \varepsilon_n) = \det\left[\mathrm{K}(\varepsilon_i, \varepsilon_j)\right]_{i,j=1,...n}, \tag{5.10}$$

where $\mathtt{K}(\varepsilon_i, \varepsilon_j)$, in the large N limit, takes the form of

$$\mathtt{K}(\varepsilon_i, \varepsilon_j) = \begin{cases} \mathtt{R}_1(\varepsilon_i) = \dfrac{N}{2\pi}\sqrt{4-\varepsilon_i^2}\,\Theta(2-|\varepsilon_i|), & i = j. \\ \\ \mathtt{K}(\varepsilon_i - \varepsilon_j) = \dfrac{N}{\pi}\dfrac{\sin\left[N\left(\varepsilon_i - \varepsilon_j\right)\right]}{N\left(\varepsilon_i - \varepsilon_j\right)}, & i \neq j. \end{cases} \tag{5.11}$$

Here $\mathtt{R}_1(\varepsilon_i)$ is the average single-particle level density and exhibits Wigner's semicircle law. We note that the 2-point many-body SFF $K(t)$ is given by the summation of the Fourier transform of the n-point single-particle energy level correlation function which is closely related to the single-particle n-point SFF, with n running over $1, ...N$. As the correlation functions can be expressed as a determinant, the eigenvalues of Haar random matrices are often said to follow a "determinantal point process" [191].

In this Chapter, we will focus on the derivation of $K_3(t)$, using a method based on the description of GUE levels as a determinantal point process [36, 191], which is converted to the exponential of a series in an effective cumulant expansion. We note that $K_1(t)$ and $K_2(t)$ can be derived through a separate resummation of the cumulant expansion, with certain elaborate approximations or a field theory argument [4]. Further, similar expressions for $K_1(t)$ and $K_2(t)$ for the Majorana SYK-2 model can be derived from saddle points in the G, Σ large-N field theory for SYK [87].

For $K_3(t)$, we are interested in the $t = \Theta(N)$ behavior. we start from the following expression, obtained using the well known technique of expressing $P(\varepsilon_1, ..., \varepsilon_N)$ as a determinant of the Hermite polynomials $H_n(x)$ [36],

$$K(t) = D^{-1}\det\left[\delta_{jk} + \mathtt{M}_{jk}(t)\right]_{j,k=1,...,N}, \tag{5.12}$$

where

$$\mathtt{M}_{jk}(t) = \mathtt{W}_{jk}\left(\frac{t^2}{N}\right)\begin{cases}(-1)^{\frac{j-k}{2}}\cos(\mu t), & j-k \text{ is even},\\ (-1)^{\frac{j-k-1}{2}}\sin(\mu t), & j-k \text{ is odd},\end{cases} \tag{5.13}$$

$$\mathtt{W}_{j\geq k}(\tau) = \sqrt{\frac{(k-1)!}{(j-1)!}}\tau^{\frac{j-k}{2}}e^{-\frac{\tau}{2}}L_{k-1}^{j-k}(\tau),$$

with $\mathtt{W}_{jk} = \mathtt{W}_{kj}$, and $L_n^\alpha(x)$ are the Laguerre polynomials.

To simplify this expression further, we require an approximate asymptotic expression for $L_n^\alpha(x)$ for large n, with large or small α. For this purpose, we use a modification of a standard result given in Ref. [192], which is described in Sec. 5.3. One principal consequence of using this expression is that each $\mathtt{M}_{jk}(t)$ acquires a cutoff in t above which it vanishes:

$$\mathtt{M}_{jk}(t) \propto \Theta\left(2N(j+k) - t^2\right). \tag{5.14}$$

Note that the maximum value of $j+k$ is $2N$, so the Θ-function ensures that all the $\mathtt{M}_{jk}(t)$ vanish for $t > 2N$, showing that the plateau occurs at $t_* = 2N$, which is also the plateau time of the corresponding single-particle SFF [154, 193].

We can further use the matrix relation $\det A = e^{\mathrm{Tr}\ln A}$ in Eq. (5.12), and expand the exponent in powers $\mathtt{M}^n$ of $\mathtt{M}$. Due to the oscillatory nature of the polynomials, only terms with even n contribute. The $n = 2$ term is simple to evaluate, and gives a monotonically increasing upper bound for $K(t)$ for $t = O(N)$ (see Sec. 5.3), whose form is given by $K_3(t)$ in Eq. (5.6c) for $t > \sqrt{2}N$. It is also seen numerically that the $n > 2$ terms appear to be negligible in this region, and $K_3(t)$ approximates $K(t)$ fairly closely.

Discussion— To study the level statistics with the help of $K(t)$, we can take the Fourier transform of Eq. (5.4) to get the joint density function of two many-body levels with separation S, summed over

the entire spectrum,

$$\tilde{R}_2(S) = \int_{-\infty}^{\infty} \frac{dt}{2\pi} D^2 K(|t|) e^{-iSt} - D\ \delta(S), \tag{5.15}$$

where in the second term, we have subtracted off the contribution from when the two levels are identical. We will now set $\mu = 0$ as it simplifies our arguments without altering their essential content (as the ramp and plateau are μ-independent). For $S \gg 1$, only the small-t behavior of $K(t)$ is relevant. We therefore use the expression $K_1(t)$ from Eq. (5.6a) with $\mu = 0$. Expanding to leading order in t, we have $K_1(t) \approx \exp\left(-\frac{N}{2}t^2\right)$, which gives

$$\tilde{R}_2(S \gg 1) \approx \frac{D^2}{\sqrt{2\pi N}} \exp\left(-\frac{S^2}{2N}\right), \tag{5.16}$$

showing that the many-body energy spectrum has a width $w \sim \sqrt{N}$.

At the scale $\Delta \sim N^{-1}$ of single-particle level spacings, we must account for the contribution from K_1 (by setting $S \approx 0$ in Eq. (5.16)) as well as the ramp and plateau. We note that for t near $2N$, we can expand the exponent in Eq. (5.6c), obtaining, to leading order,

$$K_3(t < 2N) \approx D^{-1} \exp\left[-\frac{(2N-t)^{\frac{3}{2}}}{3\pi N^{\frac{1}{2}}}\right]. \tag{5.17}$$

$K_3(t)$ is therefore comparable to D^{-1} only in a relatively small region of size $\sim N^{1/3} \ll 2N$, and is much smaller for other $t < 2N$. This shows that for $N \to \infty$ and $t = O(N)$, $K(t)$ is essentially given by a step function from 0 to L at the plateau time $t_* = 2N$, i.e. $K(t) \to D^{-1}\Theta(t - 2N)$. To compute $\tilde{R}_2(S \sim N^{-1})$ for large but finite N, we can approximate $K_3(t)$ by a step function in $0 \lesssim t < 2N$, with the width chosen to enclose the same area up to the t-axis as in Eq. (5.17),

$$K_{\text{ramp}}(t) \approx D^{-1}\Theta(t - 2N\alpha), \tag{5.18}$$

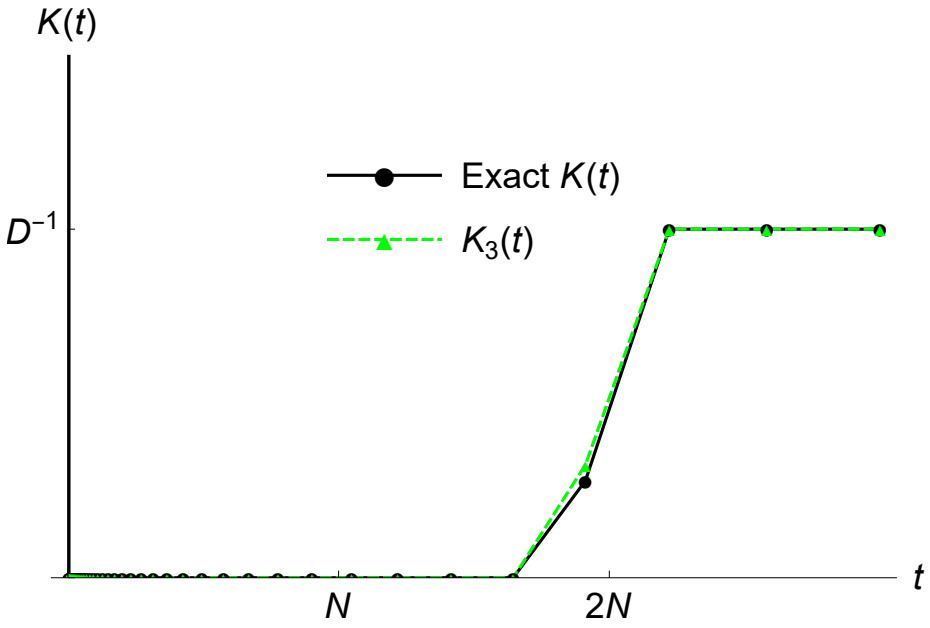


Figure 5.2: A linear scale plot of the many-body SFF of the non-interacting GUE model (Eqs. (5.2) and (5.3)) computed for $N = 400$ levels with chemical potential $\mu = 0$: numerically from Eqs. (5.12), (5.13) ("exact"), compared with the approximate expression $K_3(t)$ given by Eq. (5.6c). This is the linear counterpart to Fig. 5.1, and partly illustrates Eqs. (5.17) and (5.18).

where $\alpha = 1 - \frac{2}{3}\Gamma\left(\frac{2}{3}\right)(3\pi/N)^{\frac{2}{3}} + O(N^{-3/5})$, as determined by integrating Eq. (5.17), with $\alpha \to 1$ as $N \to \infty$.

We show in Sec. 5.4 that the generalization of $K_3(t)$ to an arbitrary (non-Gaussian) unitary random matrix ensemble [36] is the following upper bound for the many-body SFF at large t (comparable to the inverse of the scale of single-particle level spacings),

$$K(t) \leq D^{-1} \exp\left[\frac{\kappa(t) - N}{4}\right], \tag{5.19}$$

where $\kappa(t)$ is the single-particle SFF with a plateau time t_*'. For $t > t_*'$, the inequality becomes an equality, and $K(t)$ also attains a plateau. Consequently, as $0 < \left(N - \kappa(t < t_*')\right) \sim O(N)$ in the exponent, $K(t)$ can be approximated by a step function expression similar to Eq.(5.18), with $2N\alpha$ replaced by $\sim t_*'$. Therefore, the form of the local contribution to many-body level statistics from

Eq.(5.18) is not just specific to the system defined by Eq. (5.2), but should apply generally to a typical system showing unitary Wigner Dyson statistics in the single particle energy levels.

Using Eq. (5.18) in Eq. (5.15), together with a constant contribution at this scale from $K_1(t)$ gives

$$\tilde{\mathrm{R}}_2(S \sim N^{-1}) \approx \frac{D^2}{\sqrt{2\pi N}} - \frac{2N\alpha D}{\pi}\frac{\sin(2N\alpha S)}{2N\alpha S}. \tag{5.20}$$

The local contribution is the oscillatory second term, $\Delta\tilde{\mathrm{R}}_2(S) = \tilde{\mathrm{R}}_2(S) - D^2/\sqrt{2\pi N}$. This is plotted in Fig. (5.3), and compared with a numerical computation based on Eqs. (5.12) and (5.15). Equation (5.20) can be contrasted with the more familiar two-level correlation function for the single-particle levels in GUE (e.g. [28, 36, 154, 193]),

$$\mathrm{R}_2(S \sim N^{-1}) \propto 1 - \frac{\sin^2(NS)}{N^2S^2}. \tag{5.21}$$

The second term in both Eq. (5.20) and Eq. (5.21) contains the level repulsion effect, with any two levels least likely to have $S \ll N^{-1}$ at this scale. Unlike the single-particle GUE, where $\mathrm{R}_2(0) = 0$, the level repulsion for the many body case, Eq. (5.20), isn't total (i.e. $\tilde{\mathrm{R}}_2(0) \neq 0$), and in fact, negligible compared to the actual two-level density near $S = 0$. This is essentially because the ultimate origin of this level repulsion is still the single particle level spectrum. If we introduce interactions between the fermions (taking us outside the model of Eq. (5.3), however, we should expect spectral rigidity to set in closer to the Heisenberg time scale $t_{\mathrm{H}} \sim 2\pi\mathcal{N}^{\#}$ of the many-body system (typically with $\sim \sqrt{\ln D}$ spectral fluctuations relative to this scale as expected in RMT). This establishes a separation in the scale and magnitude of single-particle effects vs. truly many-body interaction effects in the many-body eigenvalue statistics.

Another interesting feature of Eq. (5.20) is that the second term can take both positive and negative values. This means that in addition to the level repulsion effect, there are less dominant

centers of level attraction i.e. values of S where levels tend to be found relative to each other with higher probability than in Poisson statistics (as determined by the asymptotic large-S value within this regime). As $\alpha \to 1$ when $N \to \infty$, these essentially occur for values of S immediately preceding the local maxima of Eq. (5.21).

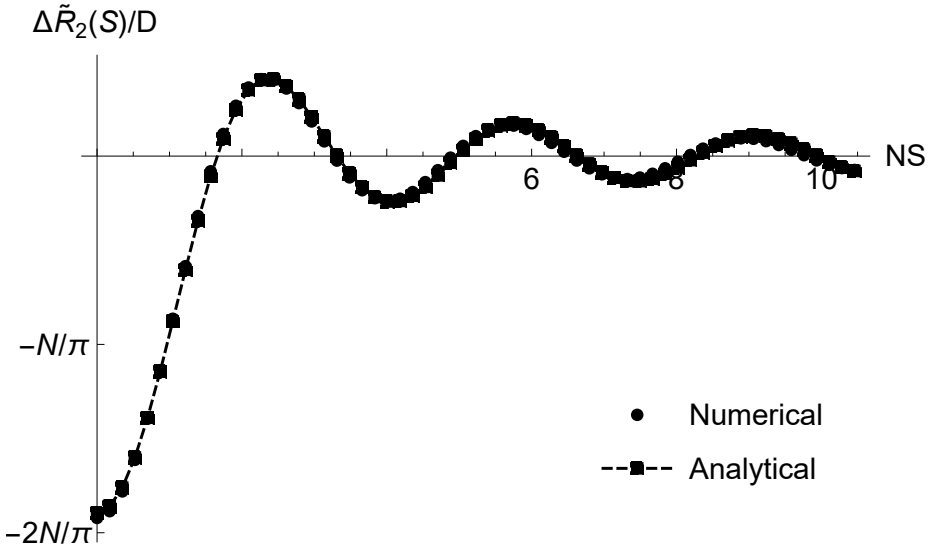


Figure 5.3: A plot of $\Delta\tilde{R}_2(S)/L$ obtained from the Fourier transform of the SFF for $N = 240$ as a function of scaled level separation NS, with the 'analytical' curve from the second term of Eq. (5.20) and the 'numerical' data obtained from Eqs. (5.12), (5.13) and (5.15).

Dynamical (non-)ergodicity in the complex SYK-2 model— It is worth briefly considering how these results impact ergodic dynamics in a single realization of the SYK-2 model. The fact that spectral statistics is essentially Poisson between t_* and t_H, and therefore for almost all time up to t_H, suggests that the ergodic behavior of this model in its full Hilbert space should be virtually comparable to that of Poisson statistics, i.e., non-ergodic. However, a single particle subspace has the statistics of a single realization of GUE, and we should expect ergodicity and aperiodicity precisely as in Sec. 2.5.4. An interesting question to consider is whether the precise quantitative form of $K(t)$ for this model has any tractable impact on observable-independent quantum dynamics in any subspace, beyond these straightforward observations.

5.3 Laguerre polynomials for Gaussian random matrices

5.3.1 The SFF in determinant form

The Wigner-Dyson (WD) distribution for N GUE levels is given by Eq. (5.8). Following Ref. [36], we express this in terms of arbitrary polynomials $C_k(x)$ of degree $k-1$ (with arbitrary coefficients, which we will later compensate for by normalizing $P(\varepsilon_1,...,\varepsilon_N)$. Introducing permutation operators I,J that act on $\{1,...,N\}$ and the symbol $\epsilon^{\{J\}}_{\{I\}} = +1,-1$ if I,J are of the same or opposite parities (respectively), we have

$$P(\varepsilon_1,...,\varepsilon_N) = \left(\prod_i e^{-N\varepsilon_i^2/2}\right)\left(\sum_{I,J}\epsilon^{\{J\}}_{\{I\}}\prod_{k=1}^{N} C_k(\varepsilon_{I_k})C_k(\varepsilon_{J_k})\right) \tag{5.22}$$

$$= \sum_{i,j}\epsilon^{\{j\}}_{\{i\}}\prod_{k=1}^{N} e^{-N\varepsilon_k^2/2}C_{i_k}(\varepsilon_k)C_{j_k}(\varepsilon_k), \tag{5.23}$$

again up to normalization. The operators i,j in the second line are the inverses of I,J from the corresponding terms in the first line.

It is convenient to choose $C_k(x) = \mathscr{H}_{k-1}\left(\sqrt{\frac{N}{2}}x\right)$, where $\mathscr{H}_k(x) = (2^k k!\sqrt{\pi})^{-1/2}H_k(x)$ are the normalized Hermite polynomials (satisfying $\int \mathrm{d}x\ e^{-x^2}\mathscr{H}_i(x)\mathscr{H}_j(x) = \delta_{ij}$). In that case, using Theorem. 5.7.1 in Ref. [36] to perform the integrals in the normalization condition,

$$\int \mathrm{d}\varepsilon_1...\mathrm{d}\varepsilon_N\ P(\varepsilon_1,...,\varepsilon_N) = 1, \tag{5.24}$$

we can show that the normalized Wigner-Dyson distribution is

$$P(\varepsilon_1,...,\varepsilon_N) = \frac{(N/2)^{N/2}}{N!}\sum_{i,j}\epsilon^{\{j\}}_{\{i\}}\prod_{k=1}^{N} e^{-N\varepsilon_k^2/2}\mathscr{H}_{i_k-1}\left(\sqrt{\tfrac{N}{2}}\varepsilon_k\right)\mathscr{H}_{j_k-1}\left(\sqrt{\tfrac{N}{2}}\varepsilon_k\right). \tag{5.25}$$

Inserting Eq. (5.25) into Eq. (5.7) and noting that all the integrals over the ε_k are of the same

form, we obtain for the SFF (with the replacement $\varepsilon_k\sqrt{N/2}\to x$ in each factor),

$$K(t)=\frac{2^{-N}}{N!}\sum_{i,j}\epsilon_{\{i\}}^{\{j\}}\prod_{k=1}^{N}\int \mathrm{d}x\; e^{-x^2}\mathscr{H}_{i_k-1}(x)\,\mathscr{H}_{j_k-1}(x)\left(1+\cos\left[\left(x\sqrt{\frac{2}{N}}-\mu\right)t\right]\right). \tag{5.26}$$

Using the orthonormality of the $\mathscr{H}_k(x)$, and defining,

$$\mathtt{M}_{jk}(t)=\int \mathrm{d}x\; e^{-x^2}\left\{\mathscr{H}_{j-1}(x)\mathscr{H}_{k-1}(x)\cos\left[\left(x\sqrt{\frac{2}{N}}-\mu\right)t\right]\right\}. \tag{5.27}$$

and identifying $(1/N!)\sum_{i,j}\epsilon_{\{i\}}^{\{j\}}A_{ij}=\det A$ for a matrix A, we obtain,

$$K(t)=2^{-N}\det\left[\delta_{jk}+\mathtt{M}_{jk}(t)\right]_{j,k=1,\ldots,N}. \tag{5.28}$$

Expanding out $\cos((\varepsilon-\mu)T)=\cos(\varepsilon T)\cos(\mu t)+\sin(\varepsilon T)\sin(\mu t)$, the resulting integrals can be evaluated with the help of results from standard tables (e.g. 7.388(6,7) in Ref. [194]), and we get

$$\mathtt{M}_{j\geq k}(t)=\mathtt{W}_{jk}\left(\frac{t^2}{N}\right)\mathtt{f}_{jk}(\mu t), \tag{5.29}$$

with $\mathtt{M}_{kj}(T)=\mathtt{M}_{jk}(T)$, where

$$\mathtt{W}_{j\geq k}(\tau)=\sqrt{\frac{(k-1)!}{(j-1)!}}\tau^{\frac{j-k}{2}}e^{-\frac{\tau}{2}}L_{k-1}^{j-k}(\tau), \tag{5.30}$$

with $L_n^{\alpha}(x)$ denoting the Laguerre polynomials, and

$$\mathtt{f}_{jk}(\mu t)=\begin{cases}(-1)^{\frac{j-k}{2}}\cos(\mu t), & j-k \text{ is even},\\ (-1)^{\frac{j-k-1}{2}}\sin(\mu t), & j-k \text{ is odd}.\end{cases} \tag{5.31}$$

Using $\det A = e^{\mathrm{Tr}\ln A}$ in Eq. (5.28) and expanding $\ln A$ in a power series, we get

$$K(t) = 2^{-N} \exp\left\{-\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \mathrm{Tr}\left[\mathtt{M}^n(t)\right]\right\}. \tag{5.32}$$

5.3.2 The asymptotic behavior of Laguerre polynomials

Now we will provide some additional background for the approximate form of Laguerre polynomials $L_n^\alpha(x)$ referred to in Sec. 5.1. Specifically, we are interested in the form for large n, with large or small α. Defining $\nu = 4n + 2\alpha + 2$, the general behavior of the Laguerre polynomials can be split into the oscillatory region, $x < \nu$, and the monotonic region, $x > \nu$ [192]. In the oscillatory region, we consider the leading term of Eq.(8) from Sec.10.15 in Ref. [192], which amounts to

$$e^{-\frac{x}{2}} x^{\frac{\alpha}{2}} L_n^\alpha(x)|_{x<\nu} \approx \sqrt{\frac{2}{\pi}} \left(\frac{\nu}{4}\right)^{\frac{\alpha}{2}} \frac{\sin\left(\varphi_n^\alpha(x)\right)}{(x(\nu-x))^{\frac{1}{4}}}, \tag{5.33}$$

where we have absorbed the overall sign into the oscillatory factor $\sin\left(\varphi_n^\alpha(x)\right)$ whose specific form is unimportant for our purposes. In the monotonic region, we have (Eq.(15) form Sec.10.15 in Ref. [192]),

$$e^{-\frac{x}{2}} x^{\frac{\alpha}{2}} L_n^\alpha(x)|_{x>\nu} \approx \sqrt{\frac{2}{\pi}} (-1)^n \frac{e^{-\frac{1}{2}\sqrt{x(x-\nu)} + \frac{\nu}{2}\cosh^{-1}\sqrt{\frac{x}{\nu}}}}{(x(x-\nu))^{\frac{1}{4}}}, \tag{5.34}$$

which corresponds to a rapid decay to zero. Both of these expressions are valid for $n \to \infty$, but without a corresponding $\alpha \to \infty$ limit.

For our purposes, we may approximate the right hand side of Eq. (5.34) by zero due to the rapid decay of the expression, and cover the entire region of $x < \nu$ and $x > \nu$ by the single approximate expression,

$$e^{-\frac{x}{2}} x^{\frac{\alpha}{2}} L_n^\alpha(x) \approx \sqrt{\frac{2}{\pi}} \left(\frac{\nu}{4}\right)^{\frac{\alpha}{2}} \frac{\sin\left(\varphi_n^\alpha(x)\right)\Theta(\nu-x)}{(x(\nu-x))^{\frac{1}{4}}}. \tag{5.35}$$

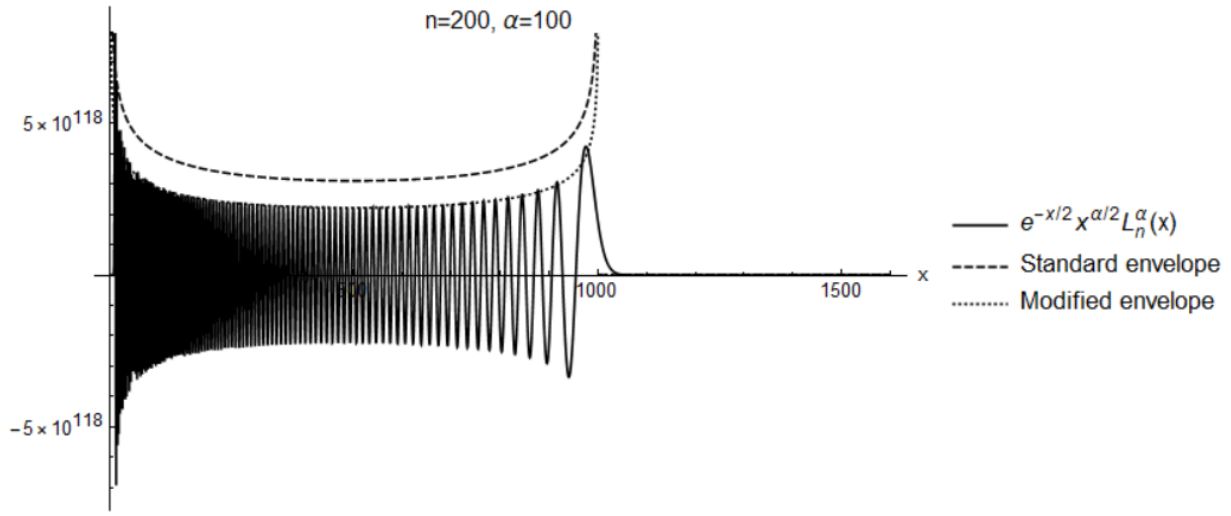


Figure 5.4: Comparison of approximations for the Laguerre polynomials $L_n^\alpha(x)$ for $n = 200, \alpha = 100$. "Standard envelope" refers to Eq. (5.35), and "Modified envelope" to Eq. (5.36), with the oscillatory factor $\sin\left(\varphi_n^\alpha(x)\right)$ dropped and the overall sign set to $+1$ in both cases.

We see numerically (cf. Fig.(5.4)) that Eq. (5.35) is not a good approximation for larger α (e.g. $\alpha = O(n)$), given that the $\alpha \to \infty$ limit was not taken in the above standard results. We will need an expression that is valid for any $\alpha \leq n$ for our application. Based on the form of Eq. (5.30), we instead try the approximate expression,

$$e^{-\frac{x}{2}} x^{\frac{\alpha}{2}} L_n^\alpha(x) \approx \sqrt{\frac{2}{\pi}\frac{(n+\alpha)!}{n!}} \frac{\sin\left(\varphi_n^\alpha(x)\right)\Theta(\nu - x)}{(x(\nu-x))^{\frac{1}{4}}}. \tag{5.36}$$

Numerically, it appears that this expression works better than Eq. (5.35) in the desired range of n, α (cf. Fig.(5.4)), and also has the advantage of greatly simplifying the evaluation of the $n = 2$ term in the exponent of Eq. (5.32).

5.3.3 Evaluating the Trace

Here, we will evaluate the trace in Eq. (5.32) for the $n = 2$ term, obtaining the late-time ramp. Due to the oscillatory term in Eq. (5.36), we can approximate all terms with odd n to zero. Also, `M` is

a symmetric matrix, and therefore has real eigenvalues. $\mathrm{Tr}[\mathtt{M}^n]$ is then the sum of the n-th power of eigenvalues, and for even n must always be positive. While it is hard to precisely evaluate the terms for $n = 4, 6...$, we can definitively say that their contribution decreases the SFF from the $n = 2$ estimate. Therefore, evaluating the $n = 2$ term alone would give us an (approximate) upper bound for the SFF.

We have

$$\mathrm{Tr}\left[\mathtt{M}^2(t)\right] = \sum_{j,k=1}^{N} \mathtt{M}_{jk}^2(t). \tag{5.37}$$

It is convenient to separate out the sum into groups of terms with even or odd $j-k$,

$$\mathrm{Tr}\left[\mathtt{M}^2(t)\right] = \sum_{\substack{j,k=1 \\ \text{even } j-k}}^{N} \mathtt{M}_{jk}^2(t) + \sum_{\substack{j,k=1 \\ \text{odd } j-k}}^{N} \mathtt{M}_{jk}^2(t), \tag{5.38}$$

As the sums range over $1, ..., N$, we expect the indices j, k to typically be large when N is large. Using Eqs. (5.29), (5.30), (5.31) to write an explicit expression for $\mathtt{M}_{jk} = \mathtt{M}_{kj}$,

$$\mathtt{M}_{j\geq k}(t) = \sqrt{\frac{(k-1)!}{(j-1)!}} \left(\frac{t^2}{N}\right)^{\frac{j-k}{2}} e^{-\frac{t^2}{2N}} L_{k-1}^{j-k}\left(\frac{t^2}{N}\right) \begin{cases} (-1)^{\frac{j-k}{2}} \cos(\mu t), & j-k \text{ is even}, \\ (-1)^{\frac{j-k-1}{2}} \sin(\mu t), & j-k \text{ is odd}, \end{cases} \tag{5.39}$$

together with the asymptotic expression for Laguerre polynomials, Eq. (5.36), gives

$$\mathtt{M}_{jk}^2(t) = \frac{2\Theta(\nu_{j+k} - \tau)}{\pi\sqrt{\tau(\nu_{j+k} - \tau)}} \sin^2\left(\varphi_{jk}(\tau)\right) \begin{cases} \cos^2(\mu t), & j-k \text{ is even}, \\ \sin^2(\mu t), & j-k \text{ is odd}, \end{cases} \tag{5.40}$$

where $\tau = t^2/N$, and $\nu_{j+k} \approx 2(j+k)$ for large j, k.

Now, we consider the first sum in Eq. (5.38), with $j-k$ restricted to be even. We transform the summation to the variables $b = (j+k)/2$ (necessarily an integer) and $c = j-k$; b here must be an

integer as $j-k$ being even requires that $j+k$ is also even. With $\nu_{j+k} = 4b$, we can write

$$\sum_{\substack{j,k=1 \\ \text{even } j-k}}^{N} \mathtt{M}_{jk}^2(t) = \sum_{\substack{b,c \\ \text{even } c}}^{N} \frac{2\Theta(4b-\tau)}{\pi\sqrt{\tau(4b-\tau)}} \sin^2\big(\varphi_{jk}(\tau)\big)\cos^2(\mu t) \tag{5.41}$$

$$\approx \sum_{\substack{b,c \\ \text{even } c}}^{N} \frac{\Theta(4b-\tau)}{\pi\sqrt{\tau(4b-\tau)}} \cos^2(\mu t), \tag{5.42}$$

where in the second line, we have assumed that $\sin^2\big(\varphi_{jk}(\tau)\big)$ oscillates several times over regions of nearly constant $(4b-\tau)^{-1/2}$ (for a given τ) as the latter is typically a slowly varying function (cf. Fig. 5.4). This allows us to replace $\sin^2\big(\varphi_{jk}(\tau)\big)$ by its mean value over an oscillation i.e. $1/2$. As the expression is now completely independent of c, the sum over c just corresponds to accounting for the number of elements with the same b i.e. the number of elements of an $N \times N$ matrix on the anti-diagonal given by $j+k = 2b$. This number is given by $1 + 2\min(b-1, N-b)$ as a function of b, which we can approximate by $2\min(b, N-b)$ as $N \gg 1$ and b is typically $O(N)$ over most of the sum. We can additionally replace the sum with an integral as the summand/integrand is not a rapidly varying function of b, getting

$$\sum_{\substack{j,k=1 \\ \text{even } j-k}}^{N} \mathtt{M}_{jk}^2(t) \approx \frac{2}{\pi\sqrt{\tau}}\cos^2(\mu t)\left[\int_0^{\frac{N}{2}} \mathrm{d}b\ \frac{b}{\sqrt{4b-\tau}}\Theta(4b-\tau) + \int_{\frac{N}{2}}^{N} \mathrm{d}b\ \frac{N-b}{\sqrt{4b-\tau}}\Theta(4b-\tau)\right]. \tag{5.43}$$

These are elementary integrals, and a straightforward calculation gives (as expressed in terms of t rather than τ),

$$\sum_{\substack{j,k=1 \\ \text{even } j-k}}^{N} \mathtt{M}_{jk}^2(t) \approx \frac{\Theta(2N-t)}{6\pi N t}\cos^2(\mu t)\Big[(4N^2-t^2)^{\frac{3}{2}} - 2(2N^2-t^2)^{\frac{3}{2}}\Theta(\sqrt{2}N-t)\Big]. \tag{5.44}$$

For the term with odd $j-k$, almost exactly the same reasoning goes through, except b must now be summed over half-integers (i.e. odd $j+k$), which is a negligible difference when one replaces

the sum with an integral in the large N limit. We therefore get

$$\sum_{\substack{j,k=1 \\ \text{odd } j-k}}^{N} \mathtt{M}_{jk}^2(t) \approx \frac{\Theta(2N-t)}{6\pi N t} \sin^2(\mu t)\Big[(4N^2-t^2)^{\frac{3}{2}} - 2(2N^2-t^2)^{\frac{3}{2}}\Theta(\sqrt{2}N-t)\Big]. \tag{5.45}$$

These are to be summed over to obtain the desired trace as in Eq. (5.38). With the trigonometric identity $\cos^2 x + \sin^2 x = 1$ removing all dependence of the expression on the chemical potential μ, we obtain

$$\mathrm{Tr}\big[\mathtt{M}^2(t)\big] \approx \frac{\Theta(2N-t)}{6\pi N t}\Big[(4N^2-t^2)^{\frac{3}{2}} - 2(2N^2-t^2)^{\frac{3}{2}}\Theta(\sqrt{2}N-t)\Big], \tag{5.46}$$

leading directly to

$$K(t) \lesssim 2^{-N} \exp\left\{-\frac{\Theta(2N-t)}{12\pi N t}\Big[(4N^2-t^2)^{\frac{3}{2}} - 2(2N^2-t^2)^{\frac{3}{2}}\Theta(\sqrt{2}N-t)\Big]\right\}, \tag{5.47}$$

We expect this expression to hold for $t > \eta_2 = O(N)$ as the approximations we have made i.e. the approximation for Laguerre polynomials in (5.36), and the assumption that our range covers several oscillations of the polynomials, are valid only for $t \sim O(N)$. If we are interested only in t near $t_* = 2N$, this upper bound also seems to serve as a good approximation to $K(t)$ (observed numerically, as mentioned in Sec. 2.1), and we get a simple form of the approach of the ramp towards the plateau:

$$K(t) \approx 2^{-N} \exp\left[-\frac{(4N^2-t^2)^{\frac{3}{2}}}{12\pi N t}\Theta(2N-t)\right] \tag{5.48}$$

This form is valid in $\sqrt{2}N < t < \infty$, owing to the dropped second term from the exponent of Eq. (5.47), and is used as $K_3(t)$ in Eq. (5.6c).

5.4 General determinantal approach for unitary class random Hamiltonians

The eigenvalue distribution for N levels in an arbitrary unitary random matrix ensemble is expected to be given by (see e.g. Chap.19 in Ref. [36]),

$$P(\varepsilon_1, ..., \varepsilon_N) \propto \left(\prod_k w(\varepsilon_k) \right) \prod_{1 \le i < j \le N} |\varepsilon_i - \varepsilon_j|^2. \tag{5.49}$$

The quadratic factors $|\varepsilon_i - \varepsilon_j|^2$ contain the WD level repulsion for the unitary class, while the weights $w(\varepsilon)$ determine the overall distribution of states. We will assume that $w(\varepsilon)$ is strictly positive in a connected interval, and zero everywhere else if this interval is finite. The special case of Gaussian weights gives the GUE distribution, Eq. (5.8). We can write the eigenvalue distribution in terms of arbitrary polynomials $C_k(x)$ of degree $k-1$ in a manner completely analogous to Eq. (5.22), obtaining

$$P(\varepsilon_1, ..., \varepsilon_N) \propto \left(\prod_k w(\varepsilon_k) \right) \left(\sum_{I,J} \epsilon_{\{I\}}^{\{J\}} \prod_{k=1}^{N} C_k(\varepsilon_{I_k}) C_k(\varepsilon_{J_k}) \right). \tag{5.50}$$

Instead of choosing Hermite polynomials, it is now convenient to choose polynomials $f_n(x)$ that are orthonormal with respect to the weight $w(x)$, i.e.

$$\int \mathrm{d}x \; w(x) f_n(x) f_n(x) = \delta_{jk}. \tag{5.51}$$

This can always be done, for instance, by starting with the functions $\{x^k\}_{k=0}^{\infty}$ and applying the Gram-Schmidt orthonormalization procedure, with the inner product of functions defined with the weight $w(x)$ in the connected interval of nonzero $w(x)$.

To calculate the SFF of the corresponding many-body system, we set the explicit chemical potential μ to zero as it can now be absorbed into the mean of $w(x)$, which we have allowed to be arbitrary.

We can then largely imitate the derivation of Eq. (5.32) for the many-body SFF, replacing the Hermite polynomials with $f_n(x)$ and obtaining

$$K(t) = 2^{-N} \exp\left\{ -\sum_{n=1}^{\infty} \frac{(-1)^n}{n} \mathrm{Tr}\left[\Lambda^n(t)\right] \right\}, \tag{5.52}$$

where we have defined the matrix Λ in analogy with $\mathtt{M}$ as

$$\Lambda_{jk}(t) = \int \mathrm{d}x \; w(x) f_j(x) f_k(x) \cos(xt). \tag{5.53}$$

In what follows, 'large t' or 'late times' in this section will always refer to when t is at the order of magnitude of (the inverse of) the scale of single-particle level spacings, as opposed to (the inverse of) the scale of variation of the single-particle density of states. The latter of these is typically smaller by a factor of $\sim N$ in the time domain.

As with Eq. (5.32), Eq. (5.52) can be identified term-by-term with a cumulant expansion (see Ref. [4]) for the ensemble, which will be useful to keep in mind for subsequent calculations. We assume that at late times (t of the order of single-particle level spacings), there is negligible contribution from terms with odd values of n, leaving us with a sum over only even values of n in Eq. (5.52). The vanishing of the $n = 1$ term at late times is immediate as it corresponds to the Fourier transform of the slowly varying single-particle density of states (e.g. [36]), and we will justify this assumption for odd $n \geq 3$ later in this section. On account of Λ_{jk} being a symmetric matrix, the traces with even n are guaranteed to be positive, so retaining any one term in the sum and dropping all the other terms gives us an upper bound for the SFF.

We again choose to focus on the $n = 2$ term for its simplicity, in which the trace can be written

as

$$\mathrm{Tr}[\Lambda^2(t)] = \int \mathrm{d}x \int \mathrm{d}y \; w(x)w(y)\left(\sum_j f_j(x)f_j(y)\right)\left(\sum_k f_k(x)f_k(y)\right)\cos(xt)\cos(yt). \tag{5.54}$$

We recognize the RMT kernel [36]

$$\mathtt{K}(x,y) = \sum \sqrt{w(x)w(y)}\sum_k f_k(x)f_k(y), \tag{5.55}$$

and note that it is an even function. Determinants constructed from $\mathtt{K}(x,y)$ with the appropriate arguments give all the single-particle correlation functions $\mathtt{R}_n$ of the ensemble, as in Eq. (5.10). We then obtain

$$\begin{aligned}\mathrm{Tr}[\Lambda^2(t)] &= \int \mathrm{d}x \int \mathrm{d}y \; \mathtt{K}(x,y)\mathtt{K}(y,x)\cos(xt)\cos(yt) \\ &= \frac{1}{2}\int \mathrm{d}x \int \mathrm{d}y \; \mathtt{K}(x,y)\mathtt{K}(y,x)\cos((x+y)t) + \frac{1}{2}\int \mathrm{d}x \int \mathrm{d}y \; \mathtt{K}(x,y)\mathtt{K}(y,x)\cos((x-y)t).\end{aligned} \tag{5.56}$$

For any unitary ensemble with the above assumptions for $w(\varepsilon)$, it appears that $\mathtt{K}(x,y)$ has the universal form (for $N \to \infty$) [36],

$$\mathtt{K}(x,y) \approx \frac{\sin(\rho(x)\pi(x-y))}{\pi(x-y)}, \tag{5.57}$$

where $\rho(x)$ is the single-particle density of states, which is ensemble-specific. Eq. (5.57) should ideally symmetrize $\rho(x)$ in x and y, but typically $\rho(x) \approx \rho(y)$ if x and y are close enough for $\mathtt{K}(x,y)$ to be non-negligible (i.e. are comparable to the scale of single-particle level spacings) and we can ignore the symmetrization. The slow variation of $\rho(x)$ at this scale also implies that for large t, $\mathtt{K}(x,y)$ has a negligible Fourier component with respect to $x+y$ or either of x or y, for fixed $x-y$. This allows

us to neglect the first term in Eq. (5.56). We can identify the second term in Eq. (5.56) as essentially the connected part of the single particle SFF,

$$\kappa_c(t) = -\int dx \int dy\ \mathrm{K}(x,y)\mathrm{K}(y,x)\cos((x-y)t), \tag{5.58}$$

which is related to the full single particle SFF $\kappa(t)$ by

$$\begin{aligned}\kappa(t) &= \int d\varepsilon_1 \int d\varepsilon_2\ \mathrm{R}_2(\varepsilon_1,\varepsilon_2)\cos((\varepsilon_1-\varepsilon_2)t) \\ &= N + |\tilde{\rho}(t)|^2 + \kappa_c(t),\end{aligned} \tag{5.59}$$

where $\tilde{\rho}(t)$ is the Fourier transform of the single-particle density of states. Using this identification in the $n=2$ term of Eq.(5.52) leads to the inequality

$$K(t) \leq 2^{-N}\exp\left[\frac{\kappa_c(t)}{4}\right] \to 2^{-N}\exp\left[\frac{\kappa(t)-N}{4}\right], \tag{5.60}$$

where we have assumed that $\tilde{\rho}(t)$ vanishes at the large times of interest, which follows from the slow variation of $\rho(x)$. This inequality generalizes Eq. (5.47) to non-Gaussian ensembles.

We will also show that the many-body SFF has the same plateau time t'_* as the single-particle SFF, and simultaneously establish that all terms with odd n in Eq. (5.52) vanish at late times. We consider the nonvanishing higher order terms in Eq. (5.52), expressed in terms of the RMT kernel in Eq. (5.55) (with $n \geq 3$),

$$\mathrm{Tr}[\Lambda^n(t)] = \int d^n x\ \mathrm{K}(x_1,x_2)\mathrm{K}(x_2,x_3)...\mathrm{K}(x_n,x_1)\prod_{j=1}^{n}\cos(x_j t), \tag{5.61}$$

with $d^n x = \prod_{i=1}^n dx_i$. Let $\tilde{\mathrm{K}}(\tau_i,x_i)$ be the Fourier transform of $\mathrm{K}(x_i,x_{i+1})$ (we adopt the convention

of a periodic subscript i.e. $x_{n+1} = x_1$ and $x_0 = x_n$) with respect to the difference $(x_i - x_{i+1})$, which still depends on x_i (as argued after Eq. (5.57), x_i is interchangeable with x_{i+1} here). We can then write,

$$\mathrm{K}(x_i, x_{i+1}) = \int \frac{\mathrm{d}\tau_i}{2\pi}\, \tilde{\mathrm{K}}(\tau_i, x_i) \cos\big(\big(x_i - x_{i+1}\big)\tau_i\big). \tag{5.62}$$

It is important to emphasize that, as with $\mathrm{K}(x_i, x_{i+1})$, $\tilde{\mathrm{K}}(\tau_i, x_i)$ does not have a high frequency/large-t Fourier component in x_i - all the dependence on $x_i - x_{i+1}$ is contained in τ_i.

Inserting Eq. (5.62) in Eq. (5.61) gives,

$$\mathrm{Tr}[\Lambda^n(t)] = \int \mathrm{d}^n x \int \frac{\mathrm{d}^n \tau}{(2\pi)^n} \left(\prod_{i=1}^{n} \tilde{\mathrm{K}}(\tau_i, x_i)\right) \left[\left\{\prod_{l=1}^{2k} \cos\big(x_l \tau_l - x_{l+1}\tau_l\big)\right\} \prod_{j=1}^{n} \cos(x_j t)\right]. \tag{5.63}$$

The product of all the different cosine factors can be expanded into a sum over cosine terms using standard trigonometry,

$$\prod_{i,j=1}^{n} \cos\big(\big(x_i - x_{i+1}\big)\tau_i\big)\cos(x_j t) = \sum_{\{\zeta_i, s_i = \pm 1\}} 2^{-2n} \cos\left[\sum_{j=1}^{n} \big(\zeta_j t + s_j \tau_j - s_{j-1}\tau_{j-1}\big) x_j\right]. \tag{5.64}$$

We can now further expand each of these cosine terms, again using standard trigonometric identities, as a sum of products of sines and cosines of each $\big[\big(\zeta_j t + s_j \tau_j - s_{j-1}\tau_{j-1}\big) x_j\big]$,

$$\cos\left[\sum_{j=1}^{n} \big(\zeta_j t + s_j \tau_j - s_{j-1}\tau_{j-1}\big) x_j\right] = \sum_{\{\ell_k = \pm 1\}} \mathscr{C}_{\ell_1 \ldots \ell_n} \prod_{j=1}^{n} \mathrm{cs}_{\ell_j}\big[\big(\zeta_j t + s_j \tau_j - s_{j-1}\tau_{j-1}\big) x_j\big], \tag{5.65}$$

where we have introduced notation for a general sinusoid, $\mathrm{cs}_{+1}(x) = \cos(x)$ and $\mathrm{cs}_{-1}(x) = \sin(x)$, and the coefficients satisfy $\mathscr{C}_{\ell_1 \ldots \ell_n} \in \{-1, 0, 1\}$ (their exact values are unimportant for our purposes). Combining Eqs. (5.63), (5.64) and (5.65), and changing the order of sums and integration over x and

τ gives

$$\mathrm{Tr}[\Lambda^n(t)] = \sum_{\{\zeta_i,s_i,\ell_i=\pm1\}} 2^{-2n}\mathscr{C}_{\ell_1\ldots\ell_n}\int\frac{\mathrm{d}^n\tau}{(2\pi)^n}\left(\prod_{j=1}^{n}\int \mathrm{d}x_j\,\tilde{\mathrm{K}}(\tau_i,x_i)\,\mathrm{cs}_{\ell_j}\left[\left(\zeta_j t + s_j\tau_j - s_{j-1}\tau_{j-1}\right)x_j\right]\right). \tag{5.66}$$

Each term in the τ-integrand of Eq. (5.66) has factorized neatly into a product of Fourier cosine or sine transforms of $\tilde{\mathrm{K}}(\tau_i,x_i)$ with respect to x_i. We have already noted that this Fourier component exists only at low frequencies - unless the coefficient of x_j in the argument of the sinusoid (cs_{ℓ_j}) is comparable to the small time scales (say, $\sim t_s$) where $\tilde{\rho}(t)$ is appreciable, the integral vanishes due to the rapidly oscillating cosine. But as t_s is typically much smaller than the plateau time t_*' by a factor of $\sim N$, we can approximate $t_s \sim 0$ at large times, requiring the argument of each sinusoid to vanish. We then obtain the condition,

$$\zeta_j t + s_j\tau_j - s_{j-1}\tau_{j-1} \approx 0,\ \forall j, \tag{5.67}$$

for at least one choice of the sign coefficients $\{\zeta_i,s_i\}$, as a necessary condition for a nonzero value for $\mathrm{Tr}[\Lambda^n(t)]$. Adding up Eq.(5.67) over all values of j gives the constraint $\sum_{j=1}^{n}\zeta_j = 0$ in non-vanishing terms. This condition (valid for $n \geq 3$) is impossible to satisfy for odd n, and together with the vanishing of the $n=1$ contribution, justifies the assumption made earlier that we can ignore the contribution from all odd values of n at late times.

To obtain the plateau time, we note that substituting Eq. (5.62) in Eq. (5.58) gives by the standard Fourier convolution relation (treating x and y as interchangeable),

$$\kappa_c(t) = \int \mathrm{d}x \int \frac{\mathrm{d}\tau}{2\pi}\tilde{\mathrm{K}}(t-\tau,x)\tilde{\mathrm{K}}(\tau,x). \tag{5.68}$$

We know that $\kappa_c(|t| > t_*') = 0$, corresponding to the plateau (with both positive and negative values of t), which requires that $\mathrm{K}(\tau > t_*'/2, s) = 0$. This means that Eq. (5.67) can be satisfied only if $|t| < t_*'$,

which implies

$$\mathrm{Tr}[\Lambda^n(t > t'_*)] = 0. \tag{5.69}$$

In summary, Eq. (5.60) and Eq. (5.69) with Eq. (5.52), taken together, yield the following expressions that apply to the late-time many-body SFF of any unitary Wigner-Dyson ensemble:

$$K(t \leq t'_*) \leq 2^{-N} \exp\left[\frac{\kappa(t) - N}{4}\right], \tag{5.70}$$

$$K(t > t'_*) = 2^{-N}, \tag{5.71}$$

where $\kappa(t)$ is the single-particle SFF with plateau time t'_*.

Chapter 6: Contrasting ergodicity, scrambling, and chaos in many-body systems: Interacting SYK-q

6.1 Overview: Dynamical consequences of the SYK spectrum

In this Chapter, we illustrate how the quantum dynamics of the *interacting* SYK-q model with Majorana fermions, for $q \geq 4$ is affected by its energy spectrum. In particular, we contrast three different properties: cyclic ergodicity, scrambling, and chaos in local operators (as measured by OTOCs). The key takeaways are the following points:

1. The SYK model is maximally chaotic in local operators [163], but takes an exponentially long time to scramble information (via entanglement generation) in large subsystems. This is due to the extensive zero-temperature entropy in its spectrum, and expands on the discussion in Chapter 3.

2. Irrespective of the slow scrambling dynamics, the SYK model is cyclic ergodic to a comparable extent to random matrices in sufficiently small energy windows (within sectors of fixed fermion number parity). The timescale t_0 of each step of the cyclic permutation, associated with these ergodic energy windows, can be significantly smaller than the scrambling time.

6.2 The Majorana SYK model

The conventional form of the SYK model [163] consists of Majorana fermions $\hat{\chi}_j$ on $2N$ sites with random all-to-all 4-body interactions with antisymmetric coefficients $J_{jk\ell m}$, where each indepen-

dent $J_{jk\ell m}$ is randomly chosen from a Gaussian distribution with variance $\langle J^2_{jk\ell m}\rangle = 3!J^2/N^3$; the Hamiltonian is:

$$\hat{H} = \sum_{1\leq j<k<\ell<m\leq 2N} J_{jk\ell m}\hat{\chi}_j\hat{\chi}_k\hat{\chi}_\ell\hat{\chi}_m. \tag{6.1}$$

A slight generalization is the SYK-q model (where q is even) with q-body random interactions with $J_{j_1\ldots j_q}$ chosen from independent Gaussians of variance $(q-1)!J^2/N^{q-1}$, given by

$$\hat{H} = -i^{q/2}\sum_{1\leq j_1<\ldots<j_q\leq 2N} J_{j_1\ldots j_q}\hat{\chi}_{j_1}\cdots\hat{\chi}_{j_q}. \tag{6.2}$$

We are only concerned with a single representative realization of the interaction coefficients $J_{j_1\ldots j_q}$ chosen from the above distribution with $J = 1$, and do not consider any form of ensemble averaging. Our uses of "SYK model" and similar phrases refer, in this Chapter, to any typical individual realization of the SYK ensemble. However, we emphasize that for the time being, ensemble averaging seems essential for a path integral treatment of SYK, from which several of its well-known properties can be analytically derived [163, 164, 179]. In any case, we will not adopt a path integral approach in this chapter, but directly consider the dynamics of quantum states.

A key feature of the SYK model is that its regularized OTOCs (essentially, combinations 4-point correlation functions of certain regularized single-site operators) grow exponentially fast at the maximum allowed rate [159], due to which the model is considered to be "maximally chaotic" and a "fast scrambler". While this statement remains true if one considers scrambling in the sense of growth of OTOCs, the scrambling of information via entanglement generation over large subsystems can be quite different as noted in Sec. 3.3.4 (see also Ref. [143]). In particular, the criterion for a rapid growth of OTOCs merely to $O(1)$ values may not have sufficient resolution to capture microscopic deviations from scrambling at a resolution comparable to D_S^{-1} for large subsystems. We illustrate this statement in detail in Sec. 6.3 below.

Another important feature is that the SYK model shows the local level statistics of the Wigner-Dyson random matrices in appropriate subspaces. This has been studied, for example, in Refs. [82,86]. In Sec. 6.4, we numerically show the link between spectral statistics and observable-independent ergodic dynamics in these subspaces. This serves to illustrate the theory of Chapter 2 in a Hamiltonian many-body system.

6.3 Slow scrambling via entanglement generation in the SYK model

In this section, we focus on the scrambling of information in the Sachdev-Ye-Kitaev (SYK) model [163, 164] discussed in Sec. 3.1 (see Figs. 6.1, 6.2 below for a quick numerical overview of the main takeaways).

As discussed in Sec. 3.1, with details provided below, the density of states of the SYK model is such that the sustained scrambling time of large subsystems is at least exponentially long in the subsystem size, while a generic strongly interacting many-body system with a Gaussian density of states and sufficiently "random" eigenstates would scramble much faster as a square root of the subsystem size. Qualitatively, this is because of slowly decaying oscillations in the SFF $K(t)$ for this model caused by sharp edges in its density of states, which through the inequality Eq. (3.20) forces the mean return probability of initial states in such subsystems to "remember" the initial state at least until these oscillations decay to a sufficiently small value.

6.3.1 Defining subsystems in the SYK model

We would like to consider information scrambling in subsystems of this model. To define suitable subsystems, it is convenient to transform from the $2N$ Majoranas to complex fermions on N sites, given

by:

$$\hat{c}_j = \frac{1}{\sqrt{2}}\left(\hat{\chi}_{2j} - i\hat{\chi}_{2j+1}\right), \tag{6.3}$$

$$\hat{c}_j^\dagger = \frac{1}{\sqrt{2}}\left(\hat{\chi}_{2j} + i\hat{\chi}_{2j+1}\right). \tag{6.4}$$

These admit a vacuum state $|0\rangle$ such that $\hat{c}_j|0\rangle = 0$ for all j, and the Hilbert space $\mathscr{H}$ is spanned by orthonormal Fock states:

$$|n_1, n_2, \ldots, n_N\rangle \equiv (\hat{c}_N^\dagger)^{n_N} \ldots (\hat{c}_2^\dagger)^{n_2} (\hat{c}_1^\dagger)^{n_1} |0\rangle, \text{ for } n_j \in \{0, 1\}, \tag{6.5}$$

where we have chosen a specific ordering of fermion operators for definiteness, and reordering them only changes the overall sign of each state. It is important to emphasize that the Hamiltonian $\hat{H}$ in Eq. (6.2) does not conserve the particle number in terms of these fermions, and is quite unlike the closely related complex SYK model [179] which does conserve fermion number. In particular, the nonconservation of fermion number allows the scrambling of subsystems comparable to infinite temperature thermalization as considered in Sec. 3.1.

Now we can formally factorize the Hilbert space, exactly like a system of qubits, into a subsystem $\mathscr{H}_S$ of dimension $D_S = 2^{N_S}$ representing N_S sites, for instance, $\{1, \ldots, N_S\}$, and a factor $\mathscr{H}_E$ of dimension $D_E = 2^{N_E}$ representing the remaining $N_E = N - N_S$ sites:

$$|n_1, n_2, \ldots, n_N\rangle = |n_1, \ldots, n_{N_S}\rangle_S \otimes |n_{N_S+1}, \ldots, n_N\rangle_E. \tag{6.6}$$

We also recall that while the fermion ladder operators $\hat{c}_j$ can not be factorized in this manner between sites due to their mutual anticommutation relations, the fermion number operators $\hat{n}_j = \hat{c}_j^\dagger \hat{c}_j$ are mutually commuting and do factorize into those with $j \leq N_S$ acting on $\mathscr{H}_S$ alone, and the remaining

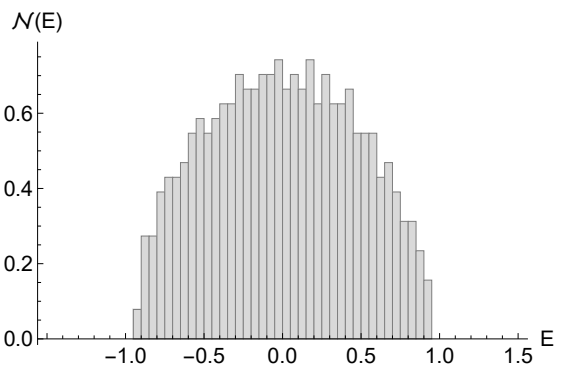
acting on $\mathscr{H}_E$ alone.

Now, we can consider $P_S(t)$ for a set of D_S initial states $\hat{\Pi}_k$ corresponding to the Fock basis states within $\mathscr{H}_S$, with the state in $\mathscr{H}_E$ being maximally mixed (representing complete uncertainty in the configuration of $\mathscr{H}_E$):

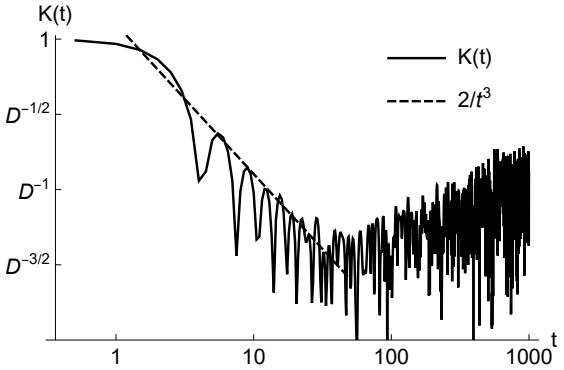
$$\hat{\Pi}_k = |n_{1(k)}, \ldots, n_{N_S(k)}\rangle_S \langle n_{1(k)}, \ldots, n_{N_S(k)}| \otimes \hat{\mathbb{1}}_E. \tag{6.7}$$

Eq. (3.20) states that $P_S(t) \geq K(t)$, and therefore, as discussed in Sec. 3.1, the scrambling time of the SYK model depends on the asymptotic structure of $K(t \to \infty)$, which we describe next.

6.3.2 Power-law decay of the SFF



(a) Density of states (per level) $\mathscr{N}(E)$, coarse grained over bins of width 0.05 in energy units where $J = 1$, for the specific realization of the SYK model considered here, showing the sharp edges.



(b) Log-log plot of SFF $K(t)$ compared with a t^{-3} decay of early oscillations after the first decay, verifying Eq. (6.8). The late-time fluctuations are the ramp and (barely visible) plateau which remain comparable to $O(D^{-1})$; these would appear much more negligible in the $D \to \infty$ limit, and have no direct effect on the scrambling of subsystems of dimension $D_S \ll D$.

Figure 6.1: Density of states and SFF for the same realization of the SYK model as in other plots; $J = 1$, $N = 10$ (with $D = 2^N = 1024$).

Being a Hamiltonian system, the SYK model satisfies $K(t) = |\widetilde{\mathscr{N}}(t)|^2$, where $\mathscr{N}(E)$ is the density of states as discussed in Sec. 3.1 [around Eq. (3.11)]. It has been noted [38, 82] that $\mathscr{N}(E)$ has sharp edges for the SYK model, which leads to a power-law asymptotic behavior of the SFF. In particular,

one obtains [38, 82] (see also Fig. 6.1)

$$K_{\mathrm{SYK}}(t \to \infty) \sim \frac{\Theta(D^{-\alpha})}{t^3} + O(D^{-1}). \tag{6.8}$$

Here, α is related to the zero-temperature entropy $\mathscr{S}_0$ of the SYK model [38] via $e^{\mathscr{S}_0} = 2^{N(1-\alpha/2)}$, and behaves as $\alpha \propto q^{-2}$ for large q with $\alpha \approx 0.66$ for $q = 4$ [163]. This power-law behavior is related to Statement 3.2 of Sec. 3.1 concerning a slower-than-exponential decay of the SFF in systems with finite spectral edges, with the slight difference that the spectral edges of the SYK model occur at asymptotically large energies in the $N \to \infty$ limit [38].

While numerical evidence for finite N (Sec. 6.3.3, Fig. 6.2) suggests that subsystems of the SYK model scramble to $P_S(t) = D_S^{-1} + o(D_S^{-1})$ for the system size considered here, we will nevertheless derive the scrambling time limit for any $P_S(t) = O(D_S^{-1})$ to maintain generality (this can be extended to any $P_S(t) = O(D_S^{-\kappa})$ with similar qualitative conclusions). By the inequality Eq. (3.20), this requires that $K(t > t_s) \leq O(D_S^{-1})$, which gives

$$\frac{\Theta(D^{-\alpha})}{t_s^3} \leq O(D_S^{-1}) \implies t_s = \Omega\left[\left(\frac{D_S}{D^\alpha}\right)^{1/3}\right] = \Omega\left[2^{(N_S - \alpha N)/3}\right]. \tag{6.9}$$

This yields Eq. (3.14) of Sec. 3.1 when combined with the observation that we need $N_S > \alpha N$ for $t_s \to \infty$ (as we used the $t \to \infty$ asymptotics of the SFF), which is the regime in which this bound provides a nontrivial constraint; thus, the subsystems of interest must be at least a finite (but increasingly small with q) fraction of the system size, similar to the requirement in the definition of scrambling adopted in Ref. [141].

As any basis for a subsystem (such as the $\hat{\Pi}_k$ in Eq. (6.7)) must involve all energies by the completeness relation ($\sum_k \hat{\Pi}_k = \hat{\mathbb{1}}$), we are forced to include the entire spectrum including the edges and cannot eliminate the power-law decay by restricting to the center of the spectrum using, e.g., a

Gaussian filter [$g(\hat{H}) \propto e^{-c\hat{H}^2}$], as done for this model in a different context in Ref. [38]. While such filtering is mathematically possible in our inequality Eq. (3.20), it is the physical choice of initial states that precludes this possibility (as noted in the discussion following Eq. (3.37)) and leads to Eq. (6.9).

We should also emphasize that this power law decay entirely concerns the initial decay ("slope") region of the SFF, and does not describe the behavior of the late-time quantum fluctuations ("ramp" or "plateau", contained in the $O(D^{-1})$ term). The late-time regions are not particularly useful in setting rigorous speed limits on scrambling, for the following reasons. Firstly, $K(t) = O(D^{-1})$ in these regions, which is much less than $P_S(t) \sim D_S^{-1}$ for the infinite temperature scrambling of any subsystem of dimension $D_S \ll D$ (for which it is sufficient that $N_E = N - N_S \gg 1$, even if $N_S/N = \Theta(1)$). Secondly, a key observation in the study of many-body chaos (e.g., Ref. [38]) is that thermalization generally occurs well before much of the "ramp" region in Hamiltonian systems including the SYK model; in fact, the time of onset of the quantum fluctuations corresponding to the ramp is comparable to the time required for complete thermalization in the full system in many cases [38], whereas smaller subsystems would thermalize much in advance of this time.

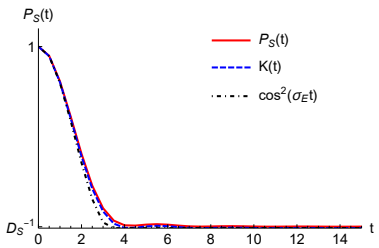
6.3.3 Scrambling in the SYK model: Numerical details

Numerics were run using an iterative procedure (with increasing N) to generate a matrix representation for the $2N$ Majoranas $\hat{\chi}_j$ described in Ref. [195], in a specific D-element orthonormal basis $\mathscr{B} = \{|B_k\rangle\}_{k=1}^{D}$, from which the matrix elements of the Hamiltonian are readily obtained using Eq. (6.2). Matrix representations of $\hat{c}_k$, $\hat{c}_k^\dagger$ are then obtained from Eqs (6.3), (6.4), for which the vacuum state (here referring to the state with no fermion occupancy, annihilated by all the $\hat{c}_j$) is given by

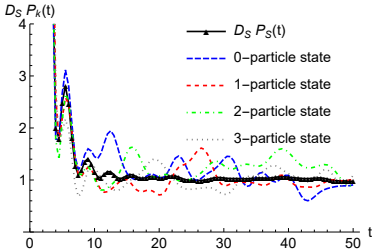
$$|0\rangle = |B_{1+\frac{D}{2}}\rangle. \tag{6.10}$$

This allows generating all the fermion Fock states in the $\mathscr{B}$-basis using Eq. (6.5). The labels 1 to N of the fermion sites is fixed by the above procedure to generate the matrix elements of $\hat{c}_k$, $\hat{c}_k^\dagger$, for which we choose the first N_S sites as our subsystem of interest as in Eq. (6.6). The initial basis states $\hat{\Pi}_k$ are given by Eq. (6.7) re-expressed in the $\mathscr{B}$-basis.

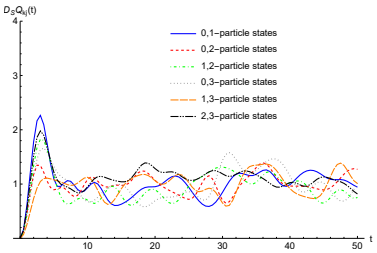
We choose $q = 4$, with $N = 10$ fermion sites (or $2N = 20$ Majoranas) and $N_S = 7$ for numerical illustration ($D = 1024$, $D_S = 128$, $D_E = 8$), with a randomly generated single realization of the disorder coefficients $J_{jk\ell m}$ in Eq. (6.1). The choice of $N_S = 7$ (so that $D_S = 1/128$) is because the first "power-law" oscillation of the SFF (after the initial decay) has a peak value of approximately $1/100$ for $N = 10$, and we require $D_S^{-1} < 1/100$ to observe a nontrivial effect of this peak in delaying scrambling to $P_S(t) = D_S^{-1}$; we note that this choice also satisfies the condition $N_S > \alpha N$ in Eq. (3.14) of Sec. 3.1, with $\alpha \approx 0.66$ for $q = 4$ [163]. Higher values of N or N_S require significantly higher computational time due to exponentially growing complexity, whereas the above values are sufficient to provide a basic illustration of the bound of Eq. (3.20).



(a) Linear-linear plot of $P_S(t)$, $K(t)$ and $\cos^2(\sigma_E t)$ (the latter for $0 \leq \sigma_E t < \pi/2$), depicting the sensitivity required for many-body thermalization in large subsystems. The first t^{-3} "recurrence" of the SFF around $t \approx 6$, whose nontrivial effect is clearly visible in the log-linear Fig. 1 of Sec. 3.1, is barely discernible in this plot.



(b) Scaled mean return probability $D_S P_S(t)$ and return probabilities $D_S P_k(t)$ for the states defined in Eq. (6.13), suggesting that subsystems in this model satisfy $P_k(t \to \infty) \to 1/D_S$, and in particular $P_S(t \to \infty) \to 1/D_S$.



(c) Scaled overlaps $D_S Q_{kj}(t)$ between specific pairs of states defined in Eq. (6.13), suggesting that $Q_{kj}(t \to \infty) \to 1/D_S$, and therefore providing evidence that scrambling to infinite temperature in the sense of Eq. (3.35) does appear to occur in this model.

Figure 6.2: Additional numerics (a) demonstrating the high sensitivity of the speed limits such as $P_S(t) \geq K(t)$ required for a many-body system, and (b, c) providing evidence for scrambling to infinite temperature in fermion subsystems of the SYK model. All plots use the same realization of the interaction strengths $J_{jk\ell m}$ (each chosen from a Gaussian distribution as described near Eq. (6.2)) used in Fig. 1 of Sec. 3.1.

Fig. 1 of Sec. 3.1 illustrates our primary inequality Eq. (3.20); here, we reproduce a linear-linear plot of the same data to highlight the sensitivity of the bound to microscopic fluctuations in the SFF and many-body thermalization [Fig. 6.2a]. We also provide additional numerical evidence in Fig. 6.2 that the behavior of the SYK model appears to be consistent with the strong definition of scrambling adopted in Sec. 3.1 based on Ref. [141] (corresponding to infinite temperature thermalization). For this, we consider the individual return probabilities,

$$P_k(t) \equiv \frac{1}{D_E} \mathrm{Tr}[\hat{\Pi}_k(t)\hat{\Pi}_k(0)], \tag{6.11}$$

and overlaps between a time-evolved state and a distinct initial state:

$$Q_{kj}(t) \equiv \frac{1}{D_E} \mathrm{Tr}[\hat{\Pi}_k(t)\hat{\Pi}_j(0)]. \tag{6.12}$$

By Eq. (3.35) (or Eq. (3.7) in Sec. 3.1) both of these quantities should become $1/D_S$ to leading order. We see this to be the case in Figs. 6.2b, 6.2c. Further, the oscillations of $D_S P_k(t)$ and $D_S Q_{kj}(t)$ are consistent with $O(D_E^{-1/2})$ (in this case, $D_E^{-1/2} = 1/\sqrt{8} \approx 0.35$), as expected for the random-matrix-like thermalization behavior in Eq. (3.10) of Sec. 3.1.

In Fig. 6.2, $P_S(t)$ refers to the mean return probability of all subsystem initial states of the form Eq. (6.7), while the specific other states involved in $P_k(t)$ and $Q_{kj}(t)$ are as follows:

$$\begin{aligned}
&\text{0-particle state}: \ \hat{\Pi}_{k0} = |0,0,0,0,0,0,0,0\rangle_S\langle 0,0,0,0,0,0,0,0| \otimes \hat{\mathbb{1}}_E, \\
&\text{1-particle state}: \ \hat{\Pi}_{k1} = |0,0,0,0,0,0,0,1\rangle_S\langle 0,0,0,0,0,0,0,1| \otimes \hat{\mathbb{1}}_E, \\
&\text{2-particle state}: \ \hat{\Pi}_{k2} = |0,0,0,0,0,0,1,1\rangle_S\langle 0,0,0,0,0,0,1,1| \otimes \hat{\mathbb{1}}_E, \\
&\text{3-particle state}: \ \hat{\Pi}_{k3} = |0,0,0,0,1,1,1\rangle_S\langle 0,0,0,0,1,1,1| \otimes \hat{\mathbb{1}}_E.
\end{aligned} \tag{6.13}$$

The terminology of "n-particle states" here only refers to the number of particles in the $N_S = 7$ sites of the subsystem, while the number of particles in the remaining $N_E = 3$ sites remains completely uncertain in these states (alternatively, each such projector corresponds to a set of all pure product states with a specific configuration in the subsystem, see Sec. 3.3.2). We note that it is possible for these states to appear to completely "mix" into each other under time evolution and thereby appear to scramble to infinite temperature (as suggested by the numerics of Fig. 6.2), due to the lack of fermion number conservation in the SYK model.

6.4 Dynamical ergodicity and spectral statistics in the SYK model

Now, we turn our attention to the dynamical ergodic properties of the SYK model, illustrating some of the ideas in Chapter 2.

6.4.1 General considerations

First, we note that Eq. (6.2) for decomposes the Hilbert space into two subspaces of fixed fermion number parity $\zeta \in \{0, 1\}$ that remain independent under time evolution [82]:

$$\mathscr{H} = \mathscr{H}_{P0} \oplus \mathscr{H}_{P1}, \tag{6.14}$$

where

$$\mathscr{H}_{P\zeta} = \text{span}\left\{ |n_1, n_2, \ldots, n_N\rangle : \left(\sum_{r=1}^{N} n_r\right) \mod 2 = \zeta \right\}. \tag{6.15}$$

This is because each term in $\mathscr{H}$ has an even number of Majorana operators $\hat{\chi}_k$ which can either create or annihilate a fermion at the site k, but not leave its number invariant. Thus, each term of $\mathscr{H}$ necessarily changes the fermion number by a multiple of 2, decoupling $\mathscr{H}_{P0}$ and $\mathscr{H}_{P1}$.

This decoupling is also associated with an antiunitary particle-hole symmetry [82], whose be-

havior leads to any of the three different types of Wigner-Dyson statistics within each $\mathscr{H}_{P\zeta}$, depending on N:

1. For odd N, small energy windows in each sector show the local level statistics of CUE.

2. For even N, the two possibilities $N \mod 4 = 0$ and $N \mod 4 = 2$ respectively lead to local COE and CSE level statistics in each sector.

For definiteness, we take odd N and work with the even parity sector $\mathscr{H}_{P0}$. The density of states within this sector is comparable to that illustrated in Fig. 6.1b — Gaussian near the center with $\sigma_E \sim O(N^{1/2})$ width (in the $N \to \infty$ limit), giving way to sharp $\sqrt{E}$ edges at energies of $O(N)$ with an extensive zero temperature entropy responsible for a subleading t^{-3} decay of the SFF within this subspace.

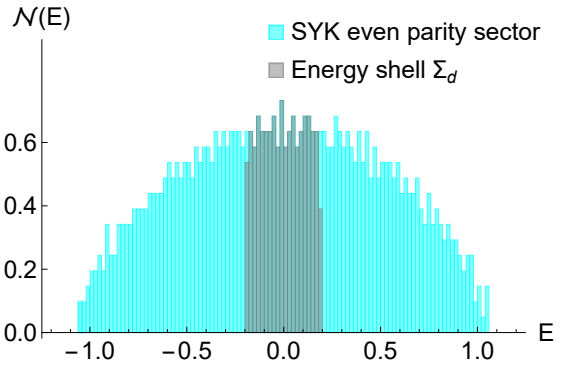
Following the discussion in Secs. 2.3 and 2.4.1, we should choose a subspace $\Sigma_d \subseteq \mathscr{H}_{P0}$ of a suitable energy width $\Delta E \lesssim 2\pi/t_{\text{ramp}}$ to examine cyclic ergodicity. However, the ramp time t_{ramp} varies over the spectrum: the power-law decay from the sharp edges has no bearing on the ramp time in the center of the spectrum. The standard approach to handle this situation [38] is to focus on the ramp time near the center of the spectrum, and multiply the spectrum with a narrow Gaussian filter of width $\sigma_g \sim O(N^{-1/2})$ to exponentially suppress the contribution from the edge. Such a Gaussian filter has the disadvantage that the effective SFF decays as a Gaussian of width $2\pi/\sigma_g$, which takes a time of at least $t_g \gtrsim \sqrt{N}/\sigma_g$ — parametrically larger than the inverse width σ_g corresponding to the energy scale of interest — to decay to the scale of $O(D^{-1})$ quantum fluctuations beyond which we can identify t_{ramp} for that region of the spectrum.

Here, we will follow an alternate strategy based on Eq. (2.41), where we choose a sharp energy window of any width $\Delta E_w \sim o(N^{1/2})$, automatically cutting off the edges of the spectrum, but looking at the SFF only at discrete times $K(t = 2\pi p/\Delta E_w)$, corresponding to the zeros of the $\text{sinc}^2(t\Delta E_w/2)$

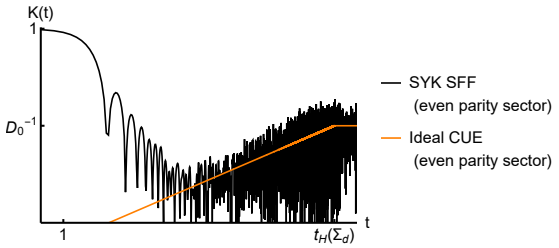
transient that results from the sharp cutoff. By avoiding the slow decay of the Gaussian filter and focusing on the perfectly predictable zeros of the sinc transient, we can identify a ramp time as small as $t_{\text{ramp}} = 2\pi/\Delta E_w \sim \omega(N^{-1/2})$. In particular, if we find that $t_{\text{ramp}} = 2\pi/\Delta E_w$ for a given choice of E_w, this marks an appropriate choice of energy window $\Delta E = \Delta E_w$ in which we can expect to see CUE statistics in discrete time steps. This procedure directly ties the lowest possible ramp time to the width ΔE_w of the energy window of interest, without the $\sqrt{N}$ enhancement of this minimum that makes a Gaussian filter suboptimal (this is illustrated in Fig. 6.4).

6.4.2 Cyclic ergodicity in the SYK model: Numerical details

For the numerical simulation, we take a single $N = 11$ realization of SYK-4, to obtain local CUE statistics in the even parity sector $\mathscr{H}_{P0}$ of dimension $D_0 = 1024$. The energy shell ΔE is chosen to be made up of the central $d = 256$ levels in the spectrum, within which the density of states appears to be reasonably uniform; see Fig. 6.3.



(a) Density of states (per level) $\mathscr{N}_0(E)$ in $\mathscr{H}_{P0}$, coarse grained over bins of width 0.05 in energy units where $J = 1$, for the specific realization of the SYK model considered here. Also depicted is the chosen energy shell Σ_d with $d = 256$.



(b) SFF (log-log) in the even parity sector $\mathscr{H}_{P0}$ compared with the ideal (ensemble-averaged) CUE SFF $D_0^{-1}\min(t/t_H, 1)$, with $t_H = 2\pi\mathscr{N}^{\#}(\Sigma_d)$ where $\mathscr{N}^{\#}(\Sigma_d)$ is the mean (total) density of states in Σ_d

Figure 6.3: Density of states, selected energy window, and SFF for the even parity sector of $N = 11$ SYK-4.

Within Σ_d, t_{ramp} is smaller than that for the full Hilbert space $\mathscr{H}$; this is because the zero-temperature entropy contribution from the spectral edges is entirely absent from Σ_d. This is shown in

Fig. 6.4.

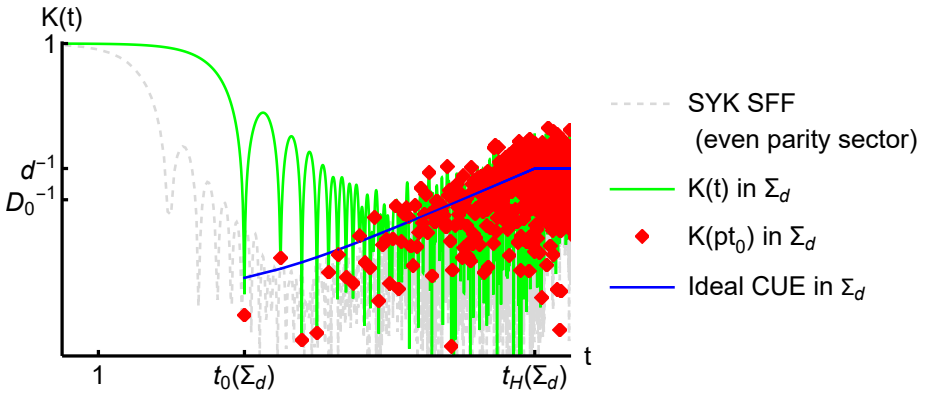


Figure 6.4: The SFF in a realization of the $N = 11$ SYK-4 model within the energy window $\Sigma_d \subset \mathscr{H}_{P0}$ of width ΔE and dimension $d = 256$ inside the even parity sector of dimension $D_0 = 1024$, both in continuous time $K(t)$ and discrete time $K(pt_0)$ with $p \in \mathbb{Z}$ and $t_0(\Sigma_d) = 2\pi/\Delta E$, compared with the ideal CUE SFF within Σ_d. The sinc^2 transient is clearly seen in the continuous time curve (see Eq. (2.41)). It is also seen that $t_{\text{ramp}}(\Sigma_d) = t_0(\Sigma_d)$, and the discrete time SFF $K(pt_0)$ is entirely consistent with the CUE expectation, suggesting that the spectrum is entirely CUE-like in Σ_d. Also shown (dashed, light gray) is the full SFF in the even parity sector which includes contributions from the spectral edge. Further, $t_{\text{ramp}}(\mathscr{H}_{P0})$ appears to be around at least $3t_0(\Sigma_d)$ in the present numerics.

Having established CUE-like behavior within Σ_d, we test for cyclic ergodicity in Σ_d with the time step $t_0(\Sigma_d)$ in Fig. 6.5. Good agreement of the persistence $z^2(p, t_0)$ is seen with the ideal Gaussian expectation for CUE (see Chapter 2),

$$z^2_{\text{CUE}}(p) = \exp\left[-\frac{2\ln d}{d^2}t^2\right], \tag{6.16}$$

which neglects $O(d^{-1})$ quantum fluctuations for simplicity.

Finally, we comment on the maximum allowable size of the energy window ΔE in which CUE behavior, and therefore cyclic ergodic and aperiodicity are likely to persist. Ref. [38] estimates that t_{ramp} at the center of the SYK spectrum grows at best as a "weak" function of N; correspondingly, $\Delta E \lesssim 2\pi/t_{\text{ramp}}$ shrinks no faster than some slow rate with N. In principle, if there are no other physical processes that restrict spectral rigidity at the center of the spectrum, CUE correlations could stretch up to nearly the entire width of the spectrum (as seen in GUE, for example), such as some

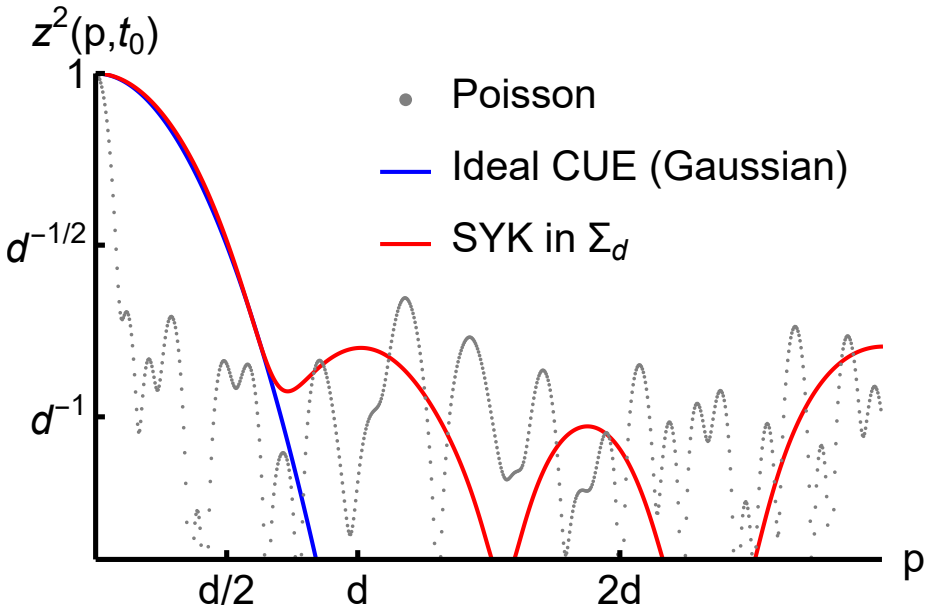


Figure 6.5: Cyclic ergodicity in a realization of the $N = 11$ SYK-4 model, for the energy shell Σ_d with $d = 256$ at the center of the even parity sector $\mathscr{H}_{P0}$ of dimension $D_0 = 1024$. The persistence for a random Poisson (uncorrelated) spectrum in $[0, 2\pi)$ is also plotted to provide a visual guide to the size of $O(d^{-1/2})$ fluctuations. Close agreement is seen with the ideal Gaussian CUE behavior of Eq. (6.16), and this is consistent with cyclic ergodicity and aperiodicity in this energy shell.

$\Delta E = o(N^{1/2})$. This means that the time step $t_0 = 2\pi/\Delta E$ for cyclic ergodicity in the center of the SYK spectrum is $t_0 = \omega(N^{-1/2})$ and at best some weakly growing function of N. This should be contrasted with the exponentially long scrambling time scale $t_s \sim \exp[\Theta(N)]$ expected for subsystems with $N_S = \Theta(N)$.

In summary, we have demonstrated slow scrambling via entanglement generation and cyclic ergodicity in the SYK model, and further that the discrete time steps for cyclic ergodicity can occur over time steps significantly shorter than the scrambling time scale of large subsystems. Overall, this behavior in the SYK model reinforces the notion that significant differences exist between various quantum dynamical phenomena associated with statistical mechanics, and the physical consequences of each must be carefully investigated in its own right.

Chapter 7: Quantization-dependent ergodicity in maximally chaotic classical maps: Quantized Bernoulli shifts

7.1 Introduction: Exceptional spectral statistics in maximally chaotic systems

7.1.1 Background and motivation

The connection between the statistics of energy levels and a variety of ergodic phenomena is a foundational problem in the study of quantum signatures of chaos [28] and the statistical mechanics of quantum many-body systems [21]. In *generic* systems with a classical limit or many-body structure, an empirically successful approach has been to look for signatures of eigenvalue statistics associated with random matrix theory (RMT) [36], in order to diagnose "ergodicity" if these are present [30–33, 35, 39–41, 73, 78–80, 82, 84, 86, 88, 154], and infer its absence otherwise [4, 34, 87, 140, 148, 196]; indeed, the presence of "ideal" RMT statistics can be shown to be sufficient (but not necessary) for an ergodic exploration of an orthonormal basis in the Hilbert space of a general quantum system [1] (see Chap. 2). However, for a complete understanding of the utility of eigenvalue statistics, it is essential to quantitatively characterize deviations from this idealized behavior, particularly to identify where such statistics no longer accurately diagnose different forms of ergodicity/thermalization.

There are a number of interesting systems where deviations from RMT have been observed that point to an increasing need to characterize non-RMT behavior [1, 34, 197–204]. Barring specific cases with alternate explanations, these deviations are generally due to *emergent* quantum symmetries not present in the classical system [28], usually connected to the classical periodic orbits, that lead

to ergodicity-breaking after the Ehrenfest time [128, 129] at which classical and quantum evolutions diverge significantly. Prominent examples include the modular multiplication [7, 199, 205] and cat maps [197,198] (which become exactly periodic in their standard quantization), and chaotic dynamics on arithmetic domains in hyperbolic surfaces [201, 202], where RMT statistics are present only if specific boundary conditions are imposed on quantization [203], being strongly violated by emergent Hecke symmetries [201] otherwise.

In this Chapter, we identify and characterize anomalies in spectral statistics of a different (and essentially opposite) nature to the above systems, originating in the quantum mechanical breaking of *discrete* symmetries that are rigorously present at the macroscopic scale. The existence and relevance of such anomalies is suggested, for instance, by studies of certain exceptional billiard systems [206–209]. Specifically, we consider quantizations of maximally chaotic quantum maps in which we show that discrete macroscopic symmetries are *not* accurately reflected in the most commonly used measures of spectral statistics: (1) the spectral form factor (SFF) [28, 40], which measures spectral rigidity over different energy scales as a function of time (namely, long-range at early times, and short-range at late times), (2) the (short-range) nearest neighbor level spacing statistics [28,30–33], and (3) the adjacent gap ratios [210] (characterizing the short-range next-nearest-neighbor statistics). The short-range statistics in particular show especially stark violations. These violations are striking in the context of the use of spectral statistics to *identify* discrete symmetries of the time evolution operator. While such diagnostics are effective in a variety of systems exhibiting block RMT behavior [73, 211–214], our results show they cannot always be relied upon, even in simple systems with a well-defined classical limit.

7.1.2 Summary of this paper

We aim to illustrate the unreliability of common spectral statistics in identifying discrete symmetries as may be present in quantized chaotic maps or many-body systems. To ensure that the systems being compared have an identical and well-understood macroscopic behavior, we consider classical maps that are *known* to have two discrete symmetries that square to unity (i.e., restore the original system on acting twice). More specifically, we study spectral statistics in different quantizations [215, 216] of the *A*-baker's maps, which are classically paradigmatic examples of ergodicity with maximally chaotic (Bernoulli) behavior [48]. Incidentally, in addition to having a classical limit, these quantizations are particularly amenable to implementation as many-body Floquet quantum circuits [217, 218]. Further, all these quantizations reduce in the classical limit to the same classical *A*-baker's maps, and thus possess the same two discrete symmetries (Sec. 7.2). These are: a canonical reflection symmetry, and an anticanonical time-reversal symmetry, which respectively correspond to a unitary reflection and antiunitary time-reversal operator on quantization.

Our main qualitative results, described more thoroughly in Sec. 7.3, are as follows. While the spectral statistics of some of these quantizations are already known to be "unusual", a key observation in this Chapter is that these unusual features can be satisfactorily organized in terms of a simple, and potentially generalizable, picture of different levels of "discrete symmetry breaking" in the spectral statistics. These anomalies are to be evaluated relative to the following general expectation based on RMT [28, 36]: the presence and absence of the time-reversal symmetry respectively correspond to COE and CUE level statistics, with the presence or absence of the reflection symmetry indicating a 2-block or 1-block structure of the associated random matrix. In particular, based on their classical symmetries, quantized *A*-baker's maps would be expected to have the spectral statistics of 2-block COE. With this context, we identify spectral anomalies of two types (Sec. 7.3.1, 7.3.2):

1. "Weak anomalies" whose primary effect is in the regime of long times corresponding to short-range energy spacings, leading to full (single-block) RMT-like behavior in the mean gap ratio statistic and nearest neighbor level spacings for large A (ranging from 1-block COE to CUE). However, the early-time SFF is consistent with the presence of the unitary reflection symmetry (2-block COE). These demonstrate that short-range measures can be misleading indicators of discrete symmetries.

2. "Strong anomalies" that affect even the regime of early times and long-range energy spacings in addition to the long-time regime as above, where even the early-time SFF shows RMT behavior consistent with the absence of unitary symmetries (1-block COE). These show that even long-range measures can be misleading in certain circumstances.

Subsequently, we study the connection between these spectral anomalies and dynamics. We show numerically that strong anomalies emerge from the inclusion of additional phases in specific quantizations [215], which have no impact in the classical limit, but occur as Berry-like phases in the semiclassical periodic orbit expression (see Ref. [5] for the latter). Further, we numerically study quantum dynamics in the Hilbert space in the sense of cyclic ergodicity [1], and find that strong anomalies appear to induce cyclic ergodicity where weak anomalies do not, verifying the direct connection between long-range spectral statistics and ergodic quantum dynamics irrespective of classical symmetries (Sec. 7.3.3). The remaining sections offer additional analytical and numerical details concerning these results.

7.2 Models

In this section, we introduce the classical and quantum systems.

Classical maps— The classical maps we consider are the A-baker's maps [48, 215, 219], which

act on the 2-torus $\mathbb{T}^2 = \mathbb{R}^2/\mathbb{Z}^2$ (identified with the unit square) via

$$(q,p) \mapsto \left(Aq - \lfloor Aq \rfloor, \frac{p + \lfloor Aq \rfloor}{A}\right), \tag{7.1}$$

for $(q,p) \in [0,1) \times [0,1)$ and $A \geq 2$ an integer. When $A = 2$, this is the same as the usual baker's map. We depict the action of the A-baker's map for $A = 3$ on the unit square in Fig. 7.1.

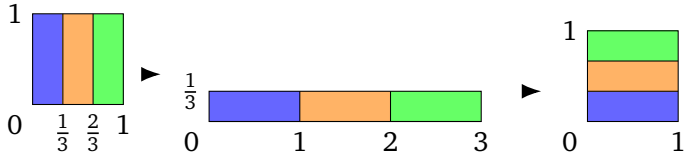


Figure 7.1: Visualization of the action of the 3-baker's map, starting from the left unit square and ending with the right unit square. The intermediate step shows the stretching, cutting, and stacking operation described by Eq. (7.1).

The classical A-baker's map is equivalent to a 2-sided Bernoulli shift [48, 215, 220] and is thus maximally chaotic. It represents a fairly "universal" model of chaotic dynamics, as any K-mixing (ergodic *and* chaotic) system with sufficiently large Kolmogorov-Sinai entropy [12] (essentially the sum of nonnegative Lyapunov exponents) $h \geq \ln A$ can be coarse-grained into a given A-baker's map (or most directly, the corresponding Bernoulli shift), by the Sinai factor theorem [221, 222]. The A-baker's maps possess two symmetries, a time-reversal (TR) symmetry $T : (q,p) \mapsto (p,q)$ and a reflection symmetry

$$R : (q,p) \mapsto (1-q, 1-p), \tag{7.2}$$

which will play key roles in our analysis. Due to the time-reversal symmetry, one expects the RMT symmetry class for the corresponding quantized systems to be that of the circular orthogonal ensemble (COE). Additionally, due to the reflection symmetry, one expects two distinct COE symmetry classes, leading to an overall behavior resembling a direct sum of two COE matrices. As we will demonstrate, however, these general expectations need not hold even approximately when the classical symmetries

are broken upon quantization.

We briefly note that while we do not exhaustively verify the absence of any further classical symmetries for $A > 2$, numerical results counting values of the classical action on orbits suggest the above symmetries are likely the only two. In any case, however, the main conclusions of this paper that spectral statistics can fail to detect symmetries would still hold.

Balazs–Voros, Saraceno, and generic quasiperiodic quantizations— For quantizing a map like the classical A-baker's map Eq. (7.1), there is no unique method; essentially one just requires the associated quantum map to be a unitary $N \times N$ matrix that reduces to the classical map in the semiclassical limit $N \to \infty$. The first quantization of the baker's map was given by Balazs and Voros in [215]; for the simplest case $A = 2$ and N even, this reads

$$\hat{B}_N = \hat{F}_N^{-1} \begin{pmatrix} \hat{F}_{N/2} & \mathbf{0} \\ \mathbf{0} & \hat{F}_{N/2} \end{pmatrix}, \tag{7.3}$$

where $\hat{F}_N$ is the $N \times N$ discrete Fourier transform (DFT) matrix defined via

$$(\hat{F}_N)_{jk} = \frac{1}{\sqrt{N}} e^{-2\pi i jk/N}, \quad j,k = 0,\dots,N-1.$$

This quantization using the standard DFT matrix is associated with periodic boundary conditions on the torus $\mathbb{T}^2$. In order to study different quantum symmetries, we will consider the natural "generic" quantization for the A-baker's map with quasiperiodic boundary conditions [215,216,223] corresponding to $\theta = (\theta_1, \theta_2) \in [0,1)^2$,

$$\mathrm{Gen}_{A,N}^{\theta_1,\theta_2} = (\hat{F}_N^{\theta_1,\theta_2})^{-1} \bigoplus_{j=0}^{A-1} \hat{F}_{N/A}^{\theta_1,\theta_2}, \tag{7.4}$$

where

$$(\hat{F}_N^{\theta_1,\theta_2})_{jk} = \frac{1}{\sqrt{N}} e^{-2\pi i(j+\theta_1)(k+\theta_2)/N} \tag{7.5}$$

is a generalized DFT matrix, and $N \in A\mathbb{N}$. The direct sum part of Eq. (7.4) produces a block diagonal matrix consisting of generalized DFT matrices $\hat{F}_{N/A}^{\theta_1,\theta_2}$.

The case $\theta_1 = \theta_2 = 0$ is the Balazs–Voros quantization of the A-baker's map, for which we may use the abbreviated label "BV" in plots or tables. The Balazs–Voros quantizations preserve an operator TR symmetry but break an operator reflection symmetry (Sec. 7.4.1), and for $A = 2$ were observed to show anomalous level spacings behavior depending on the dimension N [215].

The case $\theta_1 = \theta_2 = 1/2$ is the Saraceno quantization from [216], and corresponds to antiperiodic boundary conditions. This quantization preserves TR symmetry and moreover preserves the classical reflection symmetry, as it commutes with the microscopic reflection operator $R_N : |x\rangle \mapsto |N - x - 1\rangle$. As a consequence, this was observed for $A = 2$ to restore COE level spacing statistics within each symmetry class.

In general, for $\theta_1 \neq \theta_2$, the generic quantization in Eq. (7.4) does not appear to preserve a clear operator TR or reflection symmetry like in the Saraceno case. Possible symmetries are discussed further in Sec. 7.4.1 and Appendix D.1.

Shor baker quantization— In addition to the above generic (quasi)periodic quantizations, we consider the "Shor baker quantizations" from [7, 205]. These quantizations are part of the quantum baker's map decomposition of the modular multiplication operator in Shor's factoring algorithm [199], and can be defined as

$$\hat{S}_{A,N} = \hat{F}_N^{-1} \left(\bigoplus_{j=0}^{A-1} e^{2\pi i j^2/A} \hat{F}_{N/A}^{0,-j/A} \right), \tag{7.6}$$

where $F_{N/A}^{0,-j/A}$ denotes a generalized DFT matrix defined via Eq. (7.5). These Shor baker quantizations again appear to break both the operator TR and reflection symmetries.

Phase variants— Finally, we consider "phase variant quantizations" by adding arbitrary Berry-like phases $e^{2\pi i\alpha} = (e^{2\pi i\alpha_0}, \ldots, e^{2\pi i\alpha_{A-1}})$ to the DFT block sectors of the previous A-baker's map quantizations. These are written in the right column of Tab. 7.1. These quantizations have historically been considered as variations on the usual Balazs–Voros or Saraceno quantizations since [215], but generally are overlooked in favor of the simpler standard/phaseless quantizations. For generic or random phases, we will see that the phase variant quantizations exhibit significantly different spectral statistics than their corresponding standard/phaseless quantizations.

	Standard/Phaseless	Phase variant
Balazs–Voros	$\hat{F}_N^{-1} \bigoplus_{j=0}^{A-1} \hat{F}_{N/A}$	$\hat{F}_N^{-1} \bigoplus_{j=0}^{A-1} e^{2\pi i\alpha_j} \hat{F}_{N/A}$
Saraceno	$\left(\hat{F}_N^{\frac{1}{2},\frac{1}{2}}\right)^{-1} \bigoplus_{j=0}^{A-1} \hat{F}_{N/A}^{\frac{1}{2},\frac{1}{2}}$	$\left(\hat{F}_N^{\frac{1}{2},\frac{1}{2}}\right)^{-1} \bigoplus_{j=0}^{A-1} e^{2\pi i\alpha_j} \hat{F}_{N/A}^{\frac{1}{2},\frac{1}{2}}$
Generic $\mathrm{Gen}_A^{\theta_1,\theta_2}$	$(\hat{F}_N^{\theta_1,\theta_2})^{-1} \bigoplus_{j=0}^{A-1} \hat{F}_{N/A}^{\theta_1,\theta_2}$	$(\hat{F}_N^{\theta_1,\theta_2})^{-1} \bigoplus_{j=0}^{A-1} e^{2\pi i\alpha_j} \hat{F}_{N/A}^{\theta_1,\theta_2}$
Shor baker	$\hat{F}_N^{-1} \bigoplus_{j=0}^{A-1} e^{2\pi i j^2/A} \hat{F}_{N/A}^{0,-\frac{j}{A}}$	$\hat{F}_N^{-1} \bigoplus_{j=0}^{A-1} e^{2\pi i\alpha_j} \hat{F}_{N/A}^{0,-\frac{j}{A}}$

Table 7.1: Definitions of the different quantizations of the A-baker's map. Balazs–Voros is the same as $\mathrm{Gen}_A^{0,0}$, and Saraceno the same as $\mathrm{Gen}_A^{1/2,1/2}$. The "default" quantizations will be the standard/phaseless ones, and we may simply refer to them as the "Balazs–Voros/Saraceno/Generic/Shor baker" quantizations, while for the quantizations with arbitrary phases $e^{2\pi i\alpha}$ we will always specify that it involves the extra phases.

All of the quantizations in Tab. 7.1 are quantizations of the classical A-baker's map in the sense that they map coherent states localized in phase space near (q, p), to coherent states localized in phase space near the classical time-evolved point $\left(Aq - \lfloor Aq \rfloor, \frac{p+\lfloor Aq \rfloor}{A}\right)$ as $N \to \infty$. For details, see [224, §4] and [7, Suppl. Mat.], noting that for quasiperiodic boundary conditions one must use the appropriate quasiperiodic coherent states and generalized DFT matrix $\hat{F}_N^{\theta_1,\theta_2}$. Additionally, for the Balazs–Voros (and Saraceno) quantizations, the argument in [224, §5] proves a rigorous "Egorov property" con-

cerning time-evolution of quantum observables $\mathrm{Op}_N(a)$ corresponding to classical observables a on $\mathbb{T}^2$ supported away from classical discontinuities,

$$\|\hat{U}_N^t \,\mathrm{Op}_N(a)\hat{U}_N^{-t} - \mathrm{Op}_N(a \circ B^{-t})\| \xrightarrow{N\to\infty} 0, \tag{7.7}$$

where $\hat{U}_N$ is the quantization and B is the classical A-baker's map. The argument is insensitive to phases on the DFT blocks, so that the same rigorous correspondence holds for their corresponding phase variant quantizations. We expect the same argument (with some minor adaptations) applies to the generic quasiperiodic and Shor baker quantizations both with and without phases.

7.3 Overview of results

In this section, we explain the main results summarized in Tab. 7.2, which compares the nearest-neighbor level spacing statistics and spectral form factor behavior by quantization and presence of quantum symmetries. We provide the numerical results for the level spacing statistics, and both analytical and numerical results for the early time SFF slope. Due to the classical TR and reflection symmetries of the classical A-baker's map, one would expect its quantizations to exhibit spectral statistics similar to a 2-block COE matrix (a direct sum of two independent, equal sized COE matrices). As has been well-known since [215], this already does not hold for the level spacing statistics of the Balazs–Voros quantization with $A = 2$, which display intermediate level spacing statistics due to the mixing of symmetry sectors. But as we will see, there are several subtleties involved with the spectral statistics, and the results will depend on both the spectral statistic chosen and the particular quantization type. We emphasize the following main points.

1. Unlike the $A = 2$ case, for large A, the level spacing statistics actually do appear to follow classical RMT behavior for all considered quantizations. However, this RMT behavior can be of the wrong

symmetry class (e.g. CUE vs COE) and/or reflect the wrong number of symmetry sectors.

2. For the standard/phaseless quantizations, the early time SFF behavior correctly identifies the RMT symmetry class and symmetry sectors, even as the level spacings do not. This provides a resolution for the non-RMT level spacing statistics in [215], as well as for the wrong symmetry class behavior in the aforementioned point. Such spectral anomalies, where the long-range statistics remain reliable even as the short-range ones do not, are those we term "weak anomalies".

3. The Berry-like phases in the phase variant quantizations produce "strong spectral anomalies", where even the early time SFF misses one of the classical symmetries. We note that the reflection and TR symmetries continue to emerge in the classical limit despite these phases.

4. The presence of strong anomalies is verified numerically to be tied to ergodicity in a quantum dynamical sense of exploring an orthonormal basis in the Hilbert space [1] (see Chap. 2, irrespective of symmetries in the classical limit. However, weak anomalies do not appear to be sufficiently strong to induce ergodic dynamics in this sense.

7.3.1 Nearest-neighbor level spacing statistics

While RMT level spacing statistics are commonly used as an indicator (or even definition) of "quantum chaotic" systems [28], the A-baker's map quantizations can exhibit non-universal level spacing statistics that are strongly sensitive to the particular quantization choice. The first hint of complication is that the Balazs–Voros quantization in Eq. (7.3) ($A = 2$) was observed in [215] to have level spacing statistics that vary depending on N; they almost never look COE or block COE, which was explained as due to the quantization breaking the classical reflection symmetry in Eq. (7.2) and mixing symmetry sectors together.

		BV	Saraceno	$\text{Gen}_A^{\theta_1,\theta_2}$	Shor baker
$A=2$	Level spacings	mixed	2-COE	2-COE/mixed	mixed
	SFF slope	4	4	4	4
A large	Level spacings	COE	2-COE	CUE	CUE
	SFF slope	4	4	4	4

		BV(α)	Saraceno(α)	$\text{Gen}_A^{\theta_1,\theta_2}(\alpha)$	Shor baker(α)
$A=2$	Level spacings	COE	COE	COE/mixed	mixed
	SFF slope	2	2	2	2
A large	Level spacings	COE	COE	CUE	CUE
	SFF slope	2	2	2	2

Table 7.2: Summary of spectral statistics for the various quantizations of the A-baker's map, with the standard or phaseless quantizations in the top section, and the random phase variant quantizations in the bottom section. The columns for the generic quantization $\text{Gen}_A^{\theta_1,\theta_2}$ and its phase quantization reflect the choices $\theta = (0.2, 0.7)$ and $(0, 0.5)$ for numerics, though the SFF slope behavior we derive through the periodic orbit analysis applies to any choice of θ. As seen in the table, the level spacing statistics (and mean gap ratio) vary greatly across all quantizations, and only accurately reflect the classical symmetry sectors over all A for the standard Saraceno quantization, which preserves both classical symmetries upon quantization. The early time SFF slope successfully identifies the symmetry sectors for all standard/phaseless quantizations (top section), even when the operator does not exhibit a clear analogue of the classical symmetries. However, the SFF misses the reflection symmetry in the random phase variant quantizations in the bottom section. The entries labeled "mixed" indicate level spacings that do not adhere to a single RMT or block-RMT ensemble, and instead look somewhere inbetween ensembles.

Surprisingly, as demonstrated by Figs. 7.2 and 7.3, we find the level spacing statistics and mean gap ratio statistic (computed from the level spacings [210, 214, 225]) for the higher slope A-baker's maps begin to look very close to those of a single COE or CUE matrix as A increases, for all quantizations except the standard Saraceno quantization. Thus, for large values of A, these level spacing statistics appear RMT-like, but reflect the *wrong* symmetry classes. The effect of the classical reflection symmetry appears to completely disappear for large A (for non-Saraceno quantizations), and for some quantizations the TR symmetry separating COE from CUE is ignored as well. For reference, the mean gap ratio values for the RMT ensembles as derived in [214, 225] are provided in Tab. 7.3.

Although all quantizations share the classical limit of an A-baker's map, they exhibit a wide variety of level spacing and gap ratio statistics, ranging from the expected 2-block COE behavior, to

single block COE, to single block CUE, and to intermediate or mixed statistics inbetween two RMT ensembles. We observe that for large A, it appears these short-range spectral statistics reflect certain symmetries of the quantized operator (Sec. 7.4.1, Appendix D.1), but not necessarily those of the underlying classical map.

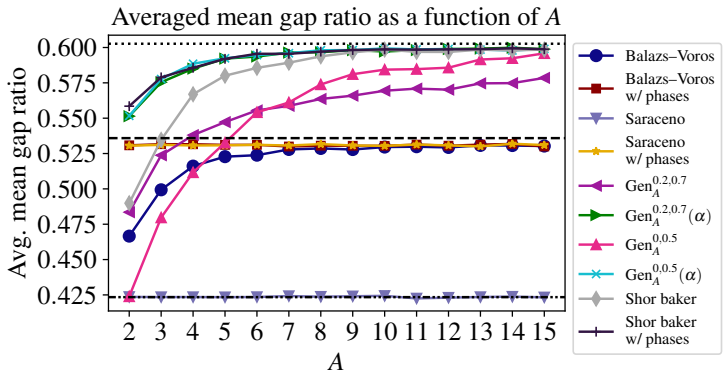


Figure 7.2: Averaged mean gap ratio for quantizations of the A-baker's map, as a function of A. Each point represents an average over 50 values of $N \in A\mathbb{N}$, starting near $N = 5000$. The horizontal lines, from top to bottom, plot the RMT reference values for CUE (dotted), COE (dashed), and 2-block COE (dash-dot-dotted). Only the standard Saraceno quantizations (downward triangle) exhibit mean gap ratios close to the expected 2-block COE value for all A.

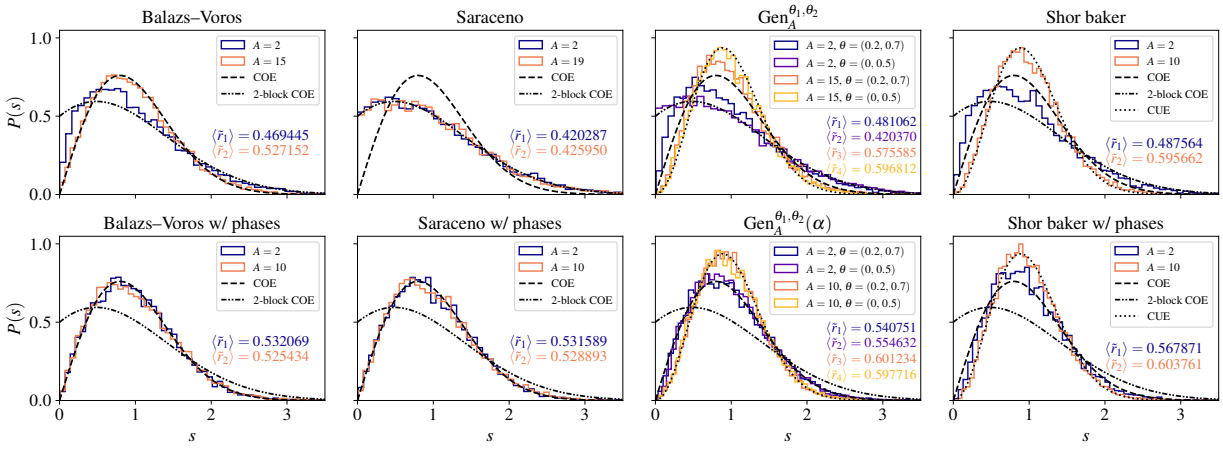


Figure 7.3: Level spacing histograms for the different quantizations of the A-baker's map, for $N = 9690$. Note the variety of behaviors–COE, 2-block COE, mixed/indeterminate, and even CUE–that can arise, despite the same classical map symmetries. Only the phaseless Saraceno quantizations and phaseless $\mathrm{Gen}_{A=2}^{0,0.5}$ quantization appear to follow the 2-block COE curve. The mean gap ratio statistic $\langle \tilde{r} \rangle$ is also computed for each quantization.

7.3.2 Spectral form factor

The spectral form factor (SFF) is the Fourier transform of the 2-point level correlation function [28, 36]. In this chapter, we will normalize it slightly differently from the previous chapters, for convenience in analyzing the ramp behavior; this will be indicated by the use of the notation $\mathrm{SFF}(t)$ as opposed to $K(t)$. For $\hat{U}_N$ an $N \times N$ unitary matrix, the SFF is given by the formula

$$\mathrm{SFF}(t) = \frac{1}{N} |\mathrm{Tr}(\hat{U}_N^t)|^2 = \frac{1}{N} \sum_{j,k=1}^{N} e^{it(\theta_j - \theta_k)}, \tag{7.8}$$

where $(\theta_j)_{j=1}^N$ are the eigenangles of $\hat{U}_N$. The normalization is chosen so that the SFF can be conveniently analyzed and compared across different values of N. For early times $t > 0$, the SFF measures long-range spectral correlations, while for larger times t, the SFF describes finer spectral correlations such as level spacings and eventually discreteness of the spectrum.

Letting $\tau = t/N$, there is the well-known formula for the COE form factor averaged over the random ensemble in the limit $N \to \infty$ [28], which for early times τ yields $\langle \mathrm{SFF}_{\mathrm{COE}}(\tau) \rangle = 2\tau + \mathcal{O}(\tau^2)$. For 2-block COE matrices, the corresponding ensemble-averaged SFF is $\langle \mathrm{SFF}_{\text{2-COE}}(\tau) \rangle = \langle \mathrm{SFF}_{\mathrm{COE}}(2\tau) \rangle$. Thus the early time SFF slope is 2 for a single COE matrix, and 4 for the 2-block COE matrix. For the A-baker's map quantizations, since we do not have an ensemble of matrices to average over, we average the SFF by averaging over neighboring points as described in Appendix D.3.

We first demonstrate that the early time (averaged) SFF resolves the two issues with the level spacing statistics for the standard/phaseless quantizations, (i) the non-universal behavior for small A of the Balazs–Voros/Generic/Shor baker quantizations, and (ii) the apparent disappearance of two distinct symmetry sectors for the same quantizations with larger A. These cases thus correspond to "weak anomalies", for which the SFF provides a satisfactory diagnostic of the spectral behavior and classical symmetries.

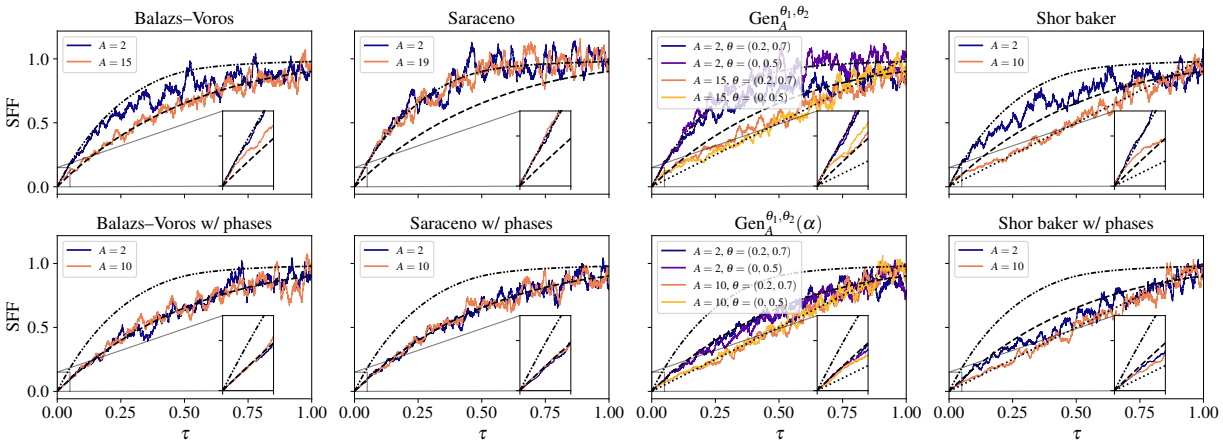


Figure 7.4: Averaged SFFs for the different quantizations of the A-baker's maps, for $N = 9690$. In the top row of standard phaseless quantizations, the very early time SFF follows the 2-block COE behavior (slope 4 at the origin), while for the bottom row of phase variant quantizations, the early time SFF has slope 2. All insets show up to $\tau = 0.05$, or up to $t = 484$ for $N = 9690$. For several of the larger A quantizations, the transition away from the early time ($\tau \approx 0$) SFF slope behavior is already visible in this window. In general, the larger time SFF, corresponding to shorter range spectral statistics like the level spacings, vary greatly depending on the particular quantization.

As shown in the top row of Fig. 7.4, for very early times τ, the SFFs for the standard phaseless quantizations follow the slope 4 reference SFF behavior for the 2-block COE, correctly reflecting the classical map symmetries. The longer time behaviors (corresponding to shorter range statistics) however vary greatly. For larger τ, the Saraceno quantizations (and $\text{Gen}_{A=2}^{0,0.5}$) continue to follow the 2-block COE SFF, as previously demonstrated for the Saraceno $A = 2$ quantization in [226], but the other standard quantizations appear to cross over to the single COE or CUE SFF at a time τ that decreases as A increases. Since the level spacing statistics are short-range, corresponding to larger τ, this faster cross-over explains the Balazs–Voros/Generic/Shor baker matrix level spacing histograms approaching those of a single COE or CUE matrix as A increases. For these cases, which describe "weak anomalies", both the RMT nature and symmetry sectors are readily apparent in the SFF, in contrast to the differing information from the level spacing statistics.

The phase variant quantizations hold the surprise, however. As shown in the bottom row of Fig. 7.4, the addition of random phases to the quantizations interferes with the classical reflection

symmetry in a way that the early time SFF fails to detect it. Instead, the early time averaged SFF has slope 2, capturing only the classical TR symmetry. (The level spacings are of even less help, as seen in Figs. 7.2 and 7.3). We remark that from Fig. 7.4, it is not entirely clear whether it is the TR or reflection symmetry that is missed by the early time SFF; the SFF for several of the quantizations follows the COE curve which strongly suggests it is the reflection symmetry that is broken in those cases, but the SFF for other quantizations crosses over to the CUE curve. From the periodic orbit analysis below, we will see that it is still the reflection symmetry that is broken at early times in all cases. From the periodic orbit analysis, we will also be able to identify the specific phases α that produce an SFF slope of 4, which is a measure zero set but contains more elements than just those corresponding to the standard/phaseless ($\alpha_j = 0$) quantizations.

In addition to the SFF plots in Fig. 7.4, we plot the best fit SFF slope over a wide range of dimensions N in Fig. 7.5. Unlike the standard/phaseless quantizations which produce SFFs with slope near 4 that accurately describe the classical symmetry sectors, the quantizations with random phases produce SFFs with slope near 2, thereby hiding the classical R symmetry.

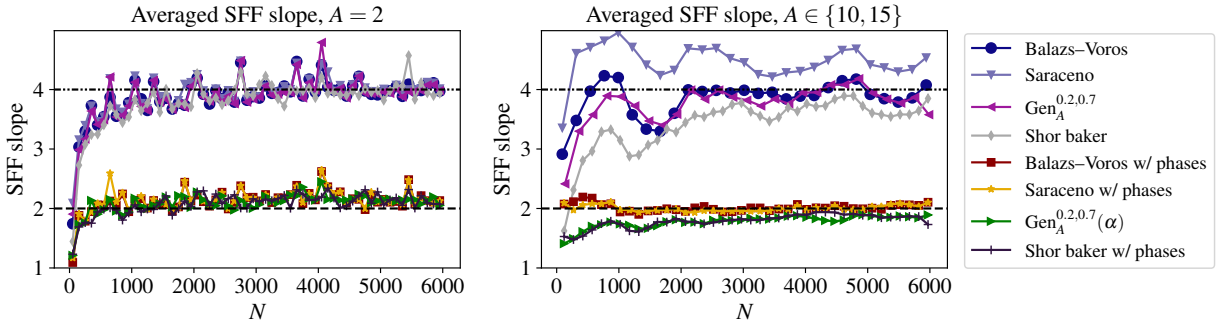


Figure 7.5: Averaged best fit early time SFF slope for $A = 2$ (left) and $A \in \{10, 15\}$ (right, $A = 15$ for standard Balazs–Voros, Saraceno, and $\mathrm{Gen}_{A=15}^{0.2,0.7}$ quantizations, and $A = 10$ for the remaining). The quantizations with random phases show a slope near 2, while those without show a slope near 4. Some of the quantizations shown share the random choice of phases. Outliers where the least squares fitting had large error were removed prior to averaging (cf. Fig. 7.9, Appendix D.3).

Overall, as summarized in Tab. 7.2, although the early time SFF slope correctly identifies both classical symmetries for the standard/phaseless quantizations ("weak anomalies"), it only captures

one classical symmetry for the phase variant quantizations ("strong anomalies"). Meanwhile the level spacings fare worse, missing either one or both classical symmetries in almost all quantizations.

7.3.3 Symmetry breaking and quantum dynamical ergodicity

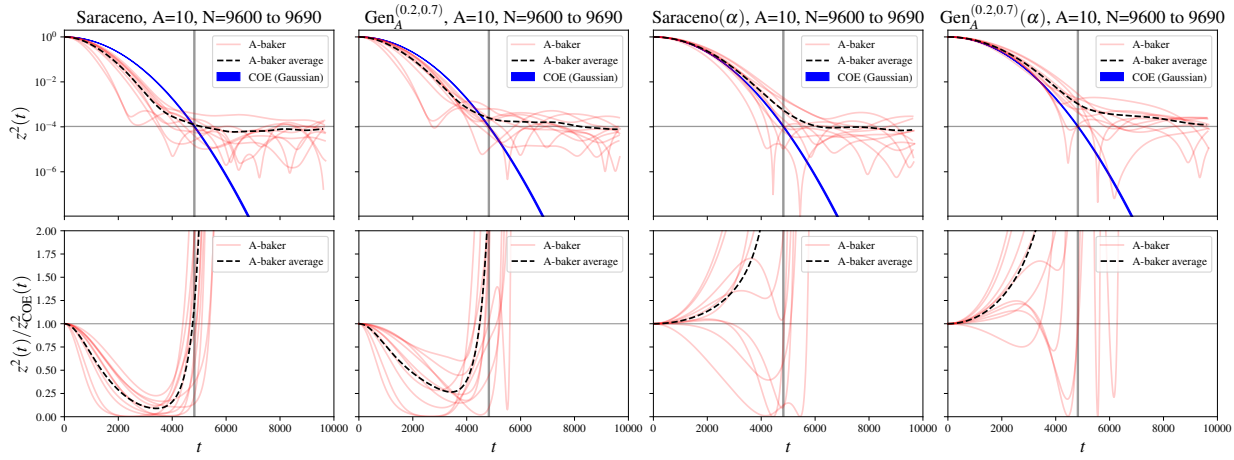


Figure 7.6: Effect of spectral anomalies on quantum cyclic ergodicity measured in terms of the persistence $z^2(t)$, via four different quantizations of $A = 10$-baker's maps with different kinds of anomalies. The depicted quantizations are Saraceno (no anomalies), $\mathrm{Gen}_A^{(0.2,0.7)}$ (weak reflection and time-reversal anomalies), Saraceno(α) (strong reflection anomaly) and $\mathrm{Gen}_A^{(0.2,0.7)}(\alpha)$ (strong reflection anomaly, weak time-reversal anomaly). All values of N from 9600 to 9690, in steps of 10, are depicted (translucent red lines) together with the average $z^2(t)$ (top row) and $z^2(t)/z^2_{\mathrm{COE}}(t)$ (bottom row) over these values of N (dashed black line) to observe the statistical trends after averaging out the strong fluctuations with N. *Top row*: The persistence $z^2(t)$ given by Eq. (7.12) is plotted as a function of time t in a log-linear scale, and compared with the COE (Gaussian) curves $z^2_{\mathrm{COE}}(t)$ denoting the ideal behavior of COE statistics given by Eq. (7.14) over these values of N [up to $O(N^{-1})$ fluctuations, which are not depicted for the COE reference]. The vertical band near the center of each plot depicts the range of $t = N/2$ over the different values of N, and the horizontal band depicts $z^2(t) = N^{-1}$ (representing the order of magnitude of $\eta^2(N) = cN^{-1}$), the value reached by the COE (Gaussian) curve at $t = N/2$ (the cutoff time for cyclic ergodicity). *Bottom row*: The ratio $z^2(t)/z^2_{\mathrm{COE}}(t)$ is plotted against t in a linear-linear scale, with the vertical band near the center again depicting the range of $t = N/2$, while the horizontal line near the center depicts a unit ratio, i.e., $z^2(t) = z^2_{\mathrm{COE}}(t)$. The rapid increase near $t = N/2$ in these plots represents the onset of $O(N^{-1})$ fluctuations as the dominant behavior of $z^2(t)$ around and beyond this time. These plots appear consistent with quantum cyclic ergodicity of the kind associated with COE [$z^2(t) \geq z^2_{\mathrm{COE}}(t)$ up to $O(N^{-1})$ fluctuations], resulting from strong anomalies (symmetry breaking in long-range measures) but not weak anomalies as explained in the text.

Having demonstrated that measures of spectral statistics can be incompatible with classical symmetries, we now consider the direct relation between spectral statistics and quantum dynamics in the

Hilbert space. This is especially of interest in illustrating the fully quantum mechanical role of spectral anomalies or deviations from ideal random matrix behavior, irrespective of symmetries in the classical limit. We will take advantage of the distinct behavior of each measure across different quantizations of the A-baker's maps to contrast the role of short-range and long-range spectral statistics in influencing quantum dynamics. In particular, we will provide numerical evidence that long-range symmetry breaking or strong anomalies are sufficient to induce ergodicity (in a sense to be clarified below) in the quantum dynamics of the system, while short-range or weak anomalies have a milder effect that may not be significant in the $N \to \infty$ limit.

For this purpose, we will consider the notion of quantum cyclic ergodicity in the Hilbert space, introduced in Ref. [1] (Chap. 2) as a direct quantum dynamical counterpart to spectral statistics. There, it was shown that the presence of sufficient long-range spectral rigidity is tied to the existence of an orthonormal basis $\{|C_k\rangle\}_{k=0}^{N-1}$ where every initial state "visits" every other state in a cyclic sequence. This form of ergodicity is appropriate for time-independent unitary systems with (quasi-)energy conservation, and differs from more direct forms related to classical ergodicity possible in open or time-dependent quantum systems [227–229]. Quantitatively, the overlap of an initial state $|C_k\rangle$ with $|C_{k+t}\rangle$ after t time-steps, called the persistence,

$$z_k^2(t) \equiv |\langle C_{k+t}|\hat{U}_N^t|C_k\rangle|^2, \tag{7.9}$$

must be larger than a cutoff $\eta^2(N) = cN^{-1}$ (where c is some $\Omega(1)$ parameter) associated with the overlap of random states for $t \in [-N/2, N/2]$, i.e.,

$$z_k^2(t) > \eta^2(N), \forall\ t \in \left[-\frac{N}{2}, \frac{N}{2}\right]. \tag{7.10}$$

Further, the "optimal" orthonormal basis in which this property is most likely to be present [in terms

of maximizing $z_k^2(1)]$ was shown to be given by the discrete Fourier transform (DFT) of the energy eigenstates:

$$|C_k\rangle = \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} e^{-2\pi i k n/N} |E_n\rangle, \tag{7.11}$$

where the energies are sorted in ascending order. In this case, $z_k^2(t) = z^2(t)$ for all k, given in terms of the energy levels by

$$z^2(t) = \left| \frac{1}{N} \sum_{n=0}^{N-1} e^{i(E_n - 2\pi n/N)t} \right|^2. \tag{7.12}$$

This is the basis we will study numerically.

For "ideal" RMT-like behavior, $z^2_{\mathrm{RMT}}(t) = \exp[-\Delta^2 t^2]$ to leading order [1] (originating in Gaussian spectral fluctuations [101, 102, 115], see also Sec. 2.5), where (specializing to even N for simplicity)

$$\Delta^2 = 2 \sum_{t=1}^{N/2} \frac{\mathrm{SFF}(t)}{N t^2} \tag{7.13}$$

gives the leading contribution to spectral fluctuations in various measures of long-range spectral rigidity such as the Dyson-Mehta Δ_3 parameter [36] or the related Δ^* [47] that measure the regularity of the spectrum. For COE, one obtains

$$z^2_{\mathrm{COE}}(t) = e^{-4t^2 \ln N/N^2}. \tag{7.14}$$

This is guaranteed to exceed $\eta^2(N) = cN^{-1}$ as per Eq. (7.10) with the slight restriction $|t| < N(1-\epsilon)/2$ (for any small $\epsilon > 0$), showing that each $|C_k\rangle$ in a system with ideal COE statistics "ergodically" visits almost all basis vectors $|C_{k-N/2}\rangle$ through $|C_{k+N/2}\rangle$ in succession. Due to the presence of time-reversal symmetry without separate sectors in COE [28], one cannot here demand ergodicity in the full interval $|t| \leq N/2$ in Eq. (7.10), corresponding to fully visiting every single basis vector, that is present [1] in CUE or a non-degenerate half of CSE.

In Fig. 7.6, the persistence $z^2(t)$ in the DFT basis for 4 different quantizations of the $A = 10$-baker map [Saraceno and $\text{Gen}_A^{(0.2,0.7)}$, with and without phases] are compared to the ideal COE persistence, to examine their quantum dynamical ergodicity relative to the behavior of COE [i.e., if $z^2(t) \geq z^2_{\text{COE}}(t) + O(N^{-1})$]. We recall that while the unitary reflection symmetry may be broken weakly or strongly in these quantizations, the antiunitary time reversal symmetry is always broken only weakly in the spectral statistics, making COE the appropriate standard for comparison. The choice of $A = 10$ statistically guarantees that the Berry-like phases α_j are generically random, as required for strong anomalies; in contrast, $A = 2$ has only one independent phase and may show a significant dependence on this phase as seen in Fig. 7.9(g). Further, due to atypical fluctuations with varying N in the level statistics of baker maps [see, e.g., Fig. 7.7 and Fig. 7.9(a)-(f)], noted as far back as Ref. [215], we consider statistical trends over 10 adjacent values of N, and additionally plot the persistence $z^2(t)$ averaged over these values of N to tame the fluctuations. This is justified for our numerics as N varies only by around 1% in our chosen range. Subsequently, we observe if the average persistence is comparable to (or is greater than) the ideal COE trend to diagnose ergodicity in the presence of a long-range time-reversal symmetry.

The numerical trends are as follows:

1. For Saraceno (no anomalies) and $\text{Gen}_A^{(0.2,0.7)}$ (weak reflection and time-reversal anomalies), $z^2(t)$ remains less than $z^2_{\text{COE}}(t)$ up to random fluctuations consistent with $O(N^{-1})$, showing compatibility with ergodicity-breaking in the presence of a long-range reflection symmetry.

2. For Saraceno with phases (strong reflection anomaly) and $\text{Gen}_A^{(0.2,0.7)}$ with phases (strong reflection anomaly and weak time-reversal anomaly), $z^2(t)$ fluctuates around $z^2_{\text{COE}}(t)$, showing compatibility with the presence of COE-type ergodicity without a long-range reflection symmetry (but with time-reversal indicated by long-range spectral statistics).

Finally, we note that in both the $\mathrm{Gen}_A^{(0.2,0.7)}$ cases (with or without phases), which possess a weak time-reversal anomaly, $z^2(t)$ oscillates around a slightly larger value than in the Saraceno cases (which have an unbroken time-reversal symmetry), though this slight increase does not statistically appear to be sufficient to induce ergodicity without strong anomalies. In fact, this slightly larger value is likely a finite size numerical effect for these values of N, stemming from the logarithmic divergence of Δ^2 with N in Eq. (7.13) for a linear ramp $\mathrm{SFF}(t) \propto t$ leading to a visible numerical contribution from the late-time regime (corresponding to the crossover in Sec. 7.3.2). However, one can show that in the $N \to \infty$ limit, as long as the SFF appreciably deviates from the early-time trend (due to weak anomalies) only for $|t| \geq cN$ in $\mathrm{SFF}(t)$, the anomalous contribution to Δ^2 is subleading compared to the early-time contribution; it is indeed for a similar reason that COE possesses logarithmically divergent ($\ln N$) spectral fluctuations [36, 40, 47] despite the SFF deviating from a linear ramp [28] for $t \sim N$. To see this quantitatively, we consider a simplified model with the interval of summation $t \in \mathscr{I} = [1, N/2]$ split into an early-time regime $\mathscr{I}_{\mathrm{UV}} = [1, cN]$ with $\mathrm{SFF}(t) = \alpha_1 t$, and a late-time regime $\mathscr{I}_{\mathrm{IR}} = (cN, N/2]$ with $\mathrm{SFF}(t) = \alpha_2 t$ for some $c \ll 1$; in this case, the leading contribution to the logarithmic divergence $\alpha_1 \ln(cN/1)$ comes entirely from the early time region, while the late-time region contributes a subleading term proportional to $\alpha_2 \ln[(N/2)/(cN)] = -\alpha_2 \ln(2c)$. Nevertheless, other effects (such as a deviation from a Gaussian profile of $z^2(t)$) are possible at larger N, and it would be interesting to explore or rule out such phenomena at values of N at least an order of magnitude larger than the present study.

In summary, our numerics for $N \approx 10^4$ in quantizations of A-baker's maps with different manifestations of spectral anomalies appear to be consistent with a direct link between long-range symmetry breaking (strong anomalies) and cyclic ergodicity, with an at best weaker effect of short-range symmetry-breaking (weak anomalies), verifying the analytical connection obtained in Ref. [1], i.e., Chap. 2, between long-range spectral statistics and quantum dynamical ergodicity.

7.4 Operator symmetries and level spacing statistics

In the remaining sections, we provide further background and details for the results in the previous section. We start with the relation between the quantizations' operator symmetries and the classical map's symmetries.

7.4.1 Operator symmetries

Classifying quantum symmetries corresponding to the classical symmetries in these models is not entirely straightforward. If one can construct a quantum version of the classical symmetry, such as in the Saraceno quantization [216], then one can say that the quantization preserves the corresponding classical symmetry. However, due to the infinite possibilities of quantum operators that can all correspond to the same classical symmetry operator in the limit $\hbar \to 0$ ($N \to \infty$), verifying that a quantization does not commute with any of those operators is much less clear. For this reason, we will discuss a limited version of the possible operator symmetries, and include more detailed analysis in Appendix D.1. These restricted definitions will still agree with those historically used to describe the symmetries of the Balazs–Voros and Saraceno quantizations [215, 216].

Quantization on the torus— To discuss the relation between the classical symmetries and operator symmetries, we first provide more background on the quantization process on the torus. For further details, see [224, 230]. Quantization on the 2-torus associates to each natural number $N \in \mathbb{N}$ and $\theta \in [0,1)^2$ an N-dimensional Hilbert space $\mathscr{H}_N(\theta)$ of quantum states. The parameter $\theta = (\theta_1, \theta_2)$ sets the quasiperiodicity requirement in position and momentum as follows. Letting $S(q,p) = e^{i(pQ-qP)/\hbar}$ denote the phase space translation operators, then the Hilbert space $\mathscr{H}_N(\theta)$ is associated with states

ψ on $\mathbb{R}$ satisfying

$$S(1,0)\psi = e^{-2\pi i\theta_1}\psi, \quad S(0,1)\psi = e^{2\pi i\theta_2}\psi,$$

for $\theta = (\theta_1, \theta_2)$. Recall the Balazs–Voros quantization corresponds to the case $\theta_1 = \theta_2 = 0$ which describes periodic states, while the Saraceno quantization corresponds to $\theta_1 = \theta_2 = 1/2$ which describes antiperiodic states. The generic quantization $\mathrm{Gen}_A^{\theta_1,\theta_2}$ corresponds to the quasiperiodic conditions described by $\theta = (\theta_1, \theta_2)$. The main consideration we need for different θ is that position representation states $|n\rangle$ and momentum representation states $|k\rangle$ are related via the generalized discrete Fourier transform $\hat{F}_N^{\theta_1,\theta_2}$ as defined in Eq. (7.5), which depends on θ. This explains why one uses the generalized DFT matrices in the Saraceno and $\mathrm{Gen}_A^{\theta_1,\theta_2}$ quantizations. The generalized DFT matrix relation between position and momentum also implies that operators on $\mathscr{H}_N(\theta)$, which are $N \times N$ matrices, are converted between position and momentum basis via conjugation by $\hat{F}_N^{\theta_1,\theta_2}$ (or its inverse).

The Shor baker quantizations involve several different generalized DFT blocks, but we will associate these quantizations with periodic boundary conditions to match the $\hat{F}_N^{-1}$ factor.

Reflection symmetry— Let B be the classical A-baker's map, and recall the classical reflection symmetry R in Eq. (7.2), which maps (q,p) to $(1-q, 1-p)$ and satisfies $RBR^{-1} = B$. Its quantum analogue R_N should then reverse, in some way, both the position states $|n\rangle$ and the momentum states $|k\rangle$, and quantizations $\hat{U}_N$ that preserve the reflection symmetry should satisfy $R_N \hat{U}_N R_N^{-1} = \hat{U}_N$.

For the Saraceno quantizations, which we will denote here by $\hat{B}_{N,A}^{\mathrm{Sar}}$, the quantum reflection is $R_N : |x\rangle \mapsto |N-1-x\rangle$, which has the same action in momentum space and commutes with $\hat{B}_{N,A}^{\mathrm{Sar}}$ since $R_N = (\hat{F}_N^{\frac{1}{2},\frac{1}{2}})^2$. One can separate the eigenvalues of $\hat{B}_{N,A}^{\mathrm{Sar}}$ according to whether its corresponding eigenstate is in the $+1$ or -1 symmetry sector of R_N, and this produces COE level spacing statistics within each symmetry sector, as explained in [216]. (See Fig. 7.8 for larger A.) Additionally, when

considering the spectrum as a whole, the two symmetry sectors of the Saraceno quantizations combine to look like that of a direct sum of two COE matrices, indicating that the two symmetry sectors behave essentially as if they are independent of each other.

On the other hand, the Balazs–Voros, generic quasiperiodic, and Shor baker quantizations do not exhibit a clear analogous reflection symmetry. We investigate possible *Fourier* reflection symmetries in Appendix D.1, and provide numerical plots demonstrating the lack of Fourier reflection symmetry for the non-Saraceno quantizations that we consider (Balazs–Voros, $\text{Gen}_A^{0.2,0.7}$, $\text{Gen}_A^{0,0.5}$, and Shor baker). While this rules out a class of reflection operators coming from the generalized DFT matrices, it does not prohibit the possibility of a different commuting reflection-like operator in the $N \to \infty$ limit.

TR symmetry— The other classical symmetry is a time reversal (TR) symmetry $T : (q, p) \mapsto (p, q)$, which satisfies $TBT^{-1} = B^{-1}$. Its quantum analogue should act on operators by switching between position and momentum basis, and mapping $i \mapsto -i$, so that quantizations $\hat{U}_N$ (in position basis) preserving TR symmetry should ideally satisfy the antiunitary relation

$$\hat{F}_N^{\theta_1,\theta_2} \hat{U}_N (\hat{F}_N^{\theta_1,\theta_2})^{-1} = (\hat{U}_N^{-1})^*, \tag{7.15}$$

where * denotes entrywise complex conjugation. We can define a quantization $\hat{U}_N$ to have an "operator TR symmetry" if it satisfies Eq. (7.15) for its corresponding boundary conditions θ. However, as for the reflection symmetry, other antiunitary operations with the same classical limit could also be a valid "quantum TR symmetry". For the quantizations we consider, the Balazs–Voros and Saraceno quantizations satisfy Eq. (7.15), while the generic quasiperiodic quantizations with $\theta_1 \neq \theta_2$ and the Shor baker quantizations do not.

7.4.2 Level spacing statistics

To investigate the level spacing statistics of an $N \times N$ unitary matrix, one orders the eigenangles θ_i, and defines the nearest neighbor level spacings (or gaps)

$$s_i = \theta_{i+1} - \theta_i, \quad i \in \mathbb{Z}/N\mathbb{Z}. \tag{7.16}$$

To compare the distribution of these level spacings to the expected universal RMT distributions, the normalization is to multiply the (s_i) by $\frac{N}{2\pi}$, which we do for all level spacing histogram plots.

The mean (adjacent) gap ratio statistic, as defined in [210], is given by

$$\langle \tilde{r} \rangle = \left\langle \min\left(\frac{s_{i+1}}{s_i}, \frac{s_i}{s_{i+1}} \right) \right\rangle_i, \tag{7.17}$$

where the average is over all $i \in \mathbb{Z}/N\mathbb{Z}$. This statistic provides a single value that can be used to compare the closeness to RMT level spacings, and does not require any normalization or unfolding of the eigenvalues. The mean gap ratios for the standard RMT ensembles in the $N \to \infty$ limit were derived in [225], and for block RMT matrices in [214]. The block RMT matrices are relevant in the presence of discrete symmetries, as one generally needs to separate eigenstates according to the symmetry sector to recover expected non-block RMT level statistics. We are primarily concerned with the circular orthogonal ensemble (COE) and circular unitary ensemble (CUE). Since the circular ensembles and Gaussian ensembles have the same local n-level correlation functions in the limit $N \to \infty$ [36], we may interchange terms such as "COE level spacings" and "GOE level spacings". We list the values of relevance to our study in Tab. 7.3.

	GOE	2-block GOE	GUE	2-block GUE	Poisson
$\langle \tilde{r} \rangle$	0.53590	0.423415	0.60266	0.422085	0.38629

Table 7.3: Mean gap ratio values for RMT ensembles, from [214, 225].

Here the 2-block GOE matrix means a direct sum of two equal sized, independent GOE matrices, and similarly for the 2-block GUE matrix.

In general, ones expects that chaotic systems with time reversal (TR) symmetry have GOE/COE spectral statistics, while those without have GUE/CUE statistics. Additionally, one expects the presence of discrete symmetries to produce block-RMT statistics, according to the number of symmetry sectors. As we saw for the A-baker's map, however, the actual level spacings' behavior can be highly variable depending on the particular quantization.

We plot in Fig. 7.7 the mean gap ratios for the different quantizations over a range of $N \in A\mathbb{N}$. As we saw for specific dimensions N in Figs. 7.2 and 7.3, out of all the quantizations in Tab. 7.2, only the Saraceno quantizations, and the generic quantization $\mathrm{Gen}_{A=2}^{0,0.5}$ (for $A = 2$ only), have mean gap ratio close to that for block COE matrices. We note that there are dips in the mean gap ratio at specific values of N, which typically correspond to powers of the slope A. For such dimensions the level spacings may look non-RMT (sometimes close to Poisson).

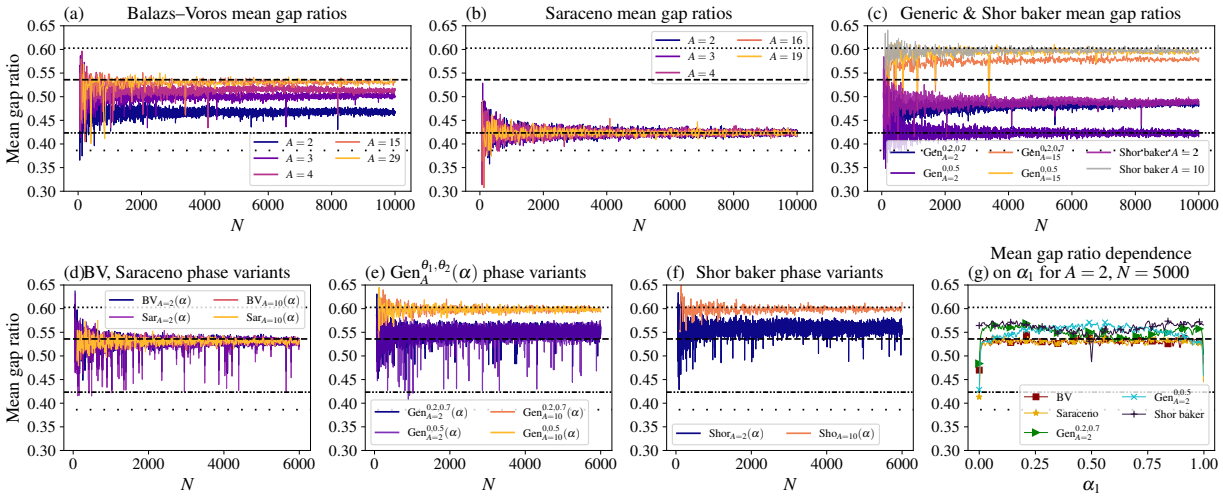


Figure 7.7: (a)–(f) Mean gap ratios for the different quantizations over $N \in A\mathbb{N}$. The horizontal lines (from top to bottom) are the RMT reference values for GUE (dotted), GOE (dash-dot-dotted), 2-block GOE (dashed), and Poisson (loosely dotted). Some of the phase variant quantizations may share the same random choice of phases. (g) Mean gap ratios for $A = 2$ and $N = 5000$ as a function of the phase $\alpha = (0, \alpha_1)$ for $\alpha_1 \in [0, 1)$ (step size 0.002). Note that $\alpha_1 = 1/2$ corresponds to the standard Shor baker quantization, while $\alpha_1 = 0$ corresponds to the standard versions of the other quantizations.

7.4.3 Approximate symmetry classes for the Balazs–Voros quantization

We now return to the Balazs–Voros-type quantizations of the A-baker's map, which we saw have level spacing statistics that can exhibit deviations from RMT and overlook the presence of classical symmetry sectors. We demonstrate how one can obtain roughly COE-like level spacings for the Balazs–Voros quantization in Eq. (7.3) ($A = 2$) by separating the eigenvalues according to *approximate* symmetry classes of their eigenstates, which was suggested as a possible method in [215]. However, we will see in Sec. 7.5 using the SFF that this separation still retains significant irregularities.

Recall the reflection operator $R_N : |x\rangle \mapsto |N - x - 1\rangle$ which commutes with the Saraceno quantization and is equal to $(\hat{F}_N^{\frac{1}{2},\frac{1}{2}})^2$. This is the permutation matrix with 1s on the top-right to bottom-left diagonal, which has the trivial block decomposition

$$R_N = \begin{pmatrix} & & & R_{N/A} \\ & & R_{N/A} & \\ & \iddots & & \\ R_{N/A} & & & \end{pmatrix}.$$

Since it commutes with the Saraceno quantization $\hat{B}^{\mathrm{Sar}}_{N,A}$, this allowed for separating the eigenstates of $\hat{B}^{\mathrm{Sar}}_{N,A}$ according to whether they fall in the $+1$ or -1 eigenspace of R_N, which recovers RMT spectral statistics.

For the Balazs–Voros quantizations, this suggests considering a similar reflection-like operator,

the permutation

$$\tilde{R}_N = (\hat{F}_N^{0,0})^2 = \begin{pmatrix} 1 & 0 & & \cdots & & 0 \\ 0 & & \cdots & & 0 & 1 \\ \cdot & & & 0 & 1 & 0 \\ \vdots & & 0 & 1 & 0 & 0 \\ & \iddots & \iddots & & & \vdots \\ 0 & 1 & 0 & \cdots & & 0 \end{pmatrix}, \tag{7.18}$$

which is a natural reflection candidate (cf. Appendix D.1) when considering states that are periodic in position and momentum (vs antiperiodic for Saraceno quantizations). The map $\tilde{R}_N$ is equal to $\hat{F}_N^2$ and sends $|x\rangle \mapsto |-x\rangle$ (taken modulo N). While $\tilde{R}_N$ does not commute with $\hat{B}_N$, it is in some sense *close* to commuting with $\hat{B}_N$. In particular, we show in Appendix D.2 that the commutator $[\hat{B}_{N,A}, \tilde{R}_N]$ has only very few non-decaying matrix elements.

Computing the overlap $\langle \varphi^{(j)} | \tilde{R}_N | \varphi^{(j)} \rangle$ for all eigenvectors $\varphi^{(j)}$ of $\hat{B}_N$, we create the two symmetry classes,

$$\begin{aligned} S_+ &= \{\varphi^{(j)} : \langle \varphi^{(j)} | \tilde{R}_N | \varphi^{(j)} \rangle \geq 0\}, \\ S_- &= \{\varphi^{(j)} : \langle \varphi^{(j)} | \tilde{R}_N | \varphi^{(j)} \rangle < 0\}. \end{aligned} \tag{7.19}$$

We can then investigate the level spacing statistics within each approximate symmetry class, which are shown (along with those for the exact Saraceno symmetry classes) in Fig. 7.8.

Approximate symmetries for $A = 2$— As seen in Fig. 7.8(c)–(d), for $A = 2$, within a single approximate symmetry class $S_\pm$, the level spacing statistics for the Balazs–Voros quantization look approximately COE. The inner products $\langle \varphi^{(j)} | \tilde{R}_N | \varphi^{(j)} \rangle$ tend to cluster near -1 and 1 (Fig. 7.8(e)), suggesting

that while not exact, $\tilde{R}_N$ is a fairly good choice of approximate symmetry. Fig. 7.8(f) plots the quantity,

$$\frac{1}{N}\sum_{j=1}^{N}\left|\,|\langle\varphi^{(j)}|\widetilde{R}_N|\varphi^{(j)}\rangle|-1\right|^2, \tag{7.20}$$

which is the mean square error of the inner product from ±1, for eigenstates of $\hat{B}_N$. Other than some outliers that appear somewhat connected to powers of A, this error is fairly constant, suggesting that the distribution shape shown for $A=2$ in Fig. 7.8(e) is likely representative for other N as well.

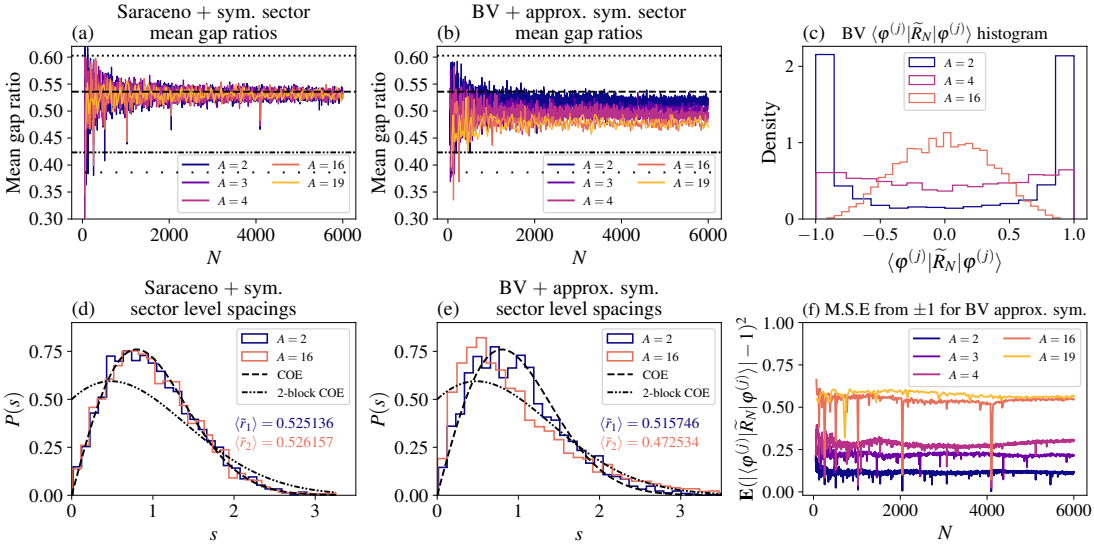


Figure 7.8: (a), (b) Mean gap ratios for the Saraceno + symmetry sectors and Balazs–Voros + approximate symmetry sectors, for $N \in A\mathbb{N}$ even. (d), (e) Level spacing histograms for the Saraceno + symmetry sector and Balazs–Voros + approximate symmetry sector for $N = 5904$. (c) Balazs–Voros inner product histogram for $N = 5904$ and $A = 2, 4, 16$. The histogram for $A = 2$ shows a strong clustering split between $+1$ and -1, but this dichotomy disappears for larger A. (f) The mean square error defined in Eq. (7.20) for the Balazs–Voros quantizations as a function of N.

We also note that attempting to use the Saraceno reflection operator $R_N : |x\rangle \mapsto |N-x-1\rangle$ here for the Balazs–Voros quantization does not appear to produce any meaningful separation, and the inner products $\langle\varphi^{(j)}|R_N|\varphi^{(j)}\rangle$ are spread within $[-1,1]$ instead of clustering near ±1.

Failure for larger A-baker's maps— For $A \geq 3$, the Saraceno quantizations of the A-baker's map continue to commute with the reflection operator R_N, and continue to exhibit level spacing statistics that look like a direct sum of two COE matrices. Thus one can try to use an approximate symmetry

for the non-symmetrized Balazs–Voros quantizations with $A \geq 3$ as well. Unlike the $A = 2$ case however, the natural approximate symmetry candidate $\tilde{R}_N$ does not produce even an approximately useful separation of eigenstates, as seen in Fig. 7.8(e). The values $\langle \varphi^{(j)} | \tilde{R}_N | \varphi^{(j)} \rangle$ no longer cluster strongly near ± 1, and separating by the sign of $\langle \varphi^{(j)} | \tilde{R}_N | \varphi^{(j)} \rangle$ does not reproduce RMT-like level statistics (Fig. 7.8(c)–(d)). Given that the unseparated eigenvalue statistics begin to look more and more like a single COE matrix as A increases, this is not that surprising.

7.5 Spectral form factor analysis

In this section, we provide more detailed analysis and plots of the spectral form factor (SFF) and its early time slope. Recall the SFF for an $N \times N$ unitary matrix is given by the formula

$$\mathrm{SFF}(t) = \frac{1}{N} |\operatorname{Tr}(U_N^t)|^2 = \frac{1}{N} \sum_{j,k=1}^{N} e^{it(\theta_j - \theta_k)}, \tag{7.21}$$

and that we set $\tau = t/N$. The formula for the ensemble-averaged COE form factor [28] is

$$\langle \mathrm{SFF}_{\mathrm{COE}}(\tau) \rangle \equiv \lim_{N \to \infty} \frac{1}{N} \mathbb{E} |\operatorname{Tr}(U_N^t)|^2 = \begin{cases} 2\tau - \tau \log(1 + 2\tau), & \tau \leq 1 \\ 2 - \tau \log\left(\frac{2\tau+1}{2\tau-1}\right), & \tau > 1 \end{cases}. \tag{7.22}$$

For the quantized baker's maps, with no ensemble to average over, we average Eq. (7.21) at time t with its nearest 2ℓ neighbors (or from time 1 to $2t-1$ if $t < \ell$), as described in more detail in Appendix D.3.

We show plots of the early time SFF slope as function of the dimension N in Fig. 7.9(a)–(f), corresponding to noisier, more detailed versions of the earlier Fig. 7.5. In general, the SFF slope computations are noisy, and even the plots in Fig. 7.9 are averaged over the nearest ∼20 neighbors, after removing outliers which did not have a low error slope fit. These outliers amount to only relatively few values of N for each quantization ($< 1\%$ for $A = 2$ quantizations, and $\sim$ 5-8% for $A = 10$ or 15 in

Fig. 7.9). As in Fig. 7.5, we see in Fig. 7.9 a clear dichotomy in the SFF slope between the standard phaseless quantizations and the phase variant quantizations. In Fig. 7.9(g), we also plot the SFF slope for $A = 2$ as a function of the phase parameter $\alpha = (0, \alpha_1)$, similarly as we did for the mean gap ratio in Fig. 7.7(g).

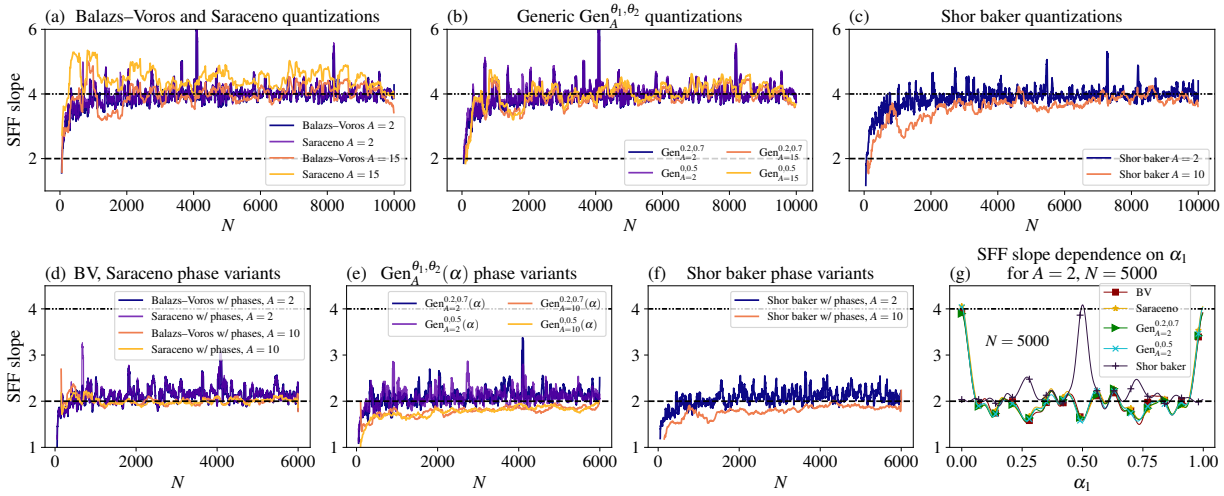


Figure 7.9: (a)–(c) Averaged early time SFF slope for the standard/phaseless quantizations, plotted as function of $N \in A\mathbb{N}$. The SFF slope values cluster near 4. Outliers (fewer than 1% of points for $A = 2$, and ∼5-8% for $A \in \{10, 15\}$) where the least squares slope fitting produced large residuals were removed before averaging. For further details, see Appendix D.3. (d)–(f) Averaged early time SFF slope for the phase variant quantizations, fewer than 1% of points removed as outliers. The SFF slope values cluster near 2. (g) SFF slope for $A = 2$ and $N = 5000$ as a function of the phase $\alpha \in [0, 1)$ (step size 0.002) for the different types of quantizations. Compare with Fig. 7.7(g).

Next, in Fig. 7.10(a) we briefly examine the SFF within an individual approximate symmetry class (Sec. 7.4.3) for the Balazs–Voros 2-baker quantization. We see that while the SFFs appear to look COE for moderately sized τ, there are irregularities near $\tau = 0$. Thus while separating by the approximate symmetry class can partially restore level spacing statistics as in Fig. 7.8, it produces long-range spectral irregularities. In contrast, for the Saraceno quantizations (not shown), the SFF for an individual symmetry class appears to follow the single COE SFF for all τ.

In Fig. 7.10(b) we also demonstrate a complication with determining the early time SFF slope. For some values of N, the SFF may show large early time irregularities. Large enough irregularities

which do not have a good least squares fit are considered outliers, and we remove such points prior to averaging and plotting in Figs. 7.5 and 7.9.

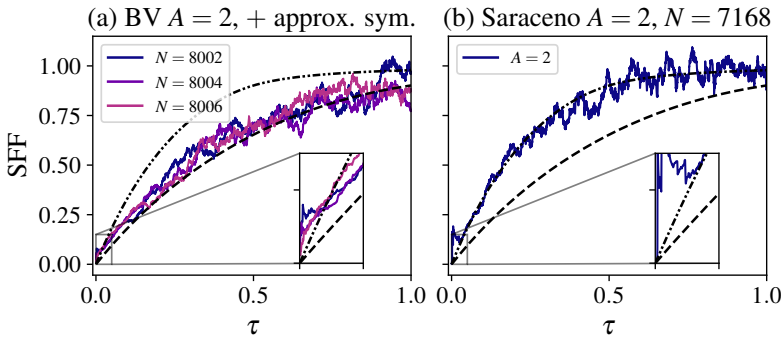


Figure 7.10: (a) Balazs–Voros SFF for + approximate symmetry classes for $N = 8002, 8004$, and 8006. The behavior for small τ shows irregularities, even though for larger times it follows the COE SFF. In contrast, the Saraceno ± 1 symmetry classes (not plotted) show single COE-like SFF. (b) Example of bad early time behavior in the SFF, for one of the rare outliers removed before averaging to produce the plots in Fig. 7.9.

We note that some of the outliers and noise are products of the averaging methods used to compute the SFF slope. While we do not optimize the averaging methods used, we choose parameters so that it becomes clear whether the slope of the early time SFF is close to 2, corresponding to the SFF for a single COE matrix, or close to 4, corresponding to the SFF for a 2-block COE matrix. Due to this choice of parameters, along with the occasional outliers, computing the SFF slope is not as convenient as computing the gap ratio statistic; however, for these models it proves to be more informative.

7.6 Conclusion

We have studied maximally chaotic quantum maps with discrete symmetries that share the same classical limit. Contrary to conventional expectations for the correspondence between discrete symmetries and spectral statistics [73, 211–214], we demonstrated that short-range spectral statistics in these models generically fail to identify discrete symmetries (weak anomalies), while long-range spectral statistics also violate these expectations in the presence of phases (strong anomalies). However, long-range spectral statistics appear more directly correlated with intrinsic quantum dynamical prop-

erties [1] in the Hilbert space. This further reinforces the notion that spectral statistics should ideally be interpreted in terms of intrinsically quantum mechanical properties, while more work is necessary to understand how they connect to macroscopic dynamics, such as in the classical limit, beyond the well-studied case of systems showing close agreement in several measures with RMT [28, 36].

One direction to explore, which may be of immediate relevance in the context of many-body statistical mechanics, is whether the introduction of simple phases — as in the case of strong anomalies studied here — could break the commonly observed correspondence [21, 140] between "macroscopic" subsystem thermalization behaviors (i.e. in a large subset of particles) and spectral signatures of ergodic phenomena. While our results already formally point to an affirmative answer, given that one can realize quantizations of A-baker's maps as many-body Floquet quantum circuits using the quantum Fourier transform and phase gates [217, 218] (with the classical $N \to \infty$ limit then corresponding to the thermodynamic limit of, e.g., many qubits), it would nevertheless be illuminating to understand the mechanisms involved (such as Berry-like phases) in a more natural setting of an interacting many-body system that does not necessarily model a classically chaotic map.

Part III

Experimental Probes of Many-Body Quantum Ergodicity

Chapter 8: Probing dynamical ergodicity and thermalization in many-body systems

8.1 Synopsis: Simultaneous probes of eigenvalue and eigenstate statistics

The previous Chapters collectively highlight that ergodicity in closed time-independent quantum systems must be understood through an elaborate lens of distinct quantum dynamical phenomena tied to different properties of the system. In particular, given Hamiltonian or Floquet evolution in any finite dimensional Hilbert space, the statistics of energy eigenvalues characterizes a dynamical form of ergodicity in the Hilbert space [1] (see Chap. 2); further, given a factorization into a subsystem and environment, the observable independent entanglement structure of eigenstates determines the possibility of ergodicity and thermalization for all observables in a subsystem [23,24,52,140]. Even where the latter eigenstate-dependent phenomenon occurs, the maximum speed at which thermalization can be achieved is again determined purely by the long-range distribution of energy eigenvalues [2,3] (see Chaps. 3 and 4).

A rather convenient feature of these results is that a single observable-independent quantity, the spectral form factor $K(t)$, suffices to partially or completely capture all aspects of eigenvalue statistics discussed in the Chaps. 2, 3 and 4. Specifically, its early-time decay sets a speed limit on thermalization, while its late time quantum fluctuations constrain spectral rigidity and, therefore, cyclic ergodicity. An experimental measurement of this quantity in a quantum system should ideally give us a reasonably comprehensive handle on its quantum ergodic properties associated with eigenvalues. Two natural questions are suggested by this observation:

1. How can we directly measure the observable-independent $K(t)$ and infer eigenvalue statistics from specific experimental observables in a generic quantum platform?

2. Given that the entanglement structure of eigenstates is also observable-independent (after the simple additional step of choosing a subsystem), can such a strategy generalize to eigenstate statistics?

In this Chapter, we describe an experimental protocol that addresses both these questions (in the affirmative), which requires only single-qubit [1] unitary gates and projective measurements in addition to implementing the dynamics of the system. This protocol is based on the framework of randomized measurements [231], which addresses the first question: The dynamics of the system is investigated via randomly generated initial states and measurements, so that the universal "observable-independent" features of the dynamics can be readily extracted without significant nonuniversal effects from specific states and measurement settings. This is accomplished by choosing an observable in which the known "random noise" from different instances of randomized measurements cancel out to 0, leaving just the observable-independent SFF.

The *same* measurement protocol can be used to compute *Partial* SFFs (PSFFs) $K_A(t)$ for a given subsystem A, from the same dataset used to compute the SFF $K(t)$. We show that PSFFs capture the observable-independent aspects of eigenstate statistics in typical quantum systems, in an analysis based on subsystem ETH [52]. Subsystem ETH expresses ETH in terms of the reduced density matrices of energy eigenstates in subsystems without relying on specific observables, and is therefore the most relevant quantitative expression of ETH for our present considerations. Taken together, the SFF $K(t)$ and the set of PSFFs $K_A(t)$ for all subsystems of interest provide a simultaneous experimental probe of signatures of dynamical and thermal ergodicity in closed quantum systems.

In the remainder of this Section (Sec. 8.1), we aim to provide a complete overview of the mea-

[1] We take qubits for definiteness, but the strategy can be generalized to higher-dimensional local Hilbert spaces like qu*d*its.

surement protocol and the sensitivity of PSFFs to eigenvalue and eigenstate statistics. Subsequently, we will focus entirely on the technical aspects of PSFFs and their illustration in quantum systems of interacting spins, as realized for instance with trapped ions [232, 233], Rydberg atoms [234] and superconducting qubits [235]. While much of this analysis assumes some form of ensemble averaging for convenience, we stress that the PSFFs may be defined for individual systems much like the SFF. For experimentally relevant details such as the proof of the measurement protocol and a detailed analysis of errors and mitigation strategies, we refer to the original paper [6]. Additionally, we note that an experimental measurement of PSFFs in systems of up to 5 superconducting qubits, simulating both Floquet and Hamiltonian evolution, has been reported in Ref. [168]. Further, an extension to open quantum systems and the corresponding experimental signatures in a 4-qubit trapped ion system will be discussed in Chap. 9.

We now turn to an overview of the main results of this Chapter. We start by recalling the standard definition of the SFF, define the PSFF and describe their estimation using randomized measurement protocol. We then illustrate the key features of the (P)SFF and demonstrate our measurement protocol using an example of a non-integrable, periodically kicked spin$-1/2$ model. We will argue on the basis of this example and show in later sections with detailed analytical and numerical calculations that the combination of the SFF and PSFF provides unique insights into the eigenvalue and eigenstate statistics of quantum many-body systems.

8.1.1 Spectral form factor

The SFF in a quantum system with time-independent Hamiltonian H and energy spectrum $\{E_j\}$ is defined as the Fourier transform of the two-point correlator of the energy level density [28]. It can

be expressed as

$$K(t) \equiv \frac{1}{D^2} \overline{\sum_{i,j} e^{i(E_i - E_j)t}} = \frac{1}{D^2} \overline{\mathrm{Tr}\left[T(t)\right] \mathrm{Tr}\left[T^\dagger(t)\right]} \,. \tag{8.1}$$

Here, we normalize $K(t)$ such that $K(0) = D^{-2}\mathrm{Tr}\left[\mathbb{1}\right]^2 = 1$, with D the Hilbert space dimension and have defined the unitary time-evolution operator $T(t) \equiv \exp(-iHt)$. The overline denotes a possible disorder or ensemble average over an ensemble of $T(t)$, which is often convenient due to non-self-averaging behavior of the SFF [110]. As in several of the previous Chapters, however, we can also consider an individual system without affecting our protocol. Replacing the energies E_i with quasi-energies, this definition carries over to Floquet models with time-periodic evolution operator $T(t = n\tau) = V^n$ ($n \in \mathbb{N}$) and V the Floquet time evolution operator for a single period τ [2].

The SFF is a probe of the statistics of energy eigenvalues in quantum systems. Lately, it has played a key role in a variety of different fields, interconnecting quantum chaos [28], quantum dynamical ergodicity [1] (Chap. 2), the quantum dynamics of black holes [38, 82, 86, 154], condensed matter systems [4, 35, 41, 83, 84, 87, 236–240], and the dynamics of thermalization [2, 3, 114] (see also Chaps. 3 and 4). In Fig. 8.1(a), we illustrate its behavior in the context of a periodically kicked spin-1/2 system. The time evolution operator T at integer multiples $n \in \mathbb{N}$ of driving period τ is given by $T(t = n\tau) = V_3^n$ with,

$$V_3 = e^{-iH^{(x)}\tau/3} e^{-iH^{(y)}\tau/3} e^{-iH^{(z)}\tau/3} \,. \tag{8.2}$$

Here, the Hamiltonians $H^{(x,y,z)}$ contain nearest-neighbor interactions with strength $J = 3\tau^{-1}$

[2] Denoting the set of eigenvalues of the Floquet operator V with $\{\exp(-iE_i\tau)\}$, the quasi-energy eigenvalues $\{E_i\}$ are only defined up to multiples of the driving frequencies $\omega = 2\pi\tau^{-1}$. We fix them to lie in the interval $[0, \omega]$.

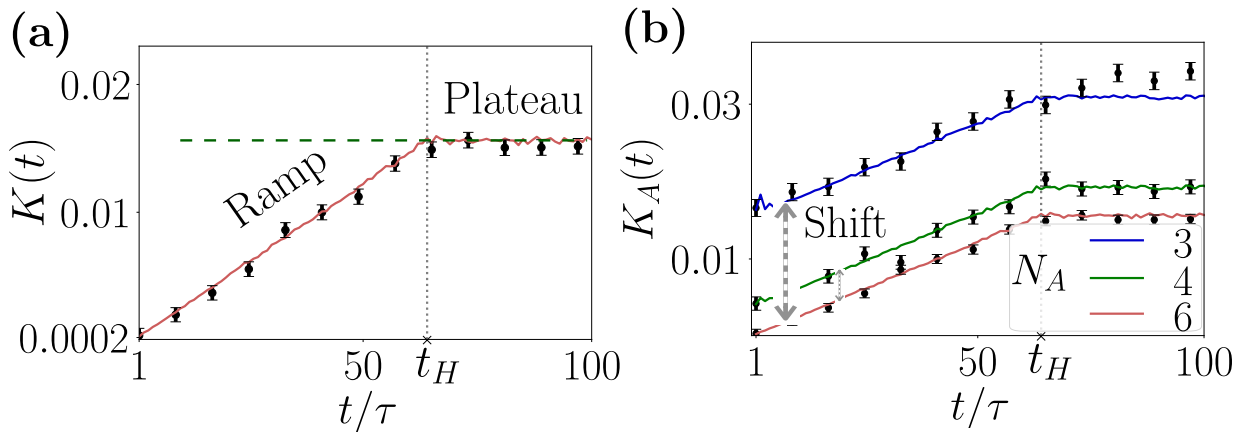


Figure 8.1: *Illustration of the characteristic properties of the SFF and PSFF using the RMT-like spin-1/2 Floquet model V_3.* (a) We display the SFF $K(t)$ for the Floquet model V_3 with $N = 6$ qubits as a function of time t. We observe characteristic features such as the ramp between $t \sim \tau$ to $t = t_{\mathrm{H}} = 2^N\tau$ and a plateau for $t > t_{\mathrm{H}}$. (b) For the PSFF $K_A(t)$ we observe ramp, plateau and, in particular, a constant, additive shift of the PSFF compared to the SFF, which depends on the subsystem size N_A of the subsystem A. We have chosen subsystems A from the middle of the total system. In both, the colored lines show the numerically calculated SFF and PSFFs, averaged over 8000 disorder realizations. In addition, we illustrate our measurement protocol (see Sec. 8.1.3) by simulating $M = 2\times 10^5$ experimental runs (single-shot randomized measurements) at each time and display the estimated SFF and PSFF as black dots with associated error bars. The dashed green line in panel (a) sketches the form of the SFF generically expected in a localized model.

and disordered transverse fields with strength $h_i^{(x,y,z)} \in [-J, J]$,

$$H^{(x,y,z)} = J\sum_{i=1}^{N-1}\sigma_i^{(x,y,z)}\sigma_{i+1}^{(x,y,z)} + \sum_{i=1}^{N} h_i^{(y,z,x)}\sigma_i^{(y,z,x)} \ ,$$

and σ^a $[a \in (x,y,z)]$ denote the Pauli matrices. We have denoted the number of spins with N such that $D = 2^N$. An ensemble average is naturally performed by averaging over many instances of $T(t = n\tau) = V_3^n$, each with local disorder potentials $h_i^{(a)}$ sampled independently from the uniform distribution on $[-J, J]$.

As shown in Fig. 8.1(a), the SFF $K(t)$ for this model and choice of parameters exhibits a period of linear growth, before transitioning to a constant at time $t/\tau \approx D = 2^N$. This *ramp-plateau* structure of the SFF is a characteristic feature of quantum RMT-like systems [28,241,242], originating from (quasi-)energy level repulsion and spectral rigidity [241], and is predicted by RMT [28, 36]. In particular,

as we briefly review in App. E.1, RMT for time evolution operators $T(t=\tau n)=V^n$, with V from the circular unitary ensemble (CUE), yields

$$K(t)=\frac{1}{D}\begin{cases} t/t_{\mathrm{H}}, & 0<t\leq t_{\mathrm{H}} \\ 1, & t>t_{\mathrm{H}}\,. \end{cases} \tag{8.3}$$

Here, the slope of the ramp and the onset of the plateau is determined by the Heisenberg (or plateau) time t_{H} which is connected to the mean inverse spacing of adjacent (quasi-) energies. It typically scales with the Hilbert space dimension $t_{\mathrm{H}}/\tau\sim D$ — for V from CUE, $t_{\mathrm{H}}/\tau=D$ [28, 38]. Thus, the SFF is expected to drop with increasing Hilbert space dimension $D=2^N$, as D^{-2} at times $1\lesssim t/\tau\ll D$ and as D^{-1} at times $t/\tau\gtrsim D$. Fig. 8.1(a) shows that the SFF $K(t)$ for the V_3 model closely follows the CUE prediction after the initial few time steps. This time after which the many-body model shows the same SFF as the one in RMT is known as the Thouless time t_{Th} [35]. We have called this the ramp time t_{ramp} in previous chapters, but the "Thouless time" terminology is especially conventional for many-body systems with some sense of spatial locality. For the model V_3 we note that $t_{\mathrm{Th}}\approx 5\tau$ (see also Sec. 8.3). Therefore, the quasi-energy eigenvalues of the Floquet operator V_3 exhibit Wigner-Dyson statistics (see also Ref. [96]).

In contrast to the example of an RMT-like system V_3 presented above, the energy eigenvalues of integrable and localized models are expected to exhibit Poisson statistics by the Berry-Tabor conjecture [34, 51, 148, 243]. While there is no known precise connection between the localization of eigenstates and Poisson eigenvalues (without possibly additional, as yet unknown, assumptions), the models we consider in this Chapter are sufficiently generic to satisfy this conjecture. We recall that Poisson statistics corresponds to a flat (ensemble-averaged) SFF without a ramp which is, after an initial transient regime, constant in time [28], $K(t\gg 0)=1/D$. This is schematically shown in Fig. 8.1(a) with green dashes. These distinct features of the SFF have been pivotal in characterizing many-body

RMT-like and localized regimes [96, 236, 237].

8.1.2 Partial Spectral Form Factor

The SFF reveals information on the statistics of (quasi-) energy eigenvalues. It is however by definition insensitive to properties of the (quasi-) energy eigenstates. In this subsection, we define the PSFF and illustrate its essential properties connected to properties of eigenvalues and eigenstates.

For a fixed subsystem $A \subseteq \mathscr{S}$ of the total system $\mathscr{S}$ with complement B ($A \cup B = \mathscr{S}$) and Hilbert space dimensions D_A and D_B respectively ($D = D_A D_B$), we define the PSFF as

$$\begin{aligned} K_{\mathrm{A}}(t) &\equiv \frac{1}{DD_A} \overline{\sum_{i,j} e^{i(E_i - E_j)t} \mathrm{Tr}_B \left[\rho_B(E_i)\rho_B(E_j)\right]} \\ &= \frac{1}{DD_A} \overline{\mathrm{Tr}_B \left[\mathrm{Tr}_A[T(t)]\mathrm{Tr}_A\left[T^\dagger(t)\right]\right]}, \end{aligned} \tag{8.4}$$

where $\rho_B(E_i) = \mathrm{Tr}_A\left[|E_i\rangle \langle E_i|\right]$ denotes the reduced density matrix obtained after partial trace of the eigenstate $|E_i\rangle$ of the Hamiltonian H (the Floquet time evolution operator V) with energy (quasi-energy) E_i. Here, the normalization of $K_A(t)$ is chosen such that $K_A(0) = \mathrm{Tr}_B\left[\mathrm{Tr}_A[\mathbb{1}]^2\right]/(DD_A) = 1$. Hence, the SFF and PSFF coincide when $A = \mathscr{S}$, i.e. $K_{\mathrm{A}=\mathscr{S}}(t) = K(t)$. We emphasize that for $A \subset \mathscr{S}$, the PSFF $K_A(t)$ contains non-trivial contributions from the eigenstates $|E_i\rangle$: We obtain terms of the form $\mathrm{Tr}(\rho_B(E_i)^2)$ and $\mathrm{Tr}(\rho_B(E_i)\rho_B(E_j))$ ($i \neq j$) which correspond to the purity and overlap of reduced eigenstates. As shown below, a measurement of the PSFF allows to extract these purities and overlaps, averaged over spectrum and ensemble, i.e. allows to characterize (second-order moments of) the statistics of eigenstates.

We remark that $K_A(t)$ has been previously discussed as a topological invariant in the classification of symmetry-protected matrix product unitaries in Ref. [244]. Its limiting cases for special subsystems (A or B consisting of a single site, in the limit of a large local Hilbert space dimension)

have been used to study matrix elements of local operators in the energy eigenbasis in 1D Floquet circuits, with comparisons to random matrix predictions for eigenstate statistics in these subsystems (as a special case of ETH) [42].

In this work, we identify a general *shift-ramp-plateau* structure of the PSFF, which reveals a direct connection to ETH contained in the subsystem dependence of the PSFF. In Fig. 8.1(b), we display the PSFF for the Floquet model (8.2) for various subsystems A, where N_A denotes number of qubits in the subsystem such that $D_A = 2^{N_A}$. We first note that the PSFF also has a ramp and plateau, similar to the full SFF. The slope of the ramp is nearly identical for the displayed subsystem sizes $N_A \gtrsim N/2 = 3$, which holds more generally for $D_A \gg 1$ in the CUE model, and the onset of the plateau in the PSFF takes place at the Heisenberg time t_{H}. Crucially, we find that, at late times comparable to the onset of the ramp, there is a subsystem dependent additive shift of the PSFF $K_A(t)$ compared to the full SFF $K(t)$.

Similar to the case of the full SFF, we can compare the behavior of the PSFF to predictions of RMT. As detailed in Sec. 8.2, we find that RMT yields for time evolution operators $T(t = \tau n) = V^n$, with V from the CUE, and sufficiently large subsystems A, B, $(D_A, D_B \gg 1)$,

$$K_A(t) = \frac{1}{D_A^2} + \frac{1}{D} \begin{cases} t/t_{\mathrm{H}}, & 0 < t \leq t_{\mathrm{H}} \\ 1, & t > t_{\mathrm{H}} \,. \end{cases} \tag{8.5}$$

As shown in Fig. 8.1(b), and analyzed in detail by further numerical studies in Sec. 8.3, the PSFF (and SFF) for the V_3 model follows closely the RMT predictions. This indicates that both *(quasi-) energy eigenvalues and eigenstates* of V_3 exhibits the Wigner-Dyson statistics of the CUE. We remark that this is consistent with previous works demonstrating that (sub-)systems of generic Floquet systems thermalize to infinite temperature states as per RMT [42, 243, 245–248].

Partial spectral form factor and eigenstate thermalization hypothesis – Using the example of an

RMT-like Floquet model, we have illustrated above the essential features of the PSFF in RMT-like quantum systems. In Sec. 8.2, we analyze its behavior in detail invoking subsystem ETH [52] for the reduced eigenstates, which is a conjecture regarding the *distribution of eigenstates* responsible for the thermal behavior (in the standard sense of ETH) of few-body observables in RMT-like systems.

By separating out the components of the reduced density matrix into maximally mixed, smooth and fluctuating parts as a function of energy, a generic late time expression for PSFF can be obtained. From here, we later conclude that the features of the ramp, plateau and shift are generic features of the PSFF in RMT-like quantum many-body systems. These features are directly connected to the spectrum and ensemble averages of the subsystem purities $\mathrm{Tr}_B(\rho_B(E)^2)$ and of the overlaps of reduced eigenstates $\mathrm{Tr}_B(\rho_B(E_i)\rho_B(E_j))$. Furthermore, the magnitudes of these features in RMT-like systems follow specific constraints when the eigenstates satisfy subsystem ETH, see Sec. 8.2.2.2. In particular, we show that this shift, connected to the average overlaps, enables the detection of thermalization of eigenstates in the framework of subsystem ETH.

Let us take for instance the shift seen in the Fig. 8.1, defined precisely in terms of the fluctuating part of the density matrix later in Sec 8.2.2. For RMT-like models, the shift can be identified as the time independent constant during the linear ramp regime, and for $D_A \ll D$ it is approximated by $K_A(t_0) - K(t_0)$ where $t_{Th} < t_0 \ll t_{\mathrm{H}}$. If the eigenstates follow ETH, it is expected that,

$$K_A(t_0) - K(t_0) \approx O\left(\frac{1}{D_A^2}\right). \tag{8.6}$$

This can be noted for the CUE in the Eqs. (8.3) and (8.5) as well as for the V_3 model in Fig. 8.1, where the shift above SFF is seen to be increasing as the N_A decreases and is found to follow Eq. (8.6) (see Sec. 8.3 for more numerical details). On the other hand, for eigenstates which do not thermalize, the time independent shift above SFF is generically much larger than $O(1/D_A^2)$.

As illustrated above, the SFF and PSFF of a quantum many-body system provide crucial insights into the statistics of energy eigenvalues and eigenstates, which results in a joint test of dynamical ergodicity and the validity of ETH. The question arises of how to probe the SFF and PSFF in today's quantum devices. In the next subsection, we present our measurement protocol which can be directly implemented in state-of-the-art quantum simulation platforms realizing lattice spin models. It builds on the toolbox of randomized measurements.

8.1.3 Randomized measurements of spectral form factors

Initially, randomized measurements have been proposed and experimentally implemented to characterize many-body quantum states [231] and (out-of-time-ordered) correlation functions of Heisenberg operators [249, 250]. Randomized measurements on quantum states exploit statistical correlations obtained between measurements obtained from different random bases. However, for measuring an object like the SFF, we need to access the full trace of the time evolution operator $T(t)$, summing contributions from all its eigenstates. Therefore, we need to devise a protocol that can measure how various initial states are propagated via $T(t)$, in a way that allows to extract the SFF from standard projective measurements. This subsection provides this protocol and the estimation formulas to achieve this. We also briefly comment on statistical errors arising from a finite number of experimental runs, which are described in detail in Ref. [6].

8.1.3.1 Description of the protocol

Before describing the experimental sequence in detail, we first outline the key idea of our protocol: As visualized in Fig. 8.2, we consider a system $\mathscr{S}$ of N qubits. The first step of our protocol is to prepare a random product state of these qubits. Next, this state is evolved with $T(t)$. Finally, a local measurement in the conjugate random product basis is performed, in order to probe how the

time-evolved state compares to the initial random product state. This is repeated for many random product states in order to sample the complete trace $\mathrm{Tr}\,[T(t)]$ of the time evolution operator and its adjoint uniformly. For instance, in the trivial case $T(t=0)=\mathbb{1}$, we obtain that the 'time-evolved' state always matches to the initial random state corresponding to $D^{-1}\mathrm{Tr}\,[T(0)]=1$. At later times t, we obtain in general a more complex statistics of measurement results from which we can extract the SFF and PSFF.

In our protocol, we note that the ensemble average over time evolution operators in the definition of SFF and PSFF can be favorably combined with the averaging over random product states and measurement bases. As detailed in the prescription of the protocol in the next paragraph, each time evolution operator can thus in practice be applied only to a single random initial product state and measured only once in the corresponding randomized basis, i.e., only a single-shot measurement for each time evolution operator is sufficient in our protocol.

In detail, the experimental recipe reads as follows: (i) We begin with a product state $\rho_0=|\mathbf{0}\rangle\langle\mathbf{0}|$ with $|\mathbf{0}\rangle\equiv|0\rangle^{\otimes N}$. (ii) On this initial state, we apply local random unitaries $U=\bigotimes_{i=1}^{N}u_i$ where u_i are the local unitaries independently sampled from a unitary 2-design [251, 252] on the local Hilbert space $\mathbb{C}^2$. Here, unitary 2-designs are ensembles of random unitaries whose first and second moments match the moments of the Haar measure on the unitary group (defining the CUE) [251, 252]. Examples of unitary 2-designs on $\mathbb{C}^2$ include the (discrete) single-qubit Clifford group as well as uniformly distributed unitary 2×2 matrices which can be sampled for instance via the algorithm presented in Ref. [253]. (iii) We evolve the system in time, i.e. apply a time evolution operator $T(t)$, which is generated by a Hamiltonian H (or Floquet operator V) with randomly sampled disorder potentials. (iv) We apply the adjoint local random unitary $U^\dagger$ resulting in the final state $\rho_f(t)=U^\dagger T(t)U\rho_0 U^\dagger T^\dagger(t)U$. (v) Lastly, we perform a single-shot measurement in the computational basis with outcome bitstring $\mathbf{s}=(s_1,\ldots,s_N)$ with $s_i\in\{0,1\}$ for $i=1,\ldots,N$. This concludes a single experimental run of our pro-

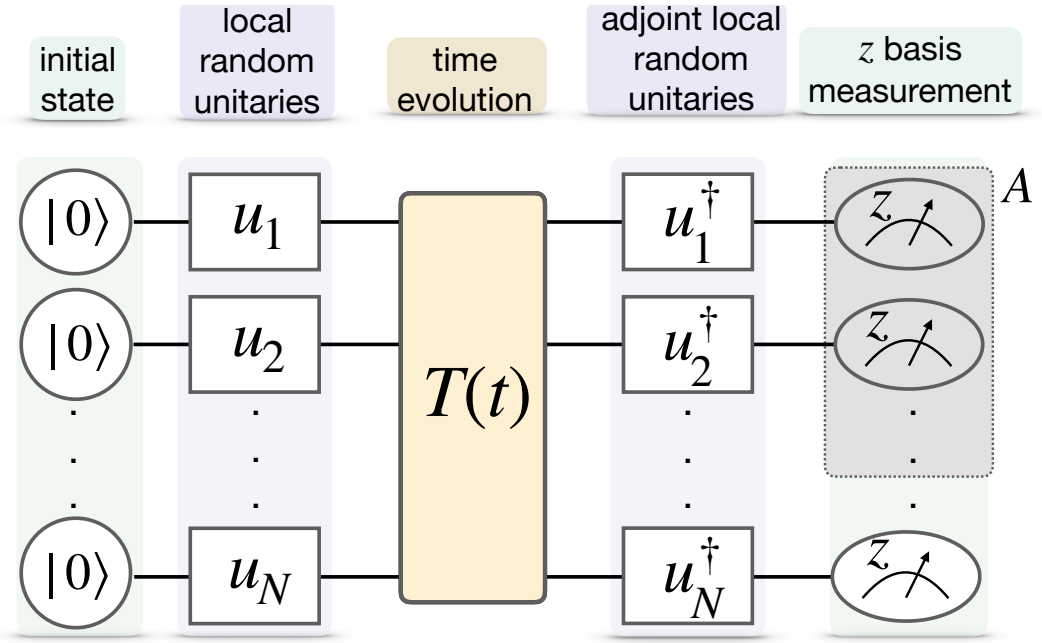


Figure 8.2: *Probing SFF and PSFF using randomized measurements.* We present our protocol for the measurement of the SFF and PSFF using statistical correlations of local random unitaries applied at different times in a single experiment. We begin with a product state $\rho_0 = |0\rangle\langle 0|^{\otimes N}$. Before and after the time evolution $T(t)$, we apply random local rotations $U = \bigotimes_i u_i$ and $U^\dagger$, respectively, where local unitaries u_i are sampled from a unitary 2—design. Here, $T(t)$ can be generated as Hamiltonian evolution, $T(t) = \exp(-iHt)$, or Floquet dynamics, $T(t = n\tau) = V^n$, $n \in \mathbb{N}$, where V is Floquet evolution operator for time period τ. In the last step, a single-shot measurement is performed in the z—basis to collect a bitstring of the form $\mathbf{s} = (s_1, s_2, ..., s_N)$ with $s_i \in \{0, 1\}$. This procedure is repeated M times and M bitstrings are collected to estimate the SFF and PSFF using Eqs. (8.7) and (8.8). The gray shaded region shows one possible choice of the subsystem A.

tocol. Steps (i)-(v) are now repeated M times with new disorder realizations and new local random unitaries such that a set of outcome bitstrings $\mathbf{s}^{(r)}$ with $r = 1, \dots M$ is collected.

8.1.3.2 Estimation formulas and illustrations

The statistics of the measured bitstrings $\mathbf{s}^{(r)}$, $r = 1, \dots M$, depends on the applied time evolution operators $T(t)$. Using the theory of unitary 2-designs, we can express the SFF as a function of this data. We define

$$\widehat{K(t)} = \frac{1}{M} \sum_{r=1}^{M} (-2)^{-|\mathbf{s}^{(r)}|} , \tag{8.7}$$

where $|\mathbf{s}| \equiv \sum_i s_i$. As shown in Ref. [6], $\widehat{K(t)}$ yields an (unbiased) estimate of the SFF for a finite number M of experimental runs and converges to $K(t)$ when $M \to \infty$.

Remarkably, from the *same* measurement data $\mathbf{s}^{(r)}$, we have also access to the PSFF $K_A(t)$ for arbitrary subsystems $A \subseteq \mathscr{S}$ via post-processing. To this end, we simply project the measured bitstrings on the subsystem A of interest, i.e., define $\mathbf{s}_A = (s_i)_{i\in A}$, and use

$$\widehat{K_A(t)} = \frac{1}{M}\sum_{r=1}^{M} (-2)^{-|\mathbf{s}_A^{(r)}|}\,, \tag{8.8}$$

which gives an (unbiased) estimate for $K_A(t)$ for finite M and converges to $K_A(t)$ when $M \to \infty$.

In Fig. 8.1(a-b), we illustrate our measurement protocol in the context of the periodically kicked spin-1/2 model V_3, Eq. (8.2). We consider a total system size of $N = 6$ qubits and present the simulated experimental results (black dots and error bars) for $K(t)$ and $K_A(t)$ using $M = 2 \times 10^5$ experimental runs for the single-shot sequence shown in Fig. 8.2 at each time t. We observe that the simulated experiment agrees with the exact numerical calculations at all times t within error bars. Here, error bars, indicating the standard error of the mean, quantify statistical errors arising from the finite measurement budget (i.e. the finite number M of simulated single-shot measurements), see next subsection.

8.1.3.3 Statistical errors and remarks

The SFF and PSFF can be accessed from the same set of measurement data via the estimators defined in Eqs. (8.7) and (8.8). Statistical errors arise in practice from a finite number M of experimental runs, and are governed by the variance of these estimators. We discuss statistical errors in detail via numerical and analytical calculations in Ref. [6], and find a typical scaling of $M \sim 10^{N_A} \approx 2^{3.32N_A}$ to access the (P)SFF of a (sub-)system of size N_A up to a fixed relative error. Such exponential scaling of the measurement effort reflects the exponential decrease of the SFF with system size [see remarks below Eq. (8.3)]. We emphasize however that this scaling of the experimental effort is substantially

better than for quantum process tomography which requires at least $\sim 2^{5N_A}$ experiments to reconstruct the full time evolution operator $T(t)$ [254]. Importantly, and in contrast to quantum process tomography, the initial state and the measurement basis coincide in our protocol.

As detailed in Ref. [6], we can further decrease the required number of experimental runs to observe the ramp and plateau of the (P)SFF, by considering a subsystem-averaged PSFF. Here, an average over PSFFs of all subsystems with a fixed size is performed. This results in a further improved signal-to-noise ratio.

Lastly, we remark that our protocol shares some similarities with randomized benchmarking [255–259], where however *global* random unitaries and their inverses are applied sequentially. In the case of randomized benchmarking the goal is to characterize noise and decoherence acting during the implementation of these global random unitaries. In contrast, with our protocol, the aim is to characterize a unitary time evolution operator $T(t)$ using *local* random unitaries $U = \bigotimes_i u_i$ applied before and after $T(t)$, which can be prepared with high fidelity [260, 261].

Organization of the Chapter— In the remainder of the manuscript, we elaborate on the contents of the above synopsis with technical details, derivations, and examples. In Sec. 8.2, we provide an in-depth theoretical analysis of the PSFF in RMT and in generic many-body models in relation to ETH. The analytic results are compared with numerics in Sec. 8.3 where we consider many-body models undergoing Floquet and Hamiltonian evolution. For the latter, we discuss both, RMT-like and localized regimes. As mentioned previously, the proof of the protocol and an analysis of statistical errors is not included here, and detailed in the original paper Ref. [6]. Lastly, we summarize in Sec. 8.4 with some concluding remarks and future directions.

8.2 Partial Spectral Form Factor: Analytical Results

In this section, we analyze the origin of the main features observed in the PSFF, namely the ramp, plateau and shift, based on analytical calculations. We provide arguments to show that the PSFF generically is a reliable probe of eigenvalue correlations characterizing RMT-like and localized regimes, signified by the presence and absence of a late time ramp-plateau structure respectively. In addition, we show that the specific features observed in the PSFF are related to the ensemble and spectrum averaged second-moments of reduced density matrices of eigenstates at different energies, and therefore provide a useful measure of eigenstate properties.

This section is organized as follows. In Sec. 8.2.1, we analyze the PSFF in standard Wigner-Dyson random matrix ensembles (see App. E.1 for a brief discussion), which we consider as a mathematically idealized prelude, in which the PSFF can be obtained exactly, to more realistic quantum RMT-like systems. These ensembles display the essential features of the PSFF and present a clear example of the roles of eigenvalue and eigenstate statistics in these features. This is followed by a discussion of more general RMT-like systems in Sec. 8.2.2, where we show using analytical arguments that the PSFF detects thermalization in the sense of ETH [18–22, 52, 66] in addition to level statistics (see also Ref. [42], that compares ETH for Floquet circuits to random matrix ensembles using the PSFF for specific subsystem sizes). We then discuss the PSFF in localized systems in Sec. 8.2.3, and summarize our main conclusions for all cases in Sec. 8.2.5.

Common to all these cases is the fact that the time-independent part of the PSFF in Eq. (8.4) is given by the plateau value, which depends only on the eigenstate purities (assuming no degeneracies) i.e. $K_A(t \to \infty) = \mathscr{P}_B/D_A$, where

$$\mathscr{P}_B = \frac{1}{D}\overline{\sum_i \mathrm{Tr}_B\left[\rho_B^2(E_i)\right]} \tag{8.9}$$

is the (spectrum- and ensemble-)averaged purity of the reduced energy eigenstates. For later reference,

we separate out this time-independent plateau value,

$$K_{\mathrm{A}}(t) = \frac{\mathscr{P}_B}{D_A} + \frac{1}{DD_A}\sum_{i\neq j}\overline{e^{i(E_i-E_j)t}\mathrm{Tr}_B\left[\rho_B(E_i)\rho_B(E_j)\right]}, \tag{8.10}$$

and note that the time-dependent second term only involves overlaps of distinct energy levels.

8.2.1 Random matrix ensembles

To understand the essential features of the PSFF we first analyze it in RMT, allowing for an exact determination of the PSFF. We choose Hamiltonians H (Floquet operators V) from the canonical Wigner-Dyson random matrix ensembles [28, 36, 262, 263], yielding time evolution operators $T(t) = \exp(-iHt)$ [$T(t = \tau n) = V^n$]. To evaluate the ensemble average in Eq. (8.10), we can utilize that for these RMT ensembles the eigenvalues and eigenstates of H (V) are uncorrelated. Thus, their ensemble average factorizes and can be performed independently. We find

$$K_A(t) = \frac{\mathscr{P}_B - \mathscr{Q}_B}{D_A} + D_B\mathscr{Q}_B K(t)\,, \tag{8.11}$$

where $\mathscr{Q}_B = (D(D-1))^{-1}\overline{\sum_{i\neq j}\mathrm{Tr}_B\left[\rho_B(E_i)\rho_B(E_j)\right]}$ and $\mathscr{P}_B$ are the averaged overlap and purities of the reduced eigenstates, respectively. We note that here the PSFF is the full SFF with a scaling factor $D_B\mathscr{Q}_B$ and a constant subsystem dependent shift $(\mathscr{P}_B - \mathscr{Q}_B)/D_A$ such that the entire time dependence of the PSFF is captured in the SFF. Therefore, the PSFF in these models preserves the characteristic ramp-plateau structure and the relevant time scales of the SFF.

As shown in Ref. [6], we can evaluate $\mathscr{P}_B$ and $\mathscr{Q}_B$ explicitly using Wigner-Dyson RMT for the eigenstates of H (V). They are functions of only the Hilbert space dimensions of subsystems A and B, i.e. $\mathscr{P}_B \equiv \mathscr{P}_B(D_A, D_B)$ and $\mathscr{Q}_B \equiv \mathscr{Q}_B(D_A, D_B)$. The precise functional form of $\mathscr{P}_B$ and $\mathscr{Q}_B$ depends on the symmetry class of the Hamiltonian H (Floquet operator V). For the case of the unitary Wigner-Dyson

ensembles, for example H from the Gaussian unitary ensemble or V from CUE, we find

$$\mathscr{P}_B = \frac{D_A + D_B}{D_A D_B + 1} \quad ; \quad \mathscr{Q}_B = \frac{D_B \left(D_A{}^2 - 1\right)}{D_A{}^2 D_B{}^2 - 1} \,. \tag{8.12}$$

The analogous expressions for orthogonal Wigner-Dyson ensembles can also be found in Ref. [6]. In both symmetry classes at $D_A, D_B \gg 1$, we have $\mathscr{P}_B - \mathscr{Q}_B \approx 1/D_A$ and $\mathscr{Q}_B \approx 1/D_B$. Thus, in this limit, the PSFF has a constant shift of $1/D_A^2$ added to the SFF and the slope of the ramp is the same as the slope of the ramp in the SFF, i.e. $K_A(t) \approx K(t) + 1/D_A^2$ [see also Eq. (8.5)].

8.2.2 General RMT-like systems

In the case of more general RMT-like systems, we begin by separating out the reduced density matrices of the energy eigenstates into smooth and fluctuating functions of energy,

$$\rho_B(E) = \frac{\mathbb{1}}{D_B} + \Delta\rho_B(E) + \delta\rho_B(E) \,. \tag{8.13}$$

Here, the first term is a constant corresponding to a maximally mixed reduced density matrix; $\Delta\rho_B(E)$ is traceless and a smooth function of E, while $\delta\rho_B(E)$ is again traceless but required to fluctuate rapidly with E. For our present purposes, it is useful to define the smooth and fluctuating parts in terms of their Fourier transforms with respect to a continuous energy variable as follows: for some cutoff time $t_\rho \ll O(D)$, we take their respective Fourier transforms to satisfy $(\Delta\tilde{\rho}_B(t))_{jk} = 0$ for $|t| > t_\rho$, and $(\delta\tilde{\rho}_B(t))_{jk} = 0$ for $|t| < t_\rho$ (with some additional details in App. E.2). The essence of the definition is that as a function of energy, the smooth part varies only over scales much larger than some energy window of size t_ρ^{-1} containing several levels, while the fluctuating part varies only over scales much smaller than t_ρ^{-1}.

We will further assume that $\delta\rho_B(E)$ behaves as if it is 'randomized' within these energy windows

over the ensemble i.e. it is uncorrelated with the smooth part and satisfies $\overline{\mathrm{Tr}_B\left[\delta\rho_B(E_i)\delta\rho_B(E_j)\right]} = \delta_{ij}\overline{\mathrm{Tr}_B\left[\delta\rho_B^2(E_i)\right]}$ for E_i, E_j closer than $\sim t_\rho^{-1}$, fluctuating around an average of zero (we do not require this behavior to persist over larger energy scales $|E_i - E_j| \gtrsim t_\rho^{-1}$). We note that this assumption is consistent with the general picture of random behavior over small energy windows in RMT-like systems [21], and we can justify it more generally (irrespective of whether the system/ensemble is RMT-like) as follows. In evaluating the SFF $K(t)$, the ensemble is usually chosen to have sufficiently large disorder so that the energy levels are randomly distributed over some large energy window, across different ensemble realizations. This is necessary to eliminate the erratic fluctuations of the SFF at large t that depend on the precise positions of levels, and obtain a smooth ensemble-averaged behavior (see e.g. Refs. [84, 110] for further discussion of this point). Our assumption is essentially that, this random redistribution of levels over different ensemble realizations extends to an energy window of $\sim t_\rho^{-1}$, effectively randomizing the fluctuations $\delta\rho_B(E)$ faster than this scale, while $\Delta\rho_B(E)$ which varies over scales larger than this energy window is not randomized in this manner. We also note that the eigenstates of a given ensemble realization themselves may additionally be random superpositions of those of a different realization, e.g. generally randomly mixing all eigenstates of the latter within the energy window in fully RMT-like systems (i.e. systems with no 'physical' conserved quantities other than energy) [18, 264–266], which gives further weight to this assumption.

8.2.2.1 Shift-ramp-plateau structure of the PSFF

Using the form in Eq. (8.13), the overlaps occurring in the definition of the PSFF in Eq. (8.4) separate out into independent contributions from each part of the reduced density matrix - the cross terms vanish, due to tracelessness for terms involving overlaps with the maximally mixed part, or due to the randomization of $\delta\rho_B(E)$ for terms involving overlaps of the smooth and fluctuating part for

$t \gg t_\rho$. We can write this as,

$$K_A(t \gg t_\rho) = K(t) + \Delta K_A(t) + \delta K_A(t)\,, \tag{8.14}$$

where $\Delta K_A(t)$ involves only overlaps of the form $\mathrm{Tr}_B\left[\Delta\rho_B(E_i)\Delta\rho_B(E_j)\right]$ and similarly, $\delta K_A(t)$ involves only those of the form $\mathrm{Tr}_B\left[\delta\rho_B(E_i)\delta\rho_B(E_j)\right]$. On decomposing $\delta K_A(t)$ in a manner analogous to Eq. (8.10), it follows that its time dependent part for $t \gg t_\rho$ (which sees contributions only from variations of the overlaps of fluctuating parts within energy windows smaller than $\sim t_\rho^{-1}$) vanishes on ensemble averaging, an important consequence of the randomization of $\delta\rho_B(E)$. This leaves only a constant contribution from the purity of the fluctuating part, $\delta K_A(t \gg t_\rho) = \delta\mathscr{P}_B/D_A$, where $\delta\mathscr{P}_B \equiv D^{-1}\overline{\sum_i \mathrm{Tr}_B\left[\delta\rho_B^2(E_i)\right]}$ (here we use 'purity' to generally mean $\mathrm{Tr}\left[x^2\right]$ for a Hermitian operator x). We see that this constant late-time shift is a generic feature of the PSFF, independent of the specific form of the full SFF $K(t)$. It merges into the plateau of the PSFF when $K(t)$ and $\Delta K_A(t)$ show only a plateau behavior - and therefore, the shift is an independent observable only if the other two terms show non-trivial time dependence at late times $t \gg t_\rho$.

We note that $\Delta K_A(t)$ is modulated only by a smooth function of two energy variables varying over scales larger than t_ρ^{-1}. For $t \gg t_\rho$, it should then essentially see the contribution to $K(t)$ from each part of the spectrum but modulated by the value of the function for nearly equal energies in that part. In App. E.2, we show this by direct calculation for a fully RMT-like system with Wigner-Dyson level statistics, obtaining a modulated linear ramp and plateau in addition to the late-time shift, for

$t \gg t_{\rm Th}, t_\rho$,

$$K_A(t \gg t_{\rm Th}, t_\rho) = \frac{\delta\mathscr{P}_B}{D_A} + \frac{1}{D}\begin{cases} (\beta_{\rm RMT}\pi D)^{-1}\gamma t\left(1 + D_B\widetilde{\Delta\mathscr{P}}_B\right) & \text{for } t \ll t_{\rm H}\,, \\ 1 + D_B\Delta\mathscr{P}_B & \text{for } t \gg t_{\rm H}\,. \end{cases} \tag{8.15}$$

Here, $\beta_{\rm RMT} = 1, 2$ respectively for the orthogonal and unitary classes, while $\gamma = \sum_i [\mathscr{N}^{\#}(E_i)]^{-1}$ is the range of energies in the spectrum with $\mathscr{N}^{\#}(E)$ representing the (smoothened) local density of states, in agreement with known results for the full SFF (see e.g. Refs. [38, 193]). To keep the expressions simple, we are ignoring corrections that are prominent near $t \sim t_{\rm H}$ [see, for instance, the exact form of the GOE SFF in Eq. (E.2)]; we focus instead on the $t \ll t_{\rm H}$ regime where the ramp appears linear for all values of $\beta_{\rm RMT}$ and profiles of $\mathscr{N}^{\#}(E)$, and the $t \gg t_{\rm H}$ regime with a constant plateau. However, both expressions are exact throughout the range of times when $\beta_{\rm RMT} = 2$ with constant density of states $\mathscr{N}^{\#}(E) = t_{\rm H}/(2\pi)$. We have also defined two ensemble-averaged quantities corresponding to slightly different spectrum averages of the purity of the smooth part, $\Delta\mathscr{P}_B = D^{-1}\overline{\sum_i \mathrm{Tr}_B\left[\Delta\rho_B^2(E_i)\right]}$ and $\widetilde{\Delta\mathscr{P}_B} = \gamma^{-1}\overline{\sum_i [\mathscr{N}^{\#}(E_i)]^{-1}\mathrm{Tr}_B\left[\Delta\rho_B^2(E_i)\right]}$, the latter including the contribution to the coefficient of the linear ramp from each part of the spectrum. We note that the purities of the smooth and fluctuating parts are (exactly) related to the overall average purity by $\mathscr{P}_B = D_B^{-1} + \Delta\mathscr{P}_B + \delta\mathscr{P}_B$, giving the expected plateau value of $\mathscr{P}_B/D_A$ in Eq. (8.15). There are also two competing time scales for the onset of the ramp, $t_{\rm Th}$ and t_ρ - the former entirely determines the behavior of $K(t)$ but the latter appears in $\Delta K_A(t)$ and $\delta K_A(t)$.

For direct comparison with numerics, it is useful to define the ensemble averaged overlap of

adjacent states, $Q_B = (D-1)^{-1} \sum_i \overline{\mathrm{Tr}_B \left[\rho_B(E_i)\rho_B(E_{i+1})\right]}$. Using Eq. (8.13), we note that,

$$\mathscr{Q}_B = \frac{1}{D_B} + \Delta\mathscr{P}_B \ ,$$
$$\mathscr{P}_B - \mathscr{Q}_B = \delta\mathscr{P}_B \ , \tag{8.16}$$

which follow from the assumption of uncorrelated $\delta\rho_B(E)$ in the ensemble, and taking $\Delta\rho_B(E_i) \simeq \Delta\rho_B(E_{i+1})$. We note that this definition of Q_B is equivalent to that in Sec. 8.2.1 for random matrix ensembles, where the ensemble averaged overlaps between distinct states are independent of their energies. Sec. 8.3 will directly use $\mathscr{P}_B$ and $\mathscr{Q}_B$, with the implicit assumption that $\widetilde{\Delta P}_B$ is of similar order of magnitude to $\Delta\mathscr{P}_B$ (due to $\mathscr{N}^{\#}(E)$ being of a similar order of magnitude throughout the spectrum) and is therefore similarly well represented by $\mathscr{Q}_B$.

8.2.2.2 Constraints from eigenstate thermalization

We have seen that at late times, the PSFF preserves the characteristic features of the SFF, such as the ramp and the Heisenberg time (as in Eq. (8.15) for fully RMT-like systems). However, there are non-negative subsystem-dependent parameters $\mathscr{P}_B$, $\delta\mathscr{P}_B$ and $\Delta\mathscr{P}_B$ ($\sim \widetilde{\Delta\mathscr{P}}_B$) that respectively influence the plateau value, the magnitude of the shift and the magnitude i.e., slope of the ramp. The purity $\mathscr{P}_B$ measures the extent of delocalization of eigenstates in a physical basis (e.g. a product basis of qubits), while we will see that $\delta\mathscr{P}_B$ and $\Delta\mathscr{P}_B$ are complementary probes of thermalization of these eigenstates. Specifically, we mean thermalization in the sense of ETH - that eigenstates corresponding to sufficiently close energies show nearly identical behavior in the dynamics of few-body observables [18–22, 66].

For our purposes, it is convenient to use subsystem ETH [52], which amounts to imposing ETH on an entire subsystem i.e. for all observables in the subsystem, and is directly expressed in terms

of reduced density matrices. It can be interpreted as the requirement of a small fluctuating part for the reduced density matrices of thermal eigenstates, as opposed to large fluctuations for non-thermal eigenstates. We can therefore apply it directly to the decomposition of reduced density matrices in Eq. (8.13). An important advantage of this version of ETH is that the dependence on subsystem size is made more explicit, whereas more conventional statements of ETH restrict themselves to few body operators, corresponding to extremely small subsystems and therefore negligible subsystem dependence. This subsystem size dependence will turn out to be the primary non-trivial indicator of the properties of eigenstates in the PSFF.

In App. E.3, we discuss the general constraints from (an extension of) subsystem ETH for eigenstates with an arbitrary extent of delocalization in a physical basis. Here, we present the results for a system with fully delocalized eigenstates, characterized by subsystem purities that follow the volume law of entanglement [140],

$$\mathscr{P}_B = D_B^{-1} + O(D_B^{-1}) + O(D_A^{-1}), \tag{8.17}$$

which cannot be less than D_B^{-1} as well as D_A^{-1}. This is the case relevant for the numerical examples of Sec. 8.3. If these eigenstates are thermal, subsystem ETH requires the smooth and fluctuating parts to satisfy,

$$\Delta\mathscr{P}_B = O(D_B^{-1}),\ \delta\mathscr{P}_B = O(D_A^{-1}). \tag{8.18}$$

Non-thermal eigenstates are characterized by much larger fluctuations, $\delta\mathscr{P}_B \gg O(D_A^{-1})$, with $\Delta\mathscr{P}_B$ being correspondingly smaller so as to satisfy the constraint $\mathscr{P}_B = D_B^{-1} + \Delta\mathscr{P}_B + \delta\mathscr{P}_B$. A narrower class of such RMT-like systems (e.g. Floquet systems) have uniformly random eigenstates that are distributed in close agreement with the standard random matrix ensembles (Sec. 8.2.1); the leading forms of the corresponding exact results in Eq. (8.12) are seen to be consistent with Eqs. (8.17),(8.18), on relating the two using Eq. (8.16). In this context, we note that Ref. [42] has observed subleading

corrections to the random matrix prediction for eigenstates in 1D Floquet quantum circuits.

8.2.3 Localized systems

Now, we consider localized systems, which generically show Poisson level statistics (i.e. uncorrelated neighboring levels) with localized non-thermal eigenstates, for strong disorder [51, 140]. Here, $K(t)$ shows only a plateau at late times, allowing us to access only the purity $\mathscr{P}_B$ through the PSFF. Fully localized states are essentially nearly pure states with $\mathscr{P}_B \sim O(1)$ (more precisely, following an area law of entanglement [140]), and additionally have large fluctuations $\delta\mathscr{P}_B \sim O(1) \leq 1 - D_B^{-1}$. In other words, fully localized states cannot thermalize, as they would have to be distributed over different physical basis states due to orthogonality. An $O(D_A^{-1})$ plateau value is, therefore, all we need to characterize the eigenstates of such systems.

On the other hand, when the eigenstates become more delocalized in the approach to an RMT-like regime, thermalization becomes a possibility. The moment any non-trivial correlations between nearby energy eigenvalues emerge in the spectrum, leading to a time dependence of $K(t)$ for $t > t_\rho$, $\delta\mathscr{P}_B$ becomes a meaningful observable in the PSFF according to the discussion following Eq. (8.14). Here, the PSFF can be used to study the extent of thermalization in addition to the delocalization of the eigenstates.

8.2.4 Illustration: Localized eigenstates with random-matrix-like eigenvalues

To provide a more explicit illustration of the sensitivity of the PSFF to eigenstate thermalization (instead of merely localization or entanglement), without as many generic randomness assumptions, we consider a simple model in which the energy eigenstates are the computational basis states $|\mathbf{s}\rangle$ themselves, while the energy eigenvalues $\{E_n\} = \{E_\mathbf{s}\}$ follow a Wigner-Dyson (e.g., CUE) distribution with a uniform density of states in $[0, 2\pi)$. Due to the independence of eigenvalues and eigenstates,

there is no mathematical restriction on the existence of such a system.

The PSFF for this system, in an arbitrarily chosen subsystem A, is given by

$$\begin{aligned} DD_A K_A(t) &= \sum_{(\mathbf{s}_A,\mathbf{s}_B),(\mathbf{s}'_A,\mathbf{s}'_B)} \langle \mathbf{s}_A,\mathbf{s}_B|e^{-iHt}|\mathbf{s}_A,\mathbf{s}'_B\rangle\langle \mathbf{s}'_A,\mathbf{s}'_B|e^{iHt}|\mathbf{s}'_A,\mathbf{s}_B\rangle \\ &= \sum_{\mathbf{s}_B}\left[\sum_{\mathbf{s}_A,\mathbf{s}'_A} e^{-i(E_{\mathbf{s}_A,\mathbf{s}_B}-E_{\mathbf{s}'_A,\mathbf{s}_B})t}\right]. \end{aligned} \tag{8.19}$$

Here, we have decomposed the computational basis into subsystems as $|\mathbf{s}\rangle = |\mathbf{s}_A\rangle_A \otimes |\mathbf{s}_B\rangle_B$.

Eigenstate thermalization within subsystem B— First, consider the situation where for a given value of $\mathbf{s}_B = \mathbf{s}_{B0}$, all the eigenvalues $E_{\mathbf{s}_A,\mathbf{s}_{B0}}$ are adjacent to each other in a subinterval $\Delta E_{B0} \subset [0, 2\pi)$ of approximate size $2\pi/D_B$. The reduced density operator ρ_B of all these eigenstates is a constant pure state within such a subinterval:

$$\rho_B(E) = |\mathbf{s}_{B0}\rangle_B\langle \mathbf{s}_{B0}|, \quad \text{for all } E \in \Delta E_{B0}. \tag{8.20}$$

This means there is no fluctuating part, $\delta\rho_B(E) = 0$ (except at specific values of E), and we should expect a vanishing shift in the PSFF. However, the purity of the smooth part (including the trace) is generically $\mathcal{Q}_B = 1$. Let us verify this expectation. Following Eq. (2.41), the contribution to the SFF from each such interval (due to the uniform density of states) resembles that of CUE at times $t = pD_B$ (where $p \in \mathbb{Z}$) with a sinc^2 transient at other times; note that the Heisenberg time is still $t_{\mathrm{H}} = D$. The contribution from within the brackets in Eq. (8.19) is then:

$$\sum_{\mathbf{s}_A,\mathbf{s}'_A} e^{-i(E_{\mathbf{s}_A,\mathbf{s}_{B0}}-E_{\mathbf{s}'_A,\mathbf{s}_{B0}})t}\Bigg|_{t<t_{\mathrm{H}}} = D_A^2 K_{\Delta E_{B0}}(t < t_{\mathrm{H}}) = D_A^2\,\mathrm{sinc}^2\left(\frac{t}{D_B}\right) + \frac{t}{D_B}. \tag{8.21}$$

The PSFF is obtained by summing over all D_B possible values of $\mathbf{s}_{B0}$ as in Eq. (8.19), which gives

$$K_A(t < t_{\mathrm{H}}) = \mathrm{sinc}^2\left(\frac{t}{D_B}\right) + \frac{t}{DD_A} \xrightarrow{t \gg DD_A^{-1/3}} \frac{t}{DD_A}. \tag{8.22}$$

Here, we have neglected the quantum fluctuations in the ramp, replacing these fluctuations with their expected average trend t. We have also implicitly identified the Thouless time $t_{\mathrm{Th}} \sim DD_A^{-1/3}$. Comparing this to the expectation $K_A(t_{\mathrm{Th}} \ll t \ll t_{\mathrm{H}}) = \delta\mathscr{P}_B/D_A + Q_B t/(DD_A)$ [see Eqs. (8.12) and (8.15)], we see that this behavior is consistent with $\delta\mathscr{P}_B = 0$ and $Q_B = 1$ as expected.

Nonthermal eigenstates within subsystem B— At the opposite extreme, consider the case adjacent computational basis states in terms of energy eigenvalues may be associated randomly with a bitstring $\mathbf{s}$. In this case, there is no eigenstate thermalization in subsystem B, and neighboring eigenstates could correspond to entirely different random pure states $|\mathbf{s}_B\rangle_B\langle\mathbf{s}_B|$ in B. While the overall purity of each eigenstate is still $\mathscr{P}_B = 1$, we expect all this purity to be fluctuating $\delta\mathscr{P}_B = \mathscr{P}_B - D_B^{-1}$ except for the maximally mixed part, with the smooth traceless part $\Delta\mathscr{P}_B$ vanishing.

To verify this claim, we note that of the D^2 energy difference terms that contribute to the full SFF, we can interpret Eq. (8.19) as randomly selecting $D_A^2 D_B$ of these terms in this case. The total ramp being $D^2K(t) = t$, each such term should be expected to contain $1/D_B$ of this contribution. At the same time, the eigenvalues corresponding to a fixed $\mathbf{s}_{B0}$ should now form an approximately Poisson spectrum of D_A eigenvalues in the full interval $[0, 2\pi)$. To leading order, we should have (taking $t \in \mathbb{N}$ to avoid the CUE transient)

$$\sum_{\mathbf{s}_A, \mathbf{s}_A'} e^{-i(E_{\mathbf{s}_A,\mathbf{s}_{B0}} - E_{\mathbf{s}_A',\mathbf{s}_{B0}})t}\Bigg|_{t\in\mathbb{N}} = O(D_A). \tag{8.23}$$

Adding up these D_B different contributions to $DD_AK(t)$ via Eq. (8.19), and including the t/D_B ramp

contribution by hand (which the above Poisson argument is not sensitive to), we get the expectation

$$K_A(t \in \mathbb{N} : t < t_{\mathrm{H}}) = O(D_A^{-1}) + \frac{t}{D^2}. \tag{8.24}$$

This heuristic is consistent with our expectations for the maximally mixed ramp and a nearly localized plateau above, in particular with $K_A(t_{\mathrm{Th}} \ll t \ll t_{\mathrm{H}}) = \delta\mathscr{P}_B/D_A + (D_B^{-1} + \Delta\mathscr{P}_B)t/(DD_A)$ [as in Eqs. (8.12) and (8.15)] after setting $\Delta\mathscr{P}_B = 0$ and $\mathscr{P}_B = D_B^{-1} + \Delta\mathscr{P}_B + \delta\mathscr{P}_B = 1$ [see Eq. (8.16)].

As both cases share Wigner-Dyson statistics for the eigenvalues and fully localized eigenstates with purity 1 in any subsystem, the different behaviors of the PSFF in the two cases, in line with our analytical expectations, demonstrates its sensitivity to the specific phenomenon of eigenstate thermalization in subsystems.

8.2.5 Summary

Let us summarize the main conclusions of this section from a unified perspective, before moving on to illustrate them with numerical examples in the next section. The PSFF in a subsystem A combines energy level statistics, as reflected in the SFF, with the purities and overlaps of the reduced energy eigenstates in the complementary subsystem B. The plateau value of the PSFF encodes the (spectrum and ensemble averaged) purity, which is $\sim O(1)$ in a fully localized regime, and small for fully delocalized states in accordance with the volume law of entanglement, Eq. (8.17). Something more interesting happens at late times if the SFF has a ramp or other time-dependent feature due to the existence of local level correlations. The PSFF inherits the ramp, but the ramp couples only to the smooth, slowly varying part of the reduced energy eigenstates. The rapidly fluctuating part is left over as a nearly time-independent shift [Eq. (8.15)].

Eigenstate thermalization is primarily encoded in the size of the fluctuating part as measured by the shift — namely, an exponential suppression of the latter with subsystem size N_A is indicative of

thermalization [Eq. (8.18)], while the lack of such a suppression translates to a failure of the eigenstates to thermalize. The smooth part is correspondingly large for thermal eigenstates and small for non-thermal eigenstates, so as to preserve the overall purity (i.e. extent of delocalization). Finally, there are special systems for which much more precise predictions for the PSFF can be theoretically derived/motivated and tested, such as RMT-like Floquet systems with their random-matrix-like eigenstates [Eqs. (8.11) and (8.12)].

Thus, the PSFF complements the SFF in analyzing late-time quantum ergodicity by being able to probe if the eigenstates satisfy ETH, in addition to (and because of) capturing information about level correlations as contained in the ramp of the SFF. In particular, we expect that it could potentially be useful in studying the joint emergence or loss of Wigner-Dyson level statistics and eigenstate thermalization (which are formally independent notions of late time quantum ergodicity) and their interdependence, across a transition or crossover between an RMT-like and non-RMT-like regime (as illustrated in a simple but artificial example in Sec. 8.2.4). This could be done by tuning the parameters of a system (say, in a quantum simulator) between such regimes, and measuring PSFFs across different choices of subsystems of different sizes - analyzing the extent of delocalization of eigenstates in the absence of a ramp via the plateau value, and additionally the extent of thermalization through the value of the shift if a ramp or other time-dependent feature is present at late times. Among the interesting possibilities that have been considered for such an intermediate regime, which could conceivably be probed with the PSFF, is the existence of so-called non-ergodic extended states [267–272] where the eigenstates are incompletely delocalized but do not thermalize, or alternatives in which the eigenstates thermalize without being fully delocalized [273].

8.3 Partial Spectral Form Factor: Numerical Results

Having discussed features of the PSFF and its connection to the SFF utilizing Wigner-Dyson random matrix ensembles and the ETH, we now present our numerical results of PSFFs in locally interacting many-body models, as realized in quantum simulators. For this purpose, we focus on two examples: the Floquet model Eq. (8.2) and the Hamiltonian model Eq. (8.25). Our results are in agreement with the analysis of the previous Sec. 8.2, in particular regarding the orders predicted for the averaged purity $\mathscr{P}_B$ and the overlap $\mathscr{Q}_B$ via Eq. (8.16). We consider the Floquet model in the RMT-like regime and the Hamiltonian model in both the RMT-like and localized regimes.

8.3.0.1 Example 1: Floquet system

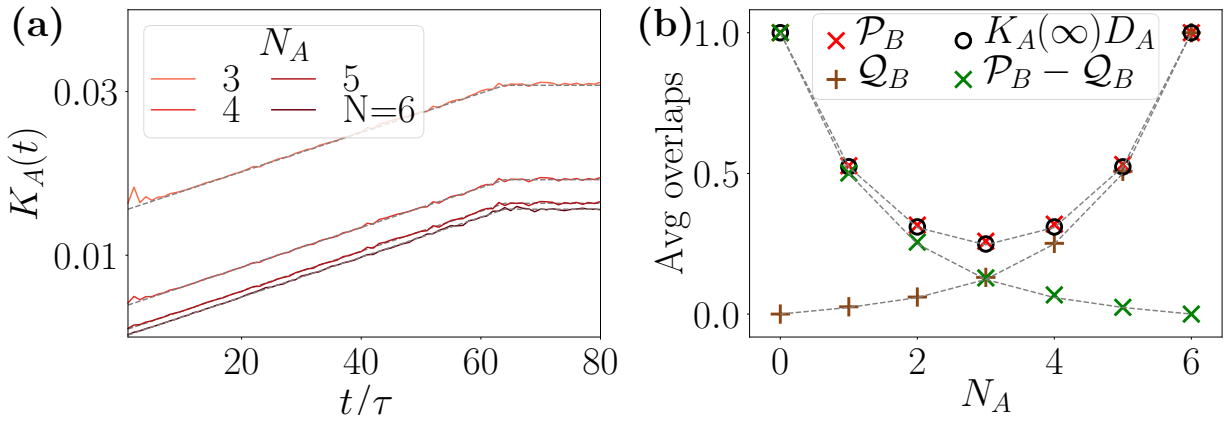


Figure 8.3: *Results for the Floquet V_3 model.* (a) The SFF and PSFF are presented for $N = 6$, $N_A = 3,\ 4,\ 5$ in red colors. In gray, we plot the same quantities in a CUE model. (b) The plateau value $K(\infty)$ multiplied with the subsystem dimension D_A is plotted in black circles and matches with the averaged purity $\mathscr{P}_B$ plotted with red crosses. The average overlap $\mathscr{Q}_B$ and the difference $\mathscr{P}_B - \mathscr{Q}_B$ are presented in brown and green respectively. We observe a perfect match with the respective quantities in CUE plotted in gray, indicating the same averaged eigenvalue and eigenstate statistics in CUE and V_3. In the numerical computation, we have taken 8000 disorder realizations to perform ensemble averaging and the subsystems A are chosen from the middle of the spin chain.

The Floquet time evolution operator V_3 has the same quasi-energy eigenvalue statistics as the CUE random matrix ensemble [96, 245]. As mentioned in Sec. 8.1 such Floquet models typically thermalize to infinite temperatures as per RMT and thus we expect the eigenstate statistics to also

be the same as in the corresponding RMT class. To show this, we present in Fig. 8.3(a) numerically obtained SFF and PSFF for a total system size of $N = 6$ and subsystem sizes $N_A = 3, 4$ and 5 for the model V_3. We plot with gray lines the corresponding $K_A(t)$ in a CUE model where the analytic forms can be exactly calculated (see Sec. 8.2.1). For the PSFF $K_A(t)$ at $N_A = 3$ and very early times, we notice that the onset of the ramp takes a few initial periods to set, but eventually the PSFF follows the CUE prediction.

The closeness between the statistics of CUE and V_3 can further be seen from the average overlaps of reduced densities of eigenstates $\mathscr{P}_B$ and $\mathscr{Q}_B$. In Fig. 8.3(b) we present the average purity and overlaps as functions of subsystem size N_A. At plateau time, $t > t_{\rm H} (= D\tau)$ the PSFF becomes $K_A(t \to \infty) = \mathscr{P}_B / D_A$, see Eq. (8.10). We plot numerically obtained $K_A(\infty) D_A$ in black circles, and the average purity $\mathscr{P}_B$ with red crosses, they confirm the analytic expectation. The average overlap $\mathscr{Q}_B$ and the difference $\mathscr{P}_B - \mathscr{Q}_B$ are plotted in brown and green circles respectively and match with the CUE data.

To conclude, the SFF, PSFF, averaged purity and overlaps match in the CUE and V_3 model and thus we expect the form of the PSFF in Eq. (8.5) to hold for the model V_3, after a small initial time period. We know from Eq. (8.12), for large Hilbert space dimensions, that $\mathscr{Q}_B \approx 1/D_B$ and $\mathscr{P}_B - \mathscr{Q}_B \approx 1/D_A$. Therefore utilizing, Eq. (8.16), we find that $\Delta\mathscr{P}_B = 0$ and $\delta\mathscr{P}_B = O(1/D_A)$ for V_3 and the RMT models. The purity of the smooth part (of the form of $\mathrm{Tr}[\Delta\rho_B^2(E)]$) appears in the ramp part of the PSFF in Eq. (8.15) and thus we note that the ramp coefficient is $\sim 1/D^2$ for $D_A \gg 1$. On the other hand, the purity of the fluctuating part (of the form of $\mathrm{Tr}[\delta\rho_B^2(E)]$) comes in the time-independent term added to the SFF in Eq. (8.15), which is to the leading orders $1/D_A^2$, as also in the CUE model [Eq. (8.5)].

8.3.0.2 Example 2: Hamiltonian system

As our second example, we consider a transverse field Ising model in presence of longitudinal local disorders,

$$H = J\left(\sum_{\substack{i,j=1\\ i<j}}^{N} \frac{1}{(i-j)^{\alpha}}\sigma_i^z\sigma_j^z + \sum_{i=1}^{N}\sigma_i^x\right) + W\sum_{i=1}^{N} h_i\sigma_i^z, \tag{8.25}$$

where h_i are drawn uniformly at random from $(-1,1)$. The coefficient J and the exponent α denote the strength and range of the interactions respectively. The disorder strength W is known to specify the nature of the dynamics; $W \sim J$ depicts RMT-like regime and $W \gg J$ corresponds to the localized regime (for a similar model see, [38]). In the App. E.4.1, we present the adjacent level gap ratio as a function of W/J and α and find that the RMT-like and localized regimes exist for short ($\alpha > 1$) as well as for long ($\alpha < 1$) range interactions. In this work, we choose $\alpha = 1.2$, and as examples of the RMT-like and localized regimes, we take $W = J$ and $W = 10J$ respectively. In contrast to the presence of the ramp and plateau in the SFF for RMT-like models, the SFF for localized models stays flat for all times $t \gg 0$. In the numerics, we will find that the PSFF preserves this flat feature of the SFF, and has a subsystem dependent shift added over the SFF, as predicted in Sec. 8.2.3. In Fig. 8.4 and 8.5 we present numerical results for the Hamiltonian model (8.25) in these two regimes. For clarity, we have used red color for the RMT-like regime ($W = J$) and blue for the localized regime ($W = 10J$). We note that the Hamiltonian of Eq. (8.25) has the time-reversal symmetry of complex conjugation in the computational (σ_i^z) basis [28, 274, 275]. An RMT-like Hamiltonian with this symmetry is known to follow the eigenvalue statistics (or the SFF) of GOE after the Thouless time $t > t_{\text{Th}}$ [21, 28, 36, 274, 275], thus we have also put the results for GOE class in gray in Fig. 8.4.

As a side remark, we emphasize at this point that the spectrum of the local Hamiltonian model, Eq. (8.25), does not have the same density of states as the GOE spectrum and thus the Hamiltonian SFF should be compared with an average of GOE SFFs, each with t_{H} determined by different parts of

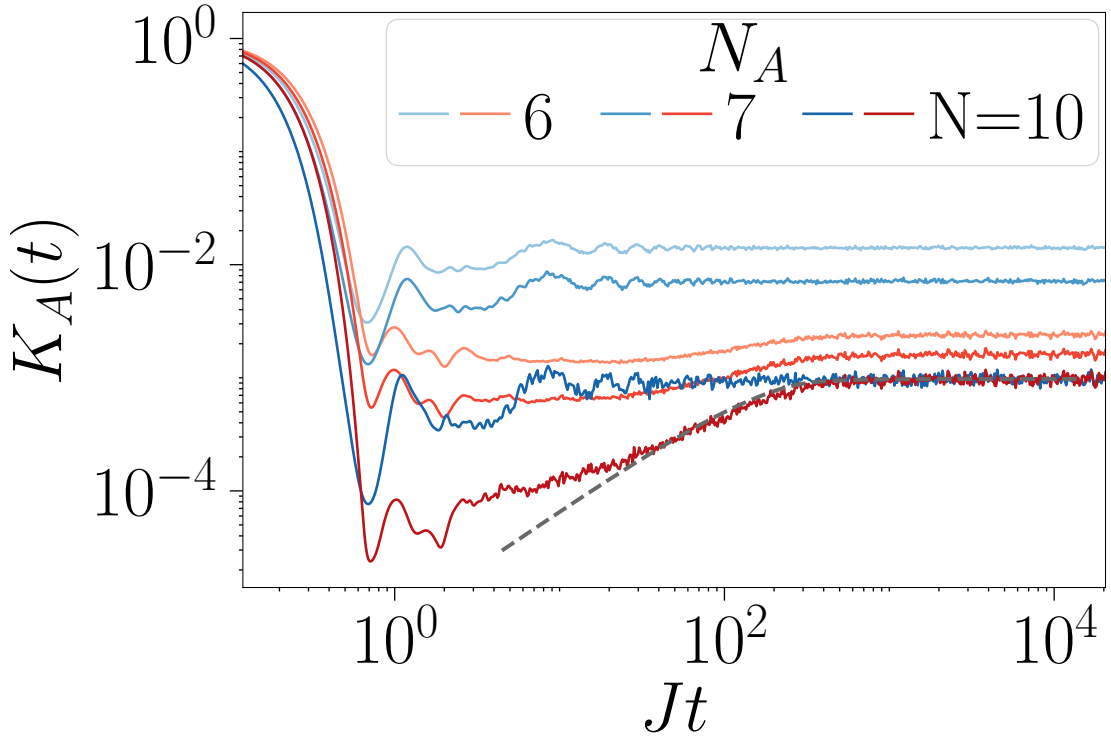


Figure 8.4: *Results for the Hamiltonian model.* In a log-log plot we present the RMT-like regime ($W = J$) in red, localized regime ($W = 10J$) in blue, and the GOE in gray. In both regimes the SFF and PSFF are plotted for (sub-)system sizes $N_A = 6$, 7 and $N_A = N = 10$. The SFF for the RMT-like regime has the characteristic ramp and plateau and follows the GOE SFF at late times. The PSFF in this regime also has the shift, ramp and plateau, we plot these in a focused linear scale plot in Fig. 8.5(a). The localized regime shows a flat SFF and PSFF for all times $t \gg 0$. The mean level spacing (i.e. the Heisenberg time) in the localized regime and GOE are numerically rescaled to match to the one in the RMT-like regime.

the Hamiltonian spectrum. Often, this is circumvented by removing the non-universal effects arising from the edges of the local Hamiltonian spectrum by using a filter function such that only the middle part of the spectrum contributes [236] or considering very large system sizes where the edge effects are effectively smaller. In our work, we focus on the measurement of RMT-like features through the observation of the ramp, plateau and the shift which can already be observed without filtering for moderate system sizes, which we focus on.

In Fig. 8.4, the SFF and PSFF are presented for the system size $N = 10$ and subsystem sizes $N_A = 6$ and 7. In order to have the same Heisenberg time t_{H}, the eigenvalues are numerically rescaled such that the average mean level spacing for $W = 10J$ match with the one for $W = J$. As a guide,

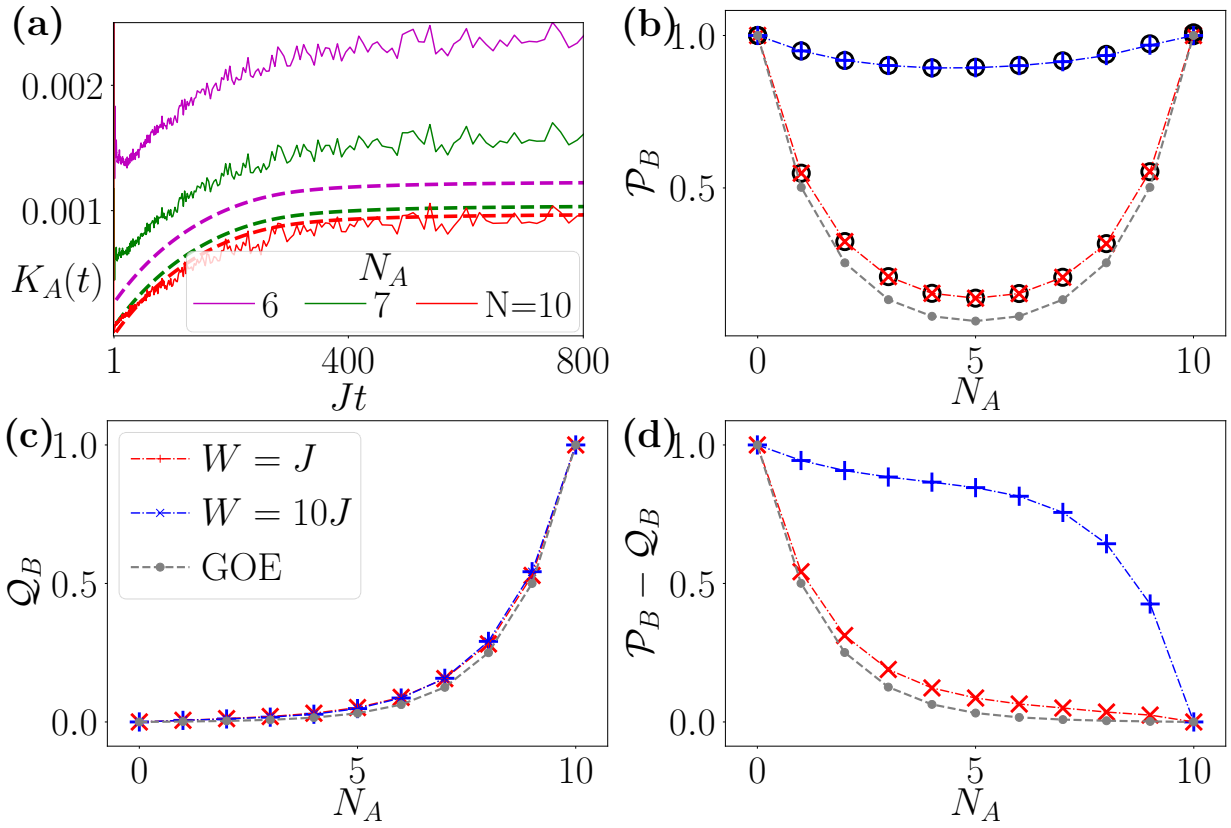


Figure 8.5: *Results for the Hamiltonian model.* (a) In linear scale we present the SFF and PSFF for the RMT-like regime ($W = J$). (Sub-)system sizes $N_A = 6$, 7 and $N_A = N = 10$ are plotted with magenta, green and red respectively for both the Hamiltonian model (with solid curves) and the GOE (with dashes). We observe differences in the PSFF for RMT-like Hamiltonian and GOE. These differences are investigated in (b), (c) and (d) through $\mathscr{P}_B$ and $\mathscr{Q}_B$. We use red color for the RMT-like regime ($W = J$) and gray for the GOE. For comparison we have also plotted these quantities in the localized regime ($W = 10J$) using blue color. (b) We plot $K_A(\infty)D_A$ using black circles which matches with the corresponding average purity $\mathscr{P}_B$ of the localized and RMT-like regime. (c) The average overlap $\mathscr{Q}_B$ for localized, RMT-like and GOE follow closely the behavior $1/D_B$. (d) The difference $\mathscr{P}_B - \mathscr{Q}_B \approx \delta\mathscr{P}_B$, which encodes the shift of the PSFF, is larger for large disorders (localized) compared to small disorders (RMT-like). In the numerical computation, we have taken 200 Hamiltonians to perform ensemble averaging and the subsystems A are chosen from the middle of the spin chain.

we have plotted in gray the GOE SFF where the $t_{\rm H}$ is determined from the full width of the RMT-like Hamiltonian spectrum and observe that the SFF for the RMT-like regime follows the GOE SFF closely. The PSFF for the RMT-like regime, shifted up compared to the SFF, also shows the ramp and plateau behavior which are seen better in a linear plot in Fig. 8.5(a). Here, focused to display RMT-like features, we have used solid lines for the RMT-like Hamiltonian and dashes for the GOE. The different subsystem sizes are shown in different colors. We note that the PSFF for the RMT-like local model and

GOE are different (see the magenta and green curves). These differences arise due to the differences in eigenstate properties of the local Hamiltonian and GOE.

Further, to concretely discuss second-moments of eigenstates, in Fig. 8.5(b) we present the averaged purity $\mathscr{P}_B$ using crossed markers. We have also plotted here the plateau values $K_A(\infty)D_A$ (in black circles) for both RMT-like and localized regimes which agree with their respective purities following $K_A(t \to \infty) = \mathscr{P}_B/D_A$ [see Eq. (8.10)]. Note that these average purities are consistent with a volume law of entanglement in the RMT-like regime, and an area law in the localized regime [140]. For the remainder of this section, it is useful to discuss the two regimes $W = J$ and $W = 10J$ separately.

For the RMT-like regime $W = J$, the average overlaps $\mathscr{Q}_B$ and $\mathscr{P}_B - \mathscr{Q}_B$ are presented in red in the bottom panel of Fig. 8.5 as functions of N_A. Assuming ETH for the RMT-like systems, we have discussed orders of magnitude of these overlaps in Sec. 8.2.2. Utilizing Eq. (8.16) we can comment on the orders of $\Delta\mathscr{P}_B$ and $\delta\mathscr{P}_B$ (see App. E.4.2 for more details on the numerical extraction of these orders). From $\mathscr{Q}_B$ [Fig. 8.5(c)], we find $\Delta\mathscr{P}_B = O(1/D_B)$ and from $\mathscr{P}_B - \mathscr{Q}_B = \delta\mathscr{P}_B$ [Fig. 8.5(d)], we find $\delta\mathscr{P}_B \sim O(1/D_A)$, confirming the ETH predictions for RMT-like systems. We verify that the value of the shift of the PSFF in the linear ramp region is given in terms of the purity of the fluctuating part i.e., by $\delta\mathscr{P}_B/D_A$ in App. E.4.3. For comparison, we have plotted the same quantities in a GOE model in gray. We note a difference between the overlaps (properties of the eigenstates) in the local RMT-like Hamiltonian and GOE, which is not surprising because the statistics of eigenstates need not be the same in the two models.

Next, we look at the orders of magnitude of the overlaps in the regime $W = 10J$, plotted in blue in the bottom panel of Fig. 8.5. Following Eq. (8.16) from the $\mathscr{Q}_B$ [Fig. 8.5(c)], we find $\Delta\mathscr{P}_B = O(1/D_B)$ and from $\mathscr{P}_B - \mathscr{Q}_B = \delta\mathscr{P}_B$ [Fig. 8.5(d)], we find $\delta\mathscr{P}_B \sim O(1) \gg O(1/D_A)$. The localized regime is not expected to satisfy ETH, and as discussed in the Sec. 8.2.3, we expect such large shift in the PSFF in localized systems. Due to larger $\delta\mathscr{P}_B$ in the localized regime, we notice a

larger overall shift of the PSFF in the localized regime, shown in blue in Fig. 8.5(b)-(d).

8.4 Conclusion and outlook

In this Chapter, we have presented randomized measurement protocols to access the statistics of energy eigenvalues and energy eigenstates of many-body quantum systems in present day quantum simulators via (partial) spectral form factors. The spectral form factor (SFF), $K(t)$ in Eq. (8.1), is known to be a key diagnostic of many-body quantum ergodicity. In RMT-like systems, it reveals universal properties of energy eigenvalue statistics and possesses a characteristic ramp-plateau structure (see Sec. 8.1.1). In addition, we have defined partial spectral form factors (PSFFs), $K_A(t)$ in Eq. (8.4), which probe both the statistics of energy eigenvalues and eigenstates (see Sec. 8.1.2). PSFFs are natural restrictions of the SFF to subsystems $A \subseteq \mathscr{S}$ of the full system $\mathscr{S}$, such that for $A = \mathscr{S}$, PSFF and SFF coincide $K_{A=\mathscr{S}}(t) = K(t)$. Utilizing random matrix theory and the eigenstate thermalization hypothesis (ETH), we have shown in Sec. 8.2, that PSFFs in generic RMT-like quantum many-body systems possess a characteristic shift-ramp-plateau structure [Eqs. (8.11) and (8.15)] and reveal crucial differences between thermal and non-thermal eigenstates in the sense of ETH. In Sec. 8.3 we investigated the PSFF numerically in examples of many-body quantum models, discussing, in particular, differences between RMT-like and localized regimes.

With our protocol to measure the SFF and PSFF in quantum simulation experiments, we have extended the toolbox of randomized measurements [231] to access genuine properties of dynamical quantum evolution, without any reference to the initial state or measured observable. We have shown that our protocol gives simultaneous access to the SFF and PSFF, thereby providing a unified testbed of the statistical properties of eigenvalues and eigenstates. Our protocol can be directly implemented in state-of-the-art quantum devices, based for instance on trapped ions [232, 233], Rydberg atoms [234] and superconducting qubits [125, 235], providing crucial experimental tools for the quantum simula-

tion of many-body quantum ergodicity and the study of thermalization in closed quantum systems.

In this Chapter, we have focused on probing the late-time features of the SFF, without emphasizing the early-time behavior. As noted in Chaps. 3 and 4, this early-time behavior ($t < t_{\rm Th}$) can be quite interesting to study when one is interested in quantum speed limits and the dynamics of thermalization. Needless to say, the measurement protocol here can be applied without modification to the early-time regime as well. However, it is not obvious at this stage what unique insights the PSFF may bring to the early-time regime. It is possible that it may provide a tighter speed limit on a subsystem than the bound in terms of the full SFF in Chap. 3; this and other possibilities would be interesting to explore further.

As noted in Sec. 8.1, this protocol has been implemented with up to 5 superconducting qubits, both for Floquet and Hamiltonian evolution [168]. In this case, decoherence is sufficiently small to be treated as experimental noise, and is corrected through the depolarization model of Ref. [6]. However, when the decoherence is extremely strong such that almost no trace of closed-system unitary dynamics remains, we will have to suitably update the signatures of ergodicity we look for, in measurements of the SFF in open systems. This is the subject of the next Chapter (Chap. 9), where we are led in this direction by experimental data gathered in a trapped ion implementation of the above protocol.

Chapter 9: Probing spectral signatures of ergodicity in open quantum systems

9.1 Introduction: Ergodic Theory in open quantum systems

Up to this point, we have primarily concerned ourselves with characterizing and detecting signatures of ergodicity in closed quantum systems undergoing Hamiltonian or Floquet dynamics. We only deviated from the closed system formulation twice: once in Chap. 3 where we showed that the speed limit set by the spectral form factor (SFF) generalizes to time-dependent quantum operations (which was later used in Chap. 4, albeit for Hamiltonian dynamics), and subsequently in Chap. 8, where we mentioned the need to account for decoherence in experimental measurements of (partial) SFFs. The latter is still geared towards closed systems, treating decoherence as an experimental error to be mitigated rather than a phenomenon of interest in its own right.

A related question is to address the role of observable-independent notions of ergodicity in open quantum dynamics that includes decoherence as a phenomenon of interest, and how it can connect to concrete, measurable statistical signatures. Most directly from an experimental standpoint, such a question is forced on us if decoherence is too strong to have reasonable expectations of error mitigation and recovering the signatures of ergodicity associated with closed systems. This is the situation in the present Chapter, in which a trapped ion implementation of the measurement protocol of Chap. 8 was faced with a high degree of decoherence, for which a theoretical approach that took open systems as its starting point was necessitated.

Here, we first expand on the nature of ergodicity in time-independent open quantum systems

by focusing on the appropriate concepts of ergodicity that can be related to the eigenvalues of the associated quantum channel or Lindbladian operator, following and slightly generalizing Ref. [227]. In contrast to the elaborate measures of spectral statistics and discrete quantum ergodic criteria that had to be introduced for autonomous Hamiltonian and Floquet dynamics in Ref. [1] (see Chap. 2), we will discuss that open system ergodicity is considerably simpler, and involves only the existence and multiplicity of eigenvalues of the quantum channel on the unit circle in the complex plane. This almost perfectly mirrors classical ergodicity in measure-preserving autonomous systems (e.g., classical Hamiltonians), involving only the degeneracy and existence of eigenvalues (and eigenvectors) of the Koopman-von Neumann unitary on the unit circle (see Sec. 2.2.1).

Subsequently, we explicitly identify the direct signatures of open system ergodicity in (generalized) spectral form factors (SFFs), which directly probe the multiplicity of the spectrum of the quantum channel. We conclude by demonstrating the experimental measurement of these signatures in a trapped ion quantum spin chain, using the randomized measurement protocol developed in Ref. [6] (see Chap. 8). The quantities being measured are, specifically, the direct generalization of the partial spectral form factor (PSFF) to quantum channels for the subsystems A of interest:

$$K_A(t) = \sum_r \mathrm{Tr}_B \left[\mathrm{Tr}_A(\hat{\mathscr{K}}_r(t)) \, \mathrm{Tr}_A(\hat{\mathscr{K}}_r^\dagger(t)) \right], \tag{9.1}$$

whose estimation can be efficiently done using our randomized measurement protocol [6]. Eq. (9.1) is the generalization to subsystems of the SFF for quantum operations considered in Ref. [2] in the context of universal speed limits on thermalization (see Chap. 3). For full Lindbladian systems, it reduces to the dissipative form factor studied in Refs. [150–153,276] (related via analytic continuation in t to the form in Ref. [133], suited for non-Hermitian spectral statistics). However, it differs from the state-dependent fidelity-based generalizations in Refs. [169,170] suited to other applications.

As has been previously noted [6, 96] (see Chap. 8), and also illustrated in this Chapter, the experimental observation of ergodic properties is naturally challenging due to the significant complexity of energy eigenstates in thermalizing systems [51, 140] as well as the observable-independence of these properties. In particular, an immediate probe of ergodicity in quantum many-body systems is faced with the challenge of exponential decrease of the relevant features of the PSFF with increasing system size N. We will therefore focus on a system of $N = 4$ qubits in our experimental measurements, which is more accessible with present-day quantum technology.

Organization of this chapter— In Sec. 9.2, we review the connection between ergodicity and eigenvalues in open quantum systems, and present its identification in the generalized PSFF. In a surprising contrast to closed systems, the SFF provides a direct probe of ergodicity and thermalization of all observables in a system undergoing open quantum dynamics. In this sense, our work demonstrates an accessible *direct* probe of the fundamental emergence of statistical ensembles, as a unique feature of open systems that can be implemented and detected in present-day quantum technology. In App. F.2 we derive analytic formulas of generalized PSFF in non-interacting and ergodic systems when there are global dephasing errors present. In Sec. 9.3 we discuss our measurements of generalized PSFF in a trapped ion setup. We have worked with $N = 4$ spin$-1/2$ particles undergoing a Floquet time-evolution. We briefly summarize our measurement results, and compare them with the theoretical expectations outlined in Sec. 9.2. Our discussion of these aspects will focus on their relevance to the ergodic theory of open quantum systems, and we refer to the Ph.D. dissertation of K. S. Collins [277] for more specific details on the experimental implementation.

9.2 Spectral signatures of open system ergodicity

Most real-world systems are open, and are necessarily subject to decoherence and dissipation with the environment in addition to internal interactions. To probe the validity of statistical mechanics

in such systems, both theoretically and experimentally, we must concern ourselves with ergodicity in open quantum systems. In this section, we describe the necessary concepts and observable experimental signatures of ergodicity for a completely positive, trace preserving (CPTP) quantum channel [50]. We assume an autonomous (time-independent) quantum channel, whose repeated action at a given time generates its action at future time steps (in discrete time), or is generated by a Lindbladian (in continuous time).

9.2.1 Ergodicity and thermalization in quantum systems

Thermalization effectively describes the ability of a system to "forget" the details of its initial state, except for some macroscopic/thermodynamic parameters such as total energy (for a closed system) or the temperature of the environment (for an open system). In a quantum system, this is most conveniently described through a thermal state $\hat{\rho}_{\text{th}}$ that depends only on these thermodynamic parameters. Given a time evolving state $\hat{\rho}(t)$ and any observable $\hat{A}$, we consider the system to thermalize in the strong sense when it becomes independent of the initial state at any long time, i.e.,

$$\lim_{t\to\infty} \text{Tr}[\hat{\rho}(t)\hat{A}] = \text{Tr}[\hat{\rho}_{\text{th}}\hat{A}], \tag{9.2}$$

to leading order (e.g., in the thermodynamic limit of many particles) for any initial state. This corresponds to the classical property of *strong mixing*; we will not consider the related property of "weak mixing" (thermalization at almost all times), which is less relevant for finite-dimensional open quantum systems but occurs, for instance, in the context of the eigenstate thermalization hypothesis for unitary systems [20, 21] (see also related discussions in Chap. 1). An important weaker form of thermalization is ergodicity, which refers to a time-averaged independence of the initial state:

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}t\, \text{Tr}[\hat{\rho}(t)\hat{A}] = \text{Tr}[\hat{\rho}_{\text{th}}\hat{A}], \tag{9.3}$$

again to leading order. The integral may be replaced by a sum for discrete-time systems. As we will discuss below, these properties have a direct relation to the eigenvalues of a quantum channel, and can be diagnosed through a measurement of the generalized SFF.

9.2.2 Ergodicity and the eigenvalues of time evolution

For both closed and open quantum systems, the eigenvalues of the appropriate time evolution operator capture all observable-independent aspects of quantum dynamics, which are intrinsic to the system and do not depend on any specific choice of observables. In this regard, the connection between quantum statistical mechanics in a closed system and the eigenvalues of the corresponding Hamiltonian or Floquet operator has been explored in some detail, most frequently by way of comparison with ensembles of random matrices [28, 36]. However, one can also directly consider the quantum dynamical effects of eigenvalue statistics, in which case one obtains a direct connection to a "discretized" form of ergodicity. In particular, the statistical distribution of (quasi-)energy eigenvalues, including random matrix signatures, can be related to a form of ergodic dynamics in the Hilbert space that entails exploring a complete orthonormal basis [1] (see Chap. 2). However, the precise conditions under which there is a correspondence between this sense of ergodicity of closed quantum systems and the behavior of observables indicated by Eqs. (9.3), (9.2) remains quite challenging to establish.

Surprisingly, the conceptual details are significantly simpler in open quantum systems, and one can directly relate Eqs. (9.3), (9.2) to simple properties of existence and non-degeneracy of certain eigenvalues of the (finite-dimensional) CPTP quantum channel that effects time evolution [227]. We will largely follow the discussion in Ref. [227], and (initially) make a distinction between two types of ergodicity for formal reasons. To be concrete, let us consider the channel

$$\hat{\rho}(t) = \mathscr{T}^t[\hat{\rho}] = \sum_{r=1}^{M_t} \hat{\mathscr{K}}_r(t)\hat{\rho}\hat{\mathscr{K}}_r^\dagger(t), \tag{9.4}$$

obtained via the repeated action over t time steps of the channel $\mathscr{T}$ at $t = 1$, with distinct eigenvalues Λ_k (i.e., where degenerate eigenvalues are not counted separately). For each Λ_k (of algebraic multiplicity $m_k \geq 1$, say), there exists a non-empty set of linearly independent eigenoperators $\hat{L}_{k,r}$ (numbering ℓ_k, the geometric multiplicity of Λ_k, such that $1 \leq \ell_k \leq m_k$) satisfying

$$\mathscr{T}[\hat{L}_{k,r}] = \Lambda_k \hat{L}_{k,r}. \tag{9.5}$$

The eigenvalues satisfy $|\Lambda_k| \leq 1$, with $\ell_j = m_j$ for all unimodular eigenvalues $|\Lambda_j| = 1$, due to the CPTP nature of the map. Further, there always exists an eigenvalue $\Lambda_0 = 1$ with an eigenoperator $\hat{L}_{0,0}$ satisfying $\hat{L}_{0,0} = \hat{L}_{0,0}^\dagger$, $\hat{L}_{0,0} \geq 0$, $\mathrm{Tr}[\hat{L}_{0,0}] = 1$, i.e. $\hat{L}_{0,0}$ is a density operator.

For this setup, the following cases are of interest:

1. Density operator ergodicity: The eigenvalue $\Lambda_0 = 1$ has only one eigenoperator ($\hat{L}_{0,0}$) that is also a density operator. This is necessary and sufficient for Eq. (9.3) to be satisfied for any operator $\hat{A}$ and initial state $\hat{\rho}(0)$, with the thermal state $\hat{\rho}_{\mathrm{th}} = \hat{L}_{0,0}$.

2. Full ergodicity: An apparently stronger form of ergodicity is when $\Lambda_0 = 1$ is a non-degenerate eigenvalue, i.e. having algebraic multiplicity $m_0 = 1$, in which case Eq. (9.3) would be satisfied for any operator $\hat{\rho}$ that is not restricted to be a density operator. However, it turns out that density operator ergodicity is sufficient (and evidently necessary) to guarantee full ergodicity [227], allowing us to use the latter to obtain a fully spectral signature of the former.

3. Thermalization: $\Lambda_0 = 1$ is the unique, non-degenerate ($m_0 = 1$) eigenvalue of unit modulus, i.e., every other eigenvalue is inside the unit disk $|\Lambda_{k\neq 0}| < 1$. This is necessary and sufficient for Eq. (9.2) to hold for any operator $\hat{A}$ and initial state $\hat{\rho}(0)$, with the thermal state $\hat{\rho}_{\mathrm{th}} = \hat{L}_{0,0}$.

We note that similar criteria also apply to emergent quantum channels in the study of subsystems of closed systems; for instance, similar eigenvalue properties have been used to characterize

local correlation functions in dual-unitary quantum circuits [92–94]. Most surprisingly, they are very reminiscent of the properties of Koopman-von Neumann (or Frobenius-Perron) unitary operators tied to ergodicity and mixing in *classical* dynamical systems [10] (see also Chap. 2), a close similarity that is not realized to a comparable level even in unitary quantum systems.

9.2.3 Signatures in the spectral form factor

The generalized SFF for a quantum channel is given by [2, 150–153]

$$K(t) \equiv \frac{1}{D^2} \operatorname{Tr}_{\mathscr{H}\otimes\mathscr{H}}(\mathscr{T}^t) = \frac{1}{D^2} \sum_{r=1}^{M_t} \left| \operatorname{Tr} \hat{\mathscr{K}}_r(t) \right|^2 . \tag{9.6}$$

In terms of the eigenvalues Λ_k, this can directly be written as

$$K(t) = \frac{1}{D^2} \sum_k m_k \Lambda_k^t . \tag{9.7}$$

The behavior of $K(t)$ as $t \to \infty$ can directly diagnose full ergodicity and thermalization in the quantum channel, these being directly reflected in the properties of eigenvalues. As discussed previously, these properties are directly related to Eqs. (9.3) and (9.2), allowing the SFF to probe these dynamical features without any specific knowledge of the eigenoperators $\hat{L}_{k,r}$. In the long-time limit, only the unimodular eigenvalues $|\Lambda_j| = 1$ contribute to the SFF:

$$K(t \to \infty) = \frac{m_0}{D^2} + \frac{1}{D^2} \sum_{\substack{j \neq 0 \\ |\Lambda_j|=1}} m_j e^{-i\lambda_j t} \tag{9.8}$$

where we have separated out the contribution from Λ_0 and written $\Lambda_j = e^{-i\lambda_j}$. We obtain:

1. Signature of non-ergodicity: By non-ergodicity, we here mean that the open system does not show full ergodicity. In this case, there is an $m_0 > 1$ algebraic multiplicity of the eigenvalue

$\Lambda_0 = 1$, and we have

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}t\; K(t) = \frac{m_0}{D^2}, \quad m_0 > 1. \tag{9.9}$$

This includes every unitary (Hamiltonian or Floquet) system, for which $m_0 = D$. Indeed, no unitary system is ergodic in the sense of Eq. (9.3) for every state and observable (e.g., when both are energy projectors), which necessitates alternate approaches [1] as constructed in Chap. 2.

2. Signature of full ergodicity: The unit eigenvalue $\Lambda_0 = 1$ is non-degenerate, $m_0 = 1$. In this case, the SFF has a constant $1/D^2$ plateau from Λ_0 plus possible sinusoidal oscillations [that average out to 0 over infinite time] from the other $|\Lambda_j| = 1$ eigenvalues (if they exist). Thus, we have

$$\lim_{T\to\infty} \frac{1}{T}\int_0^T \mathrm{d}t\; K(t) = \frac{1}{D^2}. \tag{9.10}$$

3. Signature of thermalization: All eigenvalues other than the non-degenerate Λ_0 satisfy $|\Lambda_{j\neq 0}| < 1$, with their contribution decaying in the SFF as $t \to \infty$, implying:

$$K(t \to \infty) = \frac{1}{D^2}. \tag{9.11}$$

The parallels between Eqs. (9.10), (9.11) on the one hand and Eqs. (9.3), (9.2) on the other are evident. In the context of our experiment, we will diagnose full ergodicity when $K(t)$ relaxes to $1/D^2$ at large but accessible time scales, and non-ergodicity if it relaxes to or oscillates around a value $> 1/D^2$. We do not specifically diagnose thermalization for the following reason.

The effect of ensemble averaging: It is worth briefly mentioning the effect of ensemble averaging. From Eq. (9.8) (also Eq. (9.9)), it is evident that the contribution from the unit eigenvalue $\Lambda_0 = 1$ is not significantly altered by ensemble averaging, except to the extent of averaging the algebraic multiplicity m_0 of this eigenvalue over the ensemble. In contrast, the other unimodular

eigenvalues $\Lambda_j \neq 1$ can (and typically do) interfere destructively on ensemble-averaging, and their contribution cancels out in a large ensemble. We get

$$\langle K(t \to \infty) \rangle_{\text{ensemble}} = \frac{m_0}{D^2}. \tag{9.12}$$

Thus, ensemble averaging is effectively similar to the time averaging in Eq. (9.8). For the same reason, we lose the ability to distinguish between full ergodicity with or without thermalization, Eqs. (9.10) and (9.11), as $m_0 = 1$ in both cases. Partly due to the efficiency of the measurement protocol in Ref. [6] when ensemble averaging is performed in parallel with averaging over different choices of random unitaries, we have primarily considered ensemble averaged data where we can only clearly distinguish between full ergodicity and non-ergodicity, without being able to diagnose thermalization in particular. However, for the non-ergodic case without ensemble averaging, we are able to demonstrate the presence of larger fluctuations in the transient regime of the SFF before it approaches its steady value via model-specific predictions (see App. F.1). A similar correspondence between decoherence or dissipative effects and the reduction in the size of fluctuations in a variant of the SFF has been noted, for instance, in Ref. [170].

Finally, we note that while the statistics of these eigenvalues can be studied and, in several cases, compared to non-Hermitian or non-unitary random matrix ensembles, these do not have a known connection to ergodicity or thermalization [132–134]. On the other hand, in time-dependent systems, both ergodicity [229] and generalized SFFs [2, 149] can be defined more directly but there is no (time-independent) eigenvalue statistics to characterize. Exploring the role of different diagnostics in these cases would be an interesting future direction.

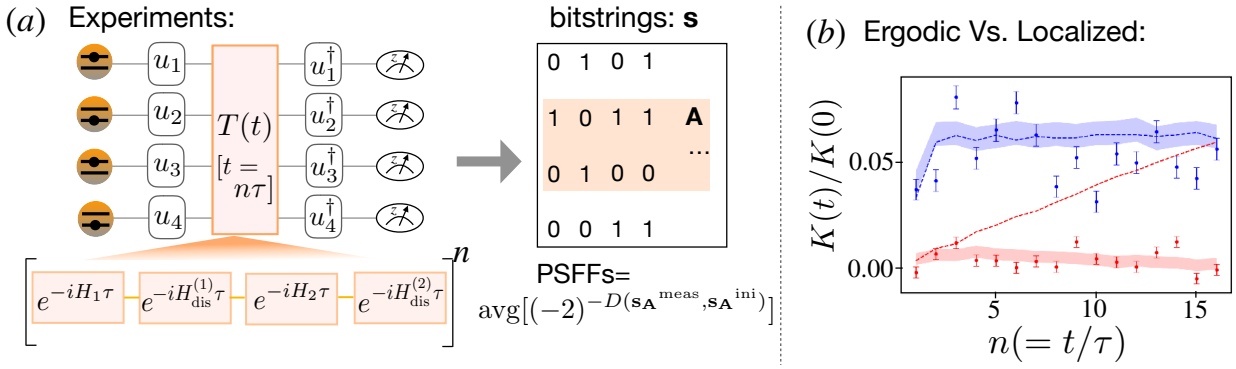


Figure 9.1: (a) The randomized measurement protocol to measure PSFF in a 4-qubit quantum simulator is shown. We perform the dynamics with many-body ergodic and localized Floquet sequence. In practice, in the experiment, the time evolution operator is a quantum channel $\mathscr{T}^t$, such that $T(t)\otimes T^\dagger(t)\to\mathscr{T}^t$ represents the actual dynamics implemented by the experiment. The protocol involves applying local random rotations to the initial state, after which the time-evolution is performed. This follows adjoint local random rotations and projective measurements. The experiment is repeated for N_r sets of random rotations and disorders, and N_m projective measurements. The collected bitstrings are then used to estimate the generalized PSFFs. (b) The ratio of measured SFF $K(t)$ and $K(0)$ is shown in dots and error bars. This normalization by the initial value accounts for SPAM errors. The shadowed curves show standard deviation of means of 20 similar experiments simulated numerically. For comparison, we plot, in dashes, SFF in the case of unitary dynamics. Red (blue) colors denote ergodic (localized) dynamics with $N_r=213$ and $N_m=100$ ($N_r=100$ and $N_m=100$) experimental runs. Reasonable agreement is seen with Eq. (9.11) for the ergodic case, and Eq. (9.9) with $m_0=D$ for the localized case.

9.3 Experimental setup and results

Here, we briefly describe the experimental setup. The generalized PSFF protocol is experimentally realized on a trapped-ion quantum simulator consisting of four $^{171}Yb^+$ ion qubits. Two different quantum many-body systems are examined: (i) one that is ergodic and (ii) one that is non-interacting. The intended unitary dynamics of both systems can be described by a periodic time-evolution operator $T(t=n\tau)=V^n$ specified at integer n multiples of the Floquet operator V and period τ. More specifically, we define an ideal (unitary) Floquet operator V as the following,

$$V_r = e^{-iH_r^{(2)}t_2}e^{-iH_2t_1}e^{-iH_r^{(1)}t_2}e^{-iH_1t_1}, \tag{9.13}$$

$$H_1 = \sum_{i<j} J_{ij}X_iX_j + B_x\sum_i X_i, \tag{9.14}$$

$$H_r^{(1)} = \sum_i g_iZ_i, \tag{9.15}$$

$$H_2 = \sum_{i<j} J_{ij}X_iX_j + B_y\sum_i Y_i, \tag{9.16}$$

$$H_r^{(2)} = \sum_i h_iZ_i, \tag{9.17}$$

where $\hbar = 1$, J_{ij} represents the long-range antiferromagnetic Ising interaction strength between the ith and jth ions, X_i, Y_i, Z_i are the Pauli operators corresponding to ion i, B_x and B_y denote the magnetic field strengths along the x and y directions, respectively while g_i and h_i specify the local disorder strengths experienced by ion i along the z direction. We define $J_0 \equiv J_{\text{avgNN}}$ as the average nearest-neighbor spin-spin interaction strength. Both the uniform and individual random magnetic field strengths are proportional to J_0 and mathematically they are expressed as: $B_x, B_y \in J_0$ and $g_i, h_i \in [0, 2J_0\xi]$ where ξ is a positive scale factor and the individual random disorders are independently and randomly sampled from the uniform distribution given.

The Floquet period is given by $\tau = 2(t_1 + t_2)$ where t_1 and t_2 correspond to the duration that Hamiltonians H_1, H_2 and $H_r^{(1)}, H_r^{(2)}$ are applied, respectively. The type of Floquet model realized can be tuned from ergodic to non-interacting by changing the value of τ. We keep the evolution time of the individual random disorder terms fixed with $t_2 \approx 1/(J_0\xi)$ and instead change the evolution time, t_1, of the long-range Ising interaction with a global magnetic field Hamiltonians. For ergodic SFF sequences, $t_1 \approx 1/J_0$ ($\tau_{\text{erg}} = \tau$) while for the effectively "non-interacting" SFF sequences, the duration of interactions is made much smaller: $t_1 \approx 1/(1000J_0)$ ($\tau_{\text{loc}} \approx 2t_2$).

Further, we note that the experiment has State Preparation and Measurement (SPAM) errors, due to which $K_A(0) < 1$ from the raw data — in fact, we see that roughly $K_A(0) \approx (0.9)^{N_A}$, with slight variations across datasets. We expect that this can be modeled by a time-independent mismatch between the computational basis of state preparation, and the computational basis of projective measurements. Correspondingly, we plot the rescaled SFF $K_A(t)/K_A(0)$ in the experimental plots, which appears to mitigate this error and match the theoretical expectations for the PSFFs.

With this setup, one of our main results, showing the respective spectral signatures of ergodicity and non-ergodicity in the two experimental Floquet sequences, is depicted in Fig. 9.1.

Part IV

Appendices

Appendix A: Calculations for dynamical quantum ergodicity from energy level statistics

A.1 Classical cyclic permutations (review)

This proof essentially follows Ref. [98]. First, we discuss the bound for cyclic ergodicity. Assume that each of the (say) M_C ergodic subsets $\{\mathscr{P}_j\}_{j=1}^{M_C}$ of $\mathscr{P}$ completely contains at least one element $C_{p(j)} \subseteq \mathscr{P}_j$ of the decomposition. As $\mathscr{T}^t C_{p(j)} \in \mathscr{P}_j$ for all t by definition, we must have $\mu[(\mathscr{T}^{[p(j+1)-p(j)]t_0} C_{p(j)}) \cap C_{p(j+1)}] = 0$. An important exception to this behavior is when $M_C = 1$, where there is no reason to impose a vanishing intersection. Thus,

$$\frac{1}{2}\sum_{j=1}^{M_C} \mu\left[(\mathscr{T}^{[p(j+1)-p(j)]t_0} C_{p(j)}) \triangle C_{p(j+1)}\right] = \frac{1}{n} M_C, \tag{A.1}$$

for $M_C \geq 2$. Now, we need to know how the error in an ℓ-step time evolution $(\mathscr{T}^{\ell t_0} C_k) \triangle C_{k+\ell}$ is related to the error $(\mathscr{T}^{t_0} C_m) \triangle C_{m+1}$ made in approximating each step. For this, we note that

$$(\mathscr{T}^{(m+1)t_0} A) - C_{m+1} \subseteq \left[(\mathscr{T}^{t_0} C_m) - C_{m+1}\right] \cup \mathscr{T}^{t_0}\left[(\mathscr{T}^{mt_0} A) - C_m\right], \ \forall\, A \subseteq \mathscr{P} \tag{A.2}$$

$$\implies \mu[(\mathscr{T}^{\ell t_0} C_k) \triangle C_{k+\ell}] \leq \sum_{m=1}^{\ell} \mu[(\mathscr{T} C_{k+m-1}) \triangle C_{k+m}], \tag{A.3}$$

where the second line follows from recursively applying the first line to $\mathscr{T}^{(m+1)t_0} C_k \triangle C_{k+q}$ with $A = C_k$. Using this in Eq. (A.1), one obtains $\overline{\epsilon}_C \geq (M_C/n)$ for $M_C \geq 2$, giving $\overline{\epsilon}_C < 2/n$ for $M_C = 1$.

Cyclic aperiodicity is more straightforward. We impose $\mu[(\mathscr{T}^{nt_0} C_k) \cap C_k] = 0$ (up to possible corrections that vanish as $n \to \infty$), which implies $\overline{\epsilon}_C \geq 1/n$ from Eq. (A.3).

A.2 Quantum cyclic permutations

A.2.1 Fastest decay of persistence

Given a cyclic permutation basis $\mathscr{C} = \{|C_j\rangle\}_{j=0}^{d-1}$, consider some initial state $|C_k\rangle$. After p steps of time evolution, it evolves into

$$\hat{U}_H(pt_0)|C_k\rangle = z_k(p,t_0)e^{i\phi_k(p,t_0)}|C_{k+p}\rangle + \sqrt{1-z_k^2(p,t_0)}|\nu_{k+p}^{(k)}\rangle, \tag{A.4}$$

where $|\nu_{k+p}^{(k)}\rangle$ is some normalized vector orthogonal to $|C_{k+p}\rangle$, and $\phi_k(p,t_0)$ is an unimportant phase. This leads to a recurrence relation for the persistence amplitudes,

$$\begin{aligned} z_k(p+1;t_0)e^{i\phi_k(p+1;t_0)} = {} & z_k(p)e^{i\phi_k(p,t_0)}\langle C_{k+p+1}|\hat{U}_H(t_0)|C_{k+p}\rangle \\ & + \sqrt{1-z_k^2(p,t_0)}\langle C_{k+p+1}|\hat{U}_H(t_0)|\nu_{k+p}^{(k)}\rangle. \end{aligned} \tag{A.5}$$

Using the triangle inequality for the magnitudes of these vectors gives

$$\begin{aligned} & \left| z_k(p,t_0)z_{k+p}(1,t_0) - \sqrt{1-z_k^2(p,t_0)}\sqrt{1-z_{k+p}^2(1,t_0)} \right| \\ & \leq z_k(p+1;t_0) \\ & \leq \left\{ z_k(p,t_0)z_{k+p}(1,t_0) + \sqrt{1-z_k^2(p,t_0)}\sqrt{1-z_{k+p}^2(1,t_0)} \right\}, \end{aligned} \tag{A.6}$$

on noting that $\hat{U}_H(t_0)|\nu_{k+p}^{(k)}\rangle$ is orthogonal to $\hat{U}_H(t_0)|C_{k+p}\rangle$, and consequently the inner product of the former with $|C_{k+p+1}\rangle$ cannot exceed $\sqrt{1-z_{k+p}^2(1,t_0)}$ in magnitude.

The above inequalities can be simplified by defining $\theta_k(p) = \arccos z_k(p,t_0) \in [0,\pi/2]$. In terms

of these variables, Eq. (A.6) becomes

$$\min\left\{\theta_k(p)+\theta_{k+p}(1),\frac{\pi}{2}\right\}\geq\theta_k(p+1)\geq|\theta_k(p)-\theta_{k+p}(1)|. \tag{A.7}$$

Summing $\theta_k(p+1)-\theta_k(p)$ from $p=p_1$ to $p=p_2-1$ gives

$$\operatorname{sgn}(p_2)\min\left\{\theta_k(p_2),\frac{\pi}{2}\right\}-\operatorname{sgn}(p_1)\min\left\{\theta_k(p_1),\frac{\pi}{2}\right\}\leq(\operatorname{sgn}(p_2)-\operatorname{sgn}(p_1))\sum_{p=p_1}^{p_2-1}\theta_{k+p}(1), \tag{A.8}$$

which becomes Eq. (2.19) when $p_1=0$. We see that the bound is saturated when $\hat{U}_H(t_0)$ at the p-th step acts like a 2D rotation by the angle $\theta_{k+p}(1)$ in the same direction as the previous steps.

A.2.2 Optimal errors for cyclic permutations

When $\hat{U}_C$ is a cycling operator, $\hat{U}_C^p$ is generally a permutation operator on $\mathscr{C}=\{|C_k\rangle\}_{k=0}^{d-1}$ that can be decomposed into a direct sum of cycling operators, each acting on a separate $[d/\mathscr{N}(d,p)]$-sized subset of $\mathscr{C}$:

$$\hat{U}_C^p=\bigoplus_{j=1}^{\mathscr{N}(d,p)}\hat{U}_{C,j}(p). \tag{A.9}$$

The number of cycling operators $\mathscr{N}(d,p)$ is given by the greatest common divisor of p and d; in particular, $\mathscr{N}(d,p)=1$ when p and d are coprime, including $p=1$. This is most easily seen in the eigenvalue structure of $\hat{U}_C^p$, which consists of $\mathscr{N}(d,p)$ identical (degenerate) sets of distinct $[d/\mathscr{N}(d,p)]$-th roots of unity. It is also convenient to consider *twisted* versions of $\hat{U}_C^p$, in which each cycle acquires an additional phase $\alpha_j(w)$:

$$w\{\hat{U}_C^p\}\equiv\bigoplus_{j=1}^{\mathscr{N}(d,p)}e^{i\alpha_j(w)}\hat{U}_{C,j}(p). \tag{A.10}$$

It is worth noting that the twisting functional w affects only the eigenvalues of $\hat{U}_C^p$, lifting the degeneracy for most values of the $\alpha_j(w)$, while preserving at least one complete orthonormal set of its

eigenvectors. Also, the p-step persistence amplitudes $z_k(p,t_0)$ are invariant under the action of w.

A.2.2.1 Optimizing the error via the trace inner product

Due to its non-negativity, the minimum persistence at a given p is bounded by the mean persistence at that time:

$$\min_{j\in\mathbb{Z}_d} z_j(p,t_0) \leq \frac{1}{d}\sum_{k=0}^{d-1} z_k(p,t_0). \tag{A.11}$$

We also have the inequality,

$$\left|\frac{1}{d}\,\mathrm{Tr}\left[w\{\hat{U}_C^p\}^\dagger \hat{U}_H(pt_0)\right]\right| \leq \frac{1}{d}\sum_{k=0}^{d-1}\left|\langle C_k|(\hat{U}_C^p)^\dagger \hat{U}_H(pt_0)|C_k\rangle\right|, \tag{A.12}$$

for any w, where the right hand side is just the mean persistence at p, expanded out.

Let us assume that for every $\mathscr{C}$ and given a p, there exists a unitary $\hat{V}_C$ and a twisting functional w such that

$$\frac{1}{d}\,\mathrm{Tr}\left[\hat{V}_C w\{\hat{U}_C^p\}^\dagger \hat{V}_C^\dagger \hat{U}_H(pt_0)\right] = \frac{1}{d}\sum_{k=0}^{d-1}\left|\langle C_k|w\{\hat{U}_C^p\}^\dagger \hat{U}_H(pt_0)|C_k\rangle\right|. \tag{A.13}$$

If this holds, then on account of Eq. (A.12) and the invariance of the $z_k(p,t_0)$ under the action of w,

$$\max_{\hat{V}\in\mathscr{U}(d)}\left|\frac{1}{d}\,\mathrm{Tr}\left[\hat{V}w\{\hat{U}_C^p\}^\dagger \hat{V}^\dagger \hat{U}_H(pt_0)\right]\right| = \max_{\text{all }\mathscr{C}}\frac{1}{d}\sum_{k=0}^{d-1} z_k(p,t_0), \tag{A.14}$$

where w is chosen so that Eq. (A.13) is satisfied for some $\mathscr{C}$ that maximizes the right hand side of Eq. (A.14). This follows as $\hat{V}_C$ becomes a special case of $\hat{V}$, and the only freedom to vary the orthonormal basis $\mathscr{C}$ is through its reorientations in Hilbert space — precisely given by all possible unitary transformations $\hat{V}\in\mathscr{U}(d)$ acting on the energy subspace Σ_d.

Now, we need to establish that Eq. (A.13) is indeed valid, and identify w. It is convenient to consider the two cases of nondegenerate and degenerate $\hat{U}_C^p$ separately.

1. **Case 1: $|p|$ and d are coprime.** In this case, $\hat{U}_C^p$ is itself a cycling operator. We separate the persistence inner product into an amplitude and phase,

$$\langle C_k|(\hat{U}_C^p)^\dagger \hat{U}_H(pt_0)|C_k\rangle = z_k(p,t_0)e^{i\phi_k(p,t_0)}. \tag{A.15}$$

Let $\overline{\phi}(p,t_0) = \sum_{k=0}^{d-1}\phi_k(p,t_0)$. Define a new cyclic permutation $\mathscr{C}'$ with basis vectors

$$|C_k'\rangle = e^{i\sum_{j=-1}^{(j+1)p=k}\left\{\phi_{jp}(p,t_0)-[\overline{\phi}(p,t_0)/d]\right\}}|C_k\rangle, \tag{A.16}$$

where $\sum_{j=-1}^{(j+1)p=k}\phi_{jp} = \phi_{-p}+\phi_0+\phi_p+\ldots+\phi_{k-p}$ is a sum over the index with steps of size p, and subtracting $\overline{\phi}(p,t_0)/d$ from each term ensures the single-valuedness of the phases in the new basis. This induces a unitary transformation $\hat{U}_C \to \hat{U}_{C'} = \hat{V}_C\hat{U}_C\hat{V}_C^\dagger$ (where $\hat{U}_{C'}$ is required to satisfy Eq. (A.15) with the $|C_k\rangle$ replaced by $|C_k'\rangle$), such that

$$\langle C_k|\hat{V}_C(\hat{U}_C^p)^\dagger \hat{V}_C^\dagger \hat{U}_H(pt_0)|C_k\rangle = z_k(p,t_0)e^{i\overline{\phi}(p,t_0)/d}. \tag{A.17}$$

We see that Eq. (A.13) is then satisfied for a twisting functional w with $\alpha_1(w) = -\overline{\phi}(p,t_0)/d$ (however, this phase is inconsequential in this case, being absorbed by the absolute value in Eq. (A.14)).

2. **Case 2: $|p|$ and d have a nontrivial common factor.** For this case, we can ensure that the analogue of Eq. (A.13) for each $[d/\mathcal{N}(d,p)]$-element cycle is satisfied following the procedure leading up to Eq. (A.17), with the total phase $\overline{\phi}(p,t_0)$ replaced by that corresponding to the respective cycle, $\overline{\phi}_j(p,t_0)$. Then, it follows that Eq. (A.13) is also satisfied overall for $\hat{U}_C^p$ with a twisting functional w given by $\alpha_j(w) = -\overline{\phi}_j(p,t_0)/d$.

Thus, from Eq. (A.14), we can maximize the mean persistence by maximizing the magnitude of

the trace

$$f_p(\hat{U}_C) = \left|\mathrm{Tr}\left[w\{\hat{U}_C^p\}^\dagger \hat{U}_H(pt_0)\right]\right| \tag{A.18}$$

with respect to reorientations $\hat{U}_C \to \hat{V}\hat{U}_C\hat{V}^\dagger$. In Sec. A.2.2.2, this maximum is shown to occur for some $\hat{U}_C$ satisfying

$$\left[\hat{U}_H(pt_0), w\{\hat{U}_C^p\}^\dagger\right] = 0, \tag{A.19}$$

as long as $f_p(\hat{U}_C) \geq \sqrt{d(d-2)}$ at some such point.

If $\hat{U}_H(pt_0)$ and $w\{\hat{U}_C^p\}$ both have nondegenerate eigenvalues, each has a unique set of d eigenvectors corresponding to the respective eigenvectors of $\hat{U}_H(t_0)$ and $\hat{U}_C$. Eq. (A.19) then implies that both sets of eigenvectors are identical, and $\hat{U}_C$ must commute with $\hat{U}_H(t_0)$ to achieve a local extremum of the mean persistence.

When there are degeneracies (in any of $\hat{U}_H(t_0)$, $\hat{U}_H(pt_0)$ or $w\{\hat{U}_C^p\}$), we can nevertheless reach a similar conclusion by infinitesimally breaking the degeneracies. We can define $\hat{U}_{H(\delta_u)} = \hat{U}_H(pt_0)e^{i\delta_u\hat{Y}}$ where $\delta_u \to 0$ and $\hat{Y}$ is any finite Hermitian operator (i.e. with finite matrix elements in any orthonormal basis), such that $\hat{U}_{H(\delta_u)}$ has nondegenerate eigenvalues when $\delta_u \neq 0$. Similarly, we define $w_{(\delta_w)}$ by $\alpha_j(w_{(\delta_w)}) = \alpha_j(w) + \delta_w\gamma_j$ with $\delta_w \to 0$, with the γ_j chosen so as to ensure the nondegeneracy of the eigenvalues of $w\{\hat{U}_C^p\}$ (essentially, infinitesimally twisting any degenerate $e^{i\alpha_j(w)}\hat{U}_{C,j}(p)$, $e^{i\alpha_k(w)}\hat{U}_{C,k}(p)$, ... relative to each other). Re-expressing Eq. (A.14) in terms of these variables, gives

$$\max_{\hat{V}\in\mathscr{U}(d)}\left|\frac{1}{d}\mathrm{Tr}\left[\hat{V}w_{(\delta_w)}\{\hat{U}_C^p\}^\dagger\hat{V}^\dagger\hat{U}_{H(\delta_u)}(pt_0)\right]\right| = \max_{\text{all }\mathscr{C}}\frac{1}{d}\sum_{k=0}^{d-1}z_k(p,t_0) + O(\delta_u,\delta_w), \tag{A.20}$$

where $O(\delta_u,\delta_w)$ consists of terms of the form $(\delta_u)^a(\delta_w)^b y_{ab}$ with $a,b \geq 1$. As with Eq. (A.19), the solution to the maximization on the left hand side must be among its local extrema, given by

$$\left[\hat{U}_{H(\delta_u)}(pt_0), w_{(\delta_w)}\{\hat{U}_C^p\}^\dagger\right] = 0. \tag{A.21}$$

Now, each nondegenerate operator $\hat{U}_{H(\delta_u)}(pt_0)$ and $w_{(\delta_w)}\{\hat{U}_C^p\}^\dagger$ has a unique set of d eigenvectors, which the above equation asserts are identical. We can choose $\hat{Y}$ and γ_j to break the degeneracy of $\hat{U}_H(pt_0)$ and $w\{\hat{U}_C^p\}$ in any desired way, i.e. to pick any complete orthonormal subset of each set of eigenvectors. By Eq. (A.20), any such choice is equally good for maximizing the mean persistence in the $\delta_u, \delta_w \to 0$ limit. In particular, we can pick $w_{(\delta_w)}\{\hat{U}_C^p\}^\dagger$ so that its eigenvectors are identical to those of $\hat{U}_C$; similarly, we can choose $\hat{Y}$ so that the eigenvectors of $\hat{U}_{H(\delta u)}(pt_0)$ are identical to any complete orthonormal set of eigenvectors of $\hat{U}_H(t_0)$. In other words, any choice of degeneracy breaking in the neighborhood of degenerate operators only infinitesimally affects the local extrema of the left hand side of Eq. (A.20).

Thus, the right hand side of Eq. (A.11) attains its global maximum when the eigenvectors of $\hat{U}_C$ are fixed to be any complete orthonormal set of eigenvectors of $\hat{U}_H(t_0)$, with the only freedom remaining in the assignment of the distinct eigenvalues of $\hat{U}_C$ to these eigenvectors. This can be concisely expressed as follows: the global maximum of the mean persistence occurs among the solutions to

$$\lim_{\delta\to 0}\left[\hat{U}_H(t_0)e^{i\delta\hat{Y}}, \hat{U}_C\right] = 0, \tag{A.22}$$

for any Hermitian $\hat{Y}$. For any $\hat{U}_C$ satisfying this property, all the $z_j(p, t_0)$ are equal at any given p. It follows that $\min_j z_j(p, t_0)$ is also maximized, and the p-step error minimized, by the same $\hat{U}_C$ that maximizes the mean persistence. From the requirement $f_p(\hat{U}_C) \geq \sqrt{d(d-2)}$, we get the condition $\varepsilon_C(p, t_0) \leq (2/d)$ on such a minimum of the error.

A.2.2.2 Local extrema, and the global maximum for large persistence amplitudes

For simplicity, let $\hat{U}_1 = w\{\hat{U}_C^p\}$ and $\hat{U}_2 = \hat{U}_H(pt_0)$. We seek stationary points of the real valued function (from Eq. (A.18))

$$\left|\mathrm{Tr}\left(\hat{U}_1^\dagger \hat{U}_2\right)\right| \tag{A.23}$$

with respect to small reorientations of $\hat{U}_1$ by $\hat{V}$, to first order. This would yield all the local maxima and minima (as well as saddle and inflection points) of the function except the global minima when the function attains the value 0, where it is not differentiable. We write $\hat{V} = e^{i\hat{X}}$ with $\hat{X}$ near 0, and require the phase of the $O(\hat{X})$ term in $\mathrm{Tr}[\hat{V}\hat{U}_1^\dagger\hat{V}^\dagger\hat{U}_2]$ to be orthogonal to the phase of the $O(1)$ term (so that the first variation corresponds only to a change in phase and not in magnitude; alternatively, one could directly extremize the square of Eq. (A.23)). This gives

$$\mathrm{Tr}\left(\hat{X}\left[\hat{U}_1^\dagger, \hat{U}_2\right]\right) = c(\hat{X})\,\mathrm{Tr}\left(\hat{U}_1^\dagger\hat{U}_2\right) \text{ for all Hermitian } \hat{X} \tag{A.24}$$

with $c(\hat{X})$ required to be a real-valued function, for the stationary points. As can be verified by imposing this for each independent degree of freedom in the matrix elements of $\hat{X}$, this requires

$$\left[e^{-i\alpha_{12}}\hat{U}_1^\dagger, \hat{U}_2\right] = \hat{F}, \tag{A.25}$$

where $\hat{F}$ is some traceless Hermitian operator, and α_{12} is the phase of $\mathrm{Tr}(\hat{U}_1^\dagger\hat{U}_2)$.

Up to this point, the unitarity of $\hat{U}_1$ and $\hat{U}_2$ played no role. Now, we use the fact that their products are unitary, and write

$$e^{-i\alpha_{12}}\hat{U}_1^\dagger\hat{U}_2 = e^{i\hat{A}_{12}}, \text{ and } \hat{U}_2 e^{-i\alpha_{12}}\hat{U}_1^\dagger = e^{i\hat{A}_{21}}, \tag{A.26}$$

for Hermitian $\hat{A}_{12}$ and $\hat{A}_{21}$. Formally defining sines and cosines of Hermitian operators through their Taylor series (which are also Hermitian), Eq. (A.25) then gives

$$\cos\hat{A}_{12} - \cos\hat{A}_{21} + i\left[\sin\hat{A}_{12} - \sin\hat{A}_{21}\right] = \hat{F}. \tag{A.27}$$

The Hermiticity of $\hat{F}$ demands that the anti-Hermitian part of the left hand side vanishes, giving

$$\sin\hat{A}_{12} = \sin\hat{A}_{21}. \tag{A.28}$$

Let $\{a(k)\}_{k=0}^{d-1}$ be the eigenvalues of $\hat{A}_{12}$ and $\hat{A}_{21}$ (which must have identical eigenvalues up to irrelevant shifts of 2π, as products of two unitaries have the same eigenvalues irrespective of the order [109]). As long as it is known that $a(k) \in [-\pi/2, \pi/2]$, the sine is invertible and $\hat{A}_{21} = \hat{A}_{12}$ (In fact, one gets $\hat{A}_{12} = \hat{A}_{21}$ and therefore $\hat{F} = 0$ for "generic" values of $a(k)$ such that $\sin\hat{A}_{21}$ is not degenerate, i.e. the set $\{a(k), \pi + a(k)\}$ is non-degenerate). Consequently,

$$\left\{a(k) \in [-\pi/2, \pi/2],\ \forall\ k\right\} \implies \left(\left[e^{-i\alpha_{12}}\hat{U}_1^\dagger, \hat{U}_2\right] = 0\right) \tag{A.29}$$

at a stationary point. The vanishing commutator on the right side of the implication is precisely the condition of Eq. (A.19).

The question of interest is now if there's a simple way to guarantee the restriction on $a(k)$ in Eq. (A.29). To see that there is, we note that $[\mathrm{Tr}(e^{i\hat{A}_{12}})] \in \mathbb{R}$ by the definition of α_{12}, which implies

$$\sum_k \cos a(k) = \mathrm{Tr}\left(e^{i\hat{A}_{12}}\right), \tag{A.30}$$

$$\sum_k \sin a(k) = 0. \tag{A.31}$$

Let us maximize the multivariable function $b[a(k)] = \sum_k \cos a(k)$ with *fixed* $a(0)$ (and free $a(k \neq 0)$) subject to the constraint in Eq. (A.31) (and implicitly, non-negativity) using e.g. the method of Lagrange multipliers. The stationary points of $b[a(k)]$ occur at

$$a(k \neq 0) = c + \pi\zeta_k, \text{ with } \zeta_k \in \{0, 1\}, \tag{A.32}$$

for some constant c. The global maximum of $b[a(k)]$ corresponds to $c \in [-\pi/2, \pi/2]$ and $\zeta_k = 0\ \forall k$. Imposing Eq. (A.31) to fix c in terms of $a(0)$, we get

$$\sum_k \cos a(k) \leq b_{\max}[a(0)] \equiv \cos a(0) + (d-1)\sqrt{1 - \frac{\sin^2 a(0)}{(d-1)^2}}. \tag{A.33}$$

This is a monotonically decreasing function of $|a(0)|$ in its full domain $[0, \pi]$ for $d \geq 2$. In particular, if $|a(0)| > \pi/2$, then it is guaranteed that $b[a(k)] < b_{\max}[\pi/2] = \sqrt{d(d-2)}$. Re-expressing $b[a(k)]$ in terms of the trace of the relevant unitaries, we then have

$$\left\{\left|\mathrm{Tr}\left(\hat{U}_1^\dagger \hat{U}_2\right)\right| \geq \sqrt{d(d-2)}\right\} \implies \left\{a(k) \in [-\pi/2, \pi/2],\ \forall\ k\right\}. \tag{A.34}$$

Combined with the implication in Eq. (A.29), it follows that maxima for which the trace is no smaller than $\sqrt{d(d-2)}$ occur for cycling operators that commute with time evolution, i.e. when $[\hat{U}_1^\dagger, \hat{U}_2] = 0$. Such commuting operators remain local extrema of the trace in other cases, but it is unclear in the present analysis if the global maximum is among them. For comparison with the following subsection, we note that $a(k) = -2\pi p \Delta_k / d$, where Δ_k are the mode fluctuations used elsewhere (see Eq. (2.34)) in the main text.

A.2.3 Decrease of persistence for small permutations of sorted energy levels

When $\Delta_n \ll d$, assuming that the energies E_n have been shifted by some additive constant so that $\sum_k \Delta_k = 0$, we have (representing d times the persistence amplitude as per Eq. (2.34))

$$\sum_{k=0}^{d-1} e^{-2\pi i \Delta_k / d} = 1 - \frac{2\pi^2}{d^2} \sum_{k=0}^{d-1} \Delta_k^2 + O(\Delta_k^3 d^{-3}). \tag{A.35}$$

For simplicity, we assume that the levels are already sorted i.e. $E_n < E_m$ when $n < m$. This further implies

$$\Delta_n - \Delta_m > -|n-m|. \tag{A.36}$$

Any permutation $q(n)$ can be broken up [36] into a set of cyclic permutations $q_r(n)$, each involving a subset of N_r levels $\mathscr{E}[q_r] = \{E_{r_k}\}_{k=0}^{N_r-1}$. For the rest of the argument, we will require (where the subtraction of r is on $\mathbb{Z}$ (linear), and not on $\mathbb{Z}_d$ (circular or modulo d))

$$|r_k - r_j| < d/2, \ \forall \ k, j \in \mathbb{Z}_{N_r}, \tag{A.37}$$

for each q_r; permutations q satisfying this are what we refer to as "small" permutations. This will ensure that Eq. (A.35) remains valid under these permutations without discrete shifts of some of the Δ_k by multiples of (2π).

The new mode fluctuations after permutation are given by

$$\Delta'_{r_k} = \Delta_{r_{k+1}} + [r_{k+1} - r_k], \tag{A.38}$$

for each cycle q_r. It follows that the mean is preserved, i.e.

$$\sum_{k=0}^{N_r-1} \Delta'_{r_k} = \sum_{k=0}^{N_r-1} \Delta_{r_k}. \tag{A.39}$$

Our goal is to show that the variance of the Δ'_k is larger than that of the Δ_k, which would translate to a decreased persistence by Eq. (A.35). We have

$$\sum_{k=0}^{N_r-1} (\Delta'_{r_k})^2 - \sum_{k=0}^{N_r-1} (\Delta_{r_k})^2 = 2 \sum_{k=0}^{N_r-1} \Delta_{r_{k+1}} [r_{k+1} - r_k] + \sum_{k=0}^{N_r-1} [r_{k+1} - r_k]^2. \tag{A.40}$$

In general, the r_{k+1} are not in any simple (e.g. ascending or descending) order. We can split each difference $[r_{k+1}-r_k]$ in the first term on the right hand side into a sum of differences of the r_ℓ lying between (and inclusive of) them:

$$\Delta_{r_{k+1}}[r_{k+1}-r_k] = \sum_{r_j \in [r_k, r_{k+1}]} \zeta_k \Delta_{r_{k+1}}(r_{j+1}-r_j), \tag{A.41}$$

where $\zeta_k = \mathrm{sgn}[r_{k+1}-r_k]$, and the r_j are chosen to be sorted according to j. On including terms with different values of k, each interval $(r_{j+1}-r_j)$ occurs in an equal number of terms with positive $\zeta_k = +1$ ($r_{k+1} \geq r_{j+1}$) and negative $\zeta_k = -1$ ($r_{k+1} \leq r_j$). We can arbitrarily pair each positive term r_+ with a negative term r_-, and use Eq. (A.36) for the difference $\Delta_{r_+} - \Delta_{r_-}$ noting that $r_+ > r_-$. This amounts to replacing the equality with $\geq$, and each $\Delta_{r_{k+1}}$ with $-r_{k+1}$, in Eq. (A.40). We therefore obtain

$$\begin{aligned} \sum_{k=0}^{N_r-1} (\Delta'_{r_k})^2 - \sum_{k=0}^{N_r-1} (\Delta_{r_k})^2 &\geq \sum_{k=0}^{N_r-1} [r_{k+1}-r_k]^2 + \left\{ -2 \sum_{k=0}^{N_r-1} r_{k+1}[r_{k+1}-r_k] \right\}, \\ \implies \sum_{k=0}^{N_r-1} (\Delta'_{r_k})^2 &\geq \sum_{k=0}^{N_r-1} (\Delta_{r_k})^2. \end{aligned} \tag{A.42}$$

The second equation follows from simplifying the first. Adding all such equations from each q_r together, we get

$$\sum_{k=0}^{d-1} e^{-2\pi i \Delta'_k/d} \leq \sum_{k=0}^{d-1} e^{-2\pi i \Delta_k/d} + O(\Delta^3 d^{-3}). \tag{A.43}$$

This shows that sorting the energy levels corresponds to the maximum persistence at $p = 1$ for a given t_0 and small Δ_k, at least among other possibilities that can be obtained as small permutations of the sorted levels. This is more like a discrete version of a local extremum. It would be interesting to check if "larger" permutations not subject to Eq. (A.37) would lead to significantly better maxima; this is unlikely to be the case without some non-intuitive conspiracy between distant energy levels.

A.3 Time dependence of persistence amplitudes

A.3.1 Error coefficient pairing in discrete sum over paths

We rewrite Eq. (2.43) for $p = 1$ as

$$\hat{U}_\Delta e^{-i\phi_\Delta(1)} = (1-\varepsilon_1)^{1/2}\left[\hat{\mathbb{1}} + g_1 \sum_{m=1}^{d-1} \nu_m(1)\hat{U}_C^m\right], \tag{A.44}$$

where $\varepsilon_p \equiv \varepsilon_C(p, t_0)$ and $g_1 = \sqrt{\varepsilon_1/(1-\varepsilon_1)}$. We note that g_1 is also the coefficient that occurs on the right hand side of Eq. (2.45). The p-th power of the error unitary is

$$\begin{aligned}\hat{U}_\Delta^p e^{-ip\phi_\Delta(1)} &= (1-\varepsilon_1)^{p/2} \sum_{s=0}^{p} \left[\binom{p}{s} g_1^s \sum_{m_1,\ldots,m_s} \nu_{m_1}(1)\ldots\nu_{m_s}(1)\hat{U}_C^{m_1+\ldots+m_s}\right], \\ &= (1-\varepsilon_1)^{p/2} \sum_{r=0}^{d-1}\left(\sum_{s=0}^{p} \Gamma_r^{(s)}\right)\hat{U}_C^r \end{aligned} \tag{A.45}$$

where $\binom{p}{s} = p!/(s!(p-s)!)$ is the binomial coefficient, and we recall that the sums are modulo d. We have also defined

$$\Gamma_r^{(s)} = \binom{p}{s} g_1^s \sum_{m_1,\ldots,m_s} \nu_{m_1}(1)\ldots\nu_{m_s}(1)\overline{\Theta}(m_1+\ldots+m_s = r), \tag{A.46}$$

with $\overline{\Theta}(x) = 1$ if x is true and 0 otherwise. Each term with fixed r in (A.45) represents a sum over paths for the transition amplitude from any $|C_k\rangle$ to $|C_{k+r}\rangle$.

Now, we apply the assumption of error coefficient pairing, by considering only terms where as many m_j as possible are each paired with a corresponding $m_k = -m_j$. For even s in Eq. (A.46), restricting to such pairings necessarily implies that $m_1 + \ldots + m_s = 0$. For odd s, it is not possible to pair all error coefficients and a free error coefficient remains, whose index must necessarily be r if the remaining coefficients are paired. Schematically (in the sense that we avoid explicitly enumerating

the possible pairings), for non-negative integer u,

$$\Gamma_r^{(2u)} \approx \delta_{r0}\left\{\binom{p}{2u} g_1^{2u} \sum_{\text{pairings}} [\nu_{m_1}(1)\nu_{-m_1}(1)]\ldots[\nu_{m_u}(1)\nu_{-m_u}(1)]\right\}, \tag{A.47}$$

$$\Gamma_r^{(2u+1)} \approx g_1 \nu_r(p-2u)\left\{\binom{p}{2u} g_1^{2u} \sum_{\text{pairings}} [\nu_{m_1}(1)\nu_{-m_1}(1)]\ldots[\nu_{m_u}(1)\nu_{-m_u}(1)]\right\}. \tag{A.48}$$

In the second line, we have accounted for $s = 2u+1$ different ways of choosing the unpaired coefficient, and used $s\binom{p}{s} = (p+1-s)\binom{p}{s-1}$.

For a given u, the sum over pairings and coefficients within the braces in Eqs. (A.47) and (A.48) are identical, irrespective of the value of r. Treating g_1 as a formally independent parameter that we can take partial derivatives with respect to, we can further replace $(p-2u)$ with $(p - g_1\vec{\partial}/\partial g_1)$ acting on its right in Eq. (A.48), which moves all the u dependence to inside the braces. For even p, this means that each sum over s in Eq. (A.45) — which is naturally restricted to even s for $r = 0$ and odd s for $r \neq 0$ after pairing — produces coefficients for all r that are identical except for the operators outside the braces in Eqs. (A.47) and (A.48). If the time dependence is sufficiently slow, the result for odd p can be extrapolated (to a good approximation) in any convenient way between those for $p \pm 1$. Thus, we have the approximate form

$$\hat{U}_\Delta^p e^{-ip\phi_\Delta(1)} \approx (1-\varepsilon_1)^{p/2}\left[\hat{\mathbb{1}} + g_1 \sum_{r=1}^{d-1} \nu_r(1)\hat{U}_C^r\left(p - g_1\frac{\vec{\partial}}{\partial g_1}\right)\right] h(p, g_1). \tag{A.49}$$

The function $h(p, g_1)$ originates in the sum over pairings within the braces of Eqs. (A.47) and (A.48); from the above expression, it is formally related to the persistence amplitude at p by

$$z(p, t_0) = (1-\varepsilon_1)^{p/2} h\left(p, \sqrt{\frac{\varepsilon_1}{1-\varepsilon_1}}\right). \tag{A.50}$$

A.3.2 Gaussian estimate

The persistence amplitude at $p+1$ can be expressed in terms of the coefficients in $\hat{U}^p_\Delta$ and $\hat{U}^1_\Delta$ as follows:

$$z(p+1,t_0) = \left|\frac{1}{d}\,\mathrm{Tr}\left[\hat{U}^1_\Delta \hat{U}^p_\Delta\right]\right| = \left|\sqrt{1-\varepsilon_1}\sqrt{1-\varepsilon_p} + \sqrt{\varepsilon_1\varepsilon_p}\sum_{r=1}^{d-1} \nu_r(1)\nu_{-r}(p)\right|. \tag{A.51}$$

Substituting the appropriate expressions for ε_p and ν_p from Eq. (A.49), we get

$$z(p+1,t_0) \approx (1-\varepsilon_1)^{1/2}\left|z(p,t_0) + \left\{(1-\varepsilon_1)^{p/2} g_1^2 \sum_{r=1}^{d-1} \nu_r(1)\nu_{-r}(1)\left(p - g_1\frac{\vec{\partial}}{\partial g_1}\right)h(p,g_1)\right\}\right| \tag{A.52}$$

Now, we assume that the second term within the absolute value is smaller than the first, and $ph \gg g_1 \partial h/\partial g_1$; both will be justified retroactively. Further defining

$$\nu_C = -\sum_{r=1}^{d-1} \nu_r(1)\nu_{-r}(1), \tag{A.53}$$

which happens to measure the goodness of the approximation in Eq. (2.48), we are led to

$$z(p+1,t_0) \approx (1-\varepsilon_1)^{1/2}\left[1 - g_1^2 p\,\nu_C\right] z(p,t_0). \tag{A.54}$$

It is now straightforward to multiply over values of p from some given $\overline{p}$ through to 1. For $\varepsilon_1 \ll 1$ and setting $\nu_C \approx 1$ as per Eq. (2.48), we get

$$z(\overline{p},t_0) \approx \exp\left[-\frac{\varepsilon_1}{2}|p| - \frac{g_1^2}{2}p^2\right]. \tag{A.55}$$

We see that the smallness of the second term in Eq. (A.52) and $ph \gg g_1 \partial h/\partial g_1$ are both satisfied

when $p \ll 1/g_1$, i.e. when the persistence amplitude is still close to 1.

A.3.3 Minimum error constraints from the SFF

Substituting the form $K(t) = \lambda t^\gamma$ in Eq. (2.52) and dropping subleading terms in $\varepsilon_1 = \varepsilon_C(1, t_0)$ gives

$$2\lambda t_0^\gamma \sum_{p=1}^{1/(M\sqrt{\varepsilon_1})} p^{\gamma-2} \lessapprox \varepsilon_1. \tag{A.56}$$

For $\gamma \in [0, 1)$, the left hand side is dominated by small p and is independent of M. Replacing $1/(M\sqrt{\varepsilon_1}) \to \infty$, we obtain

$$\varepsilon_1 \gtrapprox 2\lambda t_0^\gamma \zeta(2-\gamma), \tag{A.57}$$

where $\zeta(x)$ is the Riemann zeta function. In particular, for $\gamma = 0$ and $\lambda = 1/d$ (Poisson statistics), we have $\varepsilon \gtrapprox \pi^2/(3d) \in O(1/d)$. For $\gamma > 1$, it is instead the terms with larger p that dominate. Using the leading term in Faulhaber's formula for the sum (formula (0.121) in Ref. [194]; equivalent to replacing the sum with an integral), we have

$$2\lambda t_0^\gamma \frac{[1/(M\sqrt{\varepsilon_1})]^{\gamma-1}}{\gamma-1} \lessapprox \varepsilon_1. \tag{A.58}$$

The presence of $M \in O(1) \geq 1$ in this expression allows us to make only order of magnitude statements. We get

$$\varepsilon_1^{(1+\gamma)/2} \gtrapprox 2\lambda t_0^\gamma \frac{(\gamma-1)}{M^{\gamma-1}}, \tag{A.59}$$

which implies $\varepsilon_1 \gtrsim O(d^{-4/(\gamma+1)})$ when $\lambda \in O(d^{-2})$ and $t_0 \in O(1)$, for any $\gamma \in O(1) > 1$. The most generic case (i.e. typical for Haar random [28, 36] systems), $\gamma = 1$, is a bit more subtle. Here, it is

again the large-p terms that dominate, so we take the $\gamma \to 1$ limit of Eq. (A.58), which gives

$$\frac{\varepsilon_1}{\ln\left(\frac{1}{M\sqrt{\varepsilon_1}}\right)} \gtrapprox 2\lambda t_0. \tag{A.60}$$

This is a transcendental equation for ε_1, but we can nevertheless invert it to leading order in λ^{-1} (i.e. substituting $\varepsilon_1 = \mu(\lambda)\lambda$ and solving for μ, neglecting $\ln(\ln\lambda)$), obtaining

$$\varepsilon_1 \gtrapprox \lambda t_0 \ln\frac{1}{\lambda}. \tag{A.61}$$

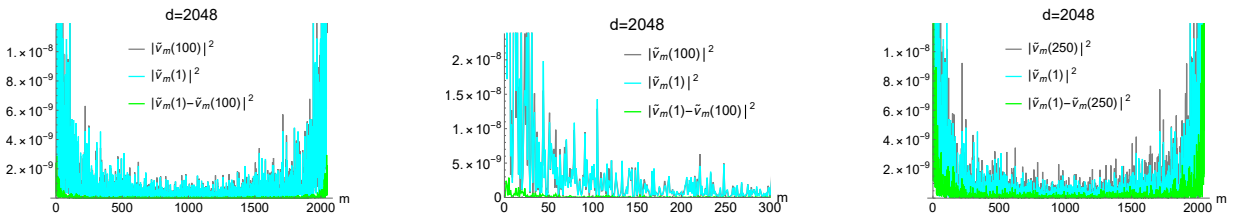
For Wigner-Dyson statistics, $\lambda \in O(d^{-2})$ and $t_0 \in O(1)$ gives $\varepsilon \geq cd^{-2}\ln d$ for a constant c.

A.3.4 Numerical evidence for error coefficient pairing

To provide numerical evidence for the pairing of error coefficients, we test the prediction of Eq. (A.49) when $g_1 \partial h/\partial g_1$ is negligible, i.e. Eq. (2.49) in the main text. More directly, we define

$$\tilde{\nu}_m(p) = \frac{1}{pz(p,t_0)}\nu_m(p). \tag{A.62}$$

Eqs. (A.49), (2.49) then imply that $\tilde{\nu}_m(p) = \tilde{\nu}_m(1)$ for any $p \ll 1/\sqrt{\varepsilon_C(1,t_0)}$. This is verified in Fig. A.1 for the ($\beta_{\text{RMT}} = 2$, $d = 2048$) CUE dataset of Fig. 2.7, for which $1/\sqrt{\varepsilon_C(1,t_0)} \approx 525$.



(a) $p = 100$. (b) $p = 100$, zoomed in. (c) $p = 250$.

Figure A.1: Comparison of $\tilde{\nu}_m(p)$ with $\tilde{\nu}_m(1)$, using magnitudes $|\tilde{\nu}_m(p)|^2$, $|\tilde{\nu}_m(1)|^2$ and residuals $|\tilde{\nu}_m(1) - \tilde{\nu}_m(p)|^2$ for $d = 2048$. The residuals are predicted to be negligible compared to the magnitudes at the same m for $p \ll 525$, which these plots are in good agreement with even when p is a considerable fraction of 525.

A.4 Non-ergodicity and spectral fluctuations in rational tori

When $\alpha = m/\ell$ is rational with coprime m and ℓ, all energy eigenstates of the form $(J_x - rm, J_y + r\ell)$, $r \in \mathbb{Z}$ for a given (J_x, J_y) have the same energy [34]. Consequently, taking into account the restriction of (J_x, J_y) to a rectangle of area d in Σ_d, the energy spectrum is $O(\sqrt{d})$-fold degenerate with $O(\sqrt{d})$ unique energy values (when $d_x = d_y$). Thus, in the calculation of the mode fluctuations $\Delta_n = [E_{q(n)} t_0 d/(2\pi)] - n$ (see Eq. (2.33)) for generic t_0, each unique value of E is associated with $O(\sqrt{d})$ different (at best, consecutive) values of the integer n due to the degeneracy, and we have $\Delta_n \gtrsim O(\sqrt{d})$ in general. This argument means that for rational tori, we should expect (based on Eqs. (2.32) and (2.34))

$$\varepsilon_C(1, t_0)[q] \sim O(d^{-2}\sigma_\Delta^2) \gtrsim O(d^{-1}), \tag{A.63}$$

which is at least of the same order of magnitude of spectral fluctuations as Poisson level statistics. Given that the 1-step error for any cyclic permutation must be $\gtrsim O(d^{-1})$ in this case by Theorem 2.6, it follows that rational tori are unlikely to possess any ergodic cyclic permutation in Σ_d. We also note that an analytically solvable illustration of this argument is given by the sorted DFT cyclic permutation for $t_0 = 2\pi/(\omega_x d_x)$ when $d_x = d_y$, in which the (wrapped) energy levels are *exactly* $\sqrt{d}$-fold degenerate.

A.5 Entanglement and semiclassical phase-space volumes

A heuristic correspondence between the entanglement of mixed states and the phase space volumes (or measures) of classical regions is as follows:

1. A classical phase space region A of measure $\mu(A) \gg 1/d$ can be represented by a mixed state $\hat{\rho}_A$ with equal eigenvalues (i.e. $\hat{\rho}_A = \hat{\Pi}_A / \operatorname{Tr}[\hat{\Pi}_A$ for some projector $\hat{\Pi}_A$), whose purity determines

the phase space volume:

$$\mu(A) \approx \frac{1}{d\,\mathrm{Tr}(\hat{\rho}_A^2)}. \tag{A.64}$$

2. Given two regions A and B, the measure of their intersection is proportional to the overlap of the corresponding mixed states:

$$\frac{\mu(A\cap B)}{\mu(A)\mu(B)} \approx d\,\mathrm{Tr}(\hat{\rho}_A\hat{\rho}_B). \tag{A.65}$$

We emphasize that there is no reason to assume the converse - not every mixed state looks like a simple phase space region in the classical limit; nevertheless, these relations are consistent with the quantum density operator measure algebra discussed in Ref. [26, 27]. The use of approximate equalities here represents the unavoidable ambiguity in quantizing a classical system (where any corrections that vanish as $d \to \infty$ are allowed). Loosely speaking, Eqs. (A.64) and (A.65) follow from considering each mixed state to be an (equal) ensemble of a subset of orthogonal pure states, where each pure state has a fixed phase space volume (in accordance with Weyl's law for the density of states [28]). In what follows, we discuss a more formal justification based on Wigner quasiprobability functions.

We consider the special problem of a particle with position coordinates $\mathbf{x} \in \mathbb{R}^N$, allowing the definition of the real-valued Wigner quasiprobability functions W (following the conventions adopted in Ref. [126]) in an effective Hamiltonian phase space $\mathscr{P} = \{(\mathbf{x},\mathbf{p})\} = \mathbb{R}^{2N}$. For a general (i.e. mixed state) density matrix $\hat{\rho}$ (in units where the reduced Planck's constant $\hbar = 1$),

$$W(\mathbf{x},\mathbf{p}) = \int_L \mathrm{d}\mathbf{y}\ \langle \mathbf{x} - \tfrac{1}{2}\mathbf{y}|\hat{\rho}|\mathbf{x} + \tfrac{1}{2}\mathbf{y}\rangle e^{i\mathbf{p}\cdot\mathbf{y}}, \tag{A.66}$$

where W is normalized according to $\int \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p}\ W(\mathbf{x},\mathbf{p}) = (2\pi)^N$. The overlap of the density matrices is

directly given by the overlap of these quasiprobabilities,

$$\mathrm{Tr}(\hat{\rho}_1\hat{\rho}_2) = \frac{1}{(2\pi)^N}\int \mathrm{d}\mathbf{x}\int \mathrm{d}\mathbf{p}\; W_1(\mathbf{x},\mathbf{p})W_2(\mathbf{x},\mathbf{p}). \tag{A.67}$$

When $W_A(\mathbf{x},\mathbf{p})$ is a uniform distribution over some region A in the phase space (at least at some level of approximation; W is in general not non-negative everywhere [127]), we have $W_A[(\mathbf{x},\mathbf{p})\in A] = [(2\pi)^N/\tilde{\mu}(A)]$, with its value being (approximately) 0 elsewhere. Here, $\tilde{\mu}(A) = \int_A \mathrm{d}\mathbf{x}\mathrm{d}\mathbf{p}$. Using this expression in Eq. (A.67), we get for any two regions A and B and the density matrices $\hat{\rho}_A$ and $\hat{\rho}_B$ corresponding to such uniform Wigner functions,

$$\frac{1}{(2\pi)^N}\mathrm{Tr}(\hat{\rho}_A^2) = \frac{1}{\tilde{\mu}(A)}, \tag{A.68}$$

$$\frac{1}{(2\pi)^N}\mathrm{Tr}(\hat{\rho}_A\hat{\rho}_B) = \frac{\tilde{\mu}(A\cap B)}{\tilde{\mu}(A)\tilde{\mu}(B)}. \tag{A.69}$$

These are equivalent to Eqs. (A.64) and (A.65) above, subject to the normalization of the measure $\tilde{\mu}$. Eqs. (A.64) and (A.65) assume that $\mu(\mathscr{P}) = 1$, but the system considered for Wigner functions has an infinite phase space with $\tilde{\mu}(\mathscr{P}) = \infty$, as well as an infinite dimensional Hilbert space with $d = \infty$. By requiring that the maximally mixed state $\hat{\rho} = \hat{\mathbb{1}}/d$ corresponds to a uniform distribution over the full phase space $\mathscr{P}$ (e.g. any projective measurement onto an orthonormal basis has equal probabilities for every outcome in the former, and the latter is equally distributed over any foliation of the phase space into surfaces of constant position coordinates e.g. $\mathbf{x}$), we heuristically obtain $\tilde{\mu}(\mathscr{P}) = (2\pi)^N d$, fixing the normalization of μ and directly giving Eqs. (A.64) and (A.65). We expect this reasoning to go through without dealing with infinities, if one can suitably define analogues of the Wigner functions restricted to energy shells of finite measure.

Appendix B: Quantum correlators, level spacings, and cyclic ergodicity

B.1 Setup: Level spacing statistics and cyclic permutations

Partly as an attempt to construct a bridge between cyclic ergodicity and correlation functions, we will consider how spectral rigidity influences 2-pt correlators of the form

$$\mathrm{Tr}[\hat{\rho}(t)\hat{A}(0)] = \sum_{n,m} \rho_{nm}(t) A_{mn}(0) e^{i(E_m - E_n)t}. \tag{B.1}$$

Here, we take $\{E_n\}_{n=0}^{d-1}$ to be a set of d energy levels in a Hilbert (sub)space of interest and consider their statistical properties. We will additionally assume that these energy levels are sorted by default:

$$(n < m) \implies E_n \leq E_m. \tag{B.2}$$

Initially, we will consider level spacing distributions for levels paired according to the number of levels between them. The r-th nearest neighbor spacings are:

$$S_n^{(r)} \equiv E_{n+r} - E_n, \tag{B.3}$$

where it is understood that $(n+r)$ is modulo d; with this definition, we have $\sum_{n\in\mathbb{Z}_d} S_n^{(r)} = 0$ for all r.

We focus on the Fourier transform of the r-th spacing distribution $\widetilde{\mathscr{P}}^{(r)}(S)$:

$$\widetilde{\mathscr{P}}^{(r)}(t) \equiv \frac{1}{d} \sum_{n \in \mathbb{Z}_d} e^{iS_n^{(r)} t} = \int \mathrm{d}S \; \widetilde{\mathscr{P}}^{(r)}(S) e^{iSt}. \tag{B.4}$$

Similar Fourier transforms have also been considered recently, for example in Ref. [278].]

In addition to the time evolution operator $\hat{U}_H(t)$, define the energy-shifted unitaries:

$$\hat{U}_{H,r}(t) \equiv \sum_{n \in \mathbb{Z}_d} e^{-iE_{n+r}t} |E_n\rangle\langle E_n|, \tag{B.5}$$

with $\hat{U}_H 0(t) = \hat{U}_H(t)$. We also recall the sorted DFT basis of energy eigenstates from Chap. 2:

$$|C_k\rangle \equiv \frac{1}{\sqrt{d}} \sum_{n \in \mathbb{Z}_d} e^{-2\pi i k n/d} |E_n\rangle, \tag{B.6}$$

which has an implicit dependence on the phases of the energy eigenstates $|E_n\rangle$ that we will neglect. It is further convenient to work with unitary "twisted" cyclic operators for this basis, defined by:

$$\hat{U}_{C,r}|C_k\rangle = e^{-2\pi i k r/d} \hat{U}_C |C_k\rangle = e^{-2\pi i k r/d} |C_{k+1}\rangle \tag{B.7}$$

where $\hat{U}_C \equiv \hat{U}_{C,0}$ simply cyclically permutes the basis vectors (with nonzero r only differing in the introduction of phases. In terms of these entities, we can write the spacing distribution as

$$\widetilde{\mathscr{P}}^{(r)}(t) = \frac{1}{d} \mathrm{Tr}\left[\hat{U}_{H,r}^\dagger(t) \hat{U}_H(t)\right] = \langle C_k | \hat{U}_{H,r}^\dagger(t) \hat{U}_H(t) | C_k \rangle, \tag{B.8}$$

for any $k \in \mathbb{Z}_d$.

Now, we make use of the fact that $[\hat{U}_{H,r}, \hat{U}_{C,r}] = 0$. For $r = 0$, this essentially corresponds to the periodic/random decomposition of the unitary in Sec. 2.5.1; however, we will not need as much

elaborate detail on the coefficients as for cyclic ergodicity, and it suffices to write:

$$\hat{U}_H(t) = \sum_{m\in\mathbb{Z}_d} h_m(t)\hat{U}_C^m. \tag{B.9}$$

We note that any unitary that is diagonal in the energy eigenbasis may be written in this form. From Eq. (B.5), it also follows that

$$\hat{U}_{H,r}(t) = \sum_{m\in\mathbb{Z}_d} h_m(t)\hat{U}_{C,r}^m, \tag{B.10}$$

with the *same* coefficients $h_m(t)$. Substituting this in Eq. (B.8), we get

$$\begin{aligned}\widetilde{\mathscr{P}}^{(r)}(t) &= \sum_{m\in\mathbb{Z}_d} |h_m(t)|^2 e^{2\pi irm/d} \\ &= \sum_{m\in\mathbb{Z}_d} |\langle C_{k+m}|\hat{U}_H(t)|C_k\rangle|^2 e^{2\pi irm/d}, \text{ for any } k.\end{aligned} \tag{B.11}$$

Stated in words, $\widetilde{\mathscr{P}}^{(r)}(t)$ represents the r-th Fourier coefficient of the probability distribution of the time-evolving initial state $\hat{U}_H(t)|C_k\rangle$ in the sorted energy DFT basis.

For the sake of generality, it is useful to consider certain generalized level spacing measures:

$$\widetilde{\mathscr{P}}^{(r,q)}(t) \equiv \frac{1}{d}\sum_{n\in\mathbb{Z}_d} e^{iS_n^{(r)}t} e^{2\pi inq/d}. \tag{B.12}$$

Here, $\widetilde{\mathscr{P}}^{(r,0)}(t) = \widetilde{\mathscr{P}}^{(r)}(t)$, and $\widetilde{\mathscr{P}}^{(0,0)}(t) = 1$; except in the case of $q = 0$ (see Eq. (B.4)), this cannot be expressed as the Fourier transform a distribution function of the S_n. Incidentally, while $\widetilde{\mathscr{P}}^{(r)}(t)$ allows one to recover the exact values of $S^{(r)}$, the set of $\widetilde{\mathscr{P}}^{(r,q)}(t)$ allows one to recover each $S_n^{(r)}$ individually, i.e. including which $S^{(r)}$ corresponds to which n. The appropriate generalization of Eq. (B.11) can be

derived similarly, and is given by:

$$\begin{aligned}\widetilde{\mathscr{P}}^{(r,q)}(t) &= \sum_{m\in\mathbb{Z}_d} h_{m+q}(t)h_m^*(t)e^{2\pi irm/d} \\ &= \sum_{m\in\mathbb{Z}_d} \langle C_{k+m+q}|\hat{U}_H(t)|C_k\rangle\langle C_k|\hat{U}_H^\dagger(t)|C_{k+m}\rangle e^{2\pi irm/d}, \text{ for any } k. \end{aligned} \tag{B.13}$$

Eq. (B.13) is the main relation of interest for the considerations of the next section.

B.2 A complete set of time evolution profiles for 2-point correlators

Here, we will express Eq. (B.1) in terms of the $\widetilde{\mathscr{P}}^{(r,q,()}t)$ [see Eq. (B.22)], using the sorted energy DFT basis. This can be done much more conveniently in the energy eigenbasis (which is certainly simpler for numerics), but we hope to gain better intuition about the connection to quantum cyclic permutations through DFT bases. Consider an *arbitrary* time-evolving observable $\hat{A}(t) = \hat{U}_H(t)\hat{A}(0)\hat{U}_H^\dagger(t)$ [i.e., in the Heisenberg picture], whose matrix element decomposition in the $|C_k\rangle$-basis at the time $t = 0$ is given by:

$$\hat{A}(0) = \sum_{m,\ell} A^{\mathscr{C}}_{m\ell}|C_m\rangle\langle C_\ell|. \tag{B.14}$$

It is convenient to Fourier transform these matrix elements as follows (note that we are treating m and $n = \ell - m \in \mathbb{Z}_d$ as the independent variables here):

$$\begin{aligned} A^{\mathscr{C}}_{m,m+n} &= \sum_{r\in\mathbb{Z}_d} a^{\mathscr{C}}_{rn} e^{2\pi irm/d}, \\ \implies \hat{A}(0) &= \sum_{r,m,n} a^{\mathscr{C}}_{rn} e^{2\pi irm/d}|C_m\rangle\langle C_{m+q}|. \end{aligned} \tag{B.15}$$

Now, consider the matrix elements of $\hat{A}(t)$ in the basis of $|C_k\rangle$:

$$\langle C_j|\hat{A}(t)|C_k\rangle = \sum_{r,m,n} a^{\mathscr{C}}_{rn} e^{2\pi i r m/d} \langle C_{m+n}|\hat{U}_H(t)|C_k\rangle \langle C_j|\hat{U}_H^\dagger(t)|C_m\rangle. \tag{B.16}$$

Using the fact that $[\hat{U}_H, \hat{U}_C] = 0$ to simultaneously shift the j and m indices in the last factor by $k-j$ (to k and $k+m-j$ respectively), we have

$$\langle C_j|\hat{A}(t)|C_k\rangle = \sum_{r,n} a^{\mathscr{C}}_{rn} \left[e^{2\pi i r j/d} \sum_{m\in\mathbb{Z}_d} e^{2\pi i r(m-j)/d} \langle C_{k+(m-j)+(n+j-k)}|\hat{U}_H(t)|C_k\rangle \langle C_k|\hat{U}_H^\dagger(t)|C_{k+(m-j)}\rangle \right]. \tag{B.17}$$

Comparing this with Eq. (B.13), we see that the second line is essentially $\widetilde{\mathscr{P}}^{(r,n+j-k)}(t)$. Writing $n+j-k=q$, we have

$$\langle C_j|\hat{A}(t)|C_k\rangle = \sum_{r,q\in\mathbb{Z}_d} a^{\mathscr{C}}_{r,q+k-j} e^{2\pi i r j/d}\ \widetilde{\mathscr{P}}^{(r,q)}(t). \tag{B.18}$$

This allows us to provide a complete characterization of all 2-point correlation functions $\mathrm{Tr}[\hat{\rho}(0)\hat{A}(t)]$ of a quantum dynamical system in terms of its level spacing distributions; here, $\hat{\rho}(0)$ may be interpreted as an initial state in the Heisenberg picture, in which case correlation functions of this form represent the expectation value of an arbitrary observable in an arbitrary initial state. We also do not impose any further restrictions such as the Hermiticity of the operators or the positivity of $\hat{\rho}$, to allow 2-point correlation functions of an arbitrary number of possibly non-Hermitian operators, such as $\mathrm{Tr}[\hat{\rho}(0)\hat{B}(0)\hat{A}(t)\hat{C}(t)]$. Using $\rho^{\mathscr{C}}_{kj}$ to denote the matrix elements of $\hat{\rho}(0)$ in the $|C_k\rangle$ basis:

$$\hat{\rho}(0) = \sum_{kj} \rho^{\mathscr{C}}_{kj} |C_k\rangle\langle C_j|, \tag{B.19}$$

we have

$$\mathrm{Tr}[\hat{\rho}(0)\hat{A}(t)] = \sum_{r,q\in\mathbb{Z}_d} \left(\sum_{j,k\in\mathbb{Z}_d} \rho^{\mathscr{C}}_{kj} e^{2\pi i j r/d} a^{\mathscr{C}}_{r,(q+k-j)} \right) \widetilde{\mathscr{P}}^{(r,q)}(t). \tag{B.20}$$

Now, expressing $a_{r,(q+k-j)}$ in terms of the $A_{m\ell}$ i.e.

$$a^{\mathscr{C}}_{r,(q+k-j)} = \frac{1}{d}\sum_{m\in\mathbb{Z}_d} A^{\mathscr{C}}_{m,(m+q+k-j)} e^{-2\pi i r m/d}, \tag{B.21}$$

we get (implicitly redefining m to $m-j$):

$$\mathrm{Tr}[\hat{\rho}(0)\hat{A}(t)] = \sum_{r,q\in\mathbb{Z}_d}\left(\frac{1}{d}\sum_{j,k,m\in\mathbb{Z}_d} \rho^{\mathscr{C}}_{k,j} A^{\mathscr{C}}_{(j+m),(k+m+q)} e^{-2\pi i r m/d}\right)\widetilde{\mathscr{P}}^{(r,q)}(t) \tag{B.22}$$

It is convenient to rewrite this using shorthand notation, after translating back to the Schrödinger picture by moving the time-dependence to $\hat{\rho}$:

$$\mathrm{Tr}[\hat{\rho}(t)\hat{A}(0)] = \sum_{r,q\in\mathbb{Z}_d} \left(\hat{\rho},\hat{A}\right)^{\mathscr{C}}_{r,q} \widetilde{\mathscr{P}}^{(r,q)}(t) \tag{B.23}$$

The key point to note here is that all the time dependence is contained in the level spacing distribution measures $\widetilde{\mathscr{P}}^{(r,q)}(t)$, and *every* 2-point correlation function can be expressed as a simple linear combination of these spacing measures. In other words, we can think of the $\widetilde{\mathscr{P}}^{(r,q)}(t)$ as providing a complete basis of possible time-evolution profiles admitted by the level statistics of a quantum dynamical system. Further, each term in (r,q) corresponds to an *independent* measure of level spacing statistics: the $(r,q=0)$ terms measure the r-th nearest neighbor level spacing distribution, while the $(r,q\neq 0)$ terms measure how the r-th neighbor level spacings are correlated with the value of energy. In principle, this provides a complete characterization of quantum dynamics in terms of level spacing statistics, and we can use this to understand *precisely* how level spacing distributions are reflected in different dynamical processes.

However, we also emphasize that there are some tradeoffs. The separation and collecting together of levels spacings into those of r-th nearest neighbors is a somewhat artificial procedure [278],

which is mathematically convenient, but not easy to find naturally in a physical system. We will favor mathematical convenience for now and assume that it is convenient to isolate the $\mathscr{P}^{(r)}(S)$ for each r.

Intuitively, the above procedure amounts to considering the energy eigenbasis expression $\mathrm{Tr}[\hat{\rho}(0)\hat{A}(t)] = \sum_{n,m} \rho_{nm} A_{mn} e^{i(E_m - E_n)t}$, collecting terms involving $e^{i(E_{n+r}-E_n)t}$ for each r (modular addition implicit) and re-expressing them in terms of the $\widetilde{\mathscr{P}}^{(r,q)}(t)$ using $e^{i(E_{n+r}-E_n)t} = \sum_q \widetilde{\mathscr{P}}^{(r,q)}(t) e^{-2\pi i n q/d}$. Doing so gives:

$$\left(\hat{\rho}, \hat{A}\right)^{\mathscr{C}}_{r,q} = \sum_n \rho_{n,n+r} A_{n+r,n} e^{-2\pi i n q/d}. \tag{B.24}$$

This expression is more convenient for numerics; however, Eq. (B.22) allows us to see more readily that the relative contribution of the level spacing distributions $\widetilde{\mathscr{P}}^{(r,0)}(t)$ (i.e., with $q = 0$) is "maximized" (in the sense that the $q \neq 0$ terms do not contribute at all) when either the initial state or the observable is diagonal in the sorted DFT basis $\mathscr{C} = \{|C_k\rangle\}_k$ [this basis is non-unique up to phase factors for the energy eigenstates, but these phase factors do not affect $\left(\hat{\rho}, \hat{A}\right)^{\mathscr{C}}_{r,q}$]. This suggests, for example, that one must work with such a basis (rather than a more physical basis obtained by a more random unitary transformation of the energy eigenstates) to see the maximum dynamical effect of spectral correlators. This provides an alternate justification for the focus on the energy DFT basis in Chap. 2 (as well as, to some extent, Chap. 3).

B.3 Typical behavior of level spacing modes in the time domain

B.3.1 General properties

To gain some intuition for the role of level spacing statistics in correlation functions, it is useful to summarize some essential aspects of the $\widetilde{\mathscr{P}}^{(r,q)}(t)$. The $q = 0$ measures $\widetilde{\mathscr{P}}^{(r,0)}(t)$ fully encode the r-th nearest neighbor level spacing distribution functions via a simple Fourier transform. However, this is insufficient to fully reconstruct the energy levels, for which one must also account for the $q \neq 0$

measures $\widetilde{\mathscr{P}}^{(r,q\neq 0)}(t)$, which are measures of the *correlation* between the r-th nearest neighbor level spacings and the order of the energy levels; in short, the $q \neq 0$ measures describe how the level spacing distribution is actually distributed between the individual energy levels. Due to the discrete nature of the spectrum, we should typically have $\widetilde{\mathscr{P}}^{(r,q)}(t \to \infty) \sim O(d^{-1/2})$, representing small quantum fluctuations.

We further have the mathematical properties:

1. Range of values: $\widetilde{\mathscr{P}}^{(r,0)}(0) = 1$, $\widetilde{\mathscr{P}}^{(r,q\neq 0)} = 0$. Additionally, $|\widetilde{\mathscr{P}}^{(r,q)}(t)| \leq 1$.

2. Conjugation symmetry: $[\widetilde{\mathscr{P}}^{(r,q)}(t)]^* = \widetilde{\mathscr{P}}^{(d-r,q)}(t)$ (due to $\mathscr{P}^{(r)}(S) = \mathscr{P}^{(d-r)}(-S)$).

3. Relation to SFF: $\sum_r \widetilde{\mathscr{P}}^{(r,0)}(t) = K(t)$, where the SFF is normalized to $K(0) = 1$.

The first property suggests that the $\widetilde{\mathscr{P}}^{(r,0)}(t)$ modes describe the "*durability*" of the $t = 0$ value of any 2-point correlation function, while the $\widetilde{\mathscr{P}}^{(r,q\neq 0)}(t)$ modes describe other forms of "noise " generated under time evolution. Due to the generic lack of correlations between spacings and the energy value[1], we should expect that $|\widetilde{\mathscr{P}}^{(r,q\neq 0)}(t)| = O(d^{-1/2})$ in most cases due to the typical random distribution of level spacings; thus, the $q \neq 0$ modes can be treated as noise for a zeroth order intuitive picture, though this argument does not rule out any correlations that may be hidden in this apparent noise (in numerical tests, we see some brief "revivals" at early times in, e.g., CUE).

The third property, describing the relation to the SFF (also noted in Ref. [278]), shows why the SFF appears in most "random" correlation functions from the point of view of this framework [see, e.g., Eq. (3.10)] the weights of each $\widetilde{\mathscr{P}}^{(r,0)}(t)$ are approximately equal (or at least, statistically equidistributed) for random correlators, reminiscent of canonical and normal typicality [17, 23, 24] (see also Sec. 1.1.2).

[1] at least for a uniform density of states.

B.3.2 Effective sinusoidal pairs with slow envelopes in each $\widetilde{\mathscr{P}}^{(r,q)}(t)$ for $q=0$

The $q=0$ modes $\widetilde{\mathscr{P}}^{(r,0)}(t)$ heuristically oscillate as two sinusoids, each with an effective period determined by the mean r-th nearest neighbor level spacing. To see this more clearly, define

$$\begin{aligned}\langle S\rangle_r^+ &\equiv \frac{1}{d-r}\int_0^\infty \mathrm{d}S\; S\; \mathscr{P}^{(r)}(S),\\ \langle S\rangle_r^- &\equiv \frac{1}{r}\int_{-\infty}^0 \mathrm{d}S\; S\; \mathscr{P}^{(r)}(S) = -\frac{d-r}{r}\langle S\rangle_r^+.\end{aligned} \tag{B.25}$$

Note that $\langle S\rangle_r^\pm$ is defined as the mean value of S in $S \gtrless 0$, because $\mathscr{P}^{(r)}(S)$ has zero mean in $S\in\mathbb{R}$ by definition. Now, we write the $S>0$ and $S<0$ components of $\mathscr{P}^{(r)}(S)$ each as smearings of delta functions at $S=\langle S\rangle_r^+$ and $S=\langle S\rangle_r^-$ respectively. To this end, introduce the smearing functions

$$\mathscr{Q}_\pm^{(r)}(S') = \mathscr{P}^{(r)}(S'+\langle S\rangle_r^\pm)\overline{\Theta}\left[\pm(S'+\langle S\rangle_r^\pm)>0\right] \tag{B.26}$$

$$\implies \mathscr{P}^{(r)}(S) = \int \mathrm{d}S'\,\left[\mathscr{Q}_+^{(r)}(S-S')\delta(S'-\langle S\rangle_r^+) + \mathscr{Q}_-^{(r)}(S-S')\delta(S'-\langle S\rangle_r^-)\right]. \tag{B.27}$$

We can therefore write

$$\widetilde{\mathscr{P}}^{(r,q)}(t) = \widetilde{\mathscr{Q}}_+^{(r,q)}(t)e^{i\langle S\rangle_r^+ t} + \widetilde{\mathscr{Q}}_-^{(r,q)}(t)e^{i\langle S\rangle_r^- t}. \tag{B.28}$$

Here, $\widetilde{\mathscr{Q}}_\pm^{(r,q)}(t)$ is the Fourier transform of $\mathscr{Q}_\pm^{(r)}(S)$ multiplied by the appropriate phases $e^{2\pi i n(S)q/d}$, where $n(S)$ is the index of the(/a) energy level with spacing(s) S. This is exact, but its value is in allowing one to interpret $\widetilde{\mathscr{P}}^{(r,0)}(t)$ in intuitive terms as an oscillation possessing two sinusoidal frequencies $\langle S\rangle_r^\pm$, each modulated by the envelope $\widetilde{\mathscr{Q}}_\pm^{(r,q)}(t)$. In particular, when any set of $\mathscr{Q}_\pm^{(r)}(S)$ are symmetric in S, $\widetilde{\mathscr{Q}}_\pm^{(r)}(t)$ is entirely real valued (and symmetric in t) and does not contribute to the oscillating phase. This tends to be the case for $1\ll r\ll d$, where $\mathscr{Q}_\pm^{(r)}(S)$ tends to be a Gaussian having zero mean with variance determined by the spectral rigidity. For asymmetric $\mathscr{Q}_\pm^{(r)}(S)$ (e.g., as is usually

the case for $r = 1$ or $r = d-1$), however, each envelope may also contribute a time dependent phase with frequency comparable to $\langle S\rangle_r^{\pm}$; nevertheless, this effect quickly becomes negligible as $r(\mathrm{mod}\, d)$ moves away from 0, and it is convenient to continue to loosely speak of modes of "frequencies" $\langle S\rangle_r^{\pm}$ even when $r(\mathrm{mod}\, d)$ is close to 0.

Note: For Floquet systems with a uniform density of states on the unit circle (in the complex plane), the two sinusoidal components in each $\widetilde{\mathscr{P}}^{(r,0)}(t)$ collapse into one with the same effective frequency due to the discretization of time, as it turns out that $\langle S\rangle_r^- + 2\pi = \langle S\rangle_r^+$ (with $\langle S\rangle_r^+ = 2\pi r/d$).

Example: A unitary cyclic permutation— A discrete cyclic permutation of an orthonormal basis provides a simple illustration of the above discussion (and one of the few analytical illustrations, given that level spacings are especially difficult to obtain analytically). The regularly spaced levels ensure that $\mathscr{Q}_+^{(r)}(t) = 1$, and we have $\widetilde{\mathscr{P}}^{(r,q=0)}(t) = e^{2\pi i r t/d}$ (in an energy shell of d energy levels, in units where it spans an energy range of 2π), while $\widetilde{\mathscr{P}}^{(r,q\neq 0)}(t) = 0$.

The important aspects to note are:

1. Relative strengths of the two frequencies:

$$\widetilde{\mathscr{Q}}_+^{(r,0)}(0) = (d-r)/d,\ \widetilde{\mathscr{Q}}_-^{(r,0)}(0) = r/d. \tag{B.29}$$

2. Conjugation symmetry of each frequency:

$$\langle S\rangle_r^+ = -\langle S\rangle_{d-r}^-. \tag{B.30}$$

3. Lowest and highest frequency oscillations in correlation functions:

   - The *lowest* frequencies of the system are contained in $r = 1$ and $r = d-1$, namely $\langle S\rangle_1^+ = -\langle S\rangle_{d-1}^-$ (around $\mathscr{N}/d$ if the density of states $\mathscr{N}$ is uniform over large energy intervals),

with strength $\widetilde{\mathscr{Q}}_{+}^{(1,0)}(0) = (d-1)/d \approx 1$. Given that the width of the distribution tends to be comparable to its mean spacing, it is hard to interpret these "lowest frequencies" as actual sinusoids.

- The *highest* frequencies in continuous time (comparable to $\mathscr{N}$) are also contained in the same modes with the negligible strength $\widetilde{\mathscr{Q}}_{-}^{(r,0)}(0) = 1/d \approx 0$. However, in an appropriate discrete time Floquet system, the highest frequencies are instead present near $r \approx d/2$ (and are comparable to $\mathscr{N}/2$) due to $\langle S \rangle_r^{\pm}$ becoming identical), with strengths $\widetilde{\mathscr{Q}}_{\pm}^{(r\approx d/2,0)}(0) \approx 1/2$.

From these observations, we conclude that the following relationship holds:

Nearest-neighbor level spacing statistics $\longleftrightarrow$ *Decay of slowest oscillation mode in correlators* (B.31)

B.3.3 Counter-intuitive slow decay of envelopes (consistent with cyclic ergodicity)

To make concrete statements, we need an estimate of the level spacing distributions $\mathscr{P}^{(r)}(S)$. We consider distributions of the form (where $r' = r$ if $0 \le r < d/2$, and $r' = r - d$ if $d/2 \le r < d$):

$$\mathscr{P}^{(r)}(S) \approx \frac{1}{\sqrt{2\pi\sigma_r^2}} \left\{ \frac{(d-r)}{d} \exp\left[-\frac{(s - \langle S \rangle_r^{+})^2}{2\sigma_r^2} \right] + \frac{r}{d} \exp\left[-\frac{(s - \langle S \rangle_r^{-})^2}{2\sigma_r^2} \right] \right\}, \text{ for } r' \gg 1. \tag{B.32}$$

For large r', Eq. (B.32) describes the most general Gaussian profile for the level spacing fluctuations $\mathscr{Q}_{\pm}^{(r)}(S)$ consistent with the two sinusoidal modes of Sec. B.3.2. Both Wigner-Dyson and Poisson statistics approximately have the Gaussian form in Eq. (B.32) (as seen in e.g. Ref. [102]; see also Chap. 2), but the variance σ_r^2 is determined by the spectral rigidity (quite independent of $\mathscr{P}^{(1)}(S)$): $\sigma_r^2(\text{Poisson}) \sim r$ while $\sigma_r^2(\text{WD}) \sim \ln r$.

Now, let us consider $r \gg 1$, $d - r \gg 1$, for which the form in Eq. (B.32) applies. Here, we get

$$\widetilde{\mathscr{P}}^{(r,0)}(t) \approx \frac{d-r}{d} e^{-\frac{\sigma_r^2 t^2}{2}} e^{i\langle S\rangle_r^+ t} + \frac{r}{d} e^{-\frac{\sigma_r^2 t^2}{2}} e^{i\langle S\rangle_r^- t}. \tag{B.33}$$

See Ref. [278] for the behavior at smaller r. Of particular note, and rather counter-intuitively, these modes decay over a time scale of $t_{\text{WD}} \sim d/\sqrt{\ln r}$ for Wigner-Dyson statistics (as $\sigma_r^2 \sim \ln r/d^2$ and the much faster $t_{\text{P}} \sim d/\sqrt{r}$ for Poisson statistics (as $\sigma_r^2 \sim r/d^2$), though the effective frequencies $\langle S\rangle_r^{\pm}$ are identical:

$$\frac{\left[\widetilde{\mathscr{P}}^{(r,0)}(t)\right]_{\text{Wigner-Dyson}}}{\left[\widetilde{\mathscr{P}}^{(r,0)}(t)\right]_{\text{Poisson}}} \approx \frac{e^{-\frac{\mu \ln r}{\beta_{\text{RMT}} d^2} t^2}}{e^{-\frac{\nu r}{d^2} t^2}} \gg 1,$$

$$\text{for } r \gg 1, (d-r) \gg 1, \text{ and } t = o(d) \text{ (i.e. before the Heisenberg time scale).} \tag{B.34}$$

Here, μ, ν are two $\Theta(1)$ constants, and β_{RMT} represents the level repulsion exponent in the Haar distribution [$\beta_{\text{RMT}} = 1, 2, 4$ respectively for COE, CUE, CSE]. The significantly slower decay of Wigner-Dyson correlation functions compared to Poisson correlation functions in this pre-Heisenberg regime of times may appear puzzling if one keeps to the usual expectation that "chaotic" systems (which tend to be associated with Wigner-Dyson statistics) should relax faster than "non-chaotic" ones (which tend to have Poisson-like statistics). This apparent puzzle is resolved if one recalls that the r-th nearest neighbor level spacing statistics most significantly influences correlation functions of observables diagonal in the energy DFT basis, rather than in more generic bases.

We can understand this better in terms of the following observations:

1. Eq. (B.33) (as a special case of Eq. (B.28)) separates each mode of correlation functions into a (fast) oscillating part and a (slow) decaying envelope modulating this part. Spectral rigidity largely influences only the envelope.

2. Fast ("chaotic") relaxation in correlation functions stems from superpositions of these oscillating parts; at early enough times, the envelope can be approximated to 1, and this is just like Fourier analysis. The appropriate coefficient of each frequency is determined solely by the eigenstates of the operators according to Eq. (B.22). We note that the $q \neq 0$ modes may also play a role in fast relaxation.

3. The correlation functions relevant for cyclic ergodicity effectively eliminate the oscillating part by introducing suitable phases, as shown below. A concrete relation between the persistence amplitudes $z(p, t_0)$ characterizing cyclic ergodicity and these level spacing modes is given by:

$$z(p, t_0) = \frac{1}{d} \sum_{r \in \mathbb{Z}_d} \widetilde{\mathscr{P}}^{(r,0)}(p t_0) e^{-2\pi i r p/d}. \tag{B.35}$$

   For the circular ensembles, the frequencies $\langle S \rangle_r^{\pm}$ are given precisely by $(2\pi r)/d$, so the phases $e^{-2\pi i r p/d}$ cancel out the oscillating parts. This leads to the long-time persistence of a cyclic permutation in spite of any early-time (eigenstate-determined) chaotic relaxation features.

So if one takes the viewpoint of *quantum cyclic ergodicity* rather than "chaotic" relaxation, it is quite natural and expected that Wigner-Dyson correlation functions have envelopes that decay slower in the appropriate basis than in Poisson statistics, to keep $z(p, t_0)$ larger for a sufficiently long time. The faster "chaotic" relaxation processes are, as noted above, determined by superpositions of the oscillatory parts (and possibly the $q \neq 0$ modes) in a form similar to conventional Fourier analysis, and crucially depend on the matrix elements of the relevant operators in the energy eigenbasis. This further substantiates the difference between dynamical ergodicity and, e.g., thermalization or relaxation processes.

Appendix C: Miscellaneous details on quantum scrambling bounds

C.1 A qualitative overview of the Hayden-Preskill argument for fast scrambling

Here, we will briefly review the original setting for motivating the fast scrambling bound $t_s \geq O(\log N)$ in an N-particle system, partly for historical completeness, and partly to emphasize that it is not an entirely quantum mechanical argument but also involves a certain (at this time, unverified) idea related to quantum gravity. We recall that one of the central questions regarding black holes is whether they respect the unitarity of quantum mechanical evolution at the singularity — in other words, can information that falls into the singularity be recovered in any way? This is not the case in standard general relativity (for infalling classical information), or even for quantum field theory in a classical curved spacetime (the usual setting for Hawking radiation), both of which are in any case expected to break down near the singularity. However, there is a hope that there will be mechanisms for avoiding the singularity and the accompanying information loss in quantum gravity. The Hayden-Preskill protocol [53] provides an effective quantum information model for one possible mechanism. While this protocol is entirely quantum mechanical, it does not by itself imply that $t_s \geq O(\log N)$; the presumed connection to quantum gravity is essential for motivating this bound through their argument.

Qualitatively, there are two ingredients at work here: (1) the quantum no-cloning theorem, which is a rigorously established result in quantum information [50], and (2) a proposed (and unverified) principle of "black hole complementarity" that offers a way to prevent information loss in black holes (see Refs. [53, 54], and the references therein). According to the latter, for the purposes

of Hawking radiation observed from far away, quantum information in a black hole may be modeled by a layer of N qubits just outside the horizon (where N is proportional to the Bekenstein-Hawking entropy, and therefore the area of the event horizon), which captures and scrambles the information in any infalling matter and eventually emits it as Hawking radiation. This picture is complementary to the experience of the infalling matter itself, whose information crosses into the horizon and cannot escape.

In essence, the infalling information is "cloned" into two copies: one just outside the horizon, and one retained by the infalling matter, violating the no-cloning theorem. It is argued that this violation of quantum mechanics inside the event horizon may be explained by a more complete theory of quantum gravity [53], and in the meantime, we will continue to externally observe consistency between quantum mechanics and the layer of qubits outside the event horizon if no external observer can observe this violation: i.e., recover the infalling information from Hawking radiation, and then enter the event horizon to also recover the original copy before it hits the singularity. Arguments incorporating the geometry of the problem near the horizon with quantum mechanical estimates [53, 54] suggest that the original copy cannot be recovered if the external observer crosses the horizon after a time $t \sim \log N$ (according to the proper time of the qubits just outside the horizon); thus, if the qubits scramble the information and emit Hawking radiation with parts of this information after $t_s > O(\log N)$, any external observer that waits to recover this information will have lost the opportunity to observe cloning within the horizon.

C.2 Bounds on the growth of OTOCs from overlaps and the SFF

In Chaps. 3 and 4, we have emphasized the difference between scrambling-via-entanglement-generation, and an exponential growth in local out-of-time-ordered correlators (OTOCs). For a time evolving Hermitian operator $\hat{A}(t)$ and another fixed Hermitian operator $\hat{B}(0)$, their OTOC is propor-

tional to the expectation value of their squared commutator in some state (which we leave implicit for now):

$$C_{A,B}(t) \equiv -\langle [\hat{A}(t), \hat{B}(0)]^2 \rangle \geq 0. \tag{C.1}$$

Its main connection to "chaos" is through the fact that classical Poisson brackets show an exponential growth $\{A(t), B(0)\}_{\text{PB}} \propto e^{\lambda t}$ in chaotic systems (including, but not solely, K-mixing systems; ergodicity is not *necessary* for this exponential sensitivity). In fully quantum many-body systems, its physical relevance is due to the commutator: if $\hat{A}$ and $\hat{B}$ are localized to different subsystems at $t = 0$, their commutator measures the extent to which $\hat{A}(t)$ spreads through the system to influence $\hat{B}$.

Given that the SFF $K(t)$ constrains the quantum dynamics of a system through its return probability $P_S(t)$ (see Chap. 3), it is natural to ask what influence this bound has on OTOCs. In this section, we will derive such a bound for the infinite temperature case of Chap. 3, which shows in essence that a certain average of OTOCs of projectors $\hat{\Pi}_k$ in a subsystem cannot grow faster than the SFF decays. We note, however, that this still does not imply any direct sensitivity of *local* OTOCs to scrambling-via-entanglement-generation, as emphasized in Sec. 3.3.4; instead, this should be considered another variant of the energy-time uncertainty principle (see also Ref. [279] for qualitatively related bounds on derivatives of OTOCs).

First, we consider two generic projectors $\hat{A} = \hat{\Pi}_a$ and $\hat{B} = \hat{\Pi}_b$. A mathematical result called Halmos' decomposition [123, 124] provides an intuitive tool to contrast the overlaps and commutators of these projectors. It states that the Hilbert space $\mathscr{H}$ can be decomposed into subspaces that represent different possibilities for the intersection of the projectors (essentially, a vector space version of a Venn diagram):

$$\mathscr{H} = (\mathscr{H}_{00} \oplus \mathscr{H}_{01} \oplus \mathscr{H}_{10} \oplus \mathscr{H}_{11}) \oplus [\mathscr{H}_1 \oplus \mathscr{H}_0]. \tag{C.2}$$

Here, $\mathscr{H}_{s_a s_b}$ represent subspaces in which projector k is not supported ($s_k = 0$) or has full support, i.e.,

acts as identity ($s_k = 1$). The subspaces $\mathscr{H}_1$ and $\mathscr{H}_0$ are equal in dimension $\dim \mathscr{H}_1 = \dim \mathscr{H}_0 = n_o$, respectively having full and no support for $\hat{\Pi}_a$, while both subspaces have partial support for $\hat{\Pi}_b$ (i.e. $\hat{\Pi}_b$ mixes the two subspaces). In particular, there exists a unitary transformation $\hat{V}$ and n_o "principal angles of overlap" $\theta_k \in (0, \pi/2)$ such that the projectors can be written as:

$$\hat{V}\hat{\Pi}_a\hat{V}^\dagger = (0 \oplus 0 \oplus \hat{\mathbb{1}} \oplus \hat{\mathbb{1}}) \oplus [\hat{\mathbb{1}} \oplus 0], \tag{C.3}$$

$$\hat{V}\hat{\Pi}_b\hat{V}^\dagger = (0 \oplus \hat{\mathbb{1}} \oplus 0 \oplus \hat{\mathbb{1}}) \oplus \begin{bmatrix} \bigoplus_k (\cos^2\theta_k)_{1,1} & \bigoplus_k (\cos\theta_k \sin\theta_k)_{1,0} \\ \bigoplus_k (\cos\theta_k \sin\theta_k)_{0,1} & \bigoplus_k (\sin^2\theta_k)_{0,0}, \end{bmatrix} \tag{C.4}$$

Here, the order of direct sums is identical to that in Eq. (C.2), while the subscripts in the matrix for $\hat{\Pi}_b$ in $[\mathscr{H}_1 \oplus \mathscr{H}_0]$ indicates the subspaces being linked by each entry. Another simplification is that when the projectors have the same dimension $\mathrm{Tr}[\hat{\Pi}_a] = \mathrm{Tr}[\hat{\Pi}_b]$ (which we will take to be the case moving forward), it follows that $\dim \mathscr{H}_{10} = \dim \mathscr{H}_{01} = n_p$. Further, we can formally expand the dimension of $\mathscr{H}_{00}$ by introducing auxiliary dimensions (if necessary) to exceed $\dim(\mathscr{H}_{11})$, so that $\mathscr{H}_{11}$ can be absorbed into $\mathscr{H}_1$ and an equal subspace of $\mathscr{H}_{00}$ is absorbed into $\mathscr{H}_0$ with principal angles $\theta_k = 0$. In this case, $\mathscr{H}_{10}$ and $\mathscr{H}_{11}$ can be merged into $\mathscr{H}_1$, and $\mathscr{H}_{01}$ with a subspace of $\mathscr{H}_{00}$ into $\mathscr{H}_0$ with $n = n_o + n_p + \dim(\mathscr{H}_{11})$ principal angles $\theta_k \in [0, \pi/2]$.

Using this decomposition, we can write simple expressions for overlaps and commutators that highlight the main quantitative difference between the two. We have

$$\mathrm{Tr}[\hat{\Pi}_a\hat{\Pi}_b] = \sum_{k=1}^{n} \cos^2\theta_k, \tag{C.5}$$

$$-\mathrm{Tr}\left([\hat{\Pi}_a, \hat{\Pi}_b]^2\right) = 2\sum_{k=1}^{n} \cos^2\theta_k \sin^2\theta_k = 2\sum_{k=1}^{n} (\cos^2\theta_k - \cos^4\theta_k). \tag{C.6}$$

Given a fixed overlap as per Eq. (C.5), we can vary the angles θ_k to extremize the squared commutator in Eq. (C.6). The latter is maximized when all the principal angles are equal (for a given overlap),

resulting in (another special case of the Cauchy-Schwarz inequality)

$$\sum_{k=1}^{n} \cos^4 \theta_k \geq \frac{1}{n} \left(\mathrm{Tr}[\hat{\Pi}_a \hat{\Pi}_b] \right)^2 . \tag{C.7}$$

We note that $n < \mathrm{Tr}[\hat{\Pi}_a] \equiv D_{\hat{\Pi}}$, which gives the bound

$$-\mathrm{Tr}\left([\hat{\Pi}_a, \hat{\Pi}_b]^2 \right) \leq 2 \left\{ \mathrm{Tr}[\hat{\Pi}_a \hat{\Pi}_b] - \frac{1}{D_{\hat{\Pi}}} \left(\mathrm{Tr}[\hat{\Pi}_a \hat{\Pi}_b] \right)^2 \right\} . \tag{C.8}$$

This bound is unaffected by the auxiliary dimensions introduced in $\mathscr{H}_{00}$, which we can now remove from the problem to restore $\mathscr{H}$ to its original size.

Now, we specialize to OTOCs as in Eq. (C.1) for the infinite temperature state (where the expectation value is taken over the maximally mixed state $\hat{\rho} = \hat{\mathbb{1}}/D$), between projectors in the complete set $\hat{\Pi}_k = |k\rangle_S \langle k| \otimes \hat{\mathbb{1}}_E$ considered in Chaps. 3 and 4:

$$C_{kj}(t) \equiv -\frac{1}{D} \mathrm{Tr}\left([\hat{\Pi}_k(t), \hat{\Pi}_j(0)]^2 \right) . \tag{C.9}$$

We should emphasize that these are highly nonlocal OTOCs, and do not measure operator spreading due to being supported on the same subsystem. In particular, both operators $\hat{\Pi}_k$ and $\hat{\Pi}_j$ are delocalized over the entire subsystem $\mathscr{H}_S$, rather than over different (but localized) subsystems. OTOCs of this type have been considered, for example, in Ref. [280]. We also recall the overlaps and return probabilities (see also Chap. 6, Eqs. (6.12) and (6.11)):

$$Q_{kj}(t) \equiv \frac{1}{D_E} \mathrm{Tr}[\hat{\Pi}_k(t) \hat{\Pi}_j(0)], \quad P_k(t) \equiv Q_{kk}(t). \tag{C.10}$$

Setting $\hat{\Pi}_a = \hat{\Pi}_k(t)$ and $\hat{\Pi}_b = \hat{\Pi}_j(0)$ in Eq. (C.8), and noting that $D_{\hat{\Pi}} = D_E$, we get

$$C_{kj}(t) \leq \frac{2}{D_S}\left\{Q_{kj}(t) - Q_{kj}^2(t)\right\}. \tag{C.11}$$

To involve the SFF $K(t)$, we will have to combine the different $C_{jk}(t)$ to obtain $P_k(t)$ via the completeness relation $\sum_{j\neq k} Q_{kj}(t) = 1 - P_k(t)$. However, we do not see a simple way of handling[1] the term with $Q_{kj}^2(t)$, and we will simply drop this term as $Q_{kj}^2(t) \geq 0$. Inserting these relations into Eq. (C.11), we get

$$\sum_{j\neq k} C_{kj}(t) \leq \frac{2}{D_S}\left[1 - P_k(t)\right]. \tag{C.12}$$

Now, we can use $D_S^{-1}\sum_k P_k(t) \geq K(t)$ to get

$$\frac{1}{D_S(D_S-1)}\sum_k\sum_{j\neq k} C_{kj}(t) \leq \frac{2}{D_S(D_S-1)}[1 - K(t)]. \tag{C.13}$$

An analogous bound for $k = j$ [except at the level of Eq. (C.11)] does not appear to be straightforward.

Eqs. (C.13) relates the mean growth of off-diagonal OTOCs of the projectors $\hat{\Pi}_k$, representing computational basis states within $\mathscr{H}_S$, to the decay of the SFF $K(t)$. However, as suggested in Sec. 3.3.4, the saturation criterion to $O(D_S^{-2})$ values (in the present normalization, the averaged off-diagonal OTOC may only be as large as $2/[D_S(D_S-1)]$) may be satisfied at $O(1)$ times as $[1-K(t)]$ is not sensitive to the asymptotic decay of $K(t)$ to leading order. Thus, this bound presently appears to be of merely formal interest. We also note that a generalization to regularized OTOCs [159] at finite temperatures may be possible (similar to the setting of Chap. 4) if a suitable extension of Halmos' decomposition exists when one of the projectors is replaced by a positive operator, e.g., $\hat{\rho}_\beta^{1/2}\hat{\Pi}_a\hat{\rho}_\beta^{1/2}$.

[1] While we can set an upper bound on this term in terms of $-[1-P_k(t)]^2$, the difficulty is in setting a further upper bound after this stage.

Appendix D: Additional details on symmetries and spectra in baker's maps

D.1 Reflection commutators

In this section, we provide numerical evidence that the (generic) generic quasiperiodic and Shor baker quantizations do not have a Fourier reflection symmetry, as defined below. We also provide numerical plots demonstrating symmetries of various eigenvectors.

We will say that a quantization $\hat{U}_N$ has a "Fourier reflection symmetry" if $\hat{U}_N$ commutes with some $\tilde{R}_N^{\omega_1,\omega_2} := (\hat{F}_N^{\omega_1,\omega_2})^2$, for $(\omega_1,\omega_2) \in [0,1)^2$, for each $N \in A\mathbb{N}$. Interestingly enough, there is a generic quantization $\mathrm{Gen}_{A=2}^{0.5,0}$ that does not commute with its "natural" reflection candidate $\tilde{R}_N^{0.5,0}$, but does commute with $\tilde{R}_N^{0,0}$, and so counts as possessing a Fourier reflection symmetry. As discussed in Sec. 7.4.1, these Fourier reflection symmetries are only a small subset of all possible quantum reflection operators.

Letting $\hat{B}_{A,N}^{\theta_1,\theta_2}$ be the generic quasiperiodic quantization for the A-baker's map, we plot the Frobenius matrix norm for a variety of commutators $[\hat{B}_{N,A}^{\theta_1,\theta_2}, \tilde{R}_N^{\omega_1,\omega_2}]$ in Fig. D.1. It appears that the Balazs–Voros quantization, most generic quasiperiodic quantizations, and the Shor baker quantization have nonzero commutators and do not possess a Fourier reflection symmetry.

D.2 Commutator for approximate symmetry

In this section, we analytically check the approximate symmetry $\tilde{R}_N$ introduced in Section 7.4.3 (Eq. (7.18)) is in some sense close to commuting with the Balazs–Voros quantization $\hat{B}_{N,A}$. More

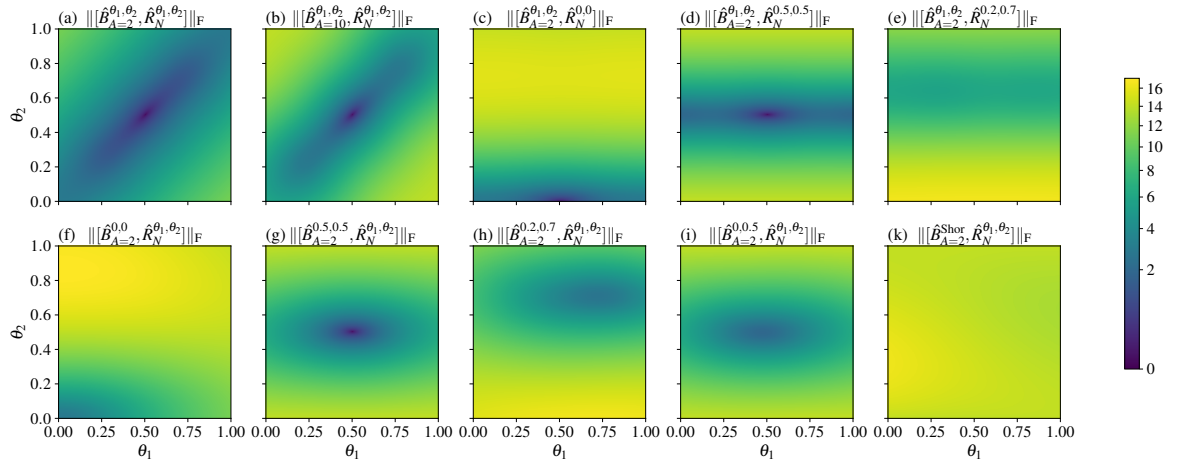


Figure D.1: (a)–(e): Plots of the Frobenius matrix norm of the commutator $[\hat{B}_{N,A}^{\theta_1,\theta_2}, \tilde{R}_N^{\omega_1,\omega_2}]$ as a function of $\theta = (\theta_1, \theta_2)$, for: (a)–(b) $\omega = \theta$, (c) $\omega = (0,0)$, (d) $\omega = (0.5, 0.5)$, and (e) $\omega = (0.2, 0.7)$. In all cases $N = 100$. In (a), (b), and (d), the commutator is zero only at $\theta_1 = \theta_2 = 1/2$, which corresponds to the Saraceno quantization. In (c), the $\mathrm{Gen}_{A=2}^{0.5,0}$ quantization is seen (perhaps surprisingly) to commute with $\hat{R}_N^{0,0}$. However, in (e) and for randomly chosen ω, it appears that $\|[\hat{B}_{A,N}^{\theta_1,\theta_2}, \tilde{R}_N^{\omega_1,\omega_2}]\|_{\mathrm{F}}$ is bounded away from zero for all θ. (f)–(k): Plots of the Frobenius matrix norm of the commutator $[\hat{U}_N, \tilde{R}_N^{\theta_1,\theta_2}]$, where $\hat{U}_N$ is a fixed quantization and $\tilde{R}_N^{\theta_1,\theta_2}$ ranges over $\theta \in [0,1)^2$. In all plots except for the Saraceno quantization in (g), the matrix norm appears bounded away from zero, indicating the quantizations should not have a Fourier reflection symmetry. Plots for $A = 10$ appear similar, and plots for the phase variant quantizations also appear bounded away from zero. In all of the above plots, the sampling mesh is size 200×200.

precisely, for $\hat{B}_{N,A}$ the Balazs–Voros quantization and $\tilde{R}_N = \tilde{R}_N^{0,0}$, we show the only possible large matrix elements $\langle x|[\hat{B}_{N,A}, \tilde{R}_N]|y\rangle$ of the commutator $[\hat{B}_{N,A}, \tilde{R}_N]$ are those (x, y) with $y \in \frac{N}{A}\mathbb{Z}$ and with x close to 0 or N and not in $A\mathbb{Z}$.

Let $a, b \in \{0, \ldots, A-1\}$ be defined so that $a\frac{N}{A} \leq y < (a+1)\frac{N}{A}$ and $b\frac{N}{A} \leq N - y \bmod N < (b+1)\frac{N}{A}$.

Using that $\tilde{R}_N|y\rangle = |N - y\rangle$ (taken modulo N), direct evaluation shows,

$$\langle x|[\hat{B}_{N,A}, \tilde{R}_N]|y\rangle = \frac{\sqrt{A}}{N} \sum_{m=0}^{N/A-1} \left[e^{2\pi i a x/A} e^{2\pi i x m/N} e^{2\pi i m y A/N} - e^{2\pi i x b/A} e^{-2\pi i x m/N} e^{-2\pi i m y A/N} \right]. \tag{D.1}$$

First, if $x + yA \in N\mathbb{Z}$, which would prevent geometric summation, then since $A|N$ we must also have $x \in A\mathbb{Z}$. Combined with $x + yA \in N\mathbb{Z}$, then Eq. (D.1) is zero in this case. For $x + yA \notin N\mathbb{Z}$, we can

evaluate,

$$\langle x|[\hat{B}_{N,A},\tilde{R}_N]|y\rangle = \frac{\sqrt{A}}{N}\Big[e^{2\pi iax/A}\frac{e^{2\pi ix/A}-1}{e^{2\pi ix/N}e^{2\pi iyA/N}-1} - e^{-2\pi ibx/A}\frac{e^{-2\pi ix/A}-1}{e^{-2\pi ix/N}e^{-2\pi iyA/N}-1}\Big], \tag{D.2}$$

which we see is zero if $x \in A\mathbb{Z}$. If $y \in \frac{N}{A}\mathbb{Z}$, then one can check that $a+b \in \{0,A\}$, and we use the bound $|e^{2\pi x/N}e^{2\pi iAy/N}-1| \geq \frac{c}{N}d(x,N\mathbb{Z})$ for a numerical constant $c>0$. This gives the bound

$$\langle x|[\hat{B}_{N,A},\tilde{R}_N]|y\rangle = \mathcal{O}\Big(\frac{\sqrt{A}}{d(x,N\mathbb{Z})}\Big), \tag{D.3}$$

which thus allows large commutator matrix elements for the A values of $y \in \frac{N}{A}\mathbb{Z}\cap[0,N-1]$ and x close to 0 or N (and not in $A\mathbb{Z}$).

If $y \notin \frac{N}{A}\mathbb{Z}$, then one can check $a+b=A-1$, and we obtain from Eq. (D.2) that

$$\langle x|[\hat{B}_{N,A},\tilde{R}_N]|y\rangle = \frac{\sqrt{A}}{N}e^{2\pi iax/A}(1-e^{2\pi ix/A}) = \mathcal{O}\Big(\frac{\sqrt{A}}{N}\Big),$$

which is small. Thus, the only possible large matrix elements of the commutator $[\hat{B}_{N,A},\tilde{R}_N]$ are those (x,y) from Eq. (D.3) with $y \in \frac{N}{A}\mathbb{Z}$ and x close to 0 or N and not in $A\mathbb{Z}$.

D.3 Details for the computation of the early time SFF slope

In this section, we provide the details for our numerical computations of the early time SFF slope. Examples of RMT behavior and (rare) bad early time behavior are shown in Fig. D.2.

1. We averaged the SFF at time t with its nearest 2ℓ neighbors (or up to time $2t-1$ if $t<\ell$), with $\ell=20$ for $N<1000$ and $\ell=40$ for $N\geq 1000$. The choice of averaging to time $2t-1$ for $t<\ell$ keeps the averaging symmetric about t.

2. We took the first f points of the above averaged SFF, where $f=20$ for $N<1000$, $f=40$ for

$1000 \leq N < 5000$, and $f = 60$ for $N \geq 5000$, and ran a least squares fit for a line through the origin to get the best slope. We also retained the scaled residual error, which is the residual error when running the least squares fit for $x \in [1, f\,] \cap \mathbb{Z}$ and $y = N$ SFF(x).

3. We removed all "outliers" which had scaled residual error over 100 (or 400 for $A = 15$, to make sure not too many points were removed). We then averaged the slopes among points within 10 units away (ignoring outliers) and plotted the resulting slopes. We note that the removed outlier points are not necessarily those with an outlier SFF slope value, but just those for which the least squares fit did not work well.

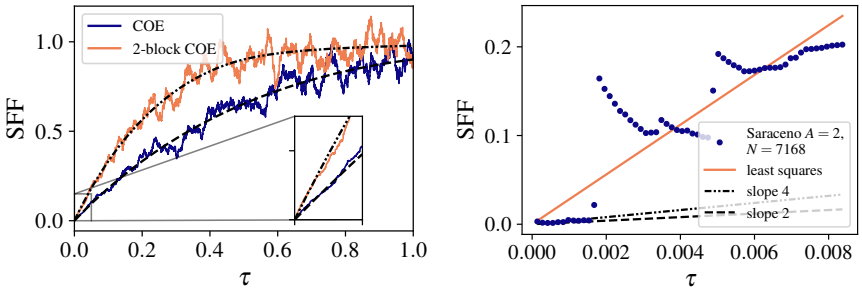


Figure D.2: (Left) SFF for random instances of a COE and a 2-block COE matrix for reference, with $N = 9690$ and $\ell = 100$. There is a clear distinction between COE and 2-block COE with this averaging method, which in particular identifies the slope near 0. (Right) Example of the least squares fit for a removed outlier of the Saraceno $A = 2$ quantization, $N = 7168$ (plotted for longer times in Fig. 7.10(b)). Removed outliers amount to only 0.86% of the values of $N \in 2\mathbb{N}$ considered for this quantization in Fig. 7.9(b).

Appendix E: Analytical and numerical details for Partial SFFs

E.1 Spectral form factor in Wigner-Dyson random matrix ensembles

In this appendix, we review the definition and essential properties of the Wigner-Dyson random matrix ensembles. Further, we recall the expressions of the SFF for Hamiltonian and Floquet dynamics modeled with random matrices from these ensembles.

The Wigner-Dyson ensembles are standard distributions of random matrices used to model some of the properties of energy or quasi-energy eigenvalues and eigenstates of RMT-like Hamiltonian and Floquet systems [28, 36, 262, 263]. We work with two classes of the Wigner-Dyson ensembles - the unitary (U) class for systems that are not time reversal invariant, and the orthogonal (O) class for some systems with time-reversal invariance (the symplectic (S) class applies to other systems with time-reversal invariance, but is not relevant for our examples). We note in particular that nonconventional time-reversal symmetries should also be considered [28], e.g., invariance under complex conjugation in some basis (which corresponds to the orthogonal class). Each class is characterized by a symmetry group comprised of the corresponding set of similarity transformations (i.e. all unitary or orthogonal transformations).

For Hamiltonian systems with time evolution operator $T(t) = \exp(-iHt)$, it is conventional to choose the Gaussian Unitary Ensemble (GUE) of Hermitian matrices or the Gaussian Orthogonal Ensemble (GOE) of real symmetric matrices to represent the Hamiltonian H of the appropriate class. In the case of periodically driven Floquet dynamics with time-evolution operator $T(t = \tau n) = V^n, n \in \mathbb{N}$,

where V is the unitary Floquet operator corresponding to a time period τ, the appropriate representative ensembles for V are the Circular Unitary Ensemble (CUE) of unitary matrices and the Circular Orthogonal Ensemble (COE) of symmetric unitary matrices. These ensembles accurately model the local eigenvalue correlations of the corresponding systems (but not necessarily global eigenvalue features larger than the inverse Thouless time scale [21, 38] e.g. the smoothened density of states), and describe an idealization of the eigenstate distribution (which is generalized by ETH [21, 52]). But for the special case of RMT-like Floquet systems, the eigenstate distribution is seen to be in close agreement with the Wigner-Dyson ensembles [42, 243, 245–248].

For these random matrix models, the spectral form factor can be calculated analytically (see for instance Ref. [193]). For completeness, we recall the well-known expressions here. For Hamiltonians H from GUE or GOE, one finds

GUE model

$$K(t) = r(t)^2 + \frac{1}{D}\begin{cases} \frac{t}{t_H} & \text{for } 0 < t \leq t_H, \\ 1 & \text{for } t > t_H, \end{cases} \tag{E.1}$$

GOE model

$$K(t) = r(t)^2 + \frac{1}{D}\begin{cases} 2\frac{t}{t_H} - \frac{t}{t_H}\log\left(1 + 2\frac{t}{t_H}\right) & \text{for } 0 < t \leq t_H, \\ 2 - \frac{t}{t_H}\log\left(\frac{2t+t_H}{2t-t_H}\right) & \text{for } t > t_H, \end{cases} \tag{E.2}$$

where $r(t) = t_H J_1(4Dt/t_H)/(2Dt)$ with J_1 denoting the Bessel's function of the first kind. The Heisenberg time t_H, connected to the inverse spacing of adjacent energy levels, depends on the width of the Gaussian distribution of the matrix elements and marks the onset time of the plateau of the SFF. For the results presented in Sec. (8.3), we fix it numerically, by matching plateau onset times for the

Hamiltonian Eq. (8.25) and the GOE model.

For the Floquet operators V from CUE or COE, one finds

CUE model

$$K(t) = \frac{1}{D} \begin{cases} \frac{t}{t_H}, & \text{for } 0 < t \leq t_H, \\ 1, & \text{for } t > t_H, \end{cases} \tag{E.3}$$

COE model

$$K(t) = \frac{1}{D} \begin{cases} \frac{2t}{t_H} - \frac{t}{t_H} \log\left(1 + 2\frac{t}{t_H}\right) & \text{for } 0 < t \leq t_H, \\ 2 - \frac{t}{t_H} \log\left(\frac{2t/t_H+1}{2t/t_H-1}\right) & \text{for } t > t_H, \end{cases} \tag{E.4}$$

Here, $t_H = D\tau$ with τ to be identified with the period of the Floquet system to be modeled.

E.2 Partial spectral form factor in general RMT-like systems

Here, we derive the typical behavior of the PSFF for ensembles of RMT-like systems, more general than random matrix ensembles, as considered in Sec. 8.2.2. As in Eq. (8.13), we decompose the reduced density matrix into a pure trace, a traceless smooth part and a traceless fluctuating part, $\rho_B(E) = D_B^{-1}\mathbb{1} + \Delta\rho_B(E) + \delta\rho_B(E)$. For the smooth part, we assume that there exists an extrapolation of each matrix element to a continuous energy variable such that for some (as yet unspecified) time $t_\rho \ll O(D)$,

$$(\Delta\tilde{\rho}_B(t))_{jk} \equiv \int \mathrm{d}E\; e^{-iEt} (\Delta\rho_B(E))_{jk} = 0, \;\forall\; |t| > t_\rho \;. \tag{E.5}$$

The remaining energy dependent part of $\rho_B(E)$ i.e. the part that oscillates rapidly and has no low frequency Fourier component (on extrapolation to continuous energy) will be taken to be the fluctuating

part,

$$(\delta\tilde{\rho}_B(t))_{jk} \equiv \int \mathrm{d}E\; e^{-iEt}(\delta\rho_B(E))_{jk} = 0,\ \forall\ |t| \leq t_\rho\ . \tag{E.6}$$

Up to this point, such a decomposition is always possible. We will additionally take t_ρ to be set by the scale of randomization in the ensemble discussed in Sec. 8.2.2, so that the fluctuating part can be identified as the part that is completely randomized in the ensemble. We note that the smooth part may fluctuate between different ensemble realizations, but can not be randomized in the same sense as the fluctuating part as it is roughly constant within an energy window of size t_ρ^{-1}. Similarly, we will not require randomization of the correlators of $\delta\rho_B(E)$ between energies further apart than $\sim t_\rho^{-1}$, for which the correlator may have to be nonvanishing to maintain zero Fourier component of the fluctuating part at $t \leq t_\rho$.

To understand the effect of this decomposition in the PSFF, we will first perform a prototype calculation with simpler notation. Consider two functions $f(E)$ and $g(E)$ of a continuous variable E, with respective Fourier transforms $\tilde{f}(t)$ and $\tilde{g}(t)$, both of which potentially vary over different realizations of the ensemble. We will eventually associate these functions with (components of) the different parts of the reduced density matrices of the energy eigenstates. Define the quantity,

$$\begin{aligned} F(t) &= \frac{1}{D^2}\overline{\sum_{j,k} e^{i(E_j-E_k)t} f(E_j)g^*(E_k)} \\ &= \frac{1}{D^2}\int \frac{\mathrm{d}t_l}{2\pi}\int \frac{\mathrm{d}t_r}{2\pi}\,\overline{\tilde{f}(t_l)\tilde{g}^*(t_r)\sum_{j,k} e^{iE_j(t+t_l)-iE_k(t+t_r)}}\ . \end{aligned} \tag{E.7}$$

Now, it is convenient to define an ensemble-averaged unequal time SFF $K(t_1,t_2) = D^{-2}\overline{\sum_{j,k} e^{iE_jt_1-iE_kt_2}}$, which reduces to $K(t)$ at equal times $t_1 = t_2 = t$. The sum of phases $D^{-2}\sum_{j,k} e^{iE_j(t+t_l)-iE_k(t+t_r)}$ in Eq. (E.7) would fluctuate strongly over different ensemble realizations at large t_1, t_2 corresponding to fluctuations of the positions of energy levels, much like the SFF without ensemble averaging [110];

if we assume the ensemble is such that these fluctuations are not correlated with those of f and g (i.e. the reduced energy eigenstates), we can perform the ensemble average over the sum of phases independently, allowing us to formally replace it with $K(t+t_l, t+t_r)$,

$$F(t) = \int \frac{\mathrm{d}t_l}{2\pi} \int \frac{\mathrm{d}t_r}{2\pi} K(t+t_l, t+t_r) \overline{\tilde{f}(t_l)\tilde{g}^*(t_r)} \,. \tag{E.8}$$

For instance, in a fully RMT-like system as we will soon specialize to, this assumption can be justified by considering the energy eigenstates in an ensemble realization as sufficiently random superpositions of those of another ensemble realization (in the spirit of Refs. [18, 264–266]), which should then be uncorrelated with the precise positions of the energy levels.

To simplify Eq. (E.8) further, we need to know the form of $K(t_1, t_2)$. For mathematical simplicity, we assume (fully RMT-like) level statistics in the unitary Wigner-Dyson class. The ensemble-averaged two level correlation function for nearby energy levels E_j,E_k (closer than $\sim t_{\mathrm{Th}}^{-1}$) in this class takes the universal form [28, 36, 193],

$$\begin{aligned} &\overline{\delta\left(E+\frac{\omega}{2}-E_j\right)\delta\left(E-\frac{\omega}{2}-E_k\right)} \\ &= [\mathscr{N}^{\#}(E)]^2 \left\{1 + \frac{\delta(\omega)}{\mathscr{N}^{\#}(E)} - \mathrm{sinc}^2\left[\omega\pi\mathscr{N}^{\#}(E)\right]\right\} \,, \end{aligned} \tag{E.9}$$

where $\mathscr{N}^{\#}(E)$ is the smoothened (continuous and ensemble-averaged) density of states, whose Fourier transform satisfies $\tilde{\mathscr{N}}^{\#}(t \gg t_{\mathrm{Th}}) \approx 0$. The ensemble averaged sum over E_j, E_k in the definition of $K(t)$ can then be replaced by an integral weighted by the two level correlation in Eq. (E.9). Using methods analogous to the calculation of $K(t)$ for this correlation function in Ref. [193], we obtain the following

late time behavior for $t_1, t_2 \gg t_{\text{Th}}$,

$$K(t_1,t_2) = \frac{1}{D^2}\begin{cases} \tilde{\mathcal{N}}^{\#}(\tau_{12}), & T_{12} > 2\pi\mathcal{N}^{\#}(E)\ \forall E, \\ \dfrac{|T_{12}|}{\beta_{\text{RMT}}\pi}\tilde{\Theta}_{\mathcal{N}^{\#}}(\tau_{12}), & T_{12} < 2\pi\mathcal{N}^{\#}(E)\ \forall E, \end{cases} \tag{E.10}$$

where $\beta_{\text{RMT}} = 2$ for the unitary Wigner-Dyson class, and we have introduced the shorthand symbols $T_{12} = (t_1+t_2)/2$, $\tau_{12} = t_2 - t_1$. $\tilde{\Theta}_{\mathcal{N}^{\#}}(t)$ is the Fourier transform of the unit step function $\Theta(\mathcal{N}^{\#}(E))$, the latter being 1 where $\mathcal{N}^{\#}(E) > 0$ and zero elsewhere. Essentially, the unequal time SFF is generally negligible for (large) unequal times, with a small spread around $t_1 = t_2$ determined by the variation of the density of states; as noted earlier, it reduces to the SFF at precisely equal times. We also identify $2\pi\mathcal{N}^{\#}(E)$ with the Heisenberg time t_H, assuming that $\mathcal{N}^{\#}(E)$ is at least of the same order of magnitude throughout the spectrum. In the orthogonal and symplectic Wigner-Dyson classes, there are significant corrections (relative to the unitary class) to the form of the equal time SFF $K(t)$ near $t \sim t_H$. But for $t \ll t_H$, virtually the same results hold with $\beta_{\text{RMT}} = 1$ for the orthogonal class and $\beta_{\text{RMT}} = 4$ for the symplectic class [193] (of course, the plateau behavior for $t \gg t_H$ is generally independent of such specifics). Analogously, we expect similar replacements (the appropriate value of β_{RMT}, and focusing on the $T_{12} \gg t_H$ and $T_{12} \ll t_H$ regimes) to work for the unequal time SFF in Eq. (E.10) as well. With this expectation, we write

$$K(t_1,t_2) = \frac{1}{D^2}\begin{cases} \tilde{\mathcal{N}}^{\#}(\tau_{12}), & T_{12} \gg t_H, \\ \dfrac{|T_{12}|}{\beta_{\text{RMT}}\pi}\tilde{\Theta}_{\mathcal{N}^{\#}}(\tau_{12}), & T_{12} \ll t_H, \end{cases} \tag{E.11}$$

for $t_1, t_2 \gg t_{\text{Th}}$ in any Wigner-Dyson symmetry class.

Using the decomposition of $\rho_B(E)$ with these definitions then gives several terms for $K_A(t)$ of the form of Eq. (E.7), where f and g independently go over each of D_B^{-1}, $\Delta\rho_B$ and $\delta\rho_B$, with an

additional trace of the product over the B subspace. Now, we will argue that all cross terms with $f \neq g$ may be taken to vanish. When $f = D_B^{-1}$, the overlap becomes $\mathrm{Tr}_B[fg] = D_B^{-1}\mathrm{Tr}_B[g]$, which is zero when $g = \Delta\rho_B, \delta\rho_B$, which are both traceless. When, say, f is $\Delta\rho_B$ and g is $\delta\rho_B$, the cross term vanishes due to the assumption that ensemble averaging randomizes $\delta\rho_B$.

Dropping the cross terms for the above reasons gives the form of Eq. (8.14), $K_A(t) = K(t) + \Delta K_A(t) + \delta K_A(t)$, where $K(t)$ is the full SFF, and

$$\Delta K_A(t) = \sum_{j,k} \frac{\overline{e^{i(E_j - E_k)t}\mathrm{Tr}_B\left[\Delta\rho_B(E_j)\Delta\rho_B(E_k)\right]}}{DD_A}, \tag{E.12}$$

$$\delta K_A(t) = \sum_{j,k} \frac{\overline{e^{i(E_j - E_k)t}\mathrm{Tr}_B\left[\delta\rho_B(E_j)\delta\rho_B(E_k)\right]}}{DD_A}. \tag{E.13}$$

In Chap. 8, it is argued that $\delta K_A(t \gg t_\rho)$ amounts to a constant shift after ensemble averaging due to the randomization of $\delta\rho_B(E)$. Here, we will complete the evaluation of $\Delta K_A(t)$ using the prototype Eq. (E.8) with $f = g = (\Delta\rho_B)_{ab}$ and the expression in Eq. (E.11) with $t_1 = t + t_l$, $t_2 = t + t_r$. As the definition of $\Delta\rho_B$ sets $t_l, t_r < t_\rho$, we have $|T_{12}| = |t| + \mathrm{sgn}(t)(t_l + t_r)/2$ at large times (i.e. $t \gg t_{\mathrm{Th}}, t_\rho$). For $t \ll t_H$ in this regime, this gives,

$$\begin{aligned}\Delta K(t : t_{\mathrm{Th}}, t_\rho \ll t \ll t_H) = \frac{1}{DD_A}\int \frac{\mathrm{d}t_l}{2\pi}\int \frac{\mathrm{d}t_r}{2\pi} \\ \left[\frac{1}{\beta_{\mathrm{RMT}}\pi}\left(|t| + \mathrm{sgn}(t)\frac{t_l + t_r}{2}\right)\tilde{\Theta}_{\mathscr{N}^\#}(t_l - t_r)\right. \\ \left.\overline{\left(\sum_{a,b}\left(\Delta\tilde{\rho}_B(t_l)\right)_{ab}\left(\Delta\tilde{\rho}_B^*(t_r)\right)_{ab}\right)}\right].\end{aligned} \tag{E.14}$$

The Hermiticity of $\Delta\rho_B$ implies that $\left(\Delta\tilde{\rho}_B(-t)\right)_{ab} = \left(\Delta\tilde{\rho}_B^*(t)\right)_{ba}$. Consequently, making the integration variable transformation $t_l \to -t_r$, $t_r \to -t_l$ in Eq. (E.14), we see that inside the parentheses in the second line the $|t|$ term is unaltered but the $\mathrm{sgn}(t)$ term transforms to its negative, while all factors

outside the parentheses remain unaltered. It follows that the contribution from the sgn(t) term actually evaluates to zero, leaving only a linear ramp term from $|t|$. For $t \gg t_H$, we directly obtain only a plateau contribution. Now, it is straightforward to Fourier transform back to the energy variable E,

$$\begin{aligned}
&\Delta K_A(t \gg t_{\mathrm{Th}}, t_\rho) \\
&= \frac{1}{DD_A} \int \mathrm{d}E \begin{cases} \overline{\mathscr{N}^{\#}(E)\mathrm{Tr}_B\left[\Delta\rho_B^2(E)\right]}\,, & t \gg t_H, \\ \dfrac{t}{\beta_{\mathrm{RMT}}\pi}\overline{\Theta(\mathscr{N}^{\#}(E))\mathrm{Tr}_B\left[\Delta\rho_B^2(E)\right]}\,, & t \ll t_H. \end{cases}
\end{aligned} \tag{E.15}$$

For ease of interpretation, we can convert E back to a discrete energy variable from its present continuous form via the following correspondence relations for sums over energy levels: $\sum_i \longleftrightarrow \int \mathrm{d}E\ \mathscr{N}^{\#}(E)$ and $\sum_i [\mathscr{N}^{\#}(E_i)]^{-1} \longleftrightarrow \int \mathrm{d}E\ \Theta(\mathscr{N}^{\#}(E))$, which become equalities on ensemble averaging. Then we get the expression,

$$\begin{aligned}
&\Delta K_A(t \gg t_{\mathrm{Th}}, t_\rho) \\
&= \frac{1}{DD_A} \begin{cases} \overline{\sum_i \mathrm{Tr}_B\left[\Delta\rho_B^2(E_i)\right]}\,, & t \gg t_H, \\ \dfrac{t}{\beta_{\mathrm{RMT}}\pi}\overline{\sum_i [\mathscr{N}^{\#}(E_i)]^{-1}\mathrm{Tr}_B\left[\Delta\rho_B^2(E_i)\right]}\,, & t \ll t_H. \end{cases}
\end{aligned} \tag{E.16}$$

Together with the expression for the full SFF [$t_1 = t_2$ in Eq. (E.11)] and the constant contribution from the fluctuating part, this directly leads to Eq. (8.15).

E.3 Constraints from eigenstate thermalization

In this Appendix, we discuss the constraints on the spectrum and ensemble averaged PSFF parameters, $\mathscr{P}_B$ (purity of reduced density matrices), $\delta\mathscr{P}_B$ (fluctuating part) and $\Delta\mathscr{P}_B$ (smooth part), as measures of the extent of delocalization and thermalization of energy eigenstates. In App. E.3.1, we

discuss these constraints based on a qualitative picture of subsystem ETH, paying particular attention to thermalization as a distinct phenomenon from delocalization. We justify this qualitative picture in the subsequent section, first in terms of a version of the original conjecture of subsystem ETH [52] for fully delocalized states in App. E.3.2.1, and argue for its extension to eigenstates of arbitrary delocalization in App. E.3.2.2.

E.3.1 PSFF as a probe of thermalization and delocalization

We begin with a qualitative discussion of thermalization (in the sense of subsystem ETH) and delocalization. We work in a 'physical basis' - one whose basis vectors are close to pure states in most physically accessible (e.g. local [51]) subsystems, such as a product basis of qubits. *Thermalization* then corresponds to a significant overlap of the macroscopic features of eigenstates of nearby energies whose individual components are sufficiently random (and therefore, macroscopically similar), whereas non-thermal behavior is seen when nearby eigenstates do not have a large overlap. This is to be distinguished from the extent of *delocalization* of an eigenstate, which is the number of bases states $\ell \leq D$ that it has a significant probability of being found in.

It is useful to introduce an effective dimension $D_A^{\text{eff}} \leq D_A, \ell$ of the Hilbert space of subsystem A, corresponding to the typical number of degrees of freedom of subsystem A over which the eigenstate is delocalized within its support in the physical basis. In particular, $D_A^{\text{eff}} = D_A$ if the eigenstates appear completely delocalized over subsystem A, and more generally D_A^{eff} is typically larger for larger D_A (up to ℓ). For instance, D_A^{eff} is a monotonically increasing function of D_A when the latter is varied by successively choosing larger subsystems A containing the previous one; additionally, it increases from $D_A^{\text{eff}} = 1$ for $D_A = 1$, to $D_A^{\text{eff}} = \ell$ for $D_A = D$. We also use the notation $O(x)$ to mean a non-negative number whose magnitude is at most of the order of magnitude of x, to leading order when $x \gg 1$. In particular, we will take $D \gg D_A, D_B \gg 1$.

Assuming that D_A^{eff} is typical for A throughout the spectrum, the purity in subsystem B satisfies,

$$\mathscr{P}_B = (D_A^{\text{eff}}/\ell) + O(D_A^{\text{eff}}/\ell) + O(1/D_A^{\text{eff}}), \tag{E.17}$$

subject to $\mathscr{P}_B \gtrsim (D_A^{\text{eff}}/\ell), (1/D_A^{\text{eff}})$. The first two terms are due to the eigenstate being delocalized in subsystem B with effective dimension (ℓ/D_A^{eff}), with the second term containing larger scale variations of its components. We will call this, the 'macroscopic' contribution, which grows with D_A^{eff}. The last term is due to the randomness of the eigenstate components i.e. the 'microscopic' contribution, which decays with D_A^{eff} (and is also typically bounded from below by $(1/D_A^{\text{eff}})$). Being a linear combination of the macroscopic and microscopic contribution, the purity shows an initial decay with D_A^{eff} for small values of the latter, and eventually a growth for larger values of $D_A^{\text{eff}} \gtrsim \sqrt{\ell}$. Both $D_A^{\text{eff}} = 1, \ell$ correspond to pure states with $\mathscr{P}_B = 1$.

The parameters $\delta\mathscr{P}_B$, $\Delta\mathscr{P}_B$ satisfy the following order-of-magnitude inequalities,

$$\delta\mathscr{P}_B \gtrsim O(1/D_A^{\text{eff}}), \tag{E.18}$$

$$D_B^{-1} + \Delta\mathscr{P}_B \lesssim (D_A^{\text{eff}}/\ell) + O(D_A^{\text{eff}}/\ell). \tag{E.19}$$

The first inequality is the statement that the fluctuating part must include at least the randomness of eigenstate components; the second says that the smooth part or overlap of such eigenstates can at most contain all their macroscopic features. They are also subject to the constraint $\mathscr{P}_B = D_B^{-1} + \Delta\mathscr{P}_B + \delta\mathscr{P}_B$, which can be interpreted in the present context as follows: the macroscopic contribution to the purity must be distributed in some manner between the smooth and fluctuating parts (with the exception of the maximally mixed part D_B^{-1}); the microscopic contribution is however completely contained in the fluctuating part.

According to ETH, the only difference between *thermal eigenstates* of nearby energies is in their

microscopic random fluctuations, with all their macroscopic features completely contained in their overlap. This means that the inequalities in Eqs. (E.18) and (E.19) are satisfied as equalities for thermal eigenstates. In particular, $\delta\mathscr{P}_B$ can only decay with increasing D_A^{eff} - a fact that is responsible for the nearly identical dynamics of observables in subsystem B (for large D_A) in such eigenstates. In contrast, *non-thermal eigenstates* have at least some of the macroscopic contribution included in the fluctuating part, and therefore satisfy Eqs. (E.18) and (E.19) much further from equality. In this case, the macroscopic contribution to the fluctuating part may even show up as a growth of $\delta\mathscr{P}_B$ with D_A^{eff} if the latter is sufficiently large (analogous to the behavior of the purity), for choices of subsystems where the incomplete overlap of neighboring eigenstates remains 'visible'. At the same time, all eigenstates trivially satisfy $\delta\mathscr{P}_B = \Delta\mathscr{P}_B = 0$ for $D_A = D$.

We conclude that $\mathscr{P}_B$ is a measure of delocalization of eigenstates, while $\delta\mathscr{P}_B$ and $\Delta\mathscr{P}_B$ are probes of thermalization. Setting $\ell = D$ gives the results discussed in Chap. 8 for RMT-like systems with fully delocalized eigenstates (Sec. 8.2.2.2). For fully localized systems, $\ell = O(1)$ gives $D_A^{\text{eff}} = O(1)$, with $\mathscr{P}_B = O(1)$ and $\delta\mathscr{P}_B = O(1) \lesssim (1 - D_B^{-1})$, automatically implying a lack of thermalization (Sec. 8.2.3). Additionally, the same results hold when the PSFF is defined only over a portion of the spectrum, where the parameters merely become averages over that portion of the spectrum. This suggests that such a filtered [38] PSFF can access equivalent information about the properties of a smaller set of eigenstates of interest.

E.3.2 Subsystem ETH constraints

E.3.2.1 Fully delocalized eigenstates

Subsystem ETH [52] is a hypothesis concerning the behavior of energy eigenstates in an RMT-like system, applying in its original version to fully delocalized eigenstates. It states that the eigenstates are of such a form as to lead to the thermal behavior of all observables on subsystem B, when it is a

physically accessible subsystem - in the sense of diagonal and off-diagonal ETH (e.g. as presented in the reviews [21, 22]). Denoting the eigenstates by $|E\rangle$, there are two statements of the hypothesis: the diagonal statement stating that the reduced density matrix $\rho_B(E) = \mathrm{Tr}_A[|E\rangle\langle E|]$ is close to some smooth density matrix $P_B(E)$ that does not vary rapidly with energy, and the off-diagonal statement requiring the reduced transition operators $q_B(E_1, E_2) = \mathrm{Tr}_A\left[|E_1\rangle\langle E_2|\right]$ with $E_1 \neq E_2$ to be small. We will adapt these statements, in their subsystem dependent version (which doesn't need the restriction $D_B \ll D_A$ to few-body subsystems), for our present context as follows:

$$\mathrm{Tr}_B\left[\left(\rho_B(E) - P_B(E)\right)^2\right] = O(D_A^{-1}), \tag{E.20}$$

$$\mathrm{Tr}_B\left[q_B^2(E_1, E_2)\right] = O(D_A^{-1}), \tag{E.21}$$

where we use the notation $x^2 = xx^\dagger$ for an operator x for simplicity. Eqs. (E.20) and (E.21) should be considered leading order constraints on the order of magnitude of these quantities when $D_A, D_B \gg 1$, as noted in Chap. 8. They are also slightly different in some minor technical details from the main statements of Ref. [52], which we will refer to as the 'original conjecture' in this appendix, and we will now comment on these differences.

We replace the density of states $\mathcal{N}^\#(E)$ with its $O(D)$ scaling behavior in all subsequent discussions, though the original conjecture is stated in terms of $\mathcal{N}^\#(E)$. This is justified by assuming an $O(1)$ spectral width for the D energy levels and that $\mathcal{N}^\#(E)$ is of a comparable order of magnitude throughout the spectrum (consistent with e.g. $t_H = O(D)$ in fully RMT-like systems). As the PSFF involves averages over the entire spectrum, it is only this scaling behavior that is of interest to us rather than $\mathcal{N}^\#(E)$-dependent variations in smaller regions of the spectrum.

The smallness of $(\rho_B - P_B)$ and q_B are enforced above by requiring the trace of their squares $\mathrm{Tr}_B\left[x^2\right]$ (which we will generally call purity) to be $O(D_A^{-1})$. However, the original conjecture is stated

in terms of the trace norm $(1/2)\mathrm{Tr}_B\left[(x^2)^{1/2}\right]$ restricted to be $O(\sqrt{D_B/D_A})$. As Ref. [52] notes, on account of the inequality $\left\{\mathrm{Tr}_B\left[(x^2)^{1/2}\right]\right\}^2 \leq D_B \mathrm{Tr}_B\left[x^2\right]$ the constraints in terms of purity would imply the original conjecture but are also slightly stronger, and it is in fact these stronger constraints that they verify numerically. We use the stronger statement because it is more convenient for our purposes, and also because there appears to be no compelling theoretical reason to rule out such stronger statements in general. For instance, Ref. [52] motivates the diagonal statement of the original conjecture in terms of the trace norm based on analogous canonical typicality [23, 281] constraints for the thermalization of Haar-random superpositions of energy eigenstates derived in Refs. [24, 282]; but in the process of the derivation in the latter, constraints in terms of purity similar to Eq. (E.20) are also seen to hold. We also note that the purity constraints remain $< O(1)$ for $D_B > D_A$, whereas the corresponding constraints on the trace norm (which cannot be greater than 1 for differences of density matrices [50]) are $> O(1)$ and therefore meaningless in this regime. The original conjecture had to restrict the subsystem-dependent form to $D_B < D_A$ (in our notation) for this reason. However, in Sec. 8.3, we find numerical support for the validity of Eqs. (E.20) and (E.21) even for $D_B > D_A$.

Finally, we note that the smooth reduced density matrix $P_B(E)$ is not precisely characterized in Ref. [52] - but it is also unnecessary to be too precise in specifying it as Eq. (E.20) is only an order-of-magnitude constraint. Here, in analogy with Eq. (E.5), we will define $P_B(E)$ to be that part of $\rho_B(E)$ that varies slower than some rate t_s,

$$P_B(E) = \int \frac{\mathrm{d}\tau}{2\pi} e^{iE\tau} \Theta\left(t_s - |\tau|\right) \int \mathrm{d}E' \; e^{-iE'\tau} \rho_B(E'), \tag{E.22}$$

effectively amounting to a weighted average of $\rho_B(E)$ over energy windows of size $\sim t_s^{-1}$. We will assume Eq. (E.20) is satisfied for any choice of t_s larger than some minimum magnitude $\sim t_{\mathrm{ETH}} \ll O(D)$ (intuitively, because the more the smooth part is allowed to fluctuate, the more closely it can

approximate $\rho_B(E)$). Then, if our ensemble is such that $t_\rho \gtrsim t_{\text{ETH}}$, we can choose $t_s = t_\rho$. This allows the identification $P_B(E) = D_B^{-1} + \Delta\rho_B(E)$ in the decomposition $\rho_B(E) = D_B^{-1} + \Delta\rho_B(E) + \delta\rho_B(E)$ of Eq. (8.13). Eq. (E.20) then gives,

$$\text{Tr}_B\left[\delta\rho_B^2(E)\right] = O(D_A^{-1}). \tag{E.23}$$

The constraint $\delta\mathscr{P}_B = O(D_A^{-1})$ then follows directly from here.

To similarly obtain a condition from Eq. (E.21) that applies directly to the PSFF, we note that this equation can be rewritten in terms of reduced density matrices of the complementary subsystem A as

$$\text{Tr}_A\left[\rho_A(E_1)\rho_A(E_2)\right] = O(D_A^{-1}). \tag{E.24}$$

On taking the ensemble average, and using the expansion of $\rho_A(E)$ in terms of its smooth and fluctuating parts, the contribution from the fluctuating part $\delta\rho_A(E)$ to the left hand side vanishes due to the randomization assumption in Sec. 8.2.2. We are then left with $D_A^{-1} + \overline{\text{Tr}_A\left[\Delta\rho_A(E_1)\Delta\rho_A(E_2)\right]}$, in which we can take $E_1 - E_2 \ll t_\rho^{-1}$ (e.g. neighboring levels) so that the second term is approximately $\overline{\text{Tr}_A\left[\Delta\rho_A^2(E_1)\right]}$. From this, we get the smooth purity constraint $\Delta\mathscr{P}_A = O(D_A^{-1})$ on taking the appropriate spectrum averages. In the context of $\Delta\mathscr{P}_B$ (and $\widetilde{\Delta\mathscr{P}}_B$) in Chap. 8, these purities are evaluated in subsystem B rather than A, and the corresponding constraints are therefore consequences of off-diagonal subsystem ETH, Eq. (E.21), applied to subsystem A instead of B.

E.3.2.2 Extension to partially delocalized eigenstates

We begin with a complementary approach to that of the previous subsection, to argue that the purity based expressions of subsystem ETH should generally hold for RMT-like systems with fully delocalized eigenstates. Consider requiring each matrix element of $\rho_B(E)$ to differ from the corresponding matrix element of $P_B(E)$ only by a small amount $O(\sqrt{D_A}/D)$, as a stronger diagonal statement that im-

plies Eq. (E.20) (a weaker, D_A-independent version of such a statement is also considered in Ref. [52]). To justify this constraint, we consider the following situation. Let $|E_1\rangle$ and $|E_2\rangle$ be two 'typical' nearby eigenstates that are completely delocalized over the D basis vectors (in some 'physical' product basis of subsystems A and B) with random (real or complex) phases. Their density operators $\rho(E_1) = |E_1\rangle\langle E_1|$ and $\rho(E_2) = |E_2\rangle\langle E_2|$ have matrix elements of the schematic form

$$\rho_{ab}(E) \sim O(D^{-1}) e^{i\phi_a - i\phi_b}. \tag{E.25}$$

The difference $\rho(E_1) - \rho(E_2)$, after a partial trace over A, can be taken to represent the fluctuations of $\rho_B(E)$ around $P_B(E)$. Given our above assumptions on the eigenstates, the matrix elements of $\rho(E_1) - \rho(E_2)$ are typically $\sim O(D^{-1})$ in magnitude with random signs or phases (i.e. with zero 2-point correlation, which crucially requires even large-scale non-uniformities in the magnitudes to agree up to random fluctuations). The sum of D_A such matrix elements in the partial trace over subsystem A then has magnitude $O(\sqrt{D_A}/D)$, justifying the above constraint. Similarly, the operator $q(E_1, E_2) = |E_1\rangle\langle E_2|$ for such eigenstates has $O(D^{-1})$ matrix elements with random phases, giving $O(\sqrt{D_A}/D)$ matrix elements after the partial trace and therefore the off-diagonal statement Eq. (E.21). Such a picture of random energy projector matrix elements of comparable magnitudes is reminiscent of Berry's conjecture for RMT-like wavefunctions [283] (as well as other related statements e.g. Refs. [18, 264–266]), which has been interpreted as the origin of eigenstate thermalization in RMT-like systems [18, 19].

Using an analogous argument for eigenstates that are not necessarily delocalized over all D basis vectors, we can clearly highlight the difference between delocalization and thermalization, and the distinct information contained in the overall purities as opposed to the smooth and fluctuating parts of the reduced density matrices. For this purpose, consider an eigenstate $|E_1\rangle$ that is randomly (but not

necessarily uniformly) distributed only over a set of $\sim \ell \leq D$ 'physical' basis vectors, with negligible support outside this set. Its density matrix $\rho(E_1)$ then has an $\ell \times \ell$ block (after suitably permuting rows and columns) of non-vanishing elements each of typical magnitude $O(\ell^{-1})$, and all elements outside this block may be taken to vanish. As always, all the diagonal elements are strictly non-negative and add to 1, while the independent off-diagonal elements could have arbitrary signs or phases (which are typically random). Thus, we have the schematic form,

$$\rho_{ab}(E_1) \sim \left[O(\ell^{-1})e^{i\phi_a - i\phi_b}\right]\overline{\Theta}(1 \leq \{p_a, p_b\} \lesssim \ell), \tag{E.26}$$

where $\overline{\Theta}(x) = 1$ if x is true and 0 otherwise, and p_k denotes the index corresponding to k after a permutation p of rows/columns.

The behavior of $\rho(E_1)$ under a partial trace depends on the choice of the subsystem A. We will choose subsystems which can be traced out by factorizing the chosen basis (which means the basis states are pure states within the subsystem). This identifies a class of subsystems which are sensitive to the specific extent of delocalization ℓ of eigenstates; in a more general basis in the Hilbert space, the eigenstates may appear delocalized by an arbitrary extent, including fully localized in the energy eigenbasis and generically fully delocalized ($\ell = D$) in a Haar random basis according to canonical typicality [23, 24]. An equivalent, more physically motivated viewpoint is that the extent of delocalization of eigenstates ℓ should be determined by their minimum such delocalization in bases comprised of nearly pure states (e.g. a product basis) in most physically accessible subsystems - so that a small subset of eigenstates may be treated as if they each have ℓ independent random components (neglecting the global constraint of orthonormality) under a (sufficiently small) partial trace.

For convenience, we first consider the case where the eigenstate looks fully delocalized in subsystem A within its support on the physical basis - in other words, the partial trace over A does not mix

the zero and nonzero elements of $\rho(E_1)$. In this appendix, we will call such a subsystem A an *unbiased* subsystem (from the point of view of the eigenstate of interest). Then $\rho_B(E_1)$ has an $\sim (\ell/D_A)\times(\ell/D_A)$ non-vanishing block with non-negative diagonal elements of magnitude $O(D_A/\ell)$, and off-diagonal elements of typical magnitude $O(\sqrt{D_A}/\ell)$ in the case of an eigenstate with random phases (as long as the partial trace combines several basis vectors where the eigenstate has comparable magnitudes). Now, we can evaluate the purity $\mathrm{Tr}_B\left[\rho_B^2(E_1)\right]$, which sees a net contribution of $O(D_A/\ell)$ from the diagonal elements and $O(D_A^{-1})$ from the off-diagonal elements. Additionally, normalization requires that the diagonal elements must add up to 1, therefore the sum of their squares is greater than or equal to $\sim \ell/D_A$ - the inverse of the number of diagonal elements. Their contribution to the purity can then be written in a more descriptive form as $[(D_A/\ell)+O(D_A/\ell)]$, giving

$$\mathrm{Tr}_B\left[\rho_B^2(E_1)\right] = (D_A/\ell)+O(D_A/\ell)+O(D_A^{-1}). \tag{E.27}$$

Thus, we can extract information about the extent of delocalization, ℓ, by looking at the subsystem size dependence of the purity. We note that the purity can also be written as $\mathrm{Tr}_A\left[\rho_A^2(E_1)\right]$ from the viewpoint of subsystem A giving an additional lower bound of D_A^{-1}, which is mostly contained in the $O(D_A^{-1})$ term for $D_A \gg 1$ (as the diagonal contribution to purity from $\rho_A(E_1)$ is primarily due to contributions from the off-diagonal elements of $\rho_B(E_1)$).

A nearby eigenstate $|E_2\rangle$ that is also distributed only across ℓ basis vectors (but not necessarily the same ones or in the same way as $|E_1\rangle$) again shows a subsystem purity of the form of Eq. (E.27). The two eigenstates thermalize if their reduced density matrices do not differ significantly, in small enough subsystems that trace out a lot of the independent eigenstate components. This would be the case if these two eigenstates are distributed across roughly the same ℓ basis vectors in a largely similar manner (up to random fluctuations). From this point of view, subsystem ETH is a qualitative

identification of the thermalization of a set of otherwise random-looking eigenstates with the extent of their 'overlap' within subsystems, rather than merely with entanglement as represented by their individual purities (the latter being the canonical typicality approach that is only sufficient for fully, uniformly delocalized random eigenstates).

We now consider two illustrative extreme cases of fully overlapping (thermal) and fully non-overlapping (non-thermal) eigenstates. In both cases, we will be interested in $\mathrm{Tr}_B\left[(\rho_B(E_1)-\rho_B(E_2))^2\right]$ as a representative of the size of the fluctuating part $[\rho_B(E)-P_B(E)]$ of reduced energy eigenstates in subsystem B, as well as the (real-valued) overlap $\mathrm{Tr}_B\left[\rho_B(E_1)\rho_B(E_2)\right]$ which is equal to the norm of off-diagonal operators $\mathrm{Tr}_A\left[q_B(E_1,E_2)q_B(E_2,E_1)\right]$ in subsystem A. These are complementary quantities, being related to the subsystem purities of the individual eigenstates via

$$\begin{aligned}&\mathrm{Tr}_B\left[(\rho_B(E_1)-\rho_B(E_2))^2\right]+2\mathrm{Tr}_B\left[\rho_B(E_1)\rho_B(E_2)\right]\\&\quad=\mathrm{Tr}_B\left[\rho_B^2(E_1)\right]+\mathrm{Tr}_B\left[\rho_B^2(E_2)\right].\end{aligned}\tag{E.28}$$

This relation quantifies the identification of thermalization with overlap.

- *Thermal eigenstates*: If $|E_1\rangle$ and $|E_2\rangle$ are distributed in a similar manner across the same basis vectors, then again has $(\rho(E_1)-\rho(E_2))$ an $\ell\times\ell$ block structure, with random $O(\ell^{-1})$ off-diagonal elements within the block. However, the diagonal elements, being differences of random $O(\ell^{-1})$ non-negative numbers, also have at most $O(\ell^{-1})$ magnitudes with random signs (if large scale non-uniformities match), and largely cancel each other out in a partial trace. After the partial trace, all matrix elements of $(\rho_B(E_1)-\rho_B(E_2))$ in an $\sim(\ell/D_A)\times(\ell/D_A)$ non-vanishing block are therefore only $O(\sqrt{D_A}/\ell)$ in magnitude, and we have

$$\mathrm{Tr}_B\left[(\rho_B(E_1)-\rho_B(E_2))^2\right]=O(D_A^{-1}),\tag{E.29}$$

consistent with diagonal ETH Eq. (E.20) in subsystem B. For the overlap $\mathrm{Tr}_B\left[\rho_B(E_1)\rho_B(E_2)\right]$, the positivity and normalization of the diagonal matrix elements of each reduced density matrix ensures that their contribution is of the form $[(D_A/\ell)+O(D_A/\ell)]$. The products of the off-diagonal matrix elements add up with random phases, leading to a negligible $O(\ell^{-1})$ contribution. We therefore have

$$\mathrm{Tr}_B\left[\rho_B(E_1)\rho_B(E_2)\right] = (D_A/\ell) + O(D_A/\ell), \tag{E.30}$$

which is the analog of off-diagonal ETH, Eq. (E.21), for subsystem A.

- *Non-thermal eigenstates*: In the non-thermal case, $|E_1\rangle$ and $|E_2\rangle$ are distributed in completely different ways, and the diagonal elements of $(\rho(E_1)-\rho(E_2))$ do not have completely random signs among elements with comparable magnitudes. Consequently, there is no longer a significant cancellation of the diagonal elements in a partial trace for a general choice of A. The fluctuating part $\mathrm{Tr}_B\left[(\rho_B(E_1)-\rho_B(E_2))^2\right]$ is then typically much larger than $O(D_A^{-1})$ with some $O(D_A/\ell)$ contribution, and the overlap is correspondingly smaller. In the extreme case of the two eigenstates being distributed across completely different basis vectors, $(\rho(E_1)-\rho(E_2))$ has two different $\ell\times\ell$ blocks, and the reduced difference in subsystem B also has the structure of two independent blocks. We then obtain behavior analogous to the subsystem purities,

$$\begin{aligned}&\mathrm{Tr}_B\left[(\rho_B(E_1)-\rho_B(E_2))^2\right]\\ &\sim 2\left[(D_A/\ell)+O(D_A/\ell)+O(D_A^{-1})\right],\end{aligned} \tag{E.31}$$

while the overlap for this case vanishes entirely,

$$\mathrm{Tr}_B\left[\rho_B(E_1)\rho_B(E_2)\right] = 0. \tag{E.32}$$

We note that these trends hold only for $D_A < \ell$, due to the assumption on subsystem A. The reduced energy eigenstates in subsystem B are pure basis states when $D_A = \ell$, and behave accordingly on a further partial trace.

The fluctuations in reduced energy eigenstates and their overlaps therefore contain information about eigenstate thermalization that is not visible to the purity alone, which is merely an indicator of eigenstate delocalization. We also see that, at least for 'typical' eigenstates, diagonal subsystem ETH should be understood (in a coarse, order of magnitude sense) as a lower bound relation, while off-diagonal subsystem ETH is a complementary upper bound relation, related through Eq. (E.28) to each other and the subsystem purities. In place of Eqs. (E.20), and (E.21), we can therefore write the more general relations for partially delocalized eigenstates,

$$\mathrm{Tr}_B\left[\left(\rho_B(E) - P_B(E)\right)^2\right] \gtrsim O(D_A^{-1}), \tag{E.33}$$

$$\mathrm{Tr}_A\left[q_A^2(E_1, E_2)\right] \lesssim (D_A/\ell) + O(D_A/\ell), \tag{E.34}$$

when A is an unbiased subsystem, with $D_A \leq \ell$. Both bounds are saturated by thermal eigenstates.

For greater completeness of the present discussion, we should account for a more typical choice of subsystem A - one that would mix the zero and non-zero elements of these eigenstate reduced density matrices on performing the partial trace over A. We will consider such a typical subsystem to have an effective dimension $D_A^{\mathrm{eff}} \leq D_A$, corresponding to the typical number of non-zero density matrix elements added together in the partial trace. This can be thought of as a generalization of the notion of effective dimension, discussed for the case of infinite dimensional Hilbert spaces in Ref. [52]. We ignore the more complicated case where the number of matrix elements added together is not approximately uniform for all nonzero matrix elements (and therefore, no effective subsystem dimension exists), with the belief that it would not significantly alter our qualitative conclusions. When

the effective dimension does exist, all the above conclusions hold for any system but with D_A replaced by the smaller quantity D_A^{eff}. As an aside, Eqs. (E.33) and (E.34) continue to hold even without this replacement, but are then not necessarily saturated by thermal eigenstates unless A is an unbiased subsystem.

As a simple example, if subsystem A is unbiased with respect to a set of eigenstates of interest, then its complementary subsystem B has effective dimension $D_B^{\text{eff}} = \ell / D_A$ (note that B is not unbiased). Using this, we can finally write off-diagonal ETH for subsystem B and diagonal ETH for subsystem A as follows,

$$\text{Tr}_B \left[q_B^2(E_1, E_2) \right] \lesssim D_A^{-1} + O(D_A^{-1}), \tag{E.35}$$

$$\text{Tr}_A \left[\left(\rho_A(E) - P_A(E) \right)^2 \right] \gtrsim O(D_A/\ell). \tag{E.36}$$

More generally, expressing Eqs. (E.33), (E.34) in terms of D_A^{eff} gives the constraints discussed in App. E.3.1.

E.4 Additional numerical results for Ising Hamiltonian dynamics

In Chap. 8, we considered the Ising Hamiltonian in Eq. (8.25) as an example of local many-body models. In this Appendix, we provide some supporting data which were used in the main section. We first begin with analyzing the interesting set of parameters for which we observe RMT-like and localized regimes in the Hamiltonian model. In Sec. 8.2, we derived the orders for the purity and overlaps of the reduced density matrices on the basis of ETH, and presented them numerically in Sec. 8.3. Here, we provide some additional information on the numerics used to extract the orders for the Hamiltonian model. In the last subsection, we numerically cross-check the shift in the PSFF data with the shift $\delta \mathscr{P}_B$ calculated using Eq. (8.16), where in the latter we directly use the reduced densities of eigenstates.

E.4.1 RMT-like and localized regimes in Ising Hamiltonian

We explain our choice of parameters in the Ising Hamiltonian Eq. (8.25). The Hamiltonian contains ZZ interactions with strength J and the range of interactions is given by α. It has a transverse field with strength J and a longitudinal local random disordered field with strength W. Our interests lie in the parameters such that the Hamiltonian dynamics is either in the RMT-like regime or in the localized regime. For this purpose, we analyze the energy level statistics, using the *adjacent energy gap ratio*. From the sorted energy eigenvalues $E_1 < E_2 < \cdots < E_D$, we compute the energy gaps

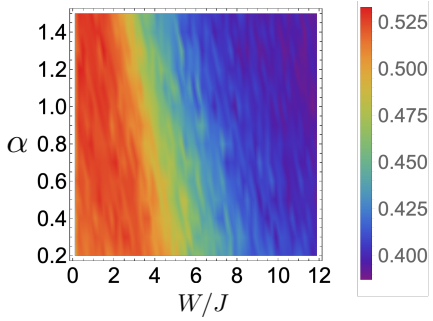


Figure E.1: *Density plot for the mean adjacent gap ratio* $\langle r_m \rangle$. The behavior of the mean adjacent gap ratio $\langle r_m \rangle$ as a function of disorder strength W/J and range of interactions α is presented. We notice the presence of RMT-like ($\langle r_m \rangle \sim 0.53$) and localized ($\langle r_m \rangle \sim 0.39$) phases for a wide range of α. We have worked with $\alpha = 1.2$, $W = J$ (RMT-like), and $W = 10J$ (localized).

$\Delta E_m = E_{m+1} - E_m$. Then we find the adjacent energy gap ratio

$$r_m = \frac{\min(\Delta E_m, \Delta E_{m+1})}{\max(\Delta E_m, \Delta E_{m+1})} \,. \tag{E.37}$$

Integrable systems are characterized by a mean ratio of $\langle r_m \rangle \approx 0.39$ whereas the RMT-like systems with time-reversal symmetry, showing GOE Wigner-Dyson energy level statistics have a mean $\langle r_m \rangle \approx 0.53$. We use this mean value of r_m to choose the parameters for the RMT-like and localized regime in our Hamiltonian model. In a density plot of the mean $\langle r_m \rangle$ as a function of W/J and α in Fig. E.1, we notice that the RMT-like and localized regimes exist for both the short ($\alpha > 1$) and the long ($\alpha < 1$)

range of interactions. In this work, to discuss the two phases, we have chosen the parameters to be $\alpha = 1.2$, $W/J = 1$ (RMT-like) and $W/J = 10$ (localized).

E.4.2 Orders of magnitude of $\Delta\mathscr{P}_B$ and $\delta\mathscr{P}_B$

The subsystem ETH specifies the orders of magnitude for $\Delta\mathscr{P}_B$ and $\delta\mathscr{P}_B$ to be $O(1/D_B)$ and $O(1/D_A)$ respectively for the RMT-like models. For the localized models, which are known to not satisfy ETH, we concluded that the shift coefficient, $\delta\mathscr{P}_B \gg O(1/D_A)$. We note that these orders for the RMT-like regime, expressed in terms of $\mathscr{P}_B$ and $\mathscr{Q}_B$ as deviations from the RMT prediction, amount to,

$$D_B\Delta\mathscr{P}_B = D_B\mathscr{Q}_B - 1 \approx O(1)$$
$$\text{and} \qquad D_A\delta\mathscr{P}_B - 1 = D_A(\mathscr{P}_B - \mathscr{Q}_B) - 1 \approx O(1). \tag{E.38}$$

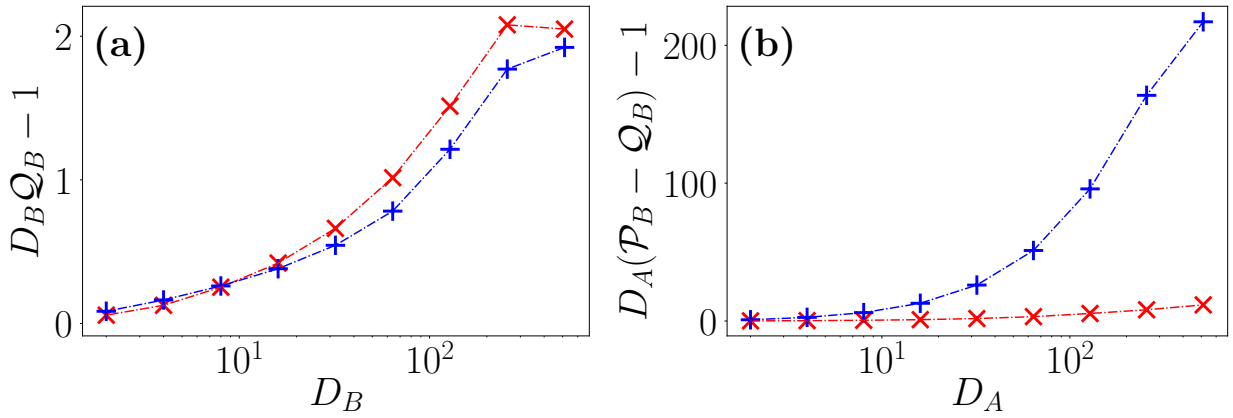


Figure E.2: *Validating the orders.* We test the validity of Eq. (E.38) for the RMT-like and localized regimes of the Hamiltonian. As expected, the RMT-like regime, in red, satisfies the predicted order and the localized regime, in blue, violates it (in the right plot for $\mathscr{P}_B - \mathscr{Q}_B$).

In Fig. E.2, we plot these quantities for the Hamiltonian model, Eq. (8.25), for a total of $N = 10$ qubits. We find that the RMT-like regime $W = J$ (in red) satisfies the ETH results whereas for the

localized regime $W = 10J$ (in blue), the shift coefficient $\mathscr{P}_B - \mathscr{Q}_B \gg 1/D_A$, as predicted in the Sec. 8.2.

E.4.3 Comparison of the PSFF shift and $\delta\mathscr{P}_B$

Here, we numerically verify the prediction of Eq. (8.15) for the constant late time shift of the PSFF in the RMT-like regime - namely, that the shift is given by $\delta\mathscr{P}_B/D_A$ in the ramp region. For this purpose, we subtract the full SFF from the PSFF at some time t_0 in the linear ramp region, satisfying $t_{\text{Th}}, t_\rho \ll t_0 \ll t_H$, which gives

$$K_A(t_0) - K(t_0) = \frac{\delta\mathscr{P}_B}{D_A} + t_0\left(\frac{\gamma}{\beta_{\text{RMT}}\pi D^2} D_B \widetilde{\Delta\mathscr{P}}_B\right). \tag{E.39}$$

This difference has two contributions - the first term is the additive shift which we are presently interested in, but the second term is due to the differing slopes of the linear ramp, from the excess purity of the smooth part of the reduced density matrix. We will now argue that it is reasonable to take

$$K_A(t_0) - K(t_0) \approx \frac{\delta\mathscr{P}_B}{D_A}, \tag{E.40}$$

for our purposes. In Eq. (E.39), by subsystem ETH, the former is $O(D_A^{-2})$ while the latter is $\sim O(D^{-1}(t_0/t_H))$ (taking $\gamma \sim O(1)$, consistent with $t_H \sim O(D)$). The second term is therefore negligible if $t_0/t_H \ll O(D_B/D_A)$. This is immediately satisfied for any t_0 in the linear ramp region if $D_A < D_B$; conversely, for a given choice of t_0, D_A can be as large as $\sim \sqrt{D t_H/t_0}$ while maintaining the validity of Eq. (E.40). As $t_0 \ll t_H$ in general, we expect Eq. (E.40) to be a reasonable approximation for a range of values of $D_A > \sqrt{D}$ as well. A minor additional effect that improves this approximation is that for large D_A, the coefficient $D_B\widetilde{\Delta\mathscr{P}}_B$ of the second term would be small, though still $O(1)$, from Fig. E.2 (as $D_B\widetilde{\Delta\mathscr{P}}_B \sim [D_B Q_B - 1]$).

On the basis of Eq. (E.40) and the relation $\delta\mathscr{P}_B = \mathscr{P}_B - \mathscr{Q}_B$ from Eq. (8.16), we compare

$D_A(K_A(t_0) - K(t_0))$ for some suitably chosen t_0 to $\mathscr{P}_B - \mathscr{Q}_B$ in Fig. E.3 and observe good agreement, especially for smaller D_A as expected. We note that this agreement is much closer than, for instance, the difference between $\mathscr{P}_B - \mathscr{Q}_B$ for the Hamiltonian system and the corresponding RMT prediction in Fig. 8.5(d), which is considerable evidence that the origin of the shift is indeed the randomization of the fluctuating part of the reduced energy eigenstates, as discussed in Sec. 8.2.2.

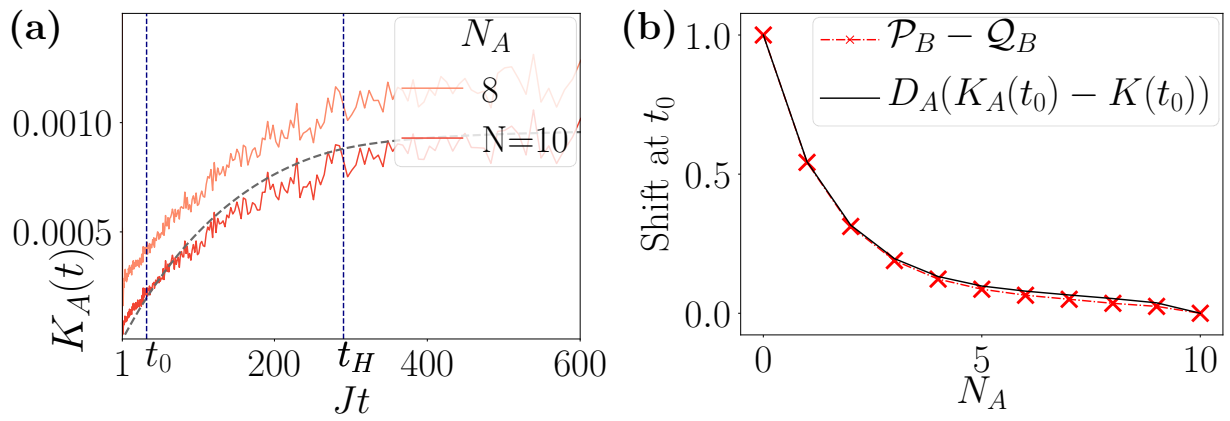


Figure E.3: *Results on the shift.* (a) Linear-linear plot of the PSFF for $N_A = 8$ and SFF, reflecting the choice of comparison time t_0 in the ramp region, with $Jt_0 = 25$; the Heisenberg time t_H (marking the onset of the plateau) and the corresponding GOE SFF (dashed curve) are also shown. (b) Comparison of the scaled PSFF shift $D_A(K_A(t_0) - K(t_0))$ with the predicted value $\delta\mathscr{P}_B = \mathscr{P}_B - \mathscr{Q}_B$ for different subsystem sizes. The total system size is $N = 10$ in both plots.

Appendix F: Calculations for Partial SFFs in open systems

The generalized SFF for a quantum operation with Kraus operators $\hat{\mathscr{K}}_r(t)$ as in Eq. (9.1) can be rewritten as:

$$K(t) = \frac{1}{D^2}\sum_{k,j}\mathrm{Tr}\left[|j\rangle\langle k|\left(\sum_r \hat{\mathscr{K}}_r(t)|k\rangle\langle j|\hat{\mathscr{K}}_r^\dagger(t)\right)\right]. \tag{F.1}$$

Here the sum is over some set of orthonormal basis states. The corresponding PSFFs are,

$$K_A(t) \equiv \frac{1}{DD_A}\sum_r \mathrm{Tr}_B\left[\mathrm{Tr}_A(\hat{\mathscr{K}}_r(t))\,\mathrm{Tr}_A(\hat{\mathscr{K}}_r^\dagger(t))\right] \tag{F.2}$$

$$= \frac{1}{DD_A}\sum_{k_A,k_B;j_A,j_B}\mathrm{Tr}\Big[|j_A,k_B\rangle\langle k_A,k_B| \left(\sum_r \hat{\mathscr{K}}_r(t)|k_A,j_B\rangle\langle j_A,k_B|\hat{\mathscr{K}}_r^\dagger(t)\right)\Big], \tag{F.3}$$

where $|k_A,k_B\rangle = |k_A\rangle\otimes|k_B\rangle$ etc. are product basis states in the subsystems A and B. What's worth noting here is that the "initial" ($|k_A,j_B\rangle\langle j_A,k_B|$) and "final" ($|j_A,k_B\rangle\langle k_A,k_B|$) states are diagonal (effectively projectors) on to the B basis states while allowing both diagonal and off-diagonal contributions in the A subspace; in contrast, each "branch" of time evolution connects basis states with the same A factor but possibly different B factors.

Given that we are working with a Floquet system that is intended to be unitary, we will assume that unitary evolution and decoherence act as alternate steps in a sequence, with the decoherence being a collective dephasing of off-diagonal terms in the computational basis. This should be exact

for the localized case where everything commutes, being diagonal in the computational basis. In the non-integrable case, this assumption makes things tractable and we expect it to capture the main qualitative features of the PSFFs.

To model dephasing, given the Floquet unitary $\hat{U}$, we assume that each repeated step of time evolution takes the following form in the computational basis [1]:

$$\sum_r \hat{\mathscr{K}_r}(1)\left(\sum_{k,j}\rho_{kj}|k\rangle\langle j|\right)\hat{\mathscr{K}_r^\dagger}(1) = \sum_k \rho_{kk}\hat{U}|k\rangle\langle k|\hat{U}^\dagger + e^{-\lambda}\sum_{k\neq j} e^{-\mu_{kj}}\rho_{kj}\hat{U}|k\rangle\langle j|\hat{U}^\dagger. \tag{F.4}$$

Here, λ represents the collective dephasing rate while μ_{kj} represents a state-dependent dephasing rate (with $\mu_{kk} = 0$) e.g. to account for a separate dephasing mechanism in each qubit in addition to collective dephasing. Note that we've assumed that dephasing acts first, followed by the unitary evolution, in our "Trotterized" model.

F.1 Localized PSFFs

The localized time evolution operator is completely diagonal in the computational basis, including the unitary and dephasing components; the assumption of Trotterized decoherence becomes exact in this case, up to mild deviations due to the weak interaction steps present in the experiment.

The localized Hamiltonian that generates each step of unitary evolution $\hat{U} = e^{-i\hat{H}}$ is (with r labeling each qubit):

$$\hat{H} = \sum_r b_r \hat{\sigma}_{z,r}. \tag{F.5}$$

(In the experiment, there are two localized steps of the same form, but they can be readily added

[1] Working directly with the Kraus operators is easier for the present analysis than using the Lindblad equation, as we're directly considering Trotterized steps of time evolution.

together to give a single effective Hamiltonian; we are also neglecting the weak interaction steps). The PSFFs can be evaluated directly by carrying out the trace in the eigenbasis of the $\hat{\sigma}_{z,r}$, while accounting for dephasing as in Eq. (F.4), with a single dephasing rate μ_r for a given qubit. The result is:

$$K_A(t) = 2^{-N_A}\left\{1 - e^{-\lambda t} + e^{-\lambda t}\prod_{r\in A}\left[1 + e^{-\mu_r t}\cos(2b_r t)\right]\right\}. \tag{F.6}$$

At $t \to \infty$, we get $K_A(t) = 2^{-N_A}$, which is the ensemble-averaged localized plateau value. This average plateau value is preserved as dephasing does not affect the diagonal elements of the density matrix in the computational basis. This marks the key difference between dephasing and global depolarization, as the latter would also make this plateau value decay to the maximally mixed value $K_A(t) \sim 1/D_A^2$.

We can further carry out an ensemble average of the b_i over the distribution relevant for the experiment: the convolution of two uniform box distributions supported on $[0,1]$ (corresponding to a triangle of width 4). Then, we get

$$\langle K_A(t)\rangle = 2^{-N_A}\left\{1 - e^{-\lambda t} + e^{-\lambda t}\prod_{r\in A}\left[1 + \frac{\cos(4t)\sin^2(2t)}{4t^2}e^{-\mu_r t}\right]\right\}, \tag{F.7}$$

$$= 2^{-N_A}\left\{1 - e^{-\lambda t} + e^{-\lambda t}\prod_{r\in A}\langle\kappa_r(t)\rangle\right\}. \tag{F.8}$$

where we have introduced the notation

$$\kappa_r(t) \equiv [1 + e^{-\mu_r t}\cos(b_r t)]. \tag{F.9}$$

One of our experimental goals is to compare ensemble-averaged and single-disorder SFFs, and show its non-self-averaging property. However, dephasing works against this goal if it is too large, by suppressing these fluctuations. We want $e^{-\lambda t}, e^{-\mu t}$ to be larger than the experimental sensitivity for at least some integers t to have detectable differences between the two cases.

We can find the variance of the SFF for the case with dephasing. For the ensemble itself, we have

$$2^{2N_A}\langle K_A^2(t)\rangle = (1-e^{-\lambda t})^2 + 2e^{-\lambda t}(1-e^{-\lambda t})\prod_{r\in A}\langle \kappa_r(t)\rangle + e^{-2\lambda t}\prod_{r\in A}\left[2\langle\kappa_r(t)\rangle + \frac{\langle\kappa_r(2t)\rangle + e^{-2\mu_r t} - 3}{2}\right]. \tag{F.10}$$

As $\kappa_r(t) \to 1$ rather fast in t, we can get a simpler expression for the plateau region by making this substitution:

$$2^{2N_A}\langle K_A^2(t)\rangle \xrightarrow[\text{plateau}]{t\to\infty} 1 - e^{-2\lambda t} + e^{-2\lambda t}\prod_{r\in A}\left(1+\frac{1}{2}e^{-2\mu_r t}\right). \tag{F.11}$$

This means that the fluctuations of the SFF would show an exponential decay. For $N_E \gg 1$ random ensemble realizations, the expected variance of the SFF is given by $\sigma^2_{K_A(t),N_E} = (\langle K_A^2(t)\rangle - \langle K_A(t)\rangle^2)/N_E$, resulting in

$$\frac{\sigma_{K_A(t),N_E}}{\langle K_A(t)\rangle} = \frac{e^{-\lambda t}}{\sqrt{N_E}}\sqrt{\left[\prod_{r\in A}\left(1+\frac{1}{2}e^{-2\mu_r t}\right)\right] - 1}. \tag{F.12}$$

For weak or no dephasing, this is $\sigma_{K_A(t),N_E} \approx 2^{-N_A}\sqrt{(3/2)^{N_A}-1}/\sqrt{N_E}$, which is much larger than what one would expect from Poisson statistics. This large variance is because the localized Hamiltonian has only N free parameters (the b_r), while genuinely uncorrelated Poisson statistics requires 2^{N_A} free parameters – one for each eigenvalue (the effect of this on other quantities has been noted, for instance, in Ref. [148]). In this sense, the measurement of this variance of SFF fluctuations can yield information on how unconstrained the ensemble is.

For strong dephasing or $t \gg \lambda^{-1}, \mu_r^{-1}$, we have $\sigma_{K_A(t),N_E}/K_A(t) \to 0$, giving just a plateau without any oscillations even for a single disorder realization. Dephasing has a similar effect to ensemble averaging (which is not surprising if one notes that ensemble averaging correlates the evolution of the system with external information about the nature of disorder, and can loosely be thought of as N_E

independent "overcomplete" Kraus unitaries); see also Ref. [170].

F.2 Periodic orbit contributions and dephasing in RMT-like (P)SFFs

For an RMT-like system e.g. with alternating, noncommuting Floquet steps, the assumption of Trotterized decoherence may well be an impractical toy model. Dephasing is likely to be interspersed among the steps (in the experiment, primarily in the interaction steps); any "effective" decoherence operator will be dressed by commutation with the unitary evolution steps, and is unlikely to remain a simple dephasing channel in the original basis. Nevertheless, the "zeroth-order" analysis of PSFFs under Trotterized dephasing in this section sets up some intuition as well as the basic elements of any more careful calculation involving PSFFs and/or decoherence.

The simplest or most economical model of unitary evolution to work with is a caricature of CUE: a random $D \times D$ matrix $\hat{V}$ whose matrix elements V_{mn} are distributed according to independent Gaussians of variance D^{-1} in the magnitude and uniformly distributed phases:

$$P(V_{mn})\, \mathrm{d}|V_{mn}|\, \mathrm{d}(\arg V_{mn}) = \left(\frac{D}{2\pi}\right)^{\frac{1}{2}} e^{-\frac{D}{2}|V_{mn}|^2}\, \mathrm{d}|V_{mn}| \cdot \frac{1}{2\pi}\, \mathrm{d}(\arg V_{mn}). \tag{F.13}$$

This is *formally* the complex Ginibre ensemble [37]; however, we emphasize that we are using the Ginibre ensemble here not in place of a (non-Hermitian) Hamiltonian $\hat{H}$ as is conventional, but the unitary $e^{-i\hat{H}\tau}$. The matrix $\hat{V}$ only satisfies the unitarity constraint $\hat{U}^\dagger \hat{U} = \hat{\mathbb{1}}$ in an ensemble-averaged sense $\langle \hat{V}^\dagger \hat{V} \rangle = \hat{\mathbb{1}}$, and higher powers $\hat{V}^k$ eventually become quite different from unitary. On the other hand, the V_{mn} satisfy Wick's theorem, which is the leading order behavior of most random matrix ensembles responsible for the ramp. Calculations of (P)SFFs with $\hat{V}$ capture the universal leading-order "periodic orbit" contributions in a very simple and general way as a consequence of Wick's theorem, as matrix elements of any typical random matrix can be approximated as independent

Gaussians (in the spirit of the treatment of ideal gases in classical statistical mechanics); in spirit, this is equivalent to Berry's "diagonal approximation". Corrections to Wick's theorem for individual ensembles (including CUE) are non-universal, subleading, and are significant only over the Heisenberg time scale (after the onset of the linear ramp). In the specific case of the exact Haar random CUE ensemble, they can be precisely obtained using Weingarten calculus, but will generally be different for other systems in the CUE class like local random Floquet circuits.

In short, the argument here is more general and faster to calculate with for any system that looks "sufficiently close" to CUE (including nonunitary ones[2]), at the expense of rigor, and can also be easily adapted to e.g. COE-like ensembles by considering symmetric random matrices with complex-valued entries. At the same time, it should reproduce the basic structure of diagrams leading to statistically universal features in more physical models, such as Refs. [35, 41, 42]. In this context, we note that Ref. [276] computes the full SFF for a local random circuit model where unitary steps act in alternation with local dissipative channels, and finds results qualitatively comparable to this section with additional model-specific effects.

F.2.1 The linear ramp in the SFF

F.2.1.1 No dephasing

The ensemble averaged SFF for successive actions of $\hat{V}$, without any is given by

$$D^2\langle K(t)\rangle = \left\langle \mathrm{Tr}(\hat{V}^t)\,\mathrm{Tr}(\hat{V}^{t\dagger})\right\rangle \tag{F.14}$$

[2]For instance, even the non-unitary $\hat{V}$ matrices from the ideal ensemble of Eq. (F.13) show a linear ramp at early times before deviating considerably in an exponential growth due to emerging nonunitarity — as seen numerically.

To obtain the linear ramp, we should look at the correlations between the matrix elements in each trace:

$$D^2\langle K(t)\rangle = \left\langle \left(\sum_{m_t,\ldots,m_1} V_{m_1 m_t} \ldots V_{m_3 m_2} V_{m_2 m_1} \right) \left(\sum_{n_t,\ldots,n_1} V^*_{n_1 n_t} \ldots V^*_{n_3 n_2} V^*_{n_2 n_1} \right) \right\rangle \tag{F.15}$$

We can think of each set of indices $(m_t,\ldots,m_1)$ as describing a closed trajectory $m_1 \to m_2 \to \ldots \to m_1 \to m_1$ in the space of indices. To leading order in D, non-self-intersecting closed trajectories dominate the sum (i.e. each sum is over D indices, while any self-intersecting trajectories correspond to dropping one of these sums as one of the two intersecting points is no longer free, and should be smaller by a factor of D), and we will focus on these trajectories.

Wick's theorem shows that the only nonvanishing correlations between $V_{m_{j+1}m_j}$ and $V^*_{n_{k+1}n_k}$ occur when $n_k = m_j$ and $n_{k+1} = m_{j+1}$. Nonvanishing contributions to Eq. (F.15) only occur when the closed trajectories of V and V^* are identical; in other words, when $(m_t,\ldots,m_1)$ is a cyclic permutation of $(n_t,\ldots,n_1)$, each of which leads to the same effective contribution:

Relative cyclic permutation symmetry $\mathbb{Z}_t$:

$$(m_t, m_{t-1}, \ldots, m_2, m_1) = (n_{t+c}, n_{t-1+c}, \ldots, n_{2+c}, n_{1+c}), \text{ for any } c \in \mathbb{Z}_t = \{0,1,\ldots,t-1\}. \tag{F.16}$$

It is useful to separate the "diagonal" case $c = 0$ and the off-diagonal case $c \neq 0$ to anticipate dephasing. There is 1 diagonal term and $(t-1)$ off-diagonal terms generated by the cyclic permutation symmetry, so we have (with $=$ replaced by $\approx$ to show that we are neglecting self-intersecting paths)

$$D^2\langle K(t)\rangle \approx \left[1_{(c=0)} + (t-1)_{(c\neq 0)}\right] \sum_{m_t=1}^{D} \cdots \sum_{m_1=1}^{D} \left\langle |V_{m_1 m_t}|^2 \ldots |V_{m_3 m_2}|^2 |V_{m_2 m_1}|^2 \right\rangle \tag{F.17}$$

$$\approx \left[1_{(c=0)} + (t-1)_{(c\neq 0)}\right] D^t \cdot D^{-t} \tag{F.18}$$

$$= \left[1_{(c=0)} + (t-1)_{(c\neq 0)}\right] = t. \tag{F.19}$$

In the second line, we used the fact that each $|V_{ab}|^2$ has expectation value D^{-1} together with the argument that all the m_j are distinct in the leading order contribution. In this picture, the linear ramp essentially originates in a relative cyclic permutation symmetry of two closed paths in the space of indices corresponding to the two traces, which is the basis of periodic orbit arguments for semiclassical dynamical systems and generic Floquet systems, as well as Weingarten calculus based treatments of random quantum circuits; the above approach gives a quick route to this picture without dealing with the specific complications of each of these applications.

F.2.1.2 Effect of dephasing

The effect of dephasing is now easy to incorporate. For simplicity, we set $\mu_r = 0$ so that only global dephasing at the rate λ remains. Using Eq. (F.4) for each Trotterized dephasing step gives

$$V_{m_{j+1},m_j} V^*_{n_{j+1},n_j} \longrightarrow \exp\Big[-\lambda(1-\delta_{m_j,n_j})\Big] V_{m_{j+1},m_j} V^*_{n_{j+1},n_j} \tag{F.20}$$

In the closed path contributions to Eq. (F.17), this means that the $c = 0$ term remains unaffected (because that corresponds to $m_j = n_j$ for all j), while the $c \neq 0$ terms have t factors of $e^{-\lambda}$. We obtain the following leading-order expression for the SFF in the presence of global dephasing:

$$\begin{aligned} D^2\langle K(t)\rangle_\lambda &= 1_{(c=0)} + e^{-\lambda t}(t-1)_{(c\neq 0)} \\ &= 1 + e^{-\lambda t}[D^2 K_{\mathrm{CUE}}(t) - 1]. \end{aligned} \tag{F.21}$$

This gives a quantitative form of the dephasing ramp for the unitary ensemble, which at first sight looks not too different from depolarization (global decay to the maximally mixed state, see e.g. Ref. [6]). For the orthogonal ensemble (i.e. symmetric $\hat{V}$ in this simplified version), the $K(1) = 2$ value stems from the $c = 0$ contribution above, as well as an additional off-diagonal contribution from the time-

reversal symmetry; this off-diagonal contribution is again suppressed by dephasing, and one obtains $D^2K(t) \sim 1 + e^{-\lambda t}(D^2K_{\mathrm{COE}}(t) - 1)$ in that case as well. As the linear-ramp terms do not contribute at $t = 1$, we should have:

$$D^2\langle K(1)\rangle_\lambda = \begin{cases} 1, & \text{for the unitary class, e.g. CUE,} \\ 1 + e^{-\lambda}, & \text{for the orthogonal class, e.g. COE.} \end{cases} \tag{F.22}$$

which is again similar to what one would see in the presence of global depolarization.

However, the key difference between dephasing and depolarization is seen in the presence of conserved quantities, where the time evolution unitary takes a block structure (SFFs of this kind have been studied in the context of, e.g., periodic orbit theory, disordered metals and spin glasses [73, 196, 211, 212, 214]). For M independent blocks of roughly the same size, each block has an effective dimension of D/M. As long as dephasing occurs in the same basis where this block structure (or partial localization) is visible, it does not suppress the $c = 0$ terms of each block (each of which is independent, with a presumed random phase, in the trace over all blocks). Here, we get for the unitary class

$$D^2\langle K(t)\rangle_\lambda = M_{(c=0)} + Me^{-\lambda t}(t-1)_{(c\neq 0)}, \tag{F.23}$$

with a $t \to \infty$ stable value of M, unlike depolarization which would again decay to a value of 1. We again emphasize that this would require the "Trotterized" dephasing operator to be in the block basis rather than the original effect, and may otherwise well be similar to depolarization in all other bases.

F.2.2 The shift and ramp in the PSFF

F.2.2.1 No dephasing

For the PSFF, the structure of trajectories is somewhat different, due to the partial trace. We continue to have closed trajectories representing periodic orbits which generate the rescaled ramp, but there are also contributions from open trajectories — these generate the shift. For time evolution generated by $\hat{V}$, the expression for the PSFF in Eq. (F.3) becomes (using slightly different notation for convenience with indices)

$$DD_A\ \langle K_A(t)\rangle = \left\langle \sum_{a,a',b,b'} (\hat{V}^t)_{ab;ab'} (\hat{V}^{t\dagger})_{a'b';a'b} \right\rangle \tag{F.24}$$

The a indices live in subsystem A, and the b indices live in subsystem B. Expanding out the products into trajectories connected by single actions of $\hat{V}$, as before, we have:

$$\begin{aligned} DD_A\ \langle K_A(t)\rangle = \Bigg\langle \sum_{\mathbf{b}_{t+1},\mathbf{b}_1} \Bigg(\sum_{\mathbf{a},\mathbf{b}} V_{a_1\mathbf{b}_{t+1};a_t b_t} V_{a_t b_t;a_{t-1}b_{t-1}} \cdots V_{a_3 b_3;a_2 b_2} V_{a_2 b_2;a_1 \mathbf{b}_1} \Bigg) \\ \Bigg(\sum_{\overline{\mathbf{a}},\overline{\mathbf{b}}} V^*_{\overline{a}_1\mathbf{b}_{t+1},\overline{a}_t\overline{b}_t} V^*_{\overline{a}_t\overline{b}_t;\overline{a}_{t-1}\overline{b}_{t-1}} \cdots V^*_{\overline{a}_3\overline{b}_3;\overline{a}_2\overline{b}_2} V^*_{\overline{a}_2\overline{b}_2 l\overline{a}_1\mathbf{b}_1} \Bigg) \Bigg\rangle \end{aligned} \tag{F.25}$$

While the a-coordinates of the trajectories close up into subsystem periodic orbits for each of the V and V^* branches, the b-coordinates form open trajectories with a common starting point $b_1, \overline{b}_1 = \mathbf{b}_1$ and ending point $b_{t+1}, \overline{b}_{t+1} = \mathbf{b}_{t+1}$.

The subset of points with $\mathbf{b}_{t+1} = \mathbf{b}_1$ give closed periodic orbits in the full system. The remaining points with $\mathbf{b}_{t+1} \neq \mathbf{b}_1$ give necessarily open trajectories. Applying Wick's theorem in the ensemble average and neglecting self-intersecting trajectories as before, we see that the contribution from open

trajectories is fully given by the "diagonal" terms with the same trajectory in both branches.

Open orbits, diagonal terms:

$$\begin{aligned}(\mathbf{b}_1 \neq \mathbf{b}_t \text{ contribution}) &= \sum_{\mathbf{b}_1=1}^{D_B} \sum_{\mathbf{b}_{t+1}\neq\mathbf{b}_1} \sum_{\mathbf{a},\mathbf{b}} \left\langle |V_{a_1\mathbf{b}_{t+1};a_t b_t}|^2 |V_{a_t b_t;a_{t-1}b_{t-1}}|^2 \dots |V_{a_3 b_3;a_2 b_2}|^2 |V_{a_2 b_2;a_1 \mathbf{b}_1}|^2 \right\rangle \\ &= D_B(D_B-1) D_A^t D_B^{t-1} D^{-t} \\ &= D_B - 1. \end{aligned} \tag{F.26}$$

This t-independent constant contribution is the shift. Being fully comprised of diagonal terms, it sees no suppression from dephasing.

Now, we look at the periodic orbit contribution with $\mathbf{b}_1 = \mathbf{b}_t$. The $\mathbb{Z}_t$ cyclic permutation symmetry of the SFF is broken even by most of the periodic terms in the PSFF — being preserved only if all the $b_j, \overline{b}_j$ are equal (so that we can set each $b_j = \mathbf{b}_1$ as we cycle around). In general, in a set $(b_t, \dots, b_2)$ in which $s \leq t$ of the b_j are equal to $\mathbf{b}_1$, i.e.,

$$b_{q_1} = \dots = b_{q_s} = \mathbf{b}_1, \tag{F.27}$$

the only non-vanishing contributions to the periodic orbit sum, by Wick's theorem, are given by those where the two branches are related by the

Broken relative cyclic permutation symmetry $\{q_1, \dots, q_s\} \subseteq \mathbb{Z}_t$:

$$(a_t b_t, \dots, a_2 b_2, a_1 b_1) = (\overline{a}_{t+c}\overline{b}_{t+c}, \dots, \overline{a}_{2+c}\overline{b}_{2+c}, \overline{a}_{1+c}\overline{b}_{1+c}), \text{ for any } c \in \{q_1, \dots, q_s\} \subseteq \mathbb{Z}_t. \tag{F.28}$$

When summing $(a_t b_t, \dots, a_2 b_2, a_1 \mathbf{b}_1)$ each over 1 to D, the probability that any $0 \leq s \leq t-1$ of the

$b_{j\geq 2}$ are equal to $\mathbf{b}_1 = \mathbf{b}_{t+1}$, while the remaining are not equal, is given by the binomial expression $\binom{t-1}{s} D_A^s (D-D_A)^{t-1-s}$; an additional factor of D should be included to account for the sum over $\mathbf{b}_1, a_1$. Within each of these cases, there is 1 diagonal contribution from having the same periodic orbit in both branches, and s off-diagonal contributions from having the V^* trajectory related to the V trajectory by the broken cyclic permutation symmetry, for a total of $(1+s)$ contributions of the same magnitude. So the periodic orbit contribution is given by (noting that the average over the t squared matrix elements of $\hat{V}$ gives D^{-t}):

Periodic orbits, diagonal and off-diagonal terms :

$$\begin{aligned}
(\mathbf{b}_1 = \mathbf{b}_t \text{ contribution}) &= \left[D \sum_{s=0}^{t-1} (1+s) \binom{t-1}{s} D_A^s (D-D_A)^{t-1-s} \right] D^{-t} \\
&= D^{-(t-1)} \left[\frac{\mathrm{d}}{\mathrm{d}\mathbf{a}} \left\{ \mathbf{a} \sum_{s=0}^{t-1} \binom{t-1}{s} (\mathbf{a} D_A)^s (D-D_A)^{t-1-s} \right\} \right]_{\mathbf{a}=1} \\
&= D^{-(t-1)} \left[\frac{\mathrm{d}}{\mathrm{d}\mathbf{a}} \left\{ \mathbf{a} (D - D_A + \mathbf{a} D_A)^{t-1} \right\} \right]_{\mathbf{a}=1} \\
&= 1_{(c=0)} + D_B^{-1} (t-1)_{(c\neq 0)}. \qquad \text{(F.29)}
\end{aligned}$$

Combining the open orbit and periodic orbit contributions, we have

$$DD_A \langle K_A(t) \rangle = \underbrace{(D_B - 1)_{\text{open orbits}} + 1_{(c=0)}}_{\text{shift}} + \underbrace{D_B^{-1} (t-1)_{(c\neq 0)}}_{\text{rescaled ramp}}. \qquad \text{(F.30)}$$

F.2.2.2 Effect of dephasing

Once again, the simplified form of dephasing considered here should affect only the off-diagonal contributions ($c \neq 0$) to the PSFF, rather than the diagonal ones (open orbits and $c = 0$). Continuing to neglect self-intersecting orbits, the modification due to dephasing is very much similar to that of

the SFF, and we have

$$DD_A \langle K_A(t)\rangle_\lambda = D_B + D_B^{-1} e^{-\lambda t}(t-1). \tag{F.31}$$

The difference between dephasing and depolarization is not quantitatively obvious for this PSFF as well, as the shift in this case is just the leading behavior for fully decohered systems. However, taking the general lesson that the PSFF shift (including any non-RMT/ETH contributions) is likely due to diagonal contributions and the ramp is due to off-diagonal correlations, we expect the general form:

$$DD_A \langle K_A(t)\rangle_\lambda = (\text{shift}) + e^{-\lambda t}(\text{rescaled ramp}), \tag{F.32}$$

in the presence of dephasing. Like in Eq. (F.23), we also expect more nontrivial PSFFs with conserved quantities (in which the coefficients of the shift and ramp should measure the average degree of ETH behavior *within each block* by the arguments of Ref. [6] (see Chap. 8), scaled by the number of blocks as in the SFF) to show an unaltered shift and decaying ramp, in contrast to fully decaying to 1 (for $DD_A K_A(t)$) in the presence of depolarization. We again emphasize that unless decoherence is well approximated by this toy model of Trotterized dephasing in the same basis as the one with deviations from ETH or conserved quantities, dephasing may not look too different from depolarization.

Bibliography

[1] Amit Vikram and Victor Galitski. Dynamical quantum ergodicity from energy level statistics. *Phys. Rev. Res.*, 5(3):033126, 2023. URL: https://doi.org/10.1103/PhysRevResearch.5.033126.

[2] Amit Vikram and Victor Galitski. Exact universal bounds on quantum dynamics and fast scrambling. *Phys. Rev. Lett.*, 132(4):040402, 2024. URL: https://doi.org/10.1103/PhysRevLett.132.040402.

[3] Amit Vikram, Laura Shou, and Victor Galitski. Proof of a universal speed limit on fast scrambling in quantum systems. *arXiv preprint arXiv:2404.15403*, 2024. URL: https://doi.org/10.48550/arXiv.2404.15403.

[4] Yunxiang Liao, Amit Vikram, and Victor Galitski. Many-body level statistics of single-particle quantum chaos. *Phys. Rev. Lett.*, 125(25):250601, 2020. URL: https://doi.org/10.1103/PhysRevLett.125.250601.

[5] Laura Shou, Amit Vikram, and Victor Galitski. Spectral anomalies and broken symmetries in maximally chaotic quantum maps. *arXiv preprint arXiv:2312.14067*, 2023. URL: https://doi.org/10.48550/arXiv.2312.14067.

[6] Lata Kh Joshi, Andreas Elben, Amit Vikram, Benoît Vermersch, Victor Galitski, and Peter Zoller. Probing many-body quantum chaos with quantum simulators. *Phys. Rev. X*, 12(1):011018, 2022. URL: https://doi.org/10.1103/PhysRevX.12.011018.

[7] Abu Musa Patoary, Amit Vikram, Laura Shou, and Victor Galitski. Chaotic roots of the modular multiplication dynamical system in Shor's algorithm. *To appear in Phys. Rev. Res. arXiv preprint arXiv:2306.16446*. URL: https://doi.org/10.48550/arXiv.2306.16446.

[8] (In preparation) Mathematical note on quantum scrambling bounds.

[9] Cornelius Lanczos. *The variational principles of mechanics*. Dover Publications, 1986.

[10] Paul R Halmos. *Lectures on ergodic theory*. Dover Publications.

[11] Yakov G. Sinai. *Introduction to ergodic theory*, volume 18. Princeton University Press.

[12] I. P. Cornfeld, S. V. Fomin, and Yakov G. Sinai. *Ergodic theory*. Springer-Verlag New York. URL: https://doi.org/10.1007/978-1-4615-6927-5.

[13] Roman Frigg, Joseph Berkovitz, and Fred Kronz. The ergodic hierarchy. *The Stanford Encyclopedia of Philosophy, Fall 2020 Edition*. URL: "https://plato.stanford.edu/archives/fall2020/entries/ergodic-hierarchy/".

[14] Giovanni Gallavotti. *Statistical mechanics: A short treatise*. Springer Science & Business Media, 1999. URL: https://doi.org/10.1007/978-3-662-03952-6.

[15] Richard Chace Tolman. *The principles of statistical mechanics*, volume 51. Dover Publications, 2010.

[16] John von Neumann. Proof of the ergodic theorem and the H-theorem in quantum mechanics. translation of: Beweis des ergodensatzes und des H-Theorems in der neuen mechanik. *Eur. Phys. J. H*, 35(2):201–237, 2010. Original (in German): Zeit. für Phys. 57, 30 (1929). URL: https://doi.org/10.1140/epjh/e2010-00007-7.

[17] Sheldon Goldstein, Joel L Lebowitz, Christian Mastrodonato, Roderich Tumulka, and Nino Zanghì. Normal typicality and von Neumann's quantum ergodic theorem. *Proc. Roy. Soc. Lond. A*, 466(2123):3203–3224, 2010. URL: https://doi.org/10.1098/rspa.2009.0635.

[18] J. M. Deutsch. Quantum statistical mechanics in a closed system. *Phys. Rev. A*, 43:2046–2049, 1991. URL: https://link.aps.org/doi/10.1103/PhysRevA.43.2046, doi:10.1103/PhysRevA.43.2046.

[19] Mark Srednicki. Chaos and quantum thermalization. *Phys. Rev. E*, 50:888–901, 1994. URL: https://link.aps.org/doi/10.1103/PhysRevE.50.888, doi:10.1103/PhysRevE.50.888.

[20] Mark Srednicki. The approach to thermal equilibrium in quantized chaotic systems. *J. Phys. A: Math. Gen.*, 32(7):1163–1175, 1999. URL: http://dx.doi.org/10.1088/0305-4470/32/7/007, doi:10.1088/0305-4470/32/7/007.

[21] Luca D'Alessio, Yariv Kafri, Anatoli Polkovnikov, and Marcos Rigol. From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics. *Adv. Phys.*, 65(3):239–362, 2016. doi:10.1080/00018732.2016.1198134.

[22] Joshua M Deutsch. Eigenstate thermalization hypothesis. *Rep. Prog. Phys.*, 81(8):082001, 2018. URL: http://dx.doi.org/10.1088/1361-6633/aac9f1, doi:10.1088/1361-6633/aac9f1.

[23] Sheldon Goldstein, Joel L Lebowitz, Roderich Tumulka, and Nino Zanghì. Canonical typicality. *Phys. Rev. Lett.*, 96(5):050403, 2006. URL: https://doi.org/10.1103/PhysRevLett.96.050403.

[24] Sandu Popescu, Anthony J Short, and Andreas Winter. Entanglement and the foundations of statistical mechanics. *Nat. Phys.*, 2(11):754–758, 2006. doi:10.1038/nphys444.

[25] Paul Adrien Maurice Dirac. *The principles of quantum mechanics*. Number 27. Oxford university press, 1981.

[26] E. B. Stechel. Quantum ergodicity and a quantum measure algebra. *J. Chem. Phys.*, 82:364. URL: https://doi.org/10.1063/1.448754.

[27] EB Stechel and EJ Heller. Quantum ergodicity and spectral chaos. *Annu. Rev. Phys. Chem.*, 35(1):563–589. URL: https://doi.org/10.1146/annurev.pc.35.100184.003023.

[28] Fritz Haake. *Quantum signatures of chaos*. Springer, Berlin, Heidelberg, 2001. URL: https://doi.org/10.1007/978-3-662-04506-0.

[29] Ze'ev Rudnick. What is quantum chaos. *Notices of the AMS*, 55(1):32–34, 2008. URL: https://www.ams.org/notices/200801/tx080100032p.pdf.

[30] Steven W McDonald and Allan N Kaufman. Spectrum and eigenfunctions for a hamiltonian with stochastic trajectories. *Phys. Rev. Lett.*, 42(18):1189. URL: https://doi.org/10.1103/PhysRevLett.42.1189.

[31] G Casati, F Valz-Gris, and I Guarnieri. On the connection between quantization of nonintegrable systems and statistical theory of spectra. *Lett. Nuovo Cimento*, 28(8):279–282.

[32] Michael V Berry. Quantizing a classically ergodic system: Sinai's billiard and the KKR method. *Ann. Phys.*, 131(1):163–216. URL: https://doi.org/10.1016/0003-4916(81)90189-5.

[33] Oriol Bohigas, Marie-Joya Giannoni, and Charles Schmit. Characterization of chaotic quantum spectra and universality of level fluctuation laws. *Phys. Rev. Lett.*, 52(1):1. URL: https://doi.org/10.1103/PhysRevLett.52.1.

[34] M. V. Berry and M. Tabor. Level clustering in the regular spectrum. *Proc. Roy. Soc. Lond. A*, 356:375–394. URL: https://royalsocietypublishing.org/doi/10.1098/rspa.1977.0140.

[35] Pavel Kos, Marko Ljubotina, and Tomaž Prosen. Many-body quantum chaos: Analytic connection to random matrix theory. *Phys. Rev. X*, 8(2):021062, 2018. URL: https://doi.org/10.1103/PhysRevX.8.021062.

[36] Madan Lal Mehta. *Random matrices*. Elsevier, 2004.

[37] Jean Ginibre. Statistical ensembles of complex, quaternion, and real matrices. *Journal of Mathematical Physics*, 6(3):440–449, 1965. URL: https://doi.org/10.1063/1.1704292.

[38] Hrant Gharibyan, Masanori Hanada, Stephen H Shenker, and Masaki Tezuka. Onset of random matrix behavior in scrambling systems. *J. High Energy Phys.*, 2018(7):1–62, 2018. URL: https://doi.org/10.1007/JHEP07(2018)124.

[39] J H Hannay and A M Ozorio de Almeida. Periodic orbits and a correlation function for the semiclassical density of states. *J. Phys. A: Math. Gen.*, 17(18):3429. URL: https://doi.org/10.1088/0305-4470/17/18/013.

[40] Michael Victor Berry. Semiclassical theory of spectral rigidity. *Proc. Roy. Soc. Lond. A*, 400(1819):229–251. URL: https://doi.org/10.1098/rspa.1985.0078.

[41] Amos Chan, Andrea De Luca, and JT Chalker. Solution of a minimal model for many-body quantum chaos. *Phys. Rev. X*, 8(4):041019, 2018. URL: https://doi.org/10.1103/PhysRevX.8.041019.

[42] S J Garratt and J T Chalker. Local pairing of Feynman histories in many-body Floquet models. *Phys. Rev. X*, 11(2):021051. URL: https://doi.org/10.1103/PhysRevX.11.021051.

[43] Benoît Collins and Piotr Śniady. Integration with respect to the Haar measure on unitary, orthogonal and symplectic group. *Commun. Math. Phys.*, 264(3):773–795. URL: https://doi.org/10.1007/s00220-006-1554-3.

[44] Georg Köstenberger. Weingarten calculus. *arXiv preprint arXiv:2101.00921*. URL: https://doi.org/10.48550/arXiv.2101.00921.

[45] Michael V Berry. Some quantum-to-classical asymptotics. In M-J Giannoni, A Voros, and J Zinn-Justin, editors, *Les Houches Lecture series LII (1989)*, pages 251–304. Elsevier Science Publishers B.V. URL: https://michaelberryphysics.files.wordpress.com/2013/07/berry227.pdf.

[46] JB French, PA Mello, and A Pandey. Statistical properties of many-particle spectra. II. two-point correlations and fluctuations. *Ann. Phys.*, 113(2):277–293. URL: https://doi.org/10.1016/0003-4916(78)90205-1.

[47] Tómas A Brody, Jorge Flores, J Bruce French, PA Mello, A Pandey, and Samuel SM Wong. Random-matrix physics: spectrum and strength fluctuations. *Rev. Mod. Phys.*, 53(3):385, 1981. URL: https://doi.org/10.1103/RevModPhys.53.385.

[48] Edward Ott. *Chaos in dynamical systems*. Cambridge University Press.

[49] P Bocchieri and A Loinger. Quantum recurrence theorem. *Phys. Rev.*, 107(2):337, 1957. URL: https://doi.org/10.1103/PhysRev.107.337.

[50] Michael A. Nielsen and Isaac L. Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2010. doi:10.1017/CBO9780511976667.

[51] Rahul Nandkishore and David A Huse. Many-body localization and thermalization in quantum statistical mechanics. *Annu. Rev. Condens. Matter Phys.*, 6(1):15–38, 2015. URL: https://doi.org/10.1146/annurev-conmatphys-031214-014726.

[52] Anatoly Dymarsky, Nima Lashkari, and Hong Liu. Subsystem eigenstate thermalization hypothesis. *Phys. Rev. E*, 97:012140, 2018. URL: https://link.aps.org/doi/10.1103/PhysRevE.97.012140, doi:10.1103/PhysRevE.97.012140.

[53] Patrick Hayden and John Preskill. Black holes as mirrors: quantum information in random subsystems. *J. High Energy Phys.*, 2007(09):120, 2007. URL: https://doi.org/10.1088/1126-6708/2007/09/120.

[54] Yasuhiro Sekino and Leonard Susskind. Fast scramblers. *J. High Energy Phys.*, 2008(10):065, 2008. URL: https://doi.org/10.1088/1126-6708/2008/10/065.

[55] Leonid Mandelstam and IG Tamm. The uncertainty relation between energy and time in non-relativistic quantum mechanics. In *Selected papers*, pages 115–123. Springer, 1991. URL: https://doi.org/10.1007/978-3-642-74626-0_8.

[56] J Anandan and Yakir Aharonov. Geometry of quantum evolution. *Phys. Rev. Lett.*, 65(14):1697, 1990. URL: https://doi.org/10.1103/PhysRevLett.65.1697.

[57] Norman Margolus and Lev B Levitin. The maximum speed of dynamical evolution. *Physica D*, 120(1-2):188–195, 1998. URL: https://doi.org/10.1016/S0167-2789(98)00054-2.

[58] Lev B Levitin and Tommaso Toffoli. Fundamental limit on the rate of quantum dynamics: the unified bound is tight. *Phys. Rev. Lett.*, 103(16):160502, 2009. URL: https://doi.org/10.1103/PhysRevLett.103.160502.

[59] Sebastian Deffner and Steve Campbell. Quantum speed limits: from Heisenberg's uncertainty principle to optimal quantum control. *J. Phys. A: Math. Theor.*, 50(45):453001, 2017. URL: https://doi.org/10.1088/1751-8121/aa86c6.

[60] Zongping Gong and Ryusuke Hamazaki. Bounds in nonequilibrium quantum dynamics. *Int. J. Mod. Phys. B.*, 36(31), 2022. URL: https://doi.org/10.1142/S0217979222300079.

[61] Marin Bukov, Dries Sels, and Anatoli Polkovnikov. Geometric speed limit of accessible many-body state preparation. *Phys. Rev. X*, 9(1):011034, 2019. URL: https://doi.org/10.1103/PhysRevX.9.011034.

[62] Ramamurti Shankar. *Principles of quantum mechanics*. Springer Science & Business Media. URL: https://doi.org/10.1007/978-1-4757-0576-8.

[63] Ryszard Horodecki, Paweł Horodecki, Michał Horodecki, and Karol Horodecki. Quantum entanglement. *Rev. Mod. Phys.*, 81(2):865, 2009. URL: https://doi.org/10.1103/RevModPhys.81.865.

[64] Efim B Rozenbaum, Leonid A Bunimovich, and Victor Galitski. Early-time exponential instabilities in non-chaotic quantum systems. *arXiv preprint arXiv:1902.05466v2*. URL: https://doi.org/10.48550/arXiv.1902.05466.

[65] Lev Davidovich Landau and Evgenii Mikhailovich Lifshitz. *Statistical Physics*, volume 5. Elsevier, 1980. URL: https://doi.org/10.1016/C2009-0-24487-4.

[66] Marcos Rigol, Vanja Dunjko, and Maxim Olshanii. Thermalization and its mechanism for generic isolated quantum systems. *Nature*, 452(7189):854–858, 2008. doi:10.1038/nature06838.

[67] Peter Reimann. Foundation of statistical mechanics under experimentally realistic conditions. *Phys. Rev. Lett.*, 101(19):190403. URL: https://doi.org/10.1103/PhysRevLett.101.190403.

[68] Anthony J Short. Equilibration of quantum systems and subsystems. *New J. Phys.*, 13(5):053009, 2011. URL: https://doi.org/10.1088/1367-2630/13/5/053009.

[69] Fausto Borgonovi, Felix M Izrailev, Lea F Santos, and Vladimir G Zelevinsky. Quantum chaos and thermalization in isolated systems of interacting particles. *Physics Reports*, 626:1–58, 2016. URL: https://doi.org/10.1016/j.physrep.2016.02.005.

[70] Dongliang Zhang, H T Quan, and Biao Wu. Ergodicity and mixing in quantum dynamics. *Phys. Rev. E*, 94(2):022150. URL: https://doi.org/10.1103/PhysRevE.94.022150.

[71] Steve Zelditch. Quantum ergodicity and mixing. *arXiv preprint math-ph/0503026*, 2005. URL: https://doi.org/10.48550/arXiv.math-ph/0503026.

[72] Nalini Anantharaman. Delocalization of Schrödinger eigenfunctions. In *Proceedings of the International Congress of Mathematicians: Rio de Janeiro 2018*, pages 341–375. World Scientific. URL: https://doi.org/10.1142/9789813272880_0016.

[73] Nathan Argaman, Yoseph Imry, and Uzy Smilansky. Semiclassical analysis of spectral correlations in mesoscopic systems. *Phys. Rev. B*, 47(8):4440, 1993. URL: https://doi.org/10.1103/PhysRevB.47.4440.

[74] Martin Sieber and Klaus Richter. Correlations between periodic orbits and their rôle in spectral statistics. *Phys. Scr.*, 2001(T90):128. URL: https://doi.org/10.1238/Physica.Topical.090a00128.

[75] Sebastian Müller, Stefan Heusler, Petr Braun, Fritz Haake, and Alexander Altland. Semiclassical foundation of universality in quantum chaos. *Phys. Rev. Lett.*, 93(1):014103. URL: https://doi.org/10.1103/PhysRevLett.93.014103.

[76] Sebastian Müller, Stefan Heusler, Petr Braun, Fritz Haake, and Alexander Altland. Periodic-orbit theory of universality in quantum chaos. *Phys. Rev. E*, 72(4):046207. URL: https://doi.org/10.1103/PhysRevE.72.046207.

[77] Klaus Richter, Juan Diego Urbina, and Steven Tomsovic. Semiclassical roots of universality in many-body quantum chaos. *J. Phys. A: Math. Theor.*, 55(45):453001. URL: https://doi.org/10.1088/1751-8121/ac9e4e.

[78] Olivier Giraud, Jens Marklof, and Stephen O'Keefe. Intermediate statistics in quantum maps. *J. Phys. A: Math. Gen.*, 37(28):L303. URL: https://doi.org/10.1088/0305-4470/37/28/L01.

[79] Črt Lozej, Giulio Casati, and Tomaž Prosen. Quantum chaos in triangular billiards. *Phys. Rev. Res.*, 4(1):013138, 2022. URL: https://doi.org/10.1103/PhysRevResearch.4.013138.

[80] Jiaozi Wang, Giuliano Benenti, Giulio Casati, and Wen-ge Wang. Statistical and dynamical properties of the quantum triangle map. *J. Phys. A: Math. Theor*, 55:234002. URL: https://doi.org/10.1088/1751-8121/ac6a93.

[81] Giulio Casati and Tomaž Prosen. Triangle map: A model of quantum chaos. *Phys. Rev. Lett.*, 85(20):4261. URL: https://doi.org/10.1103/PhysRevLett.85.4261.

[82] Jordan S Cotler, Guy Gur-Ari, Masanori Hanada, Joseph Polchinski, Phil Saad, Stephen H Shenker, Douglas Stanford, Alexandre Streicher, and Masaki Tezuka. Black holes and random matrices. *J. High Energy Phys.*, 2017(5):1–54, 2017. URL: https://doi.org/10.1007/JHEP05(2017)118.

[83] Amos Chan, Andrea De Luca, and JT Chalker. Spectral statistics in spatially extended chaotic quantum many-body systems. *Phys. Rev. Lett.*, 121(6):060601, 2018. URL: https://doi.org/10.1103/PhysRevLett.121.060601.

[84] Bruno Bertini, Pavel Kos, and Tomaž Prosen. Exact spectral form factor in a minimal model of many-body quantum chaos. *Phys. Rev. Lett.*, 121(26):264101, 2018. URL: https://doi.org/10.1103/PhysRevLett.121.264101.

[85] Aaron J Friedman, Amos Chan, Andrea De Luca, and JT Chalker. Spectral statistics and many-body quantum chaos with conserved charge. *Phys. Rev. Lett.*, 123(21):210603, 2019. URL: https://doi.org/10.1103/PhysRevLett.123.210603.

[86] Phil Saad, Stephen H Shenker, and Douglas Stanford. A semiclassical ramp in SYK and in gravity. *arXiv preprint arXiv:1806.06840*, 2018. URL: https://doi.org/10.48550/arXiv.1806.06840.

[87] Michael Winer, Shao-Kai Jian, and Brian Swingle. Exponential ramp in the quadratic Sachdev-Ye-Kitaev model. *Phys. Rev. Lett.*, 125(25):250602, 2020. URL: https://doi.org/10.1103/PhysRevLett.125.250602.

[88] Yunxiang Liao and Victor Galitski. Universal dephasing mechanism of many-body quantum chaos. *Phys. Rev. Res.*, 4(1):L012037, 2022. URL: https://doi.org/10.1103/PhysRevResearch.4.L012037.

[89] Yunxiang Liao and Victor Galitski. Emergence of many-body quantum chaos via spontaneous breaking of unitarity. *Phys. Rev. B*, 105(14):L140202, 2022. URL: https://doi.org/10.1103/PhysRevB.105.L140202.

[90] Yunxiang Liao and Victor Galitski. Effective field theory of random quantum circuits. *Entropy*, 24(6):823. URL: https://doi.org/10.3390/e24060823.

[91] Michael Winer, Richard Barney, Christopher L Baldwin, Victor Galitski, and Brian Swingle. Spectral form factor of a quantum spin glass. *J. High Energy Phys.*, 2022(9):1–47. URL: https://doi.org/10.1007/JHEP09(2022)032.

[92] Bruno Bertini, Pavel Kos, and Tomaž Prosen. Exact correlation functions for dual-unitary lattice models in 1+ 1 dimensions. *Phys. Rev. Lett.*, 123(21):210601, 2019. URL: https://doi.org/10.1103/PhysRevLett.123.210601.

[93] Pieter W Claeys and Austen Lamacraft. Ergodic and nonergodic dual-unitary quantum circuits with arbitrary local Hilbert space dimension. *Phys. Rev. Lett.*, 126(10):100603, 2021. URL: https://doi.org/10.1103/PhysRevLett.126.100603.

[94] S Aravinda, Suhail Ahmad Rather, and Arul Lakshminarayan. From dual-unitary to quantum Bernoulli circuits: Role of the entangling power in constructing a quantum ergodic hierarchy. *Phys. Rev. Res.*, 3(4):043034, 2021. URL: https://doi.org/10.1103/PhysRevResearch.3.043034.

[95] Mauro Schiulaz, E Jonathan Torres-Herrera, and Lea F Santos. Thouless and relaxation time scales in many-body quantum systems. *Phys. Rev. B*, 99(17):174313. URL: https://doi.org/10.1103/PhysRevB.99.174313.

[96] Denis V Vasilyev, Andrey Grankin, Mikhail A Baranov, Lukas M Sieberer, and Peter Zoller. Monitoring quantum simulators via quantum nondemolition couplings to atomic clock qubits. *PRX Quantum*, 1(2):020302, 2020. URL: https://doi.org/10.1103/PRXQuantum.1.020302.

[97] Anatole B. Katok and Anatolii M. Stepin. Approximation of ergodic dynamic systems by periodic transformations. *Dokl. Akad. Nauk SSSR*, 171(6):1268–1271. URL: https://www.mathnet.ru/eng/dan32766.

[98] Anatole B. Katok and Anatolii M. Stepin. Approximations in ergodic theory. *Russ. Math. Surv.*, 22(5):77. URL: https://doi.org/10.1070/rm1967v022n05abeh001227.

[99] A B Katok, Ya G Sinai, and A M Stepin. Theory of dynamical systems and general transformation groups with invariant measure. *J. Math. Sci.*, 7(6):974–1065. URL: https://doi.org/10.1007/BF01223133.

[100] Mahendra Ganpatrao Nadkarni. *Spectral theory of dynamical systems*. Springer, Singapore. URL: https://doi.org/10.1007/978-981-15-6225-9.

[101] Rainer Aurich, Jens Bolte, and Frank Steiner. Universal signatures of quantum chaos. *Phys. Rev. Lett.*, 73(10):1356. URL: https://doi.org/10.1103/PhysRevLett.73.1356.

[102] R Aurich, A Bäcker, and F Steiner. Mode fluctuations as fingerprints of chaotic and non-chaotic systems. *Int. J. Mod. Phys. B*, 11(07):805–849. URL: https://doi.org/10.1142/S0217979297000459.

[103] Vladimir Abramovich Rokhlin. Lectures on the entropy theory of measure-preserving transformations. *Russ. Math. Surv.*, 22(5):1.

[104] Mario Castagnino and Olimpia Lombardi. Towards a definition of the quantum ergodic hierarchy: Ergodicity and mixing. *Physica A*, 388(4):247–267. URL: https://doi.org/10.1016/j.physa.2008.10.019.

[105] A Iwanik. Cyclic approximation of irrational rotations. *Proc. Am. Math. Soc.*, 121(3):691–695. URL: https://doi.org/10.1090/S0002-9939-1994-1221724-X.

[106] Godfrey Harold Hardy, Edward Maitland Wright, et al. *An introduction to the theory of numbers*. Oxford University Press.

[107] RV Chacon. Approximation and spectral multiplicity. In *Contributions to Ergodic Theory and Probability*, pages 18–27. Springer. URL: https://doi.org/10.1007/BFb0060642.

[108] Adam R Brown and Leonard Susskind. Second law of quantum complexity. *Phys. Rev. D*, 97(8):086015, 2018. URL: https://doi.org/10.1103/PhysRevD.97.086015.

[109] Roger A Horn and Charles R Johnson. *Matrix analysis*. Cambridge University Press. URL: https://doi.org/10.1017/CBO9780511810817.

[110] R E Prange. The spectral form factor is not self-averaging. *Phys. Rev. Lett.*, 78(12):2280, 1997. URL: https://doi.org/10.1103/PhysRevLett.78.2280.

[111] JT Chalker, Igor V Lerner, and Robert A Smith. Random walks through the ensemble: linking spectral statistics with wave-function correlations in disordered metals. *Phys. Rev. Lett.*, 77(3):554, 1996. URL: https://doi.org/10.1103/PhysRevLett.77.554.

[112] Michael Winer and Brian Swingle. Hydrodynamic theory of the connected spectral form factor. *Phys. Rev. X*, 12(1):021009, 2022. URL: https://doi.org/10.1103/PhysRevX.12.021009.

[113] Prot Pakoński, Karol Życzkowski, and Marek Kuś. Classical 1D maps, quantum graphs and ensembles of unitary matrices. *J. Phys. A: Math. Gen.*, 34(43):9303. URL: https://doi.org/10.1088/0305-4470/34/43/313.

[114] Peter Reimann. Typical fast thermalization processes in closed many-body systems. *Nat. Comm.*, 7(1):10821, 2016. doi:10.1038/ncomms10821.

[115] Ovidiu Costin and Joel L Lebowitz. Gaussian fluctuation in random matrices. *Phys. Rev. Lett.*, 75(1):69. URL: https://doi.org/10.1103/PhysRevLett.75.69.

[116] Freeman J Dyson and Madan Lal Mehta. Statistical theory of the energy levels of complex systems. IV. *J. Math. Phys.*, 4(5):701–712. URL: https://doi.org/10.1063/1.1704008.

[117] Herbert Goldstein, Charles Poole, and John Safko. *Classical Mechanics*. Pearson, 3 edition.

[118] Luigi Chierchia and Michela Procesi. Kolmogorov–Arnold–Moser (KAM) theory for finite and infinite dimensional systems. In R Meyers, editor, *Encyclopedia of Complexity and Systems Science*. Springer, Berlin, Heidelberg. URL: https://doi.org/10.1007/978-3-642-27737-5_302-3.

[119] Zh Cheng and Joel L Lebowitz. Statistics of energy levels in integrable quantum systems. *Phys. Rev. A*, 44(6):R3399. URL: https://doi.org/10.1103/PhysRevA.44.R3399.

[120] Pavel M Bleher, Zheming Cheng, Freeman J Dyson, and Joel L Lebowitz. Distribution of the error term for the number of lattice points inside a shifted circle. *Commun. Math. Phys.*, 154(3):433–469. URL: https://projecteuclid.org/journals/communications-in-mathematical-physics/volume-154/issue-3/Distribution-of-the-error-term-for-the-number-of-lattice/cmp/1104253074.full.

[121] A Pandey, O Bohigas, and MJ Giannoni. Level repulsion in the spectrum of two-dimensional harmonic oscillators. *J. Phys. A.: Math. Gen.*, 22(18):4083. URL: https://doi.org/10.1088/0305-4470/22/18/039.

[122] Akhilesh Pandey and Ramakrishna Ramaswamy. Level spacings for harmonic-oscillator systems. *Phys. Rev. A*, 43(8):4237. URL: https://doi.org/10.1103/PhysRevA.43.4237.

[123] Paul R Halmos. Two subspaces. *Trans. Am. Math. Soc.*, 144:381–389. URL: https://doi.org/10.2307/1995288.

[124] Albrecht Böttcher and Ilya M Spitkovsky. A gentle guide to the basics of two projections theory. *Linear Algebra Its Appl.*, 432(6):1412–1459. URL: https://doi.org/10.1016/j.laa.2009.11.002.

[125] Xiao Mi, Pedram Roushan, Chris Quintana, Salvatore Mandrà, Jeffrey Marshall, Charles Neill, Frank Arute, Kunal Arya, Juan Atalaya, Ryan Babbush, et al. Information scrambling in quantum circuits. *Science*, 374(6574):1479–1483, 2021. URL: https://doi.org/10.1126/science.abg5029.

[126] Anatoli Polkovnikov. Phase space representation of quantum dynamics. *Ann. Phys.*, 325(8):1790–1852. URL: https://doi.org/10.1016/j.aop.2010.02.006.

[127] Robin L Hudson. When is the Wigner quasi-probability density non-negative? *Rep. Math. Phys.*, 6(2):249–252. URL: https://doi.org/10.1016/0034-4877(74)90007-X.

[128] D. Shepelyansky. Ehrenfest time and chaos. *Scholarpedia*, 15(9):55031. doi:10.4249/scholarpedia.55031.

[129] B. V. Chirikov, F. M. Izrailev, and D. L. Shepelyansky. Dynamical stochasticity in classical and quantum mechanics. *Sov. Scient. Rev. C*, 2:209.

[130] Fabio Anza, Christian Gogolin, and Marcus Huber. Eigenstate thermalization for degenerate observables. *Phys. Rev. Lett.*, 120(15):150603. URL: https://doi.org/10.1103/PhysRevLett.120.150603.

[131] Robert Alicki and Mark Fannes. *Quantum Dynamical Systems*. Oxford University Press. URL: https://doi.org/10.1093/acprof:oso/9780198504009.001.0001.

[132] Lucas Sá, Pedro Ribeiro, and Tomaž Prosen. Complex spacing ratios: A signature of dissipative quantum chaos. *Phys. Rev. X*, 10(2):021019. URL: https://doi.org/10.1103/PhysRevX.10.021019.

[133] Jiachen Li, Tomaž Prosen, and Amos Chan. Spectral statistics of non-Hermitian matrices and dissipative quantum chaos. *Phys. Rev. Lett.*, 127(17):170602. URL: https://doi.org/10.1103/PhysRevLett.127.170602.

[134] Antonio M García-García, Lucas Sá, and Jacobus JM Verbaarschot. Symmetry classification and universality in non-Hermitian many-body quantum chaos by the Sachdev-Ye-Kitaev model. *Phys. Rev. X*, 12(2):021040. URL: https://doi.org/10.1103/PhysRevX.12.021040.

[135] GP Brandino, J-S Caux, and RM Konik. Glimmers of a quantum KAM theorem: insights from quantum quenches in one-dimensional Bose gases. *Phys. Rev. X*, 5(4):041043. URL: https://doi.org/10.1103/PhysRevX.5.041043.

[136] Gal Ness, Andrea Alberti, and Yoav Sagi. Quantum speed limit for states with a bounded energy spectrum. *Phys. Rev. Lett.*, 129(14):140403, 2022. URL: https://doi.org/10.1103/PhysRevLett.129.140403.

[137] Anatoli Polkovnikov, Krishnendu Sengupta, Alessandro Silva, and Mukund Vengalattore. Colloquium: Nonequilibrium dynamics of closed interacting quantum systems. *Rev. Mod. Phys.*, 83(3):863, 2011. URL: https://doi.org/10.1103/RevModPhys.83.863.

[138] Henrik Wilming, Thiago R de Oliveira, Anthony J Short, and Jens Eisert. Equilibration times in closed quantum many-body systems. In *Thermodynamics in the Quantum Regime*, pages 435–455. Springer, 2018. URL: https://doi.org/10.1007/978-3-319-99046-0_18.

[139] Pavel Kos, Tomaž Prosen, and Bruno Bertini. Thermalization dynamics and spectral statistics of extended systems with thermalizing boundaries. *Phys. Rev. B*, 104(21):214303, 2021. URL: https://doi.org/10.1103/PhysRevB.104.214303.

[140] Dmitry A Abanin, Ehud Altman, Immanuel Bloch, and Maksym Serbyn. Colloquium: Many-body localization, thermalization, and entanglement. *Rev. Mod. Phys.*, 91(2):021001, 2019. URL: https://doi.org/10.1103/RevModPhys.91.021001.

[141] Nima Lashkari, Douglas Stanford, Matthew Hastings, Tobias Osborne, and Patrick Hayden. Towards the fast scrambling conjecture. *J. High Energy Phys.*, 2013(4):1–33, 2013. URL: https://doi.org/10.1007/JHEP04(2013)022.

[142] Gregory Bentsen, Yingfei Gu, and Andrew Lucas. Fast scrambling on sparse graphs. *Proc. Natl. Acad. Sci. U.S.A.*, 116(14):6689–6694, 2019. URL: https://doi.org/10.1073/pnas.1811033116.

[143] Andrew Lucas. Quantum many-body dynamics on the star graph. *arXiv preprint arXiv:1903.01468*, 2019. URL: https://doi.org/10.48550/arXiv.1903.01468.

[144] Elliott H Lieb and Derek W Robinson. The finite group velocity of quantum spin systems. *Commun. Math. Phys.*, 28:251–257, 1972. URL: https://doi.org/10.1007/BF01645779.

[145] Bruno Nachtergaele, Yoshiko Ogata, and Robert Sims. Propagation of correlations in quantum lattice systems. *J. Stat. Phys.*, 124(1):1–13, 2006. URL: https://doi.org/10.1007/s10955-006-9143-6.

[146] Matthew B Hastings and Tohru Koma. Spectral gap and exponential decay of correlations. *Commun. Math. Phys.*, 265(3):781–804, 2006.

[147] Donald E Knuth. Big omicron and big omega and big theta. *ACM Sigact News*, 8(2):18–24, 1976. URL: https://doi.org/10.1145/1008328.1008329.

[148] Abhishodh Prakash, JH Pixley, and Manas Kulkarni. Universal spectral form factor for many-body localization. *Phys. Rev. Res.*, 3(1):L012019, 2021. URL: https://doi.org/10.1103/PhysRevResearch.3.L012019.

[149] Pavel Kos, Bruno Bertini, and Tomaž Prosen. Chaos and ergodicity in extended quantum systems with noisy driving. *Phys. Rev. Lett.*, 126(19):190601, 2021. URL: https://doi.org/10.1103/PhysRevLett.126.190601.

[150] Tankut Can. Random Lindblad dynamics. *J. Phys. A: Math. Theor.*, 52(48):485302, 2019. URL: https://doi.org/10.1088/1751-8121/ab4d26.

[151] Zhenyu Xu, Aurelia Chenu, Tomaž Prosen, and Adolfo del Campo. Thermofield dynamics: Quantum chaos versus decoherence. *Phys. Rev. B*, 103(6):064309, 2021. URL: https://doi.org/10.1103/PhysRevB.103.064309.

[152] Julien Cornelius, Zhenyu Xu, Avadh Saxena, Aurélia Chenu, and Adolfo del Campo. Spectral filtering induced by non-Hermitian evolution with balanced gain and loss: Enhancing quantum chaos. *Phys. Rev. Lett.*, 128(19):190402, 2022. URL: https://doi.org/10.1103/PhysRevLett.128.190402.

[153] Kohei Kawabata, Anish Kulkarni, Jiachen Li, Tokiro Numasawa, and Shinsei Ryu. Dynamical quantum phase transitions in SYK Lindbladians. *arXiv preprint arXiv:2210.04093*, 2022. URL: https://doi.org/10.48550/arXiv.2212.00605.

[154] Jordan Cotler, Nicholas Hunter-Jones, Junyu Liu, and Beni Yoshida. Chaos, complexity, and random matrices. *J. High Energy Phys.*, 2017(11):1–60, 2017. URL: https://doi.org/10.1007/JHEP11(2017)048.

[155] Jordan Cotler and Nicholas Hunter-Jones. Spectral decoupling in many-body quantum chaos. *J. High Energy Phys.*, 2020(12):1–62, 2020. URL: https://doi.org/10.1007/JHEP12(2020)205.

[156] A del Campo, J Molina-Vilaplana, and J Sonner. Scrambling the spectral form factor: unitarity constraints and exact results. *Phys. Rev. D*, 95(12):126008, 2017. URL: https://doi.org/10.1103/PhysRevD.95.126008.

[157] Ernest O Tuck. On positivity of Fourier transforms. *Bull. Aust. Math. Soc.*, 74(1):133–138, 2006. URL: https://doi.org/10.1017/S0004972700047511.

[158] Bertrand G Giraud and Robi Peschanski. On the positivity of Fourier transforms. *arXiv preprint arXiv:1405.3155*, 2014. URL: https://doi.org/10.48550/arXiv.1405.3155.

[159] Juan Maldacena, Stephen H Shenker, and Douglas Stanford. A bound on chaos. *J. High Energy Phys.*, 2016(8):1–17, 2016. URL: https://doi.org/10.1007/JHEP08(2016)106.

[160] Pablo Martinez-Azcona and Aurélia Chenu. Analyticity constraints bound the decay of the spectral form factor. *Quantum*, 6:852, 2022. URL: https://doi.org/10.22331/q-2022-11-03-852.

[161] Luciano Fonda, GC Ghirardi, and Alberto Rimini. Decay theory of unstable quantum systems. *Rep. Prog. Phys.*, 41(4):587, 1978. URL: https://doi.org/10.1088/0034-4885/41/4/003.

[162] Javad Mashreghi, Fedor Nazarov, and Victor Havin. Beurling–Malliavin multiplier theorem: the seventh proof. *St. Petersburg Math. J.*, 17(5):699–744, 2006. URL: https://doi.org/10.1090/S1061-0022-06-00926-5.

[163] Juan Maldacena and Douglas Stanford. Remarks on the Sachdev-Ye-Kitaev model. *Phys. Rev. D*, 94:106002, Nov 2016. URL: https://doi.org/10.1103/PhysRevD.94.106002.

[164] Alexei Kitaev and S Josephine Suh. The soft mode in the Sachdev-Ye-Kitaev model and its gravity dual. *J. High Energy Phys.*, 2018(5):183, 2018. URL: https://doi.org/10.1007/JHEP05(2018)183.

[165] Antoine Georges, Olivier Parcollet, and Subir Sachdev. Quantum fluctuations of a nearly critical Heisenberg spin glass. *Phys. Rev. B*, 63(13):134406, 2001. URL: https://doi.org/10.1103/PhysRevB.63.134406.

[166] Chi-Fang Chen, Andrew Lucas, and Chao Yin. Speed limits and locality in many-body quantum dynamics. *Rep. Prog. Phys.*, 2023. URL: https://doi.org/10.1088/1361-6633/acfaae.

[167] Yoshifumi Nakata and Masaki Tezuka. Hayden-Preskill recovery in Hamiltonian systems. *arXiv preprint arXiv:2303.02010*, 2023. URL: https://doi.org/10.48550/arXiv.2303.02010.

[168] Hang Dong, Pengfei Zhang, Ceren B Dag, Yu Gao, Ning Wang, Jinfeng Deng, Xu Zhang, Jiachen Chen, Shibo Xu, Ke Wang, et al. Measuring spectral form factor in many-body chaotic and localized phases of quantum processors. *arXiv preprint arXiv:2403.16935*, 2024. URL: https://doi.org/10.48550/arXiv.2403.16935.

[169] Apollonas S Matsoukas-Roubeas, Tomaž Prosen, and Adolfo del Campo. Quantum chaos and coherence: Random parametric quantum channels. *arXiv preprint arXiv:2305.19326*, 2023. URL: https://doi.org/10.48550/arXiv.2305.19326.

[170] Apollonas S Matsoukas-Roubeas, Mathieu Beau, Lea F Santos, and Adolfo del Campo. Unitarity breaking in self-averaging spectral form factors. *Phys. Rev. A*, 108(6):062201, 2023. URL: https://doi.org/10.1103/PhysRevA.108.062201.

[171] Guy Bunin, Laura Foini, and Jorge Kurchan. Fisher zeroes and the fluctuations of the spectral form factor of chaotic systems. *arXiv preprint arXiv:2207.02473*, 2022. URL: https://doi.org/10.48550/arXiv.2207.02473.

[172] Alaina M Green, A Elben, C Huerta Alderete, Lata Kh Joshi, Nhung H Nguyen, Torsten V Zache, Yingyue Zhu, Bhuvanesh Sundar, and Norbert M Linke. Experimental measurement of out-of-time-ordered correlators at finite temperature. *Phys. Rev. Lett.*, 128(14):140601, 2022. URL: https://doi.org/10.1103/PhysRevLett.128.140601.

[173] Lars Ahlfors and Maurice Heins. Questions of regularity connected with the Phragmén-Lindelöf principle. *Ann. of Math. (2)*, 50:341–346, 1949. doi:10.2307/1969459.

[174] W. K. Hayman. Questions of regularity connected with the Phragmén-Lindelöf principle. *J. Math. Pures Appl. (9)*, 35:115–126, 1956.

[175] Ju. V. Linnik and I. V. Ostrovsiĭ. *Decomposition of random variables and vectors*, volume Vol. 48 of *Translations of Mathematical Monographs*. American Mathematical Society, Providence, RI, 1977. Translated from the Russian.

[176] W. K. Hayman. *Subharmonic functions. Vol. 2*, volume 20 of *London Mathematical Society Monographs*. Academic Press, Inc. [Harcourt Brace Jovanovich, Publishers], London, 1989.

[177] Paul Koosis. *The logarithmic integral. I*, volume 12 of *Cambridge Studies in Advanced Mathematics*. Cambridge University Press, Cambridge, 1988. doi:10.1017/CBO9780511566196.

[178] I. V. Ostrovskiĭ. A certain class of characteristic functions. *Trudy Mat. Inst. Steklov.*, 111:195–207, 1970.

[179] Yingfei Gu, Alexei Kitaev, Subir Sachdev, and Grigory Tarnopolsky. Notes on the complex Sachdev-Ye-Kitaev model. *J. High Energy Phys.*, 2020(2):1–74, 2020. URL: https://doi.org/10.1007/JHEP02(2020)157.

[180] Subir Sachdev. Holographic metals and the fractionalized Fermi liquid. *Phys. Rev. Lett.*, 105(15):151602, 2010. URL: https://doi.org/10.1103/PhysRevLett.105.151602.

[181] Subir Sachdev. Bekenstein-Hawking entropy and strange metals. *Phys. Rev. X*, 5(4):041025, 2015. URL: https://doi.org/10.1103/PhysRevX.5.041025.

[182] Jorge Kurchan. Time-reparametrization invariance: from glasses to toy black holes. *arXiv preprint arXiv:2401.03186*, 2024. URL: https://doi.org/10.48550/arXiv.2401.03186.

[183] Tiff Brydges, Andreas Elben, Petar Jurcevic, Benoît Vermersch, Christine Maier, Ben P Lanyon, Peter Zoller, Rainer Blatt, and Christian F Roos. Probing Rényi entanglement entropy via randomized measurements. *Science*, 364(6437):260–263, 2019. URL: https://doi.org/10.1126/science.aau4963.

[184] Linda E Reichl. *A modern course in statistical physics*. John Wiley & Sons, 2016.

[185] John David Jackson. *Classical electrodynamics*. John Wiley & Sons, 1998.

[186] Elias M. Stein and Rami Shakarchi. *Complex analysis*, volume 2 of *Princeton Lectures in Analysis*. Princeton University Press, Princeton, NJ, 2003.

[187] Ruel V. Churchill and James Ward Brown. *Complex variables and applications*. McGraw-Hill Book Co., New York, fourth edition, 1984.

[188] Michael Reed and Barry Simon. *Methods of modern mathematical physics. II. Fourier analysis, self-adjointness*. Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London, 1975.

[189] The mpmath development team. *mpmath: a Python library for arbitrary-precision floating-point arithmetic (version 1.3.0)*, 2023. https://mpmath.org/.

[190] Pak Hang Chris Lau, Chen-Te Ma, Jeff Murugan, and Masaki Tezuka. Randomness and chaos in qubit models. *Physics Letters B*, 795:230–235, 2019. URL: https://doi.org/10.1016/j.physletb.2019.05.052.

[191] Gernot Akemann, Jinho Baik, and Philippe Di Francesco. *The Oxford handbook of random matrix theory*. Oxford University Press, 2011. URL: https://doi.org/10.1093/oxfordhb/9780198744191.001.0001.

[192] Harry Bateman. *Higher Transcendental Functions [Volumes I-III]*, volume 2. McGraw-Hill Book Company, 1953. URL: https://authors.library.caltech.edu/records/cnd32-h9x80.

[193] Junyu Liu. Spectral form factors and late time quantum chaos. *Phys. Rev. D.*, 98(8):086026, 2018. URL: https://doi.org/10.1103/PhysRevD.98.086026.

[194] Izrail Solomonovich Gradshteyn and Iosif Moiseevich Ryzhik. *Table of integrals, series, and products*. Academic press. URL: https://doi.org/10.1016/C2010-0-64839-5.

[195] Gábor Sárosi. AdS_2 holography and the SYK model. *arXiv preprint arXiv:1711.08482*, 2017. URL: https://doi.org/10.48550/arXiv.1711.08482.

[196] Richard Barney, Michael Winer, Christopher L. Baldwin, Brian Swingle, and Victor Galitski. Spectral statistics of a minimal quantum glass model. *SciPost Phys.*, 15:084, 2023. URL: https://scipost.org/10.21468/SciPostPhys.15.3.084, doi:10.21468/SciPostPhys.15.3.084.

[197] JH Hannay and Michael V Berry. Quantization of linear maps on a torus-fresnel diffraction by a periodic grating. *Phys. D: Nonlinear Phenom.*, 1(3):267–290, 1980. doi:10.1016/0167-2789(80)90026-3.

[198] Jonathan P Keating. The cat maps: quantum mechanics and classical motion. *Nonlinearity*, 4(2):309, 1991. URL: https://doi.org/10.1088/0951-7715/4/2/006.

[199] Peter W Shor. Algorithms for quantum computation: discrete logarithms and factoring. In *Proceedings 35th annual symposium on foundations of computer science*, pages 124–134. IEEE, 1994. doi:10.1109/SFCS.1994.365700.

[200] EB Bogomolny, U Gerland, and C Schmit. Models of intermediate spectral statistics. *Phys. Rev. E*, 59(2):R1315, 1999. URL: https://doi.org/10.1103/PhysRevE.59.R1315.

[201] E.B. Bogomolny, B. Georgeot, M.-J. Giannoni, and C. Schmit. Arithmetical chaos. *Phys. Rep.*, 291(5–6):219–324, December 1997. doi:10.1016/S0370-1573(97)00016-1.

[202] W. Luo and P. Sarnak. Number variance for arithmetic hyperbolic surfaces. *Comm. Math. Phys.*, 161(2):419–432, 1994. URL: http://projecteuclid.org/euclid.cmp/1104269909, doi:10.1007/BF02099785.

[203] Petr Braun and Fritz Haake. Level statistics in arithmetical and pseudo-arithmetical chaos. *J. Phys. A: Math. Theor.*, 43(26):262001, 2010. URL: https://doi.org/10.1088/1751-8113/43/26/262001.

[204] Suman Das, Sumit K Garg, Chethan Krishnan, and Arnab Kundu. What is the simplest linear ramp? *arXiv preprint arXiv:2308.11704*, 2023. URL: https://doi.org/10.48550/arXiv.2308.11704.

[205] Arul Lakshminarayan. Modular multiplication operator and quantized baker's maps. *Phys. Rev. A*, 76(4):042330, 2007. URL: https://doi.org/10.1103/PhysRevA.76.042330.

[206] F. Leyvraz, C. Schmit, and T. H. Seligman. Anomalous spectral statistics in a symmetrical billiard. *J. Phys. A*, 29(22):L575–L580, 1996. doi:10.1088/0305-4470/29/22/004.

[207] J. P. Keating and J. M. Robbins. Discrete symmetries and spectral statistics. *J. Phys. A*, 30(7):L177–L181, 1997. doi:10.1088/0305-4470/30/7/006.

[208] Boris Gutkin. Dynamical 'breaking' of time reversal symmetry. *J. Phys. A*, 40(31):F761–F769, 2007. doi:10.1088/1751-8113/40/31/F02.

[209] C. Dembowski, H.-D. Gräf, A. Heine, H. Rehfeld, A. Richter, and C. Schmit. Gaussian unitary ensemble statistics in a time-reversal invariant microwave triangular billiard. *Phys. Rev. E*, 62:R4516–R4519, Oct 2000. URL: https://link.aps.org/doi/10.1103/PhysRevE.62.R4516, doi:10.1103/PhysRevE.62.R4516.

[210] Vadim Oganesyan and David A. Huse. Localization of interacting fermions at high temperature. *Phys. Rev. B*, 75:155111, Apr 2007. URL: https://link.aps.org/doi/10.1103/PhysRevB.75.155111, doi:10.1103/PhysRevB.75.155111.

[211] Norbert Rosenzweig and Charles E Porter. "Repulsion of energy levels" in complex atomic spectra. *Physical Review*, 120(5):1698, 1960. URL: https://doi.org/10.1103/PhysRev.120.1698.

[212] Michael V Berry and Marko Robnik. Semiclassical level spacings when regular and chaotic orbits coexist. *J. Phys. A: Math. Gen.*, 17(12):2413, 1984. URL: https://doi.org/10.1088/0305-4470/17/12/013.

[213] Uzy Smilansky, Steven Tomsovic, and Oriol Bohigas. Spectral fluctuations and transport in phase space. *J. Phys. A*, 25(11):3261–3273, 1992. URL: http://stacks.iop.org/0305-4470/25/3261, doi:10.1088/0305-4470/25/11/029.

[214] Olivier Giraud, Nicolas Macé, Éric Vernier, and Fabien Alet. Probing symmetries of quantum many-body systems through gap ratio statistics. *Phys. Rev. X*, 12:011006, Jan 2022. URL: https://link.aps.org/doi/10.1103/PhysRevX.12.011006, doi:10.1103/PhysRevX.12.011006.

[215] N. L. Balazs and A. Voros. The quantized baker's transformation. *Ann. Physics*, 190(1):1–31, 1989. doi:10.1209/0295-5075/4/10/001.

[216] M. Saraceno. Classical structures in the quantized baker transformation. *Ann. Physics*, 199(1):37–60, 1990. doi:10.1016/0003-4916(90)90367-W.

[217] Todd A Brun and Rüdiger Schack. Realizing the quantum baker's map on a NMR quantum computer. *Phys. Rev. A*, 59(4):2649, 1999. URL: https://doi.org/10.1103/PhysRevA.59.2649.

[218] Yaakov S Weinstein, Seth Lloyd, Joseph Emerson, and David G Cory. Experimental implementation of the quantum baker's map. *Phys. Rev. Lett.*, 89(15):157902, 2002. doi:10.1103/PhysRevLett.89.157902.

[219] Nalini Anantharaman and Stéphane Nonnenmacher. Entropy of semiclassical measures of the Walsh-quantized baker's map. *Ann. Henri Poincaré*, 8(1):37–74, 2007. doi:10.1007/s00023-006-0299-z.

[220] Tanja Eisner, Bálint Farkas, Markus Haase, and Rainer Nagel. *Operator theoretic aspects of ergodic theory*, volume 272 of *Graduate Texts in Mathematics*. Springer, Cham, 2015. doi:10.1007/978-3-319-16898-2.

[221] Yakov Grigor'evich Sinai. A weak isomorphism of transformations with invariant measure. In *Dokl. Akad. Nauk SSSR*, volume 147, pages 797–800. Russian Academy of Sciences, 1962. URL: https://www.mathnet.ru/eng/dan27300.

[222] Eli Glasner. *Ergodic theory via joinings*. Number 101. American Mathematical Soc., 2003. URL: https://doi.org/10.1090/surv/101.

[223] M. Saraceno and A. Voros. Towards a semiclassical theory of the quantum baker's map. *Phys. D*, 79(2-4):206–268, 1994. doi:10.1016/S0167-2789(05)80007-7.

[224] S. De Bièvre and M. Degli Esposti. Egorov theorems and equidistribution of eigenfunctions for the quantized sawtooth and baker maps. *Ann. Inst. H. Poincaré Phys. Théor.*, 69(1):1–30, 1998. URL: http://www.numdam.org/item?id=AIHPA_1998__69_1_1_0.

[225] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux. Distribution of the ratio of consecutive level spacings in random matrix ensembles. *Phys. Rev. Lett.*, 110:084101, Feb 2013. URL: https://link.aps.org/doi/10.1103/PhysRevLett.110.084101, doi:10.1103/PhysRevLett.110.084101.

[226] Patrick W. O'Connor and Steven Tomsovic. The unusual nature of the quantum baker's transformation. *Ann. Physics*, 207(1):218–264, 1991. URL: https://www.sciencedirect.com/science/article/pii/000349169190184A, doi:10.1016/0003-4916(91)90184-A.

[227] Daniel Burgarth, Giulio Chiribella, Vittorio Giovannetti, Paolo Perinotti, and Kazuya Yuasa. Ergodic and mixing quantum channels in finite dimensions. *New Journal of Physics*, 15(7):073045, 2013. URL: https://doi.org/10.1088/1367-2630/15/7/073045.

[228] Satvik Singh, Nilanjana Datta, and Ion Nechita. Ergodic theory of diagonal orthogonal covariant quantum channels. *arXiv preprint arXiv:2206.01145*, 2022. URL: https://doi.org/10.48550/arXiv.2206.01145.

[229] Saúl Pilatowsky-Cameo, Ceren B Dag, Wen Wei Ho, and Soonwon Choi. Complete Hilbert-space ergodicity in quantum dynamics of generalized Fibonacci drives. *arXiv preprint arXiv:2306.11792*, 2023. URL: https://doi.org/10.48550/arXiv.2306.11792.

[230] A. Bouzouina and S. De Bièvre. Equipartition of the eigenfunctions of quantized ergodic maps on the torus. *Comm. Math. Phys.*, 178(1):83–105, 1996. URL: http://projecteuclid.org/euclid.cmp/1104286555.

[231] Andreas Elben, Steven T Flammia, Hsin-Yuan Huang, Richard Kueng, John Preskill, Benoît Vermersch, and Peter Zoller. The randomized measurement toolbox. *Nature Reviews Physics*, 5(1):9–24, 2023. URL: https://doi.org/10.1038/s42254-022-00535-2.

[232] R Blatt and C F Roos. Quantum simulations with trapped ions. *Nat. Phys.*, 8(4):277–284, 2012. doi:10.1038/nphys2252.

[233] C. Monroe, W.C. Campbell, L.-M. Duan, Z.-X. Gong, A.V. Gorshkov, P.W. Hess, R. Islam, K. Kim, N.M. Linke, G. Pagano, et al. Programmable quantum simulations of spin systems with trapped

ions. *Rev. Mod. Phys.*, 93(2), Apr 2021. URL: http://dx.doi.org/10.1103/RevModPhys.93.025001.

[234] Antoine Browaeys and Thierry Lahaye. Many-body physics with individually controlled Rydberg atoms. *Nat. Phys.*, 16(2):132–142, 2020. doi:10.1038/s41567-019-0733-z.

[235] Morten Kjaergaard, Mollie E Schwartz, Jochen Braumüller, Philip Krantz, Joel I.-J. Wang, Simon Gustavsson, and William D Oliver. Superconducting Qubits: Current State of Play. *Ann. Rev. Cond. Matter Phys.*, 11(1):369–395, mar 2020. doi:10.1146/annurev-conmatphys-031119-050605.

[236] Jan Šuntajs, Janez Bonča, Toma ž Prosen, and Lev Vidmar. Quantum chaos challenges many-body localization. *Phys. Rev. E*, 102:062144, Dec 2020. URL: https://link.aps.org/doi/10.1103/PhysRevE.102.062144, doi:10.1103/PhysRevE.102.062144.

[237] D.A. Abanin, J.H. Bardarson, G. De Tomasi, S. Gopalakrishnan, V. Khemani, S.A. Parameswaran, F. Pollmann, A.C. Potter, M. Serbyn, and R. Vasseur. Distinguishing localization from chaos: Challenges in finite-size systems. *Ann. Phys.*, 427:168415, 2021. doi:10.1016/j.aop.2021.168415.

[238] Nivedita, Henry Shackleton, and Subir Sachdev. Spectral form factors of clean and random quantum ising chains. *Phys. Rev. E*, 101:042136, Apr 2020. URL: https://link.aps.org/doi/10.1103/PhysRevE.101.042136, doi:10.1103/PhysRevE.101.042136.

[239] Piotr Sierant, Dominique Delande, and Jakub Zakrzewski. Thouless time analysis of anderson and many-body localization transitions. *Phys. Rev. Lett.*, 124:186601, May 2020. URL: https://link.aps.org/doi/10.1103/PhysRevLett.124.186601, doi:10.1103/PhysRevLett.124.186601.

[240] Piotr Sierant, Maciej Lewenstein, and Jakub Zakrzewski. Polynomially filtered exact diagonalization approach to many-body localization. *Phys. Rev. Lett.*, 125:156601, Oct 2020. URL: https://link.aps.org/doi/10.1103/PhysRevLett.125.156601, doi:10.1103/PhysRevLett.125.156601.

[241] Luc Leviandier, Maurice Lombardi, Rémi Jost, and Jean Paul Pique. Fourier transform: A tool to measure statistical level properties in very complex spectra. *Phys. Rev. Lett.*, 56:2449–2452, Jun 1986. URL: https://link.aps.org/doi/10.1103/PhysRevLett.56.2449.

[242] Thomas Guhr, Axel Müller–Groeling, and Hans A. Weidenmüller. Random-matrix theories in quantum physics: common concepts. *Phys. Rep.*, 299(4-6):189–425, Jun 1998. URL: http://dx.doi.org/10.1016/S0370-1573(97)00088-4.

[243] Pedro Ponte, Anushya Chandran, Z Papić, and Dmitry A Abanin. Periodically driven ergodic and many-body localized quantum systems. *Ann. Phys.*, 353:196–204, 2015. doi:10.1016/j.aop.2014.11.008.

[244] Zongping Gong, Christoph Sünderhauf, Norbert Schuch, and J. Ignacio Cirac. Classification of Matrix-Product Unitaries with Symmetries. *Phys. Rev. Lett.*, 124(10):100402, 2020. doi:10.1103/PhysRevLett.124.100402.

[245] Nicolas Regnault and Rahul Nandkishore. Floquet thermalization: Symmetries and random matrix ensembles. *Phys. Rev. B*, 93(10):1–5, 2016. doi:10.1103/PhysRevB.93.104203.

[246] Luca D'Alessio and Marcos Rigol. Long-time behavior of isolated periodically driven interacting lattice systems. *Phys. Rev. X*, 4:041048, Dec 2014. URL: https://link.aps.org/doi/10.1103/PhysRevX.4.041048, doi:10.1103/PhysRevX.4.041048.

[247] Achilleas Lazarides, Arnab Das, and Roderich Moessner. Equilibrium states of generic quantum systems subject to periodic driving. *Phys. Rev. E*, 90:012110, Jul 2014. URL: https://link.aps.org/doi/10.1103/PhysRevE.90.012110, doi:10.1103/PhysRevE.90.012110.

[248] Hyungwon Kim, Tatsuhiko N. Ikeda, and David A. Huse. Testing whether all eigenstates obey the eigenstate thermalization hypothesis. *Phys. Rev. E*, 90:052105, Nov 2014. URL: https://link.aps.org/doi/10.1103/PhysRevE.90.052105, doi:10.1103/PhysRevE.90.052105.

[249] B. Vermersch, A. Elben, L. M. Sieberer, N. Y. Yao, and P. Zoller. Probing Scrambling Using Statistical Correlations between Randomized Measurements. *Phys. Rev. X*, 9(2):21061, 2019. doi:10.1103/PhysRevX.9.021061.

[250] Manoj K. Joshi, Andreas Elben, Benoît Vermersch, Tiff Brydges, Christine Maier, Peter Zoller, Rainer Blatt, and Christian F. Roos. Quantum Information Scrambling in a Trapped-Ion Quantum Simulator with Tunable Range Interactions. *Phys. Rev. Lett.*, 124(24):240505, 2020. doi:10.1103/PhysRevLett.124.240505.

[251] Christoph Dankert, Richard Cleve, Joseph Emerson, and Etera Livine. Exact and approximate unitary 2-designs and their application to fidelity estimation. *Phys. Rev. A*, 80:012304, Jul 2009. doi:10.1103/PhysRevA.80.012304.

[252] D. Gross, K. Audenaert, and J. Eisert. Evenly distributed unitaries: On the structure of unitary designs. *J. Math. Phys.*, 48(5):052104, 2007. doi:10.1063/1.2716992.

[253] Francesco Mezzadri. How to generate random matrices from the classical compact groups, 2006. arXiv:math-ph/0609050.

[254] Giacomo Torlai, Christopher J. Wood, Atithi Acharya, Giuseppe Carleo, Juan Carrasquilla, and Leandro Aolita. Quantum process tomography with unsupervised learning and tensor networks, 2020. arXiv:2006.02424.

[255] Joseph Emerson, Robert Alicki, and Karol Życzkowski. Scalable noise estimation with random unitary operators. *J. Opt. B*, 7(10):S347, 2005. doi:10.1088/1464-4266/7/10/021.

[256] Symmetrized characterization of noisy quantum processes. *Science*, 317(5846):1893–1896, jul 2007. URL: http://dx.doi.org/10.1126/science.1145699.

[257] E. Knill, D. Leibfried, R. Reichle, J. Britton, R. B. Blakestad, J. D. Jost, C. Langer, R. Ozeri, S. Seidelin, and D. J. Wineland. Randomized benchmarking of quantum gates. *Phys. Rev. A*, 77:012307, Jan 2008. URL: https://link.aps.org/doi/10.1103/PhysRevA.77.012307, doi:10.1103/PhysRevA.77.012307.

[258] Easwar Magesan, Jay M. Gambetta, B. R. Johnson, Colm A. Ryan, Jerry M. Chow, Seth T. Merkel, Marcus P. da Silva, George A. Keefe, Mary B. Rothwell, Thomas A. Ohki, et al. Efficient measurement of quantum gate error by interleaved randomized benchmarking. *Phys. Rev. Lett.*, 109(8):080505, Aug 2012. URL: http://dx.doi.org/10.1103/PhysRevLett.109.080505.

[259] Alexander Erhard, Joel J. Wallman, Lukas Postler, Michael Meth, Roman Stricker, Esteban A. Martinez, Philipp Schindler, Thomas Monz, Joseph Emerson, and Rainer Blatt. Characterizing large-scale quantum computers via cycle benchmarking. *Nat. Commun.*, 10(1):5347, Nov 2019. URL: http://dx.doi.org/10.1038/s41467-019-13068-7.

[260] Tiff Brydges, Andreas Elben, Petar Jurcevic, Benoît Vermersch, Christine Maier, Ben P. Lanyon, Peter Zoller, Rainer Blatt, and Christian F. Roos. Probing Rényi entanglement entropy via randomized measurements. *Science*, 364(6437):260–263, 2019. doi:10.1126/science.aau4963.

[261] Andreas Elben, Benoît Vermersch, Rick Van Bijnen, Christian Kokail, Tiff Brydges, Christine Maier, Manoj K. Joshi, Rainer Blatt, Christian F. Roos, and Peter Zoller. Cross-Platform Verification of Intermediate Scale Quantum Devices. *Phys. Rev. Lett.*, 124(1):10504, 2020. doi:10.1103/PhysRevLett.124.010504.

[262] E.P. Wigner. Characteristic vectors of bordered matrices with infinite dimensions. JSTOR: Annals of Mathematics, Second Series, Vol. 62, No. 3 (Nov., 1955), pp. 548-564. *Ann. Math.*, 62(3):548–564, 1955. URL: http://www.jstor.org/discover/10.2307/1970079?uid=3738936{&}uid=2{&}uid=4{&}sid=21104185891721.

[263] Freeman J Dyson. Statistical Theory of the Energy Levels of Complex Systems. I. *J. Math. Phys.*, 3(1):140–156, jan 1962. doi:10.1063/1.1703773.

[264] J M Deutsch. Thermodynamic entropy of a many-body energy eigenstate. *New J. Phys.*, 12(7):075021, jul 2010. doi:10.1088/1367-2630/12/7/075021.

[265] Tsung-Cheng Lu and Tarun Grover. Renyi entropy of chaotic eigenstates. *Phys. Rev. E*, 99:032111, Mar 2019. URL: https://link.aps.org/doi/10.1103/PhysRevE.99.032111, doi:10.1103/PhysRevE.99.032111.

[266] Chaitanya Murthy and Mark Srednicki. Structure of chaotic eigenstates and their entanglement entropy. *Phys. Rev. E*, 100:022131, Aug 2019. URL: https://link.aps.org/doi/10.1103/PhysRevE.100.022131, doi:10.1103/PhysRevE.100.022131.

[267] A. De Luca, B. L. Altshuler, V. E. Kravtsov, and A. Scardicchio. Anderson localization on the bethe lattice: Nonergodicity of extended states. *Phys. Rev. Lett.*, 113:046806, Jul 2014. URL: https://link.aps.org/doi/10.1103/PhysRevLett.113.046806, doi:10.1103/PhysRevLett.113.046806.

[268] VE Kravtsov, IM Khaymovich, E Cuevas, and M Amini. A random matrix model with localization and ergodic transitions. *New J. Phys.*, 17(12):122002, 2015. doi:10.1088/1367-2630/17/12/122002.

[269] Davide Facoetti, Pierpaolo Vivo, and Giulio Biroli. From non-ergodic eigenvectors to local resolvent statistics and back: A random matrix perspective. *EPL (Europhysics Letters)*, 115(4):47003, 2016. doi:10.1209/0295-5075/115/47003.

[270] B. L. Altshuler, E. Cuevas, L. B. Ioffe, and V. E. Kravtsov. Nonergodic phases in strongly disordered random regular graphs. *Phys. Rev. Lett.*, 117:156601, Oct 2016. URL: https://link.aps.org/doi/10.1103/PhysRevLett.117.156601, doi:10.1103/PhysRevLett.117.156601.

[271] VE Kravtsov, BL Altshuler, and LB Ioffe. Non-ergodic delocalized phase in anderson model on bethe lattice and regular graph. *Annals of Physics*, 389:148–191, 2018. doi:10.1016/j.aop.2017.12.009.

[272] T. Micklitz, Felipe Monteiro, and Alexander Altland. Nonergodic extended states in the sachdev-ye-kitaev model. *Phys. Rev. Lett.*, 123:125701, Sep 2019. URL: https://link.aps.org/doi/10.1103/PhysRevLett.123.125701, doi:10.1103/PhysRevLett.123.125701.

[273] Felipe Monteiro, Masaki Tezuka, Alexander Altland, David A. Huse, and Tobias Micklitz. Quantum ergodicity in the many-body localization problem. *Phys. Rev. Lett.*, 127:030601, Jul 2021. URL: https://link.aps.org/doi/10.1103/PhysRevLett.127.030601, doi:10.1103/PhysRevLett.127.030601.

[274] Y Avishai, J Richert, and R Berkovits. Level statistics in a Heisenberg chain with random magnetic field. *Phys. Rev. B*, 66(5):052416, 2002. URL: https://doi.org/10.1103/PhysRevB.66.052416.

[275] Winton G Brown, Lea F Santos, David J Starling, and Lorenza Viola. Quantum chaos, delocalization, and entanglement in disordered Heisenberg models. *Phys. Rev. E*, 77(2):021106, 2008. URL: https://doi.org/10.1103/PhysRevE.77.021106.

[276] Takato Yoshimura and Lucas Sá. Robustness of quantum chaos and anomalous relaxation in open quantum circuits. *arXiv preprint arXiv:2312.00649*, 2023. URL: https://doi.org/10.48550/arXiv.2312.00649.

[277] Katherine Sky Collins. *Analog-Digital Quantum Simulations with Trapped Ions*. Doctoral dissertation, University of Maryland, College Park, 2023. URL: https://doi.org/10.13016/pjvv-wrfd.

[278] Ruth Shir, Pablo Martinez-Azcona, and Aurélia Chenu. Full range spectral correlations and their spectral form factors in chaotic and integrable models. *arXiv preprint arXiv:2311.09292*, 2023. URL: https://doi.org/10.48550/arXiv.2311.09292.

[279] Zohar Nussinov and Saurish Chakrabarty. Exact universal chaos, speed limit, acceleration, planckian transport coefficient,"collapse" to equilibrium, and other bounds in thermal quantum systems. *Ann. Phys.*, 443:168970, 2022. URL: https://doi.org/10.1016/j.aop.2022.168970.

[280] Fausto Borgonovi, Felix M Izrailev, and Lea F Santos. Timescales in the quench dynamics of many-body quantum systems: Participation ratio versus out-of-time ordered correlator. *Phys. Rev. E*, 99(5):052143, 2019. URL: https://doi.org/10.1103/PhysRevE.99.052143.

[281] Jochen Gemmer, Alexander Otte, and Günter Mahler. Quantum approach to a derivation of the second law of thermodynamics. *Phys. Rev. Lett.*, 86(10):1927. URL: https://doi.org/10.1103/PhysRevLett.86.1927.

[282] Sandu Popescu, Anthony J Short, and Andreas Winter. The foundations of statistical mechanics from entanglement:Individual states vs. averages. *arXiv preprint arXiv:quant-ph/0511225*. URL: https://doi.org/10.48550/arXiv.quant-ph/0511225.

[283] M V Berry. Regular and irregular semiclassical wavefunctions. *J. Phys. A*, 10(12):2083–2091, dec 1977. doi:10.1088/0305-4470/10/12/016.