Chaos and Randomness in Strongly-Interacting Quantum Systems

Thesis by

Nicholas R. Hunter-Jones

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy in Physics

California Institute of Technology

Pasadena, California

2018

(Defended May 22, 2018)

Acknowledgments

First and foremost I would like to thank my advisor, John Preskill, for his guidance and support over the past few years. I arrived at Caltech wanting to work on high-energy theory and over the course of the first few years took something of a random walk through theoretical physics, having now found a happy home in quantum information. I can only imagine that his patient willingness to accommodate my evolving interest in a broad and disparate panoply of topics stems from a deep and detailed mental atlas of the charted parts of science. I hope to one day understand this sonorous picture and how my interests are truly connected.

Second, I would like to thank Beni Yoshida for in many ways being an unofficial second advisor. I've benefited from patient advice and learned an incredible amount working with him. His ease of insight and intuition have shaped how I think about physics. I'd also like to thank Yoni BenTov for guidance and for pushing me to be precise and think carefully. I'd further like to thank the members of my thesis committee: Fernando Brandão, Alexei Kitaev, and David Simmons-Duffin. All three have a broad and rich knowledge of physics and I'm lucky to have had the opportunity to interact with them.

There have been many people, both at Caltech and outside, from whom I've benefited greatly through discussions about physics, advice about doing physics, and general encouragement. I'd especially like to thank Jason Alicea, Paweł Caputa, Xie Chen, Cliff Cheung, Elizabeth Crosson, Glen Evenbly, Dorota Grabowska, Sergei Gukov, Kristan Jensen, Richard Kueng, Raghu Mahajan, Alex Maloney, Henry Maxfield, Hirosi Ooguri, Fernando Pastawski, Dan Roberts, Burak Şahinoğlu, John Schwarz, Thomas Vidick, and Sasha Zhiboedov. Many thanks to the collaborators with whom I had the pleasure to work in my earlier days at Caltech: Sean Carroll, Ning Bao, Charles Cao, Aidan Chatwin-Davies, Jason Pollack, and Grant Remmen. And a special thanks to Jordan Cotler and Junyu Liu for very enjoyable collaboration on some of the work that makes up this thesis.

I'd also like to thank the Caltech grad students both in the high-energy and quantum information groups, from whom I've learned a lot: Michael Beverland, Thom Bohdanowicz, Matt Fishman, Enrico Herrmann, Matt Heydeman, Murat Kologlu, Petr Kravchuk,

Alex Kubica, Tristan McKinney, Jenia Mozgunov, Ingmar Saberi, Minyoung You, Nicole Yunger Halpern, and especially Dave Aasen, for friendship and inspiring many interests within physics, and Alex Turzillo for friendship, hiking and summer travels, and also inspiring many interests within physics.

I have benefited tremendously from years of teaching at Caltech. Teaching consistently and rapidly exposes the gaps and weaknesses in one's knowledge of a subject. From hydrodynamics to string theory to QFT, every course I've TA'ed at Caltech has given me a much deeper and more intuitive understanding of physics. Thanks to the professors who taught the courses I TA'ed (especially Sergei), and thanks to all the brilliant Caltech students who kept asking questions.

I'd also like to acknowledge support from the IQIM through the NSF and Gordon and Betty Moore Foundation, from the DOE, and from the Simons Foundation through the "It from Qubit" collaboration. I've also been very lucky to have traveled extensively during my time as a grad student. I'd especially like to thank the Perimeter Institute, the Yukawa Institute for Theoretical Physics in Kyoto, TASI and the University of Colorado Boulder, and the Cargèse summer institute for their hospitality and financial support.

Grad school is not without its stresses and hardships. There have been many people over the past few years whose friendship has been a sinew of stability. I will refrain from listing everyone, but know that your friendship meant a lot. I'd especially like to thank Amy for her love and partnership; it has been immensely sustaining.

Lastly, I'd like to thank my parents, Ian and Lynette, and my sister Bridget, for their unrelenting, unabating, and unwavering love and support. My family instilled in me a passion for science, learning, and the pursuit of knowledge that has undoubtedly pushed me to where I am today. To them I owe everything.

Abstract

Quantum chaos entails an entropic and computational obstruction to describing a system and thus is intrinsically difficult to characterize. An understanding of quantum chaos is fundamentally related to the mechanism of thermalization in many-body systems and the quantum nature of black holes. In this thesis we adopt the view that quantum information theory provides a powerful framework in which to elucidate chaos in strongly-interacting quantum systems.

We first push towards a more precise understanding of chaotic dynamics by relating different diagnostics of chaos, studying the time-evolution of random matrix Hamiltonians, and quantifying random matrix behavior in physical systems. We derive relations between out-of-time ordered correlation functions, spectral quantities, and frame potentials to relate the scrambling of quantum information, decay of correlators, and Haar-randomness. We give analytic expressions for these quantities in random matrix theory to explore universal aspects of late-time dynamics. Motivated by our random matrix results, we define k-invariance in order to capture the onset of random matrix behavior in physical systems.

We then refine our diagnostics in order to study chaotic systems with symmetry by considering Haar-randomness with respect to quotients of the unitary group, and in doing so we generalize our quantum information machinery. We further consider extended random matrix ensembles in the context of strongly-interacting quantum systems dual to black holes. Lastly, we study operator growth in classes of random quantum circuits.

Published Contents

A large portion of this thesis is adapted from papers that have appeared as preprints on the arXiv and as publications in journals. Both the DOI number linking to the journal and a link to the arXiv version are given in each published reference. For all publications, all authors contributed equally.

Chapter 2 is essentially the same as:

- J. Cotler, N. Hunter-Jones, J. Liu, B. Yoshida, "Chaos, Complexity, and Random Matrices," *JHEP* **11** (2017) 048, arXiv:1706.05400 [hep-th].

Chapter 3 is partially based on work in progress with Beni Yoshida, which will appear elsewhere in a future publication:

- N. Hunter-Jones and B. Yoshida, "Late-time chaos and symmetry," *To appear.*

Chapter 4 is essentially the same as:

- N. Hunter-Jones, J. Liu, "Chaos and random matrices in supersymmetric SYK," *JHEP* **05** (2018) 202, arXiv:1710.08184 [hep-th].

Chapter 5 is partially based on work in progress, which will appear elsewhere in a future publication:

- N. Hunter-Jones, "Operator growth in random circuits with symmetry," *To appear.*

Contents

Chapter 1

Introduction

Strongly-interacting quantum systems are inherently difficult to describe. The lack of a deep understanding of chaotic dynamics is at the heart of many open questions in theoretical physics. Any progress towards defining quantum chaos should shed light on both thermalization and transport in quantum many-body systems and black hole information loss. As such, the resurgence of interest in quantum chaos over the past few years has emerged as a confluence of ideas from quantum field theory, condensed matter physics, and quantum gravity, incorporating many old ideas from each subfield and inspiring many new seemingly disparate avenues of inquiry. Broadly speaking, the ambitious goal of this line of research is to answer the following question: How can we characterize quantum chaos in physical systems?

In this thesis we will use tools and ideas from quantum information theory to study and understand chaos in quantum many-body systems. Quantum chaos is a generic feature of strongly-interacting systems, but much of our understanding is phenomenological and precise definitions are not immediately clear. Having the spectral statistics of a random matrix and chaotically decaying correlation functions are seen as a strong indication of quantum chaos, but what exactly is being diagnosed by these diagnostics demands clarification.

The approach taken in this thesis stems from the observation that one can learn a lot about the universal aspects of a system by asking how it processes quantum information. For instance, Hayden and Preskill [1] considered a simple random unitary model of a black hole to demonstrate how rapidly information is scrambled and emitted. The suggestion that

black holes are the fastest scramblers [2, 3] has lead to new measures of chaos in quantum systems [4, 5, 6] and soluble models of black holes and holography [6, 7]. Connections have been explored with more traditional notions of chaos [8, 9], and quantum information ideas have begun to shed light on the connections between chaos, scrambling, and random matrices [10, 11, 12].

Over the last decade, it has become evident that quantum information provides a formidable conceptual framework for thinking about quantum field theory and holography. Examples include entropic derivations of c-theorems, constraints from holographic entanglement entropy, quantum error correction in bulk reconstruction of local operators, the averaged null-energy condition from relative entropy, as well as constructing tensor networks to realize salient features of the bulk geometry. More broadly, quantum information is a cornerstone of modern condensed matter and many-body physics, where entanglement and complexity have played an essential role in the classification of topological phases, the development of tensor network methods, and an understanding of thermalization.

With our motivation in mind, we will first overview a few symptoms of chaotic dynamics. This will set the stage for the ideas we explore in the bulk of this work. We then provide a brief summary of the work done in each chapter of this thesis.

1.1 Symptoms of chaotic dynamics

Interest in quantum chaos has existed for some number of years, where the working definition of quantum chaos seems to vary between decades and subfields. Very broadly speaking, the picture evolved from quantizing a classically chaotic system to studying universal correlations in the spectrum to, more recently, studying correlation functions in thermal states.

Before embarking on a historical tour, there are a few admissions we must make. First, although connections do exist in specific few-body systems, quantum chaos is different from classical chaos. As was emphasized in [13], there is a dishonesty even in referring to the object of study as 'quantum chaos' because it suggests quantum phenomena related to classically chaotic behavior.

Classical chaos refers to an exponential sensitivity of phase space trajectories to small

perturbations in initial conditions, but (at least naively) the classical notion of chaos does not apply to quantum systems: Schrödinger's equation is linear, unlike the nonlinear equations in classical mechanics and fluid dynamics that given rise to unpredictable behavior in deterministic systems. Moreover, the overlap of quantum states evolved with the same Hamiltonian is constant in time and uncertainty relations obfuscate the notion of sensitivity to small changes in phase space coordinates. This is all to say that quantum chaos is different from classical chaos.

There is also another dishonesty in calling certain quantities 'probes' or 'diagnostics' of chaos, because it makes it sound like there is a definition of the thing you are trying to diagnose. As we mentioned above, a precise definition of quantum chaos in many-body systems remains elusive. In this thesis, we will continue to refer to these measures as 'symptoms' of chaos, with the implication that they are symptoms of (possibly many) underlying physical phenomena in quantum systems which we still do not quite understand.

We will first overview some symptoms of quantum chaos, i.e. quantities we can compute that make manifest the underlying chaotic dynamics. This includes operator growth, random matrix statistics, chaotic correlation functions, scrambling of quantum information, and Haar-randomness. This list is far from exhaustive, nor are the notions distinct. Here we simply focus on a few of the ideas that will come up repeatedly in this thesis. Furthermore, we will try and be precise, but our goal will be to present an intuitive picture. As such, many details and subtleties will be elided, but will be addressed more thoroughly in the body of the thesis.

Operator growth

Although the subject of quantum chaos really started with the study of universality from random matrices, the first diagnostic of chaos we will discuss is operator growth, as it is in some sense the most intuitive. The growth of operators in local quantum spin systems has a long history, where nested commutation with the Hamiltonian can be used to bound the propagation of information [14, 15], but our discussion here is motivated a little differently and applies more generally to non-local Hamiltonians, largely following [16, 10].

Consider a few-body operator $\mathcal{O}$ acting only on one (or some non-extensive) number of

sites, where $\mathcal{O}$ is not necessarily geometrically local. Working in the Heisenberg picture, the time-evolved operator,

$$\mathcal{O}(t) = e^{iHt}\mathcal{O}e^{-iHt}\,, \tag{1.1}$$

can be expanded and written as a sum of nested commutators

$$\mathcal{O}(t) = e^{iHt}\mathcal{O}e^{-iHt} = \mathcal{O} - it[H,\mathcal{O}] - \frac{t^2}{2}[H,[H,\mathcal{O}]] + \frac{it^3}{3!}[H,[H,[H,\mathcal{O}]]]\ldots \tag{1.2}$$

If the Hamiltonian is built out of few-body interactions, then the nested commutators will generate larger and larger operators. To be more precise, we want to consider the growing operator written in a basis of operators. For a system of n qubits, where $\mathcal{H} = (\mathbb{C}^2)^{\otimes n}$ and $d = \dim \mathcal{H} = 2^n$, we can always write an operator as a sum of Pauli strings

$$\mathcal{O}(t) = \sum_p \gamma_p(t)\mathcal{O}_p\,, \tag{1.3}$$

where $\{\mathcal{O}_p\}$ is the set of Pauli strings formed by tensoring together Pauli operators for each qubit. Pauli strings are a complete orthonormal basis for the algebra of operators on $\mathcal{H}$, with their own inner product $\frac{1}{d}\mathrm{Tr}\,\mathcal{O}_p\mathcal{O}_q = \delta_{p,q}$. More generally, we could consider systems built out of qudits or fermions, where each also gives a basis of operators for $\mathcal{H}$, but we assume you are given a natural tensor decomposition of your Hilbert space in terms of the local degrees of freedom at hand. For simplicity, we will continue to refer to the elements in the basis of operators as Pauli strings.

It is important to note that the operator norm is conserved under unitary time evolution

$$\frac{1}{d}\mathrm{Tr}\left(\mathcal{O}(t)^\dagger\mathcal{O}(t)\right) = \frac{1}{d}\mathrm{Tr}\left(\mathcal{O}^\dagger\mathcal{O}\right) = 1\,, \tag{1.4}$$

for an appropriately normalized operator. This fact, and the orthonormality of Pauli strings, gives that the sum of the squares of the coefficients γ_p is conserved

$$\sum_p |\gamma_p(t)|^2 = 1\,, \tag{1.5}$$

a constraint on operator growth simply arising from unitary dynamics.

It is in this sense $|\gamma_p(t)|^2$ is the probability of finding an operator $\mathcal{O}_p$ in the growing operator at a time t. We can also talk about the weight of an operator on strings of a given length, simply by summing all coefficients for Pauli strings of length ℓ, $wt_\ell(\mathcal{O}(t)) = \sum_{|p|=\ell} |\gamma(t)|^2$. From Eq. (1.5), we have that the sum of the weights from length 1 to length n Pauli strings must be one. For a growing operator $\mathcal{O}(t)$ in generic chaotic systems, the weight of the operator will shift to longer and longer Pauli strings, eventually saturating so that almost all of the weight of the operator is on strings of length n, the size of the system.

If our H is the Hamiltonian of a spin system and $\mathcal{O}$ is a local Pauli operator, say Z_j the Pauli Z operator on site j, then repeated commutation will generate longer and longer Pauli strings. For instance, consider the canonical example of a 2-local spin chain Hamiltonian on n spins with Ising-type nearest neighbor interactions and parallel and transverse fields $H = -\sum_i Z_i Z_{i+1} - h\sum_i X_i - g\sum_i Z_i$. We now look at the growth of the operator Z_j from the nested commutators in Eq. (1.2). The first commutator of H with the operator Z_j generates Y_j. The second commutator gives terms X_j, Z_j, $X_j Z_{j+1}$, and $Z_{j-1} X_j$, where we now get contributions from length 2 Pauli strings. As we time evolve, and the dominant behavior of the operator $Z_j(t)$ shifts to higher nested commutators, the weight of the operator shifts to longer Pauli strings.

This is the simplest manifestation of the butterfly effect: small perturbations growing to affect the system. A simple local operator time evolved with a chaotic Hamiltonian will quickly grow to an operator the size of the system.

Chaos and random matrices

The diagnostic most firmly established in the history of quantum chaos is rooted in random matrix theory. Random matrix theory arose from nuclear physics as a statistical approach to understanding the energy levels of complex atomic nuclei [17, 18]. Having the spectral statistics of a random matrix has even been proposed as the defining feature of a quantum chaotic system [8]. Very roughly speaking, the intuition is that the structure of high energy states of a chaotic Hamiltonian is so complicated that certain properties simply behave as if the Hamiltonian were itself a random matrix. More precisely speaking, there are universal

aspects of the eigenvalue correlations in a chaotic Hamiltonian.

The simplest universal features of chaotic spectra are nearest-neighbor correlations. If you compute the differences between neighboring eigenvalues and make a histogram plot, then for a chaotic system the resulting distribution is Wigner-Dyson, which goes to zero for small energy differences, i.e. the energy levels 'repel' and eigenvalues of a chaotic system are not degenerate. Again, the rough intuition is that we need many symmetries in order to tune eigenvalues to sit on top of one another, and an integrability's worth to have an extensive number of degeneracies. This phenomenon of level repulsion was famously observed in heavy atomic nuclei and has since been understood in a broad array of quantum chaotic systems; see [19, 20, 21] and references therein.

However, universal correlations in chaotic spectra extend far beyond neighboring eigenvalues. Eigenvalues in the spectra of chaotic systems experience long-range repulsion, which causes the eigenvalues to be anticorrelated and gives rise to spectral rigidity. An incredibly useful quantity that allows us to see universal behavior in the spectral statistics is the spectral form factor $\mathcal{R}_2(\beta, t)$, which can be defined in terms of the analytically continued partition function $Z(\beta, t) = \mathrm{Tr}(e^{-\beta H - itH})$ as

$$\mathcal{R}_2(\beta, t) \equiv \left\langle Z(\beta, t) Z^*(\beta, t) \right\rangle . \tag{1.6}$$

Here $\langle \cdot \rangle$ denotes the average over an ensemble over Hamiltonians—for instance, a disordered quantum system or an ensemble of random matrices. We will address this further when the quantity is discussed later in the thesis. This quantity is equivalently defined as the Fourier transform of the 2-point spectral correlation function $\langle \rho(\lambda)\rho(\lambda') \rangle$, where $\rho(\lambda)$ is the spectral density. The spectral form factor has been studied extensively in the random matrix literature [18, 22, 23] and especially [24, 25, 26], but has appeared more recently in the context of the Sachdev-Ye-Kiteav model [9].

We plot the spectral form factor as defined above for a generic chaotic system in Fig. 1.1. As we time-evolve the quantity from its initial value, it will exponentially decay and then might experience a power-law decay, although the details of this depend on the system (specifically on the spectral density). At an intermediate time-scale called the dip time, we

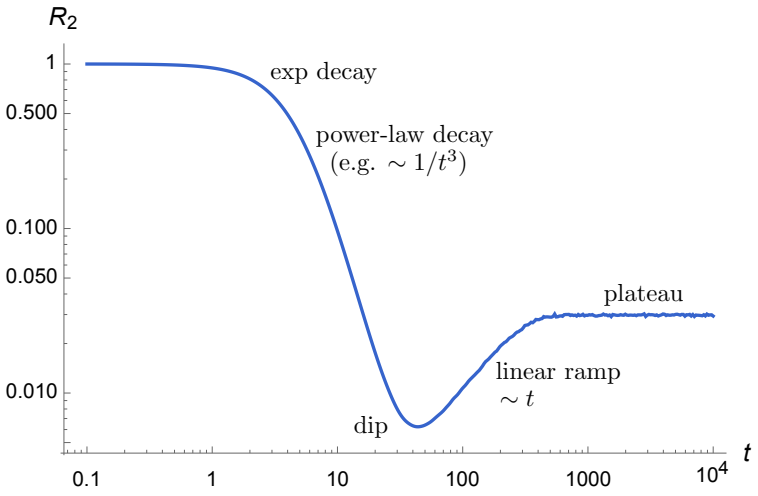


Figure 1.1: The 2-point spectral form factor $\mathcal{R}_2$, where we have denoted some of the important behaviors in a chaotic quantum system.

transition into a linearly increasing ramp, then level-off into a plateau. The late-time plateau is the result of eigenvalue repulsion: late times corresponds to small energy differences, so in the absence of degeneracies the form factor becomes constant once we start probing scales smaller than the mean spacing. The linear ramp is the consequence of long-range spectral correlations, a universal signature of random matrix statistics.

There are a few ways of deriving this linear ramp in random matrix theory; one is from the logarithmic repulsion of eigenvalues. In a saddle-point approximation, this gives rise to a linear growth at early times. The linear growth more generally arises from the sine-kernel [22] from which we compute the spectral 2-point functions. We introduce and discuss this further in Ch. 2, but the point is that a linear growth in the connected piece of the spectral 2-point function is an important signature of random matrix statistics in the spectrum and a universal feature of quantum chaotic systems.

Out-of-time order correlation functions

A recent surge of interest in quantum chaos has revolved around the out-of-time-ordered correlation function (OTOC), a 4-point function of a pair of operators evaluated in thermal states

$$\langle AB(t)AB(t)\rangle_\beta\,, \tag{1.7}$$

where A and B are few-body operators and $B(t) = e^{-iHt}Be^{-iHt}$. Black holes are chaotic and fast scrambling systems [1, 2]. It was further understood by Shenker and Stanford [4, 5] as well as Kitaev [6] that black holes are maximally chaotic in the sense that a bound on the early-time behavior of the OTOC is saturated [27]. The OTOC is actually a somewhat old idea, dating back to work by Larkin and Ovchinnikov [28], who considered electron/impurity scattering in superconductors and wrote down a Green's function that captured something chaotic about the nature of the scattering, relating it to diverging semiclassical trajectories for the electrons. We now understand this Green's function in terms of the augmented contour prescription behind the OTOC.

The OTOC has been used to probe quantum chaos in many different systems. It can be computed in holographic systems as a bulk scattering process in a shockwave background [4, 5], with a universal exponential early-time growth governed by the exponent $\lambda_L = 2\pi/\beta$. The OTOC can also be computed in holographic CFTs where a contour prescription for out-of-time ordering allows one to extract the same universal exponent from the vacuum Virasoro block [29]. As we mentioned, a bound on this chaotic exponent $\lambda_L \leq 2\pi/\beta$ [27] means that black holes (and holographic systems) are maximally chaotic.

Separately, Kitaev proposed a solvable model of strongly-interacting Majorana fermions that reproduces many features of gravitational systems and black holes [6], namely an emergent reparamentrization invariance and a zero-temperature entropy. Futhermore, an explicit calculation of the OTOC shows that the theory also saturates the chaos bound [6, 7, 30].

In general, the out-of-time-ordered correlation function will decay in a chaotic system. In a large N system, this is characterized by an early-time exponential growth of corrections to the initial value, suppressed by a small parameter, say $1/N$ or G_N, and growing exponentially in time with an exponent λ_L. At a time of order $t \sim \beta \log N$, these suppressed terms become large and the OTOC decays. At very late times, the OTOC decays to something small. We plot these features in Fig. 1.2.

The decay of the OTOC can be understood as operator growth [16]. We described before that an operator evolved by a chaotic Hamiltonian grows to longer and longer Pauli strings. For a local operator B, consider the commutator of the time-evolved operator $B(t)$ with another operator A which has support on some distant site: $[A, B(t)]$. For QFTs the

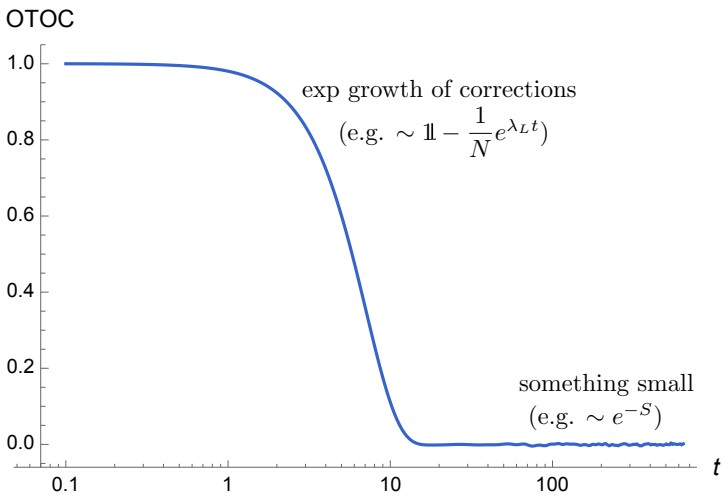


Figure 1.2: The out-of-time-ordered correlation function (OTOC), where we denote the expected behaviors in a chaotic quantum system.

commutator is zero at time $t = 0$ as space-like separated operators commute. For spin systems, Paulis at distant sites with non-overlapping support commute. As we start time-evolving, the operator $B(t)$ grows to extend over the system as its weight shifts to longer and longer Pauli strings. The failure of A to commute with B at some time is captured by the decay of the OTOC. Consider the squared commutator

$$\langle [A, B(t)]^2 \rangle_\beta \,, \tag{1.8}$$

where we take the Hermitian conjugate. Expanding the quantity out into four 4-point functions, we find that two of them are in the out-of-time ordering and two are trivial

$$\langle [A, B(t)]^2 \rangle_\beta = \langle AB(t)B(t)A \rangle_\beta + \langle B(t)AAB(t) \rangle_\beta - \langle AB(t)AB(t) \rangle_\beta - \langle B(t)AB(t)A \rangle_\beta \,, \tag{1.9}$$

where we are assuming that the operators A and B are Hermitian. If A and B are Pauli operators, then the first two terms are simply 1. More generally, in a generic strongly-coupled many-body system or QFT, the first two correlators will quickly decay to their disconnected components $\langle AA \rangle \langle BB \rangle$. For the sake of intuition, it is helpful to think of these terms as the norm of a perturbed thermal state, i.e. $\| B(t)A \, |\beta\rangle \|$, which becomes constant as the perturbation allays. Whereas the OTO terms are more like the inner product of two

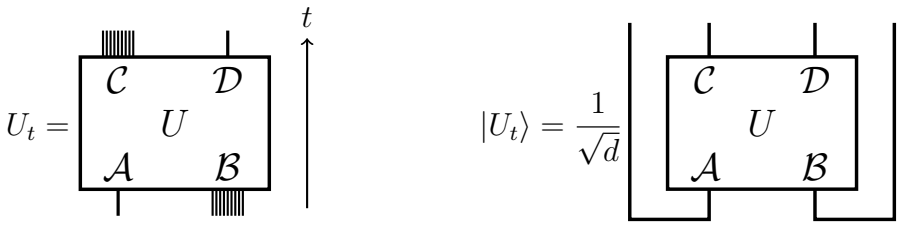


Figure 1.3: On the left: evolution by a unitary operator, or more generally a channel, denoting the input and output subsystems. On the right: the state representation of the unitary on the doubled Hilbert space.

different states $AB(t)\,|\beta\rangle$ and $B(t)A\,|\beta\rangle$, where we act with A in different ways to affect the growth of the time-evolved operator $B(t)$. The statement that these terms decay is a manifestation of the butterfly effect, acting with A has drastically changed the state and the inner product is small. Thus, operator growth is one of the many things captured by the OTOC.

Scrambling of quantum information

Scrambling is the statement that local quantum information becomes delocalized under unitary evolution and spreads over the entire system [1, 2], and information about local degrees of freedom can no longer be accessed by local measurements. The unitary evolution U_t might be time-evolution by a Hamiltonian, application of a quantum circuit, and more generally might be a quantum channel. Scrambling is usually discussed with respect to a specific initial state. The most common notion of scrambling is that, starting with a simple state $|\psi\rangle$, the evolved state $U_t\,|\psi\rangle$ looks nearly maximally mixed on arbitrary subsystems (which constitute less than half of the total system).

Consider a unitary that acts on many qubits (or qudits), with total Hilbert space dimension d. As shown in Fig. 1.3, we divide the input degrees of freedom into subsystems $\mathcal{A}$ and $\mathcal{B} = \mathcal{A}^c$, and the output into subsystems $\mathcal{D}$ and $\mathcal{C} = \mathcal{D}^c$, where $\mathcal{B}$ and $\mathcal{D}$ are somewhat larger than their complements. Scrambling is the statement that we cannot reconstruct the local information in $\mathcal{A}$ with any arbitrary subsystem $\mathcal{D}$ in the output, i.e. the state is scrambled as we cannot learn about it by performing local measurements.

Let's be a little more precise. We want to talk about the scrambling properties of the

chaotic unitary itself. Following [10], we employ the state representation of the operator $|U\rangle$ [31, 32], considering the operator U as a state on the doubled Hilbert space

$$|U_t\rangle = \mathbb{I} \otimes U_t \, |\text{EPR}\rangle \ , \quad \text{where} \quad |\text{EPR}\rangle = \frac{1}{\sqrt{d}} \sum_j |j\rangle \, |j\rangle \ . \tag{1.10}$$

The state representation of our operator U_t with the subsystems labeled is shown in Fig. 1.3. We can now talk about the entropies of subsystems in the 4-partitite state $|U_t\rangle$. We say that the information is delocalized by the unitary if the mutual information $I(\mathcal{A}, \mathcal{D})$ is small, and thus we cannot learn about local information in $\mathcal{A}$ by acting locally on the output. Moreover, we say that the chaotic unitary has scrambled the quantum information in the input if the mutual information $I(\mathcal{A}, \mathcal{CD})$ is large, i.e. the information about $\mathcal{A}$ has spread nonlocally over the entire output. Equivalently, the mutual information $I(\mathcal{A}, \mathcal{BD})$ becoming large implies we can reconstruct the state once it has scrambled over the system by accessing the other input degrees of freedom in $\mathcal{B}$.

We would like to highlight one relation between the decay of 2-point functions and delocalization of quantum information, as well as the decay of 4-point functions and the scrambling of quantum information. This is primarily due to the ideas and techniques developed in [10], but appeared as discussed here in [12] and reproduced in this thesis in Sec. 2.B. We will simply summarize here and refer the reader to these papers for details. Moreover, the physics here was essentially understood and explained in [1].

Again, consider a unitary operator which takes input subsystems $\mathcal{A}$ and $\mathcal{B}$ to output subsystems $\mathcal{C}$ and $\mathcal{D}$, where the subsystems $\mathcal{A}$ and $\mathcal{D}$ have dimensions $d_\mathcal{A}$ and $d_\mathcal{D}$. Employing the state representation of the operator U and averaging over a basis of operators on the subsystems $\mathcal{A}$ and $\mathcal{D}$, we can relate the 2-point functions to the mutual information as

$$\int d\mathcal{O}_\mathcal{A} d\mathcal{O}_\mathcal{D} \left| \langle \mathcal{O}_\mathcal{A} \mathcal{O}_\mathcal{D}(t) \rangle \right|^2 = \frac{1}{d_\mathcal{A}^2 d_\mathcal{D}^2} e^{I^{(2)}(\mathcal{A},\mathcal{D})} \, , \tag{1.11}$$

where $I^{(2)}(\mathcal{A}, \mathcal{D})$ is the Rényi-2 mutual information, $I^{(2)}(\mathcal{A}, \mathcal{D}) \equiv S^{(2)}_\mathcal{A} + S^{(2)}_\mathcal{D} - S^{(2)}_\mathcal{AD}$, in terms of reduced density matrices in the Choi state $|U\rangle$. For local operators on subsystems $\mathcal{A}$ and $\mathcal{D}$ the decay of 2-point functions occur after the thermalization time $t \sim \beta$. Here the

decay of 2-point functions means that the mutual information between $\mathcal{A}$ and $\mathcal{D}$ systems becomes small and the information has delocalized. In the language of [1], this is Bob failing to reconstruct Alice's state on $\mathcal{A}$ with only the Hawking radiation $\mathcal{D}$.

Now we consider the 4-point function averaged over the operators on subsystems $\mathcal{A}$ and $\mathcal{D}$. Again using the state representation of the operator, we can relate the OTO 4-point functions to the mutual information as

$$\int d\mathcal{O}_{\mathcal{A}} d\mathcal{O}_{\mathcal{D}} \, \langle \mathcal{O}_{\mathcal{A}} \mathcal{O}_{\mathcal{D}}(t) \mathcal{O}_{\mathcal{A}} \mathcal{O}_{\mathcal{D}}(t) \rangle = e^{-I^{(2)}(\mathcal{A}, \mathcal{B}\mathcal{D})} \,, \tag{1.12}$$

where now we find the Rényi mutual information between the subsystems $\mathcal{A}$ and $\mathcal{B}\mathcal{D}$. For local operators on subsystems $\mathcal{A}$ and $\mathcal{D}$, we expect that for chaotic systems, the OTOCs will decay after a scrambling time $t \sim \beta \log N$, the mutual information between $\mathcal{A}$ and $\mathcal{B}\mathcal{D}$ becomes large. This means that the local information in $\mathcal{A}$ has scrambled over the entire system. This also has the interpretation that Bob can reconstruct Alice's state on $\mathcal{A}$ by accessing both the early-time radiation $\mathcal{B}$ and the emitted quanta $\mathcal{D}$. The chaotic decay of 2-point functions and OTO 4-point functions imply that a chaotic unitary can be viewed as a quantum error-correcting code, where the quantum information on the $\mathcal{A}$ subsystem is delocalized and encoded nonlocally in the $\mathcal{C}$ and $\mathcal{D}$ subsystems. When the unitary is viewed as a four-partite state, that the mutual information between $\mathcal{A}$ and any two of the other subsystems is the same, implies we can tolerate the erasure of any one subregion and still reconstruct the state on $\mathcal{A}$.

It is in this sense a chaotic unitary, and black holes, may be seen as a quantum error-correcting code, and is one way to relate the chaotic decay of correlation functions to the delocalization and scrambling of quantum information.

Haar-randomness

The last measure of chaotic dynamics we will discuss, and one that plays a primary role in this thesis, is the randomness of the unitary evolution itself. Consider the unitary time evolution by a chaotic Hamiltonian $e^{-iHt} \in U(d)$. To extract universal aspects of late-time dynamics, we often average over the unitary group [1, 11]. Roughly speaking, the intuition

should be that chaotic time-evolution spreads randomly over the unitary group and, at late-times, we can extract universal behavior by averaging over $U(d)$. More precisely, we are interested an ensemble of unitaries generated by (an ensemble of) Hamiltonians at a time t

$$\mathcal{E}_t = \{e^{-iHt}, \quad H \in \mathcal{E}_H\}, \tag{1.13}$$

where the $\mathcal{E}_H$ might be a disordered spin-system, the Sachdev-Ye-Kitaev model, or a random matrix ensemble. We want to understand how random the ensemble $\mathcal{E}_t$ with respect to the Haar measure, the invariant measure on the unitary group. As Haar-random unitaries require exponential circuit complexity, we do not expect that physical time-evolution becomes exactly Haar-random. But we can quantify how random the ensemble is with the notion of a k-design.

A unitary k-design is a subset of the unitary which reproduces averages over the unitary group, capturing the first k moments of the Haar ensemble. As measure of randomness we will discuss many times in this thesis is the frame potential $\mathcal{F}_{\mathcal{E}}^{(k)}$, a quantity we can compute for any ensemble of unitaries. $\mathcal{F}_{\mathcal{E}}^{(k)}$ is lower-bounded by the Haar value, with equality if and only if we form a k-design, and in this sense defines a distance to randomness. We are interested in the time-scales at which the ensemble $\mathcal{E}_t$ becomes Haar-random and forms a k-design.

Being Haar-random relates to other measures of chaos we have discussed so far. A 2-design looks maximally mixed on subsystems, as for Haar-random unitaries in [33], and scrambles by achieving decoupling in the sense of [1]. Moreover, by Haar-averaging over random operators and random Hamiltonians, we will be able to relate OTOCs, the spectral form factor, and the frame potential, as we will explore in more detail in Ch. 2.

Outlook

As we stressed before, the above list of symptoms of chaos is certainly not exhaustive; we have not mentioned entanglement production and the growth of entanglement entropy or complexity, although both will be discussed in this thesis. Nor are the symptoms distinct; many of the diagnostics can be understood as related to one another, as we have already

discussed. Understanding the relations between these symptoms of chaos is part of the goal of this thesis. Moreover, we have not mentioned other aspects of real-time dynamics in strongly-interacting many-body systems, namely, thermalization and transport. There is still much to be done in order to understand the connection between chaos, transport, and thermalization. Nevertheless, it seems likely that quantum information will continue to play a central role in our understanding of quantum matter and quantum gravity by shedding light on chaotic dynamics. Being of the opinion that some of the deepest insights in physics come from a chaotic union of subfields, we view the liminal nature of the field as exciting.

1.2 Outline of the thesis

Motivated by many of the ideas discussed in this introduction, we will use tools and ideas from quantum information theory to study the universal aspects of chaos and randomness in strongly-interacting quantum systems.

In Chapter 2, we relate different diagnostics of chaos, study the time-evolution of random matrix Hamiltonians, and define a measure of random matrix behavior in physical systems. We will connect the notions of chaotic correlation functions, randomness, information scrambling, complexity, and random matrix behavior by deriving relations between probes of chaos, namely: OTOCs, spectral correlators, and the frame potential—a measure of Haar-randomness. We also consider the time-evolution by Gaussian random matrix Hamiltonians and analytically compute the spectral form factors, OTOCs, and frame potentials to quantify chaos and scrambling. We give a precise sense in which early-time random matrix evolution is unphysical, showing that random matrix evolution fails to capture operator growth and Lyapunov behavior, both expected features of chaotic Hamiltonians. Of more interest, we find that at an intermediate time-scale random matrix evolution forms a unitary k-design, but at late times deviates from this and becomes less Haar-random. This motivated us to introduce k-invariance, defined as the difference between the frame potential of an ensemble of unitary time-evolutions and its Haar-invariant counterpart. This is a computable measure of late-time chaos for physical systems, where the difference becoming small indicates the onset of a random matrix description.

In Chapter 3, we extend many of the tools developed in the previous chapter to consider chaotic systems with symmetry. Specifically, we consider Haar-average over different quotients of the unitary group making use of Weingarten calculus for more general Lie groups and compact symmetric spaces. The averages over the different spaces correspond to the universal aspects of different symmetry classes of Hamiltonians. We compute the OTOCs and frame potentials of the different symmetry classes to describe the universal aspects of late-time chaos in systems with symmetry. We also generalize the notion of a k-design to different Lie groups and compact symmetric spaces, making precise what it means to be random with respect to a subgroup or quotient of the unitary group. We further generalize k-invariance to ensembles with symmetry. As a check, we study the time evolution of time-reversal invariant random matrix Hamiltonians and identify the timescales at which they form symmetric k-designs. We also consider subsystems of random states with symmetry. Lastly, we numerically investigate k-invariance in spin-systems and in SYK, a first step towards making precise the connection between the time scales at which we achieve k-invariance and scrambling.

In Chapter 4, we extend our study of quantum chaos to systems with symmetry in a different way, considering extended random matrix ensembles. The motivation is to investigate chaos in supersymmetric SYK, where we compute the form factors and frame potentials for Wishart random matrices and find agreement with the model. This highlights a distinction between early-time chaos in the chaotic decay of OTO correlators, and late-time chaos in terms of Haar-randomness and scrambling.

In Chapter 5, we explore a simple application of the tools developed earlier in this thesis by considering quantum circuits built out of random unitaries drawn from quotients of the full unitary group. Random quantum circuits are simple minimal models to understand operator growth and the emergence of dissipative hydrodynamics. We derive the transition probabilities of the Markov process governing operator growth in five classes of symmetric random circuits and then compute the butterfly velocities and diffusion constants for a spreading operator by solving a simple random walk in each class of circuits.

Chapter 2

Chaos, complexity, and random matrices

This chapter is essentially the same as

- J. Cotler, N. Hunter-Jones, J. Liu, B. Yoshida, "Chaos, Complexity, and Random Matrices," *JHEP* **11** (2017) 048, arXiv:1706.05400 [hep-th].

Abstract

In this chapter, we consider time evolution by Gaussian Unitary Ensemble (GUE) Hamiltonians and analytically compute out-of-time-ordered correlation functions (OTOCs) and frame potentials to quantify scrambling, Haar-randomness, and circuit complexity. While our random matrix analysis gives a qualitatively correct prediction of the late-time behavior of chaotic systems, we find unphysical behavior at early times including an $\mathcal{O}(1)$ scrambling time and the apparent breakdown of spatial and temporal locality. The salient feature of GUE Hamiltonians which gives us computational traction is the Haar-invariance of the ensemble, meaning that the ensemble-averaged dynamics look the same in any basis. Motivated by this property of the GUE, we introduce k-invariance as a precise definition of what it means for the dynamics of a quantum system to be described by random matrix theory. We envision that the dynamical onset of approximate k-invariance will be a useful tool for capturing the transition from early-time chaos, as seen by OTOCs, to late-time chaos, as seen by random matrix theory.

2.1 Introduction

Quantum chaos is a general feature of strongly-interacting systems and has recently provided new insight into both strongly-coupled many-body systems and the quantum nature of black holes. Even though a precise definition of quantum chaos is not at hand, understanding how chaotic dynamics process quantum information has proven valuable. For instance, Hayden and Preskill [1] considered a simple model of random unitary evolution to show that black holes rapidly process and scramble information. The suggestion that black holes are the fastest scramblers in nature [2, 3] has led to a new probe of chaos in quantum systems, namely the 4-point out-of-time-order correlation function (OTOC). Starting with the work of Shenker and Stanford [4, 5], it was shown [27] that black holes are maximally chaotic in the sense that a bound on the early time behavior of the OTOC is saturated. Seperately, Kitaev proposed a soluble model of strongly-interacting Majorana fermions [34, 35], which reproduces many features of gravity and black holes, including the saturation of the chaos bound [6, 7]. The Sachdev-Ye-Kitaev model (SYK) has since been used as a testing ground for questions about black hole information loss and scrambling.

In recent work, [9] found evidence that the late time behavior of the SYK model can be described by random matrix theory, emphasizing a dynamical perspective on more standard notions of quantum chaos. Random matrix theory (RMT) has its roots in nuclear physics [17, 18] as a statistical approach to understand the spectra of heavy atomic nuclei, famously reproducing the distribution of nearest neighbor eigenvalue spacings of nuclear resonances. Random matrix theory's early success was later followed by its adoption in a number of subfields, including large N quantum field theory, string theory, transport in disordered quantum systems, and quantum chaos. Indeed, random matrix eigenvalue statistics have been proposed as a defining characteristic of quantum chaos, and it is thought that a generic classically chaotic system, when quantized, has the spectral statistics of a random matrix ensemble consistent with its symmetries [8].

Current thinking holds that both spectral statistics and the behavior of the OTOC serve as central diagnostics of chaos, although the precise relation between the two is unclear. OTOCs have recently been studied using techniques from quantum information theory, and

it was found that their decay as a function of time quantifies scrambling [10] and randomness [11]. The goal of this chapter is to connect various concepts as a step towards a quantum information-theoretic definition of quantum chaos that incorporates scrambling, chaotic correlation functions, complexity, approximate randomness, and random matrix universality.

As alluded to above, an important first step to bridge early-time chaos and late-time dynamics is to understand the relation between the OTOC and the spectral statistics. We derive an explicit analytical formula relating certain averages of OTOCs and spectral form factors which holds for arbitrary quantum mechanical systems. A simple corollary is that spectral form factors can be approximated by OTOCs defined with respect to random (typically non-local) operators, highlighting the fact that spectral statistics are good probes of macroscopic thermodynamic properties, but may miss important microscopic physics such as early-time chaos. We also compute correlation functions for an ensemble of Hamiltonians given by the Gaussian Unitary Ensemble (GUE), and find that 4-point OTOCs decay faster than 2-point correlators contrary to findings for local quantum Hamiltonians [27]. Due to the basis independence of the GUE, averaged correlation functions do not depend on sizes of operators, and thus can be expressed solely in terms of spectral form factors. Furthermore, we find that correlators for GUE Hamiltonians do not even depend on the time-ordering of operators. These results imply that the GUE ignores not only spatial but also temporal locality.

Another important question is to understand the approach to entropic (as well as quantum complexity) equilibrium via pseudorandomization at late times in strongly coupled systems. We consider the ensemble of unitaries generated by fixed GUE Hamiltonians, namely

$$\mathcal{E}_t^{\mathrm{GUE}} = \left\{ e^{-iHt},\ \text{for } H \in \mathrm{GUE} \right\}, \tag{2.1}$$

and study its approach to Haar-randomness by computing frame potentials which quantify the ensemble's ability to reproduce Haar moments. We find that the ensemble forms an approximate k-design at an intermediate time scale, but then deviates from a k-design at late times. These results highlight that the k-design property fails to capture late time behavior of correlation functions. An interesting application of unitary k-designs is that

Haar-randomness is a probe of quantum complexity. We apply techniques from [11] to lower bound the quantum circuit complexity of time evolution by GUE Hamiltonians and find a quadratic growth in time.

In order to make precise claims about the behavior of OTOCs and frame potentials for GUE Hamiltonians, we need explicit expressions for certain spectral quantities. Accordingly, we compute the 2-point and 4-point spectral form factors for the GUE at infinite temperature, as well as the 2-point form factor at finite temperature. We then use these expressions to discuss time scales for the frame potentials. We also analytically compute the late-time value of the k-th frame potential for arbitrary k.

Under time evolution by strongly-coupled systems, correlations are spread throughout the system and the locality of operators as well as time-ordering appear to be lost from the viewpoint of correlation functions, as implied by the late-time universality of random matrix theory. Also motivated by the k-design property's failure to capture late-time chaos (i.e., $\mathcal{E}_t^{\mathrm{GUE}}$ fails to be Haar-random at late times), we propose a new property called k-invariance, which may provide a better probe of chaos at both early and late times. The property of k-invariance characterizes the degree to which an ensemble is Haar-invariant, meaning that the ensemble is invariant under a change of basis. When the dynamics becomes approximately Haar-invariant, correlation functions can be captured solely in terms of spectral functions, which signifies the onset of an effective random matrix theory description. We thus provide an information theoretically precise definition of what it means for a system's dynamics to be described by random matrix theory. Specifically, we say that an ensemble of Hamiltonian time evolutions $\mathcal{E}_t$ is described by random matrix theory at times greater than or equal to t with respect to $2k$-point OTOCs when $\mathcal{E}_t$ is approximately k-invariant with respect to its symmetry class, for example the symmetry class of either the unitary, orthogonal, or symplectic groups.

The chapter is organized as follows: In Section 2.2, we provide a brief overview of random matrix theory and explicitly compute the spectral form factors for the GUE at infinite and finite temperature. In Section 2.3, we compute correlation functions for the GUE, including the OTOC, and demonstrate that they can be expressed in terms of spectral correlators as well. In Section 2.4, we compute frame potentials for the GUE, and extract the timescales

when it becomes an approximate k-design both at finite and infinite temperatures. We show that the frame potentials can be also expressed as products of sums of spectral correlators. In Section 2.5, we discuss complexity bounds and complexity growth for the GUE. In Section 2.6, we discuss Haar-invariance as a diagnostic of delocalization of spatial degrees of freedom and random matrix universality at late times. We conclude with a discussion in Section 2.7. The appendices contain a review of various information-theoretic definitions of scrambling in the literature, a discussion of information scrambling in black holes, more details of our random matrix calculations, and numerics.

2.2 Form factors and random matrices

For a long time, the spectral statistics of a random matrix were seen as a defining feature of quantum chaos. More recently, it has been proposed that the late time behavior of certain strongly coupled theories with large numbers of degrees of freedom also exhibit a dynamical form of random matrix universality at late times [9]. The central object of study in this recent work is the 2-point spectral form factor,[1] which is defined in terms of the analytically continued partition function

$$\mathcal{R}_2(\beta, t) \equiv \big\langle |Z(\beta, t)|^2 \big\rangle, \quad \text{where} \quad Z(\beta, t) \equiv \mathrm{Tr}\big(e^{-\beta H - iHt}\big) \tag{2.2}$$

and where $\langle \cdot \rangle$ denotes the average over an ensemble of Hamiltonians. In SYK as well as standard RMT ensembles, the 2-point spectral form factor decays from its initial value and then climbs linearly back up to a floor value at late times. The early time decay of the form factor is called the slope, the small value at intermediate times is called the dip, the steady linear rise is called the ramp, and the late time floor is called the plateau. In Fig. 2.1 we observe these features in SYK with $N = 26$ Majoranas, which has GUE statistics at late

[1] One motivation for studying this object is a simple version of the information loss problem in AdS/CFT [36], where the apparent exponential decay of 2-point correlation functions in bulk effective field theory contradicts the finite late-time value of $e^{-\mathcal{O}(S)}$ implied by the discreteness of the spectrum. As we shall see in the next section, the 2-point form factor is equivalent to the average of 2-point correlation functions. More recently, chaos and information loss in correlation functions and form factors have also been studied in holographic CFTs [29, 37, 38, 39, 40].

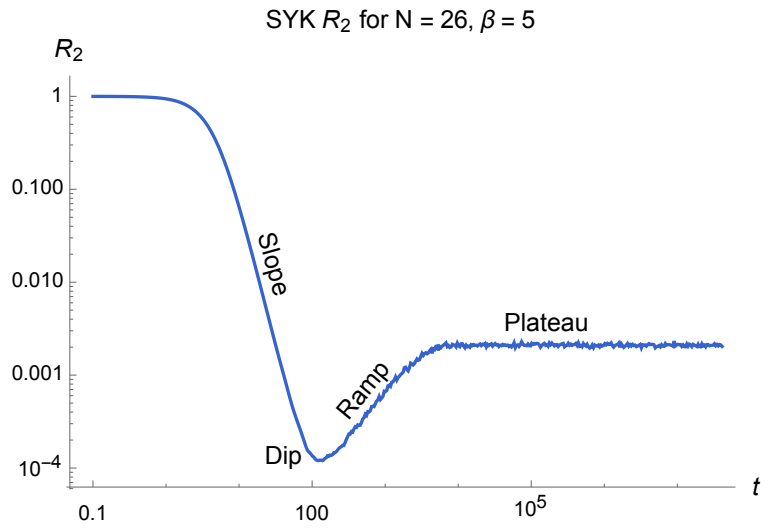


Figure 2.1: The 2-point spectral form factor for SYK with $N = 26$ Majoranas at inverse temperature $\beta = 5$, computed for 1000 random samples. The slope, dip, ramp, and plateau are labeled.

times.[2] Furthermore, it was found that in SYK, time scales and many features of the slope, dip, ramp and plateau agree with predictions from RMT.

In this section, we briefly review random matrix theory. Further, we study the 2-point spectral form factor for the GUE at both infinite and finite temperature, compute its analytic form, and extract its dip and plateau times and values.[3] In addition, we compute the 4-point form factor and extract relevant time scales and values. We find that the late-time rise in the 4-point form factor is quadratic in t, in contrast to the linear rise in the 2-point form factor. The expressions derived in this section will give us analytic control over the correlation functions and frame potentials discussed in later sections. For a detailed treatment of the random matrix ensembles, we refer the reader to [23, 43, 19].

2.2.1 Random matrix theory

The Gaussian Unitary Ensemble GUE(d, μ, σ) is an ensemble of $d \times d$ random Hermitian matrices, where the off-diagonal components are independent complex Gaussian random

[2]For SYK with N Majoranas, particle-hole symmetry dictates the symmetry class of the spectrum, where $N \pmod 8 \equiv 2$ or 6 corresponds to GUE statistics [41]. Furthermore, the spectral density of SYK and its relation to random matrices has also been discussed in [42].

[3]We consider the GUE since it corresponds to the least restrictive symmetry class of Hamiltonians. The generalization of our analysis to the GOE or GSE is left for future work.

variables $N(\mu,\sigma)_{\mathbb{C}}$ with mean μ and variance σ^2, and the diagonal components are independent real Gaussian random variables $N(\mu,\sigma)_{\mathbb{R}}$. It is common in the math literature to work with GUE$(d,0,1)$ which has zero mean and unit variance, but we will instead use the normalization GUE$(d,0,1/\sqrt{d})$ so that the eigenvalues do not scale with the system size.[4] The probability density function of the ensemble has a Gaussian form

$$P(H) \propto e^{-\frac{d}{2}\mathrm{Tr}H^2}\,, \tag{2.3}$$

up to a normalizing factor. As the GUE is invariant under unitary conjugation $H \to UHU^\dagger$, the integration measure $dH = d(UHU^\dagger)$ is likewise invariant. The probability measure $P(H)\,dH$ on the ensemble integrates to unity.

Instead of integrating over dH directly, it is convenient to change variables to eigenvalues and diagonalizing unitaries. Up to a normalizing constant C defined in Eq. (2.166) in App. 2.C, the measure becomes

$$dH = C\,|\Delta(\lambda)|^2 \prod_i d\lambda_i dU\,, \tag{2.4}$$

where dU is the Haar measure on the unitary group $U(d)$ and $\Delta(\lambda)$ is the Vandermonde determinant

$$\Delta(\lambda) = \prod_{i>j}(\lambda_i - \lambda_j)\,. \tag{2.5}$$

The joint probability distribution of eigenvalues is

$$P(\lambda_1,\ldots,\lambda_d) = Ce^{-\frac{d}{2}\sum_i \lambda_i^2}|\Delta(\lambda)|^2\,, \tag{2.6}$$

and is symmetric under permutations of its variables. For simplicity, we define a measure $D\lambda$ which absorbs the Gaussian weights, eigenvalue determinant, and constant factors. We

[4]The reason for using the normalization GUE$(d,0,1/\sqrt{d})$ instead of GUE$(d,0,1)$ is as follows: With the standard normalization GUE$(d,0,1)$, the energy spectrum ranges from $-2\sqrt{d}$ to $2\sqrt{d}$. This implies that by applying a local operator, one may change the energy of the system by $\mathcal{O}(\sqrt{d})$. With the physical normalization GUE$(d,0,1/\sqrt{d})$, the energies lie within the range -2 to 2, and local operators act with $\mathcal{O}(1)$ energy. See [44] for discussions on normalizing q-local Hamiltonians.

integrate over the GUE in the eigenvalue basis as

$$\langle O(\lambda)\rangle_{\text{GUE}} \equiv \int D\lambda\, O(\lambda) \quad \text{where} \quad \int D\lambda = C \int \prod_i d\lambda_i |\Delta(\lambda)|^2 e^{-\frac{d}{2}\sum_i \lambda_i^2} = 1\,. \tag{2.7}$$

The probability density of eigenvalues $\rho(\lambda)$, where

$$\int d\lambda\, \rho(\lambda) = 1\,, \tag{2.8}$$

can be written in terms of the joint eigenvalue probability density by integrating over all but one argument

$$\rho(\lambda) = \int d\lambda_1 \dots d\lambda_{d-1} P(\lambda_1, \dots, \lambda_{d-1}, \lambda)\,. \tag{2.9}$$

The spectral n-point correlation function, i.e. the joint probability distribution of n eigenvalues, $\rho^{(n)}$ is defined as

$$\rho^{(n)}(\lambda_1, \dots, \lambda_n) \equiv \int d\lambda_{n+1} \dots d\lambda_d P(\lambda_1, \dots, \lambda_d)\,. \tag{2.10}$$

With these definitions at hand, we quote a few central results. In the large d limit, the density of states for the Gaussian ensembles gives Wigner's famous semicircle law,

$$\rho(\lambda) = \frac{1}{2\pi}\sqrt{4-\lambda^2} \quad \text{as} \quad d \to \infty\,, \tag{2.11}$$

where the semicircle diameter is fixed by our chosen eigenvalue normalization. Also in the large d limit, the spectral 2-point function

$$\rho^{(n)}(\lambda_1, \lambda_2) = \int d\lambda_3 \dots d\lambda_d P(\lambda_1, \dots, \lambda_d)\,, \tag{2.12}$$

can be expressed in terms of a disconnected piece and a squared sine kernel as [23]

$$\rho^{(2)}(\lambda_1, \lambda_2) = \frac{d^2}{d(d-1)}\,\rho(\lambda_1)\rho(\lambda_2) - \frac{d^2}{d(d-1)}\frac{\sin^2\big(d(\lambda_1-\lambda_2)\big)}{\big(d\pi(\lambda_1-\lambda_2)\big)^2}\,. \tag{2.13}$$

2.2.2 Spectral form factors

The 2-point spectral form factor for a single Hamiltonian H is given in terms of the analytically continued partition function $Z(\beta,t) = \mathrm{Tr}\,(e^{-\beta H - iHt})$ as

$$\mathcal{R}_2^H(\beta,t) \equiv Z(\beta,t)Z^*(\beta,t) = \mathrm{Tr}\,(e^{-\beta H - iHt})\mathrm{Tr}\,(e^{-\beta H + iHt})\,. \tag{2.14}$$

Similarly, the spectral form factor averaged over the GUE is denoted by

$$\mathcal{R}_2(\beta,t) \equiv \big\langle Z(\beta,t)Z^*(\beta,t)\big\rangle_{\mathrm{GUE}} = \int D\lambda \sum_{i,j} e^{i(\lambda_i-\lambda_j)t} e^{-\beta(\lambda_i+\lambda_j)}\,, \tag{2.15}$$

which is the Fourier transform of the spectral 2-point function. At infinite temperature $\beta = 0$, the Fourier transform of the density of states is just $Z(t) = \mathrm{Tr}\,(e^{-iHt})$, the trace of unitary time evolution. Using the semicircle law, we take the average of $Z(t)$ at large d

$$\langle Z(t)\rangle_{\mathrm{GUE}} = \int D\lambda \sum_i e^{-i\lambda_i t} = d\int_{-2}^{2} d\lambda\, \rho(\lambda) e^{-i\lambda t} = \frac{dJ_1(2t)}{t}\,, \tag{2.16}$$

where $J_1(t)$ is a Bessel function of the first kind. The function $J_1(2t)/t$ is one at $t = 0$ and oscillates around zero with decreasing amplitude that goes as $\sim 1/t^{3/2}$, decaying at late times. At infinite temperature, the 2-point spectral form factor for the GUE is

$$\mathcal{R}_2(t) = \big\langle Z(t)Z^*(t)\big\rangle_{\mathrm{GUE}} = \int dH\, \mathrm{Tr}\big(e^{-iHt}\big)\,\mathrm{Tr}\big(e^{iHt}\big) = \int D\lambda \sum_{i,j} e^{i(\lambda_i-\lambda_j)t}\,. \tag{2.17}$$

More generally, we will also be interested in computing $2k$-point spectral form factors

$$\mathcal{R}_{2k}(t) = \Big\langle \big(Z(t)Z^*(t)\big)^k \Big\rangle_{\mathrm{GUE}} = \int D\lambda \sum_{i'\mathrm{s},j'\mathrm{s}} e^{i(\lambda_{i_1}+\ldots+\lambda_{i_k}-\lambda_{j_1}-\ldots-\lambda_{j_k})t}\,, \tag{2.18}$$

the Fourier transform of the spectral $2k$-point function $\rho^{(2k)}$.[5] Although the form factors can be written exactly at finite d, our analysis will focus on analytic expressions that capture

[5] In the random matrix literature, the 2-point form factor is often defined as the Fourier transform of the *connected* piece of the spectral 2-point correlation function, where the connected piece of the spectral $2k$-point function is often referred to as the $2k$-level cluster function. Our definition for the $2k$-point spectral form factor $\mathcal{R}_{2k}$ includes both connected *and* disconnected pieces.

the large d behavior.[6]

Note that in [9], 2-point form factors were normalized via dividing by $Z(\beta)^2$. At infinite temperature, this simply amounts to dividing by d^2, but at finite temperature the situation is more subtle. As we will comment on later, the correct object to study is the quenched form factor $\langle Z(\beta,t)Z^*(\beta,t)/Z(\beta)^2\rangle$, but since we only have analytic control over the numerator and denominator averaged separately, we instead work with the unnormalized form factor $\mathcal{R}_2$ as defined above.

2-point spectral form factor at infinite temperature

Here we calculate the 2-point form factor at $\beta = 0$. Working at large d, we can evaluate $\mathcal{R}_2$ by first pulling out the contribution from coincident eigenvalues

$$\mathcal{R}_2(t) = \int D\lambda \sum_{i,j} e^{i(\lambda_i-\lambda_j)t} = d + d(d-1)\int d\lambda_1 d\lambda_2\, \rho^{(2)}(\lambda_1,\lambda_2)e^{i(\lambda_1-\lambda_2)t}\,. \tag{2.19}$$

In the large d limit, we can make use of the sine kernel form of the 2-point function Eq. (2.13). Using Eq. (2.16), we integrate the first term, a product of 1-point functions, and find

$$\int d\lambda_1 d\lambda_2\, \rho(\lambda_1)\,\rho(\lambda_2)\, e^{i(\lambda_1-\lambda_2)t} = \frac{J_1^2(2t)}{t^2}\,. \tag{2.20}$$

In order to integrate the sine kernel, we make the change of variables:

$$u_1 = \lambda_1 - \lambda_2 \quad \text{and} \quad u_2 = \lambda_2\,, \tag{2.21}$$

which allows us to rewrite the integral

$$d^2\int d\lambda_1 d\lambda_2 \frac{\sin^2\big(d(\lambda_1-\lambda_2)\big)}{\big(d\pi(\lambda_1-\lambda_2)\big)^2}\, e^{i(\lambda_1-\lambda_2)t} = d^2\int du_2 \int du_1 \frac{\sin^2(du_1)}{d\pi u_1^2}\, e^{iu_1 t}\,. \tag{2.22}$$

Having decoupled the variables, in order to integrate over u_1 and u_2, we must employ a short distance cutoff. We develop a certain approximation method which we refer to as the 'box

[6] In addition to relating the form factor to the fidelty of certain states, [45] also studies the 2-point spectral form factor for the GUE, computing an analytic form at finite d and discussing the dip and plateau.

approximation,' and explain its justification in App. 2.C. Specifically, we integrate u_1 from 0 to u_2, and integrate u_2 from $-\pi/2$ to $\pi/2$,

$$d^2 \int du_1 du_2 \frac{\sin^2(du_1)}{d\pi u_1^2} e^{iu_1 t} = d \begin{cases} 1 - \frac{t}{2d}\,, & \text{for} \quad t < 2d \\ 0\,, & \text{for} \quad t > 2d \end{cases} . \tag{2.23}$$

Note that in the random matrix theory literature, a common treatment [24] is to approximate the short-distance behavior of $\rho^{(2)}(\lambda_1, \lambda_2)$ by adding a delta function for coincident points $\lambda_1 = \lambda_2$ and inserting a 1-point function into the sine kernel. For $\mathcal{R}_2$ this gives the same result as the approximation above, but this short-distance approximation does not generalize to higher k-point form factors, as discussed in App. 2.C. The 2-point form factor we compute is[7]

$$\mathcal{R}_2(t) = d^2 r_1^2(t) - d r_2(t) + d\,, \tag{2.24}$$

where we define the functions

$$r_1(t) \equiv \frac{J_1(2t)}{t}\,, \quad \text{and} \quad r_2(t) \equiv \begin{cases} 1 - \frac{t}{2d}\,, & \text{for} \quad t < 2d \\ 0\,, & \text{for} \quad t > 2d \end{cases} . \tag{2.25}$$

As was discussed in [9], we can extract the dip and plateau times and values from $\mathcal{R}_2$. From the ramp function r_2, we observe that the plateau time is given by

$$t_p = 2d\,, \tag{2.26}$$

where after the plateau time, the height of the function $\mathcal{R}_2$ is the constant d. This value can also be derived by taking the infinite time average of $\mathcal{R}_2$.

The other important time scale is the dip time t_d, which we can estimate using the

[7] We emphasize that this function relied on an approximation and while it captures certain desired behavior, it should not be viewed as exact. In App. 2.D we provide numerical checks and discuss an improvement of the ramp function $r_2(t)$.

asymptotic form of the Bessel function at large t, which gives

$$r_1(t) \approx \frac{1}{t}\frac{\cos(2t - 3\pi/4)}{\sqrt{\pi t}}\,, \tag{2.27}$$

oscillating at times $\sim \mathcal{O}(1)$ with decaying envelope $\sim t^{-3/2}$. While the first dip time is $\mathcal{O}(1)$, we will be interested in the dip time as seen by the envelope, especially because the oscillatory behavior disappears at finite temperature (see Fig. 2.3). Solving for the minimum of the envelope of $\mathcal{R}_2$, we find

$$t_d \approx \sqrt{d}\,, \tag{2.28}$$

up to order one factors. The true minimum of the envelope and ramp is $(6/\pi)^{1/4}\sqrt{d} \approx 1.18\sqrt{d}$, but in light of the approximations we made, and the fact that the precise ramp behavior is somewhat ambiguous, we simply quote the dip time as $t_d \approx \sqrt{d}$. At t_d, we find the dip value $\mathcal{R}_2(t_d) \approx \sqrt{d}$. We plot the 2-point form factor for different dimensions d in Fig. 2.2.

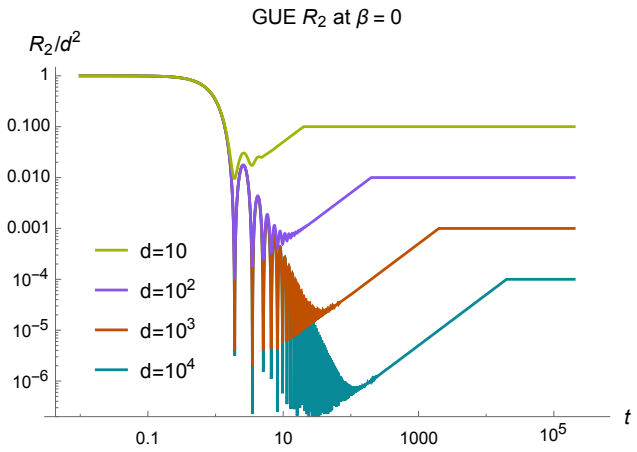


Figure 2.2: The 2-point spectral form factor at infinite temperature, as given in Eq. (2.24), plotted for various values of d and normalized by the initial value d^2. We observe the linear ramp and scaling of the dip and plateau with d.

The oscillations in the early time slope behavior of the form factor simply arise from the oscillatory behavior of the Bessel function, i.e. the zeros of $r_1(t)^2$.

2-point spectral form factor at finite temperature

Recall that spectral 2-point function at finite temperature is defined as

$$\mathcal{R}_2(t,\beta) \equiv \big\langle Z(t,\beta)Z^*(t,\beta)\big\rangle_{\text{GUE}} = \int D\lambda \sum_{i,j} e^{i(\lambda_i-\lambda_j)t} e^{-\beta(\lambda_i+\lambda_j)}\,.$$

As described in App. 2.C, we insert the spectral 2-point function $\rho^{(2)}$ and, using the short-distance kernel, find $\mathcal{R}_2(t,\beta)$ in terms of the above functions:

$$\mathcal{R}_2(t,\beta) = d^2 r_1(t+i\beta)r_1(-t+i\beta) + d r_1(2i\beta) - d r_1(2i\beta) r_2(t)\,. \tag{2.29}$$

First we comment on the validity of the approximations used in the finite temperature case. The first and third terms of Eq. (2.29), dominating at early and late times respectively, are computed from the 1-point function. Therefore, the expression captures the early time, slope, and plateau behaviors. The dip and ramp behavior, encoded in the r_2 term, are more subtle. The expression correctly captures the slope of the ramp, but deviates from the true ramp at large β. We will discuss this more in App. 2.C, but here only discuss quantities around the dip for small β, where Eq. (2.29) is a good approximation.

The ramp function r_2, which is the same as at infinite temperature, gives the plateau time

$$t_p = 2d\,. \tag{2.30}$$

For convenience we define the function $h_1(\beta) \equiv J_1(2i\beta)/i\beta$, which is real-valued in β.[8] The initial value and plateau value are thus given by

$$\mathcal{R}_2(0) = (h_1(\beta))^2 d^2\,, \qquad \mathcal{R}_2(t_p) = h_1(2\beta)d\,. \tag{2.31}$$

To find the dip time, we make use of the asymptotic expansion of the Bessel function as

$$d^2 r_1(t+i\beta)r_1(-t+i\beta) \sim \frac{d^2}{2\pi t^3}\big(\cosh(4\beta) - \sin(4t)\big) \approx \frac{d^2}{\pi t^3}\cosh^2(2\beta)\,. \tag{2.32}$$

[8] For instance, to emphasize its real-valuedness, we could equivalently write $h_1(\beta)$ as a regularized hypergeometric function $h_1(\beta) \equiv {}_0\widetilde{F}_1(2;\beta^2)$.

Finding the minimum of the expression gives the dip time

$$t_d = h_2(\beta)\sqrt{d} \quad \text{where} \quad h_2(\beta) \approx \left(1 + \frac{\beta^2}{2} + O(\beta^4)\right) , \tag{2.33}$$

and evaluating $\mathcal{R}_2$ at the dip gives

$$\mathcal{R}_2(t_d) \approx h_3(\beta)\sqrt{d} \quad \text{where} \quad h_3(\beta) \approx \left(1 + \frac{5\beta^2}{2} + \mathcal{O}(\beta^4)\right) , \tag{2.34}$$

up to order one factors. While we could write down full expressions for the dip time h_2 and dip value h_3 in terms of the Bessel function, we only trust Eq. (2.29) in this regime for small β, and thus report the functions perturbatively.

The 2-point form factor is plotted in Fig. 2.3 for various values of d and β. While increasing the dimension d lowers the dip and plateau values and delays the dip and plateau times, decreasing temperature raises the dip and plateau values and delays the dip times. We also note that lowering the temperature smooths out oscillations from the Bessel function.[9] After normalizing $\mathcal{R}_2(\beta, t)$ by its initial value, the late-time value is $\simeq 2^{-S^{(2)}}$ where $S^{(2)}$ is the thermal Rényi-2 entropy.

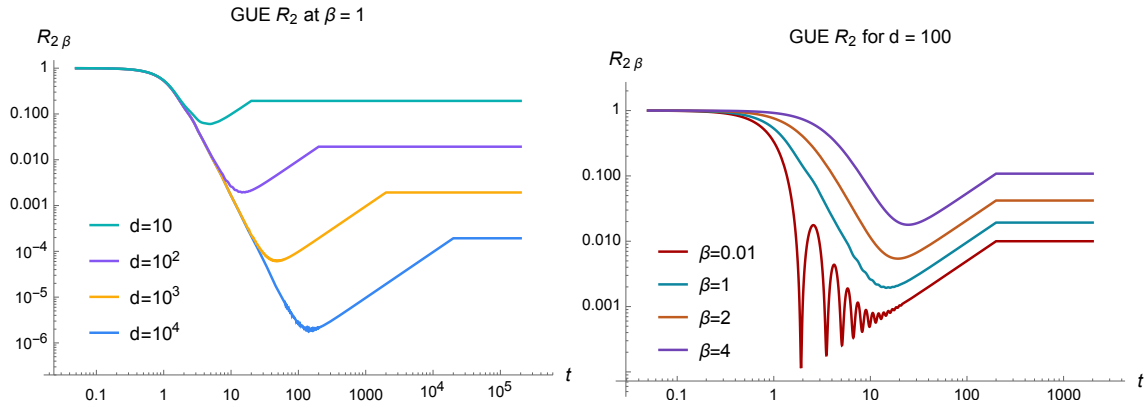


Figure 2.3: The 2-point spectral form factor at finite temperature as per Eq. (2.29), on the left plotted at different values of d, and on the right plotted at different temperatures, normalized by the initial value. We see that the dip and plateau both scale with β and d and that lowering the temperature smooths out the oscillations in $\mathcal{R}_2$.

[9] While the oscillatory behavior still persists at finite temperature, the width of the dips become very sharp as we increase β and thus the oscillations are not observed when plotted. Furthermore, if we average over a small time window, the oscillations are also smoothed out.

2.2.3 4-point spectral form factor at infinite temperature

We can also compute the 4-point form factor at infinite temperature, defined as

$$\mathcal{R}_4(t) \equiv \left\langle Z(t)Z(t)Z^*(t)Z^*(t)\right\rangle_{\text{GUE}} = \int D\lambda \sum_{i,j,k,\ell} e^{i(\lambda_i+\lambda_j-\lambda_k-\lambda_\ell)t}\,. \tag{2.35}$$

As we explain in App. 2.C, we compute $\mathcal{R}_4$ by replacing $\rho^{(4)}$ by a determinant of sine kernels and carefully integrating each term using the box approximation. The result is

$$\mathcal{R}_4(t) = d^4 r_1^4(t) + 2d^2 r_2^2(t) - 4d^2 r_2(t) - 7d r_2(2t) + 4d r_2(3t) + 4d r_2(t) + 2d^2 - d\,, \tag{2.36}$$

given in terms of the functions $r_1(t)$ and $r_2(t)$ defined above. The initial value of $\mathcal{R}_4$ is d^4. Given the dependence on the ramp function, the plateau time is still $t_p = 2d$. The plateau value $2d^2 - d$ matches the infinite time average of Eq. (2.35). The dip time is found again by considering the leading behavior of $\mathcal{R}_4$ and expanding the Bessel functions

$$\mathcal{R}_4 \approx d^4 \frac{J_1^4(2t)}{t^4} + \frac{t}{2}(t-2) \sim \frac{d^4}{t^6\pi^2} + \frac{t}{2}(t-2)\,. \tag{2.37}$$

Solving for the minimum, we find the dip time

$$t_d \approx \sqrt{d}\,, \tag{2.38}$$

where at the dip time $\mathcal{R}_4(t_d) \approx d$. We plot the $\mathcal{R}_4(t)$ for various values of d in Fig. 2.4.

Let us summarize the time scales and values for the form factors considered above:

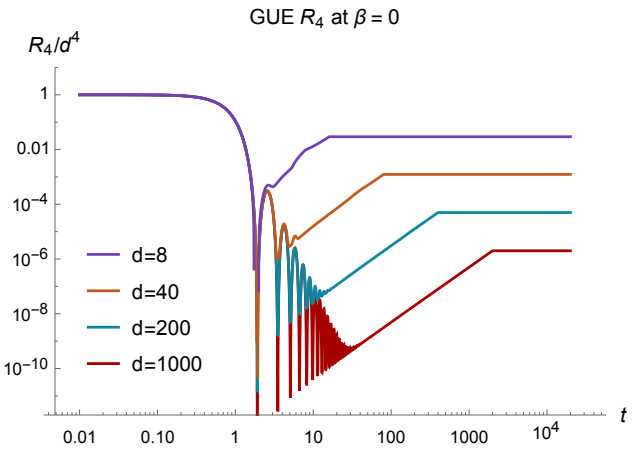


Figure 2.4: The GUE 4-point spectral form factor at infinite temperature, plotted for different values of d and normalized by their initial values. We observe the scaling of the dip and plateau, and the quadratic rise $\sim t^2$.

form factor	time scale	time	value
$\mathcal{R}_2(t)$	initial	0	d^2
	dip	$\sqrt{d}$	$\sqrt{d}$
	plateau	$2d$	d
$\mathcal{R}_2(t,\beta)$	initial	0	$h_1^2(\beta)d^2$
	dip	$h_2(\beta)\sqrt{d}$	$h_3(\beta)\sqrt{d}$
	plateau	$2d$	$h_1(2\beta)d$
$\mathcal{R}_4(t)$	initial	0	d^4
	dip	$\sqrt{d}$	d
	plateau	$2d$	$2d^2$

The β–dependent functions were defined above.

With an understanding of the first few form factors, we briefly describe the expected behavior for $2k$-point form factors $\mathcal{R}_{2k}(t)$ (with $k \ll d$). Initially, $\mathcal{R}_{2k}$ decays from d^{2k} as $\sim J_1^{2k}(2t)/t^{2k}$, reaching the dip at time $t_d \approx \sqrt{d}$ where $\mathcal{R}_{2k}(t_d) \approx d^{k/2}$. The $\sim t^k$ growth after the dip levels off at the plateau time $2d$, with plateau value $\sim kd^k$.

Given that we employed some approximation to compute the form factors, we perform numerical checks for the expressions above in App. 2.D. At both infinite and finite temperature, we correctly capture the time scales, early time decay, dip behavior, and the late-time plateau, but find slight deviations from the analytic prediction for the ramp. We discuss this

and possible improvements to the ramp function in App. 2.D.

Later we will study frame potentials which diagnose whether an ensemble forms a k-design. We will find that the frame potentials for the ensemble of unitaries generated by the GUE can be written in terms of the spectral form factors discussed here, thereby allowing us to extract important time scales pertaining to k-designs.

2.3 Out-of-time-order correlation functions

2.3.1 Spectral form factor from OTOCs

Although quantum chaos has traditionally focused on spectral statistics, recent developments from black hole physics and quantum information theory suggest an alternative way of characterizing quantum chaos via OTOCs [1, 4, 27, 10]. In this subsection, we bridge the two notions by relating the average of $2k$-point OTOCs to spectral form factors. We work at infinite temperature ($\beta = 0$), but note that by distributing operator insertions around the thermal circle, the generalization to finite temperature is straightforward. The results in this subsection are not specific to GUE and are applicable to any quantum mechanical system.

Consider some Hamiltonian H acting on an $d = 2^n$-dimensional Hilbert space, i.e. consisting of n qubits. We start by considering the 2-point autocorrelation function $\langle A(0)A^\dagger(t)\rangle$, time evolved by H. We are interested in the averaged 2-point function:

$$\int dA \langle A(0)A^\dagger(t)\rangle \equiv \frac{1}{d}\int dA \ \mathrm{Tr}(Ae^{-iHt}A^\dagger e^{iHt}) \,, \tag{2.39}$$

where $\int dA$ represents an integral with respect to a unitary operator A over the Haar measure on $U(2^n)$. We note that since the 2-point Haar integral concerns only the first moment of the Haar ensemble, we can instead average over the ensemble of Pauli operators[10]

$$\int dA \langle A(0)A^\dagger(t)\rangle = \frac{1}{d^3}\sum_{j=1}^{d^2} \mathrm{Tr}(A_j e^{-iHt}A_j^\dagger e^{iHt}) \,, \tag{2.40}$$

[10]This is because the Pauli operators form a 1-design.

where A_j are Pauli operators and $d^2 = 4^n$ is the number of total Pauli operators for a system of n qubits. To derive the spectral form factor, we will need the first moment of the Haar ensemble

$$\int dA\, A^j_k A^{\dagger \ell}{}_m = \frac{1}{d}\delta^j_m \delta^\ell_k\,, \quad \text{or equivalently} \quad \int dA\; AOA^\dagger = \frac{1}{d}\mathrm{Tr}(O)\mathbb{I}\,. \tag{2.41}$$

Applying Eq. (2.41) to Eq. (2.39), we obtain

$$\int dA \langle A(0)A^\dagger(t)\rangle = \frac{|\mathrm{Tr}(e^{-iHt})|^2}{d^2} = \frac{\mathcal{R}^H_2(t)}{d^2}\,. \tag{2.42}$$

where $\mathcal{R}^H_{2k}(t) \equiv |\mathrm{Tr}(e^{-iHt})|^{2k}$ is the same as $\mathcal{R}_{2k}(t)$ from before, but written for a single Hamiltonian H instead of averaged over the GUE. Thus, the 2-point form factor is proportional to the averaged 2-point function.

This formula naturally generalizes to $2k$-point OTOCs and $2k$-point form factors. Consider $2k$-point OTOCs with some particular ordering of operators

$$\langle A_1(0)B_1(t)\cdots A_k(0)B_k(t)\rangle \quad \text{where} \quad A_1B_1\cdots A_kB_k = \mathbb{I}\,. \tag{2.43}$$

Operators which do not multiply to the identity have zero expectation value at $t = 0$, and the value stays small as we time-evolve. We are interested in the average of such $2k$-point OTOCs. By using Eq. (2.41) $2k-1$ times, we obtain

$$\int dA_1\cdots dB_{k-1}dA_k \langle A_1(0)B_1(t)\cdots A_k(0)B_k(t)\rangle = \frac{|\mathrm{Tr}(e^{-iHt})|^{2k}}{d^{2k}} = \frac{\mathcal{R}^H_{2k}(t)}{d^{2k}}\,, \tag{2.44}$$

where $B_k = A^\dagger_k \cdots B^\dagger_1 A^\dagger_1$. Thus, higher-point spectral form factors can be also computed from OTOCs. In fact, by changing the way we take an average, we can access various types of form factors. For instance, let us consider OTOCs $\langle A_1(0)B_1(t)\cdots A_k(0)B_k(t)\rangle$ with $B_j = A^\dagger_j$. We then have

$$\int dA_1 dA_2 \cdots dA_k \langle A_1(0)A^\dagger_1(t)\cdots A_k(0)A^\dagger_k(t)\rangle = \frac{\mathrm{Tr}(e^{-iHt})^k \mathrm{Tr}(e^{iHkt})}{d^{k+1}}\,. \tag{2.45}$$

The fact that the expression on the right-hand side is asymmetric is because the operator $A_1(0)A_1^\dagger(t)\cdots A_k(0)A_k^\dagger(t)$ is not Hermitian.[11]

These expressions not only provide a direct link between spectral statistics and physical observables, but also give a practical way of computing the spectral form factor. If one wishes to compute or experimentally measure the 2-point form factor $\mathcal{R}_2(t)$, one just needs to pick a random unitary operator A and study the behavior of the 2-point correlator $\langle A(0)A^\dagger(t)\rangle$. In order to obtain the exact value of $\mathcal{R}_2(t)$, we should measure $\langle A(0)A(t)\rangle$ for all possible Pauli operators and take their average. Yet, it is possible to obtain a pretty good estimate of $\mathcal{R}_2(t)$ from $\langle A(0)A(t)\rangle$ with only a few instances of unitary operator A. Consider the variance of $\langle A(0)A(t)\rangle$,

$$\Delta\langle A(0)A^\dagger(t)\rangle^2_{\text{avg}} \equiv \int dA|\langle A(0)A^\dagger(t)\rangle|^2 - \Big|\int dA\langle A(0)A^\dagger(t)\rangle\Big|^2 . \tag{2.46}$$

If the variance is small, then the estimation by a single A would suffice to obtain a good estimate of $\mathcal{R}_2(t)$. Computing this, we obtain

$$\Delta\langle A(0)A^\dagger(t)\rangle^2_{\text{avg}} \sim \mathcal{O}\Big(\frac{1}{d^2}\Big) . \tag{2.47}$$

This implies that the estimation error is suppressed by $1/d$. By choosing a Haar unitary operator A (or 2-design operator, such as a random Clifford operator), one can obtain a good estimate of $\mathcal{R}_2(t)$.

A check in a non-local spin system

To verify Eq. (2.42) and the claim that the variance of the 2-point functions is small, consider a random non-local (RNL) spin system with the Hamiltonian given as the sum over all 2-body operators with random Gaussian couplings $J_{ij\alpha\beta}$ [46]:

$$H_{\text{RNL}} = -\sum_{i,j,\alpha,\beta} J_{ij\alpha\beta} S_i^\alpha S_j^\beta , \tag{2.48}$$

[11] BY learned Eq. (2.45) from Daniel Roberts.

where i, j sum over the number of sites and α, β sum over the Pauli operators at a given site. Such Hamiltonians have a particularly useful property where locally rotating the spins of H_{RNL} with couplings $J_{ij\alpha\beta}$ creates another Hamiltonian H'_{RNL} with different couplings $J'_{ij\alpha\beta}$. More precisely, if we consider an ensemble of such 2-local Hamiltonians;

$$\mathcal{E}_{\text{RNL}} = \{H_{\text{RNL}},\ \text{for } J_{ij\alpha\beta} \in \text{Gaussian}\} \tag{2.49}$$

the ensemble is invariant under conjugation by any 1-local Clifford operator

$$\mathcal{E}_{\text{RNL}} = V\mathcal{E}_{\text{RNL}}V^\dagger\,, \qquad V \in \text{1-body Clifford}\,. \tag{2.50}$$

Here a Clifford operator refers to unitary operators which transform a Pauli operator to a Pauli operator. For this reason, the 2-point correlation function $\langle A(0)A^\dagger(t)\rangle_{\mathcal{E}_{\text{RNL}}}$ depends only of the weight of Pauli operator A:

$$\langle A(0)A^\dagger(t)\rangle_{\mathcal{E}_{\text{RNL}}} = c_m\,, \quad \text{where } A \text{ is an } m\text{-body Pauli operator} \tag{2.51}$$

and where $\langle\,\cdot\,\rangle_{\mathcal{E}_{\text{RNL}}}$ denotes the ensemble (disorder) average. Thus, this system is desirable for studying the weight dependence of 2-point correlation functions.

As mentioned above, we can write the average over 2-point correlation functions as the average over all Paulis as

$$\int dA\langle A(0)A^\dagger(t)\rangle = \frac{1}{4^n}\sum_{A\in\text{Pauli}}\langle A(0)A^\dagger(t)\rangle = \frac{\mathcal{R}_2^{H_{\text{RNL}}}(t)}{d^2}\,, \tag{2.52}$$

time evolving with H_{RNL}. Numerically, for a single instance of H_{RNL}, we find that the average over all 2-point functions of Pauli operators gives $\mathcal{R}_2$ as expected. In Fig. 2.5, for $n = 5$ sites and averaged over 500 random instances of H_{RNL} to suppress fluctuations, we plot $\mathcal{R}_2$ along side all 2-point functions of Pauli operators. We observe that correlation functions depend only on the weight of A, with the higher weight Pauli operators clustered around $\mathcal{R}_2$. The arrangement of the 2-point functions for Paulis of different weight depends on the number of sites n. But for $n = 5$, the even and odd weight Paulis are respectively below and above

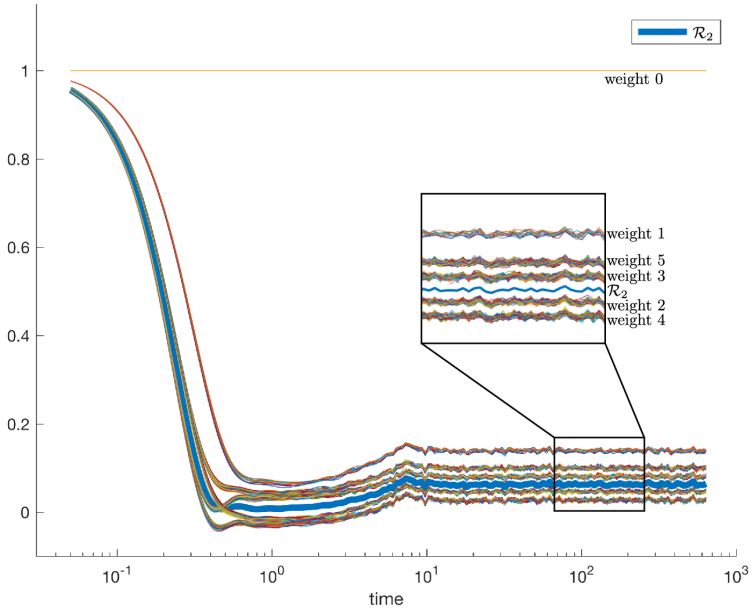


Figure 2.5: The 2-point form factor and the 2-point functions $\langle A_j A_j(t)\rangle$ of Pauli operators for H_{RNL} for $n = 5$ sites and averaged over 500 samples. The thick blue line is $\mathcal{R}_2/d^2$ and surrounding bands of lines are all 1024 Pauli 2-point functions of different weight.

$\mathcal{R}_2$ at later times and weight 2 and 3 Paulis are the closest to $\mathcal{R}_2$. We will comment on the size dependence of correlators in Sec. 2.6.

The conclusion is that we can choose a few random Paulis, and by computing 2-point functions, quickly approximate $\mathcal{R}_2$. We also checked that by increasing the number of spins, the variance becomes small and 2-point functions become closer to $\mathcal{R}_2$.

Operator averages and locality

Let us pause for a moment and discuss the meaning of considering the operator average from the perspective of spatial locality in quantum mechanical systems. In deriving the above exact formulae relating the spectrum and correlators, we considered the average of OTOCs over all the possible Pauli operators. For a system of n qubits, a typical Pauli operator has support on $\simeq 3n/4$ qubits because there are four one-body Pauli operators, $\mathbb{I}, X, Y, Z$. It is essential to recognize that the average of correlation functions is dominated by correlations of non-local operators with big supports covering the whole system. Thus, the spectral statistics have a tendency to ignore the spatial locality of operators in correlation

functions.[12]

In fact, the spectral statistics ignore not only spatial locality but also temporal locality of operators. Namely, similar formulas can be derived for correlation functions with various ordering of time. For instance, consider the following 4-point correlation function:

$$\langle A(0)B(t)C(2t)D(t)\rangle \tag{2.53}$$

where the C operator acts at time $2t$ instead of 0 such that the correlator is not out-of-time-ordered. Computing the average of the correlator with $ABCD = \mathbb{I}$, we obtain

$$\int dAdBdC\langle A(0)B(t)C(2t)D(t)\rangle = \frac{\mathcal{R}_4(t)}{d^4}\,, \tag{2.54}$$

which is exactly the same result as the average of 4-point OTOCs in Eq. (2.44). Indeed, time-ordering is washed away since GUE Hamiltonians cause a system to rapidly delocalize, thus destroying all local temporal correlations.

In strongly coupled systems with local Hamiltonians, correlation functions behave rather differently depending on the time ordering of operators, as long as the time gaps involved are small or comparable to the scrambling time [28, 4, 5, 6]. This observation hints that the spectral statistics are good probes of correlations at long time scales, but may miss some important physical signatures at shorter time scales, such as the exponential growth of OTOCs with some Lyapunov exponent.

2.3.2 OTOCs in random matrix theory

Next, we turn our attention to correlators averaged over random matrices, analytically computing the 2-point correlation functions and 4-point OTOCs for the GUE. We begin with the 2-point correlation functions for the GUE

$$\langle A(0)B(t)\rangle_{\text{GUE}} \equiv \int dH\langle A(0)B(t)\rangle \quad \text{where} \quad B(t) = e^{-iHt}B(0)e^{iHt}\,, \tag{2.55}$$

[12]Signatures of the locality of an individual Hamiltonian may be seen in properties of its spectrum, as argued in [47].

where $\int dH$ represents an integral over Hamiltonians H drawn from the GUE. Since the GUE measure dH is invariant under unitary conjugation $dH = d(UHU^\dagger)$ for all U, we can express the GUE average as

$$\langle A(0)B(t)\rangle_{\rm GUE} = \iint dH dU \left\langle AUe^{-iHt}U^\dagger BUe^{iHt}U^\dagger \right\rangle \tag{2.56}$$

by inserting $U, U^\dagger$ where dU is the Haar measure. Haar integrating, we obtain

$$\langle A(0)B(t)\rangle_{\rm GUE} = \langle A\rangle\langle B\rangle + \frac{\mathcal{R}_2(t)-1}{d^2-1}\langle\langle AB\rangle\rangle\,, \qquad \langle\langle AB\rangle\rangle \equiv \langle AB\rangle - \langle A\rangle\langle B\rangle \tag{2.57}$$

where $\langle\langle AB\rangle\rangle$ denotes the connected correlator. If A, B are non-identity Pauli operators, we have

$$\begin{aligned}\langle A(0)B(t)\rangle_{\rm GUE} &= \frac{\mathcal{R}_2(t)-1}{d^2-1} \qquad (A=B)\\ &= 0 \qquad\qquad\quad (A\neq B)\,.\end{aligned} \tag{2.58}$$

If $\mathcal{R}_2(t) \gg 1$, we have

$$\langle A(0)A^\dagger(t)\rangle_{\rm GUE} \simeq \frac{\mathcal{R}_2(t)}{d^2} \tag{2.59}$$

for any non-identity Pauli operator A. It is worth emphasizing the similarity between Eq. (2.59) and Eq. (2.42). Recall that Eq. (2.42) was derived by taking an average over all Pauli operators A and is valid for any quantum mechanical system while Eq. (2.59) was derived without any additional assumption on the locality of Pauli operator A. Namely, the key ingredient in deriving Eq. (2.59) was the Haar-invariance of the GUE measure dH. The resemblance of Eq. (2.59) and Eq. (2.42) implies that the GUE is suited for studying physical properties of chaotic Hamiltonians at macroscopic scales such as thermodynamic quantities.

Next, we compute the 4-point OTOCs for the GUE

$$\langle A(0)B(t)C(0)D(t)\rangle_{\rm GUE}\,. \tag{2.60}$$

Inserting $U, U^\dagger$, we must compute the fourth Haar moment

$$\langle A(0)B(t)C(0)D(t)\rangle_{\text{GUE}} = \iint dHdU \left\langle AUe^{-iHt}U^\dagger BUe^{iHt}U^\dagger CUe^{-iHt}U^\dagger DUe^{iHt}U^\dagger \right\rangle . \tag{2.61}$$

We can avoid dealing directly with the $(4!)^2$ terms generated by integrating here and focus on the leading behavior. Assuming that A, B, C, D are non-identity Pauli operators, we obtain

$$\langle A(0)B(t)C(0)D(t)\rangle_{\text{GUE}} \simeq \langle ABCD\rangle \frac{\mathcal{R}_4(t)}{d^4}. \tag{2.62}$$

Thus, OTOCs are almost zero unless $ABCD = \mathbb{I}$.[13,14] A similar analysis allows us to obtain the following result for $2k$-point OTOCs:

$$\langle A_1(0)B_1(t)\dots A_k(0)B_k(t)\rangle_{\text{GUE}} \simeq \langle A_1B_1\dots A_kB_k\rangle \frac{\mathcal{R}_{2k}(t)}{d^{2k}}. \tag{2.63}$$

The above equation is nonzero when $A_1B_1\dots A_kB_k = \mathbb{I}$. Again, note the similarity between Eq. (2.63) and Eq. (2.44). Recall that in order to derive Eq. (2.44), we took an average over OTOCs with $A_1B_1\dots A_kB_k = \mathbb{I}$. This analysis also supports our observation that the GUE tends to capture global-scale physics very well.

Similar calculations can be carried out for correlation functions with arbitrary time-ordering. For m-point correlators, at the leading order, we have

$$\langle A_1(t_1)A_2(t_2)\dots A_m(t_m)\rangle_{\text{GUE}} \simeq \langle A_1\dots A_m\rangle \frac{1}{d^m}\text{Tr}(e^{-it_{12}H})\text{Tr}(e^{-it_{23}H})\dots\text{Tr}(e^{-it_{m1}H}) , \tag{2.64}$$

where $t_{ij} = t_j - t_i$. Namely, we have:

$$\langle A(0)B(t)C(2t)D(t)\rangle_{\text{GUE}} \simeq \langle ABCD\rangle \frac{\mathcal{R}_4(t)}{d^4} . \tag{2.65}$$

So, for the GUE, $\langle A(0)B(t)C(2t)D(t)\rangle_{\text{GUE}} \simeq \langle A(0)B(t)C(0)D(t)\rangle_{\text{GUE}}$. This implies that

[13] In fact, one can prove that the GUE averaged OTOCs are exactly zero if $ABCD$ is non-identity Pauli operator for all times.

[14] For analysis related to Eq. (2.62) in the context of SYK, see [48].

the GUE does not care if operators in the correlator are out-of-time-ordered or not, ignoring both spatial and temporal locality.

Careful readers may have noticed that the only property we used in the above derivations is the unitary invariance of the GUE ensemble. If one is interested in computing correlation functions for an ensemble of Hamiltonians which are invariant under conjugation by unitary operators, then correlation functions can be expressed in terms of spectral form factors. Such techniques have been recently used to study thermalization in many-body systems, see [49] for instance. We discuss this point further in Sec. 2.6.

2.3.3 Scrambling in random matrices

Finally, we discuss thermalization and scrambling phenomena in random matrices by studying the time scales for correlation functions to decay.

We begin with 2-point correlators and thermalization. In a black hole (or any thermal system), quantum information appears to be lost from the viewpoint of local observers. This apparent loss of quantum information is called *thermalization*, and is often associated with the decay of 2-point correlation functions $\langle A(0)B(t)\rangle$ where A and B are some local operators acting on subsystems $\mathcal{H}_A$ and $\mathcal{H}_B$ which local observers have access to. In the context of black hole physics, $\mathcal{H}_A$ and $\mathcal{H}_B$ correspond to infalling and outgoing Hawking radiation and such 2-point correlation functions can be computed from the standard analysis of Hawking and Unruh [50, 51]. 2-point correlation functions of the form $\langle A(0)B(t)\rangle$ have an interpretation as how much information about initial perturbations on $\mathcal{H}_A$ can be detected from local measurements on $\mathcal{H}_B$ at time t. A precise and quantitative relation between quantum information (mutual information) and 2-point correlation functions is derived in Appendix 2.B. The upshot is that the smallness of $\langle A(0)B(t)\rangle$ implies the information theoretic impossibility of reconstructing from Hawking radiation (defined on $\mathcal{H}_B$) an unknown quantum state (supported on $\mathcal{H}_A$) that has fallen into a black hole.

Is the GUE a good model for describing thermalization? For the GUE, we found $\langle A(0)B(t)\rangle \simeq \mathcal{R}_2(t)/d^2$ for non-identity Pauli operators with $AB = \mathbb{I}$. Since the early

time behavior of $\mathcal{R}_2(t)$ factorizes and is given by

$$\langle A(0)A^\dagger(t)\rangle_{\rm GUE} \simeq \frac{J_1(2t)^2}{t^2}\,, \tag{2.66}$$

the time scale for the decay of 2-point correlation functions, denoted by t_2, is $\mathcal{O}(1)$. This is consistent with our intuition from thermalization in strongly coupled systems where $t_2 \simeq \beta$. As such, quantum information appears to be lost in $\mathcal{O}(1)$ time for local observers in systems governed by GUE Hamiltonians.

Next, let us consider 4-point OTOCs and scrambling. To recap the relation between OTOCs and scrambling in the context of black hole physics, consider a scenario where Alice has thrown an unknown quantum state into a black hole and Bob attempts to reconstruct Alice's quantum state by collecting the Hawking radiation. Hayden and Preskill added an interesting twist to this classic setting of black hole information problem by assuming that the black hole has already emitted half of its contents and Bob has collected and stored early radiation in some quantum memory he possesses. The surprising result by Hayden and Preskill is that, if time evolution $U = e^{-iHt}$ is approximated by a Haar random unitary operator, then Bob is able to reconstruct Alice's quantum state by collecting only a few Hawking quanta [1]. This mysterious phenomenon, where a black hole reflects a quantum information like a mirror, relies on scrambling of quantum information where Alice's input quantum information is delocalized over the whole system [10]. The definition of scrambling can be made precise and quantitative by using quantum information theoretic quantities as briefly reviewed in App. 2.A and App. 2.B.

The scrambling of quantum information can be probed by the decay of 4-point OTOCs of the form $\langle A(0)B(t)A^\dagger(0)B^\dagger(t)\rangle$ where A, B are some local unitary operators. An intuition is that an initially local operator $B(0)$ grows into some non-local operator under time evolution via conjugation by e^{-iHt}, and OTOCs measure how non-locally $B(t)$ has spread. For this reason, the time scale t_4 when OTOCs start decaying is called the scrambling time.

Having reviewed the concepts of scrambling and OTOCs, let us study scrambling in random matrices. For the GUE, we found $\langle A(0)B(t)C(0)D(t)\rangle \simeq \mathcal{R}_4(t)/d^4$ for non-identity Pauli operators with $ABCD = \mathbb{I}$. Since one can approximate $\mathcal{R}_4$ as $\mathcal{R}_4(t) \simeq \mathcal{R}_2(t)^2$ at early

times, we obtain

$$\langle A(0)B(t)C(0)D(t)\rangle_{\rm GUE} \simeq \frac{J_1(2t)^4}{t^4}\,. \tag{2.67}$$

This implies that the decay time scale of 4-point OTOCs is $t_4 \simeq \frac{1}{2}t_2$, which is $\mathcal{O}(1)$ and is faster than the decay time of 2-point correlation functions. This behavior is in strong contrast with behaviors in chaotic systems studied in the context of black hole physics. Namely, in holographic large-N CFTs with classical gravity duals, the decay times are

$$t_2 \simeq \beta\,, \qquad t_4 \simeq \beta \log N^2 \tag{2.68}$$

with $t_4 \gg t_2$. Also, the scrambling time $t_4 \sim \mathcal{O}(1)$ violates a bound on quantum signalling which would hold for quantum systems with local interactions [1, 3]. The pathology can be also seen from the viewpoint of black hole information problems. If black hole dynamics is modeled by the time evolution of some Hamiltonian sampled from GUE random matrices, then the scrambling time for OTOC decay is $\mathcal{O}(1)$. So Bob might be able to reconstruct Alice's quantum state in $\mathcal{O}(1)$ time. If Bob jumps into the black hole after decoding Alice's quantum state, Alice can send a quantum message with $\mathcal{O}(1)$ energy to Bob and verify the quantum cloning.

Another difference between GUEs and actual chaotic systems can be seen from the behaviors of correlators of the form $\langle A(0)B(t)C(2t)D(t)\rangle$. In the previous subsection, we showed that $\langle A(0)B(t)C(2t)D(t)\rangle \simeq \langle A(0)B(t)C(0)D(t)\rangle$. In strongly chaotic large-N systems, we expect the following behaviors [6, 27]:

$$\langle A(0)B(t)A(0)B(t)\rangle = 1 - \frac{1}{N}\,e^{\lambda t}\,, \qquad \beta \ll t \ll \beta \log N. \tag{2.69}$$

$$\langle A(0)B(t)C(2t)B(t)\rangle = \langle A\rangle\langle B\rangle\langle C\rangle\langle B\rangle\,, \qquad t \simeq \beta. \tag{2.70}$$

Thus these two types of correlators should behave in a rather different manner.

These discrepancies clearly highlight the failure of GUE to capture early-time quantum chaos behavior which is present in realistic strongly-coupled systems. What was wrong about random matrices? Recent developments from black hole physics teach us that the butterfly

effect in chaotic systems stems from delocalization of quantum information where initially local operators grow into non-local operators. However, for the GUE, the system does not distinguish local and non-local operators. To be concrete, let A_{local} be some one-qubit Pauli operator, and $A_{\text{non-local}} = UA_{\text{local}}U^\dagger$ be some non-local operator created by conjugating A_{local} via some non-local unitary U. Due to the Haar invariance of the GUE measure, we have

$$\langle A_{\text{local}}(0)A_{\text{local}}(t)\rangle_{\text{GUE}} = \langle A_{\text{non-local}}(0)A_{\text{non-local}}(t)\rangle_{\text{GUE}}\,. \tag{2.71}$$

As this argument suggests, the GUE is a good description of quantum systems which have no notion of locality. After the scrambling time, we expect that an initially local operator $A_{\text{local}}(0)$ will time evolve to $A_{\text{local}}(t)$ which has support on the whole system, and the notion of locality is lost (or at least obfuscated) after the scrambling time. We thus expect that $\langle A_{\text{local}}(0)A_{\text{local}}(t)\rangle_{\text{GUE}}$ will be a good description of two-point correlation functions after the scrambling time. Similarly, the GUE does not distinguish time-ordering as seen from $\langle A(0)B(t)C(2t)D(t)\rangle \simeq \langle A(0)B(t)C(0)D(t)\rangle$. This implies that, at late time scales when the GUE becomes a good description, the system forgets the locality of time. In this sense, the GUE captures physics of quantum chaos after the locality of spacetime is forgotten. We will elaborate on this issue in Sec. 2.6.

2.4 Frame potentials and random matrices

In discussions of black hole information loss, we often approximate the chaotic internal dynamics of a black hole as evolution by a Haar random unitary [1, 4], and talk about typical black hole states as random pure states generated by Haar unitaries [33]. While it is impractical to generate a Haar random unitary operator – due to its exponential quantum circuit complexity, as noted by [1] – it often suffices to sample from an ensemble that only reproduces the first few moments of the Haar ensemble. [11] made significant progress in quantifying chaos in OTOCs by relating the late-time decay of $2k$-point OTOCs to the k-th frame potential, measuring the distance to Haar-randomness.[15]

[15] Also of interest, [52] recently discussed scrambling and randomness and showed that the Rényi k-entropies averaged k-designs are typically near maximal.

One efficient way of generating a unitary k-design is to employ random local quantum circuits where one applies random two-qubit unitary gates at each unit time [53, 1, 54] and the ensemble monotonically becomes a k-design as time evolves. Motivated by tensor network descriptions of the AdS/CFT correspondence [55, 56], random local quantum circuits have been used as a toy model of the Einstein-Rosen bridge and the dynamics of the two-sided AdS black hole [10]. While such toy models are successful in capturing key qualitative features such as fast scrambling and complexity growth, their dynamics is not invariant under time translations. A natural question is to ask if systems of time-independent Hamiltonians are able to form k-designs or not.

In this section we study time-evolution by the ensemble of GUE Hamiltonians and quantify its approach to Haar-randomness by asking when it forms a unitary k-design. We consider the ensemble of unitary time evolutions at a fixed time t, with Hamiltonians drawn from the GUE

$$\mathcal{E}_t^{\text{GUE}} = \left\{ e^{-iHt}, \text{ for } H \in \text{GUE} \right\}. \tag{2.72}$$

As the frame potential quantifies the ensemble's ability to reproduce Haar moments, i.e. form a k-design, we will be interested in the time scales at which we approach "Haar values." Making use of the spectral form factors computed for the GUE, we derive explicit expressions for the frame potentials and extract the key time scales. We find that the GUE ensemble forms an approximate k-design after some time scales, but then deviates from being a k-design.

2.4.1 Overview of QI machinery

We begin by introducing the formalism of unitary k-designs and defining the frame potential. Consider a finite dimensional Hilbert space $\mathcal{H}$ of dimension d. We are primarily interested in ensembles of unitary operators $\mathcal{E} = \{p_i, U_i\}$, where the unitary U_i appears with some probability p_i. A familiar ensemble might be the Haar ensemble. The Haar ensemble is the unique left and right invariant measure on the unitary group $U(d)$, where

$$\int_{\text{Haar}} dU = 1\,, \quad \int_{\text{Haar}} dU\, f(U) = \int_{\text{Haar}} dU\, f(VU) = \int_{\text{Haar}} dU\, f(UV)\,, \tag{2.73}$$

for some function f and for all $V \in U(d)$. Taking k copies of $\mathcal{H}$, we can consider an operator O acting on $\mathcal{H}^{\otimes k}$, i.e. $O \in \mathcal{A}(\mathcal{H}^{\otimes k})$ the algebra of operators on the Hilbert space. The k-fold channel of O with respect to Haar is[16]

$$\Phi^{(k)}_{\text{Haar}}(O) \equiv \int_{\text{Haar}} dU\, (U^{\otimes k})^\dagger O U^{\otimes k}\,. \tag{2.74}$$

Given an ensemble of unitary operators $\mathcal{E} = \{p_i, U_i\}$, we might ask how Haar-random it is. More specifically, we should ask to what extent our ensemble reproduces the first k moments of the Haar ensemble, a notion quantified by unitary k-designs.[17] The k-fold channel with respect to the ensemble $\mathcal{E}$ is

$$\Phi^{(k)}_{\mathcal{E}}(O) \equiv \int_{U\in\mathcal{E}} dU (U^{\otimes k})^\dagger O U^{\otimes k}\,, \tag{2.75}$$

written here for a continuous ensemble. We say that an ensemble $\mathcal{E}$ is a unitary k-design if and only if

$$\Phi^{(k)}_{\mathcal{E}}(O) = \Phi^{(k)}_{\text{Haar}}(O)\,, \tag{2.76}$$

meaning we reproduce the first k moments of the Haar ensemble. But it does not make sense to compute the k-fold channels and check this equality for all operators in the algebra. Thus, we want a quantity which measures how close our ensemble is to being Haar-random. The frame potential, defined with respect to an ensemble as [57]

$$\mathcal{F}^{(k)}_{\mathcal{E}} = \int_{U,V\in\mathcal{E}} dU dV \left|\text{Tr}(U^\dagger V)\right|^{2k}\,, \tag{2.77}$$

measures Haar-randomness in the sense that it tells us how close the ensemble is to forming a unitary k-design. More precisely, it measures the 2-norm distance between the k-fold channel $\Phi^{(k)}_{\mathcal{E}}$ with respect to the ensemble $\mathcal{E}$, and the k-fold twirl $\Phi^{(k)}_{\text{Haar}}$ with respect to the Haar ensemble. The frame potential will be a central object of study in this section.

[16] The k-fold channel of O is also referred to in the literature as the k-fold twirl of O.

[17] Note that in the quantum information literature, these are often referred to as unitary t-designs. But here t will always denote time.

The k-th frame potential for the Haar ensemble is given by

$$\mathcal{F}^{(k)}_{\text{Haar}} = k! \quad \text{for} \quad k \leq d\,. \tag{2.78}$$

Furthermore, for any ensemble $\mathcal{E}$ of unitaries, the frame potential is lower bounded by the Haar value

$$\mathcal{F}^{(k)}_{\mathcal{E}} \geq \mathcal{F}^{(k)}_{\text{Haar}}\,, \tag{2.79}$$

with equality if and only if $\mathcal{E}$ is a k-design. In particular, the deviation from the Haar value $\mathcal{F}^{(k)}_{\mathcal{E}} - \mathcal{F}^{(k)}_{\text{Haar}}$ corresponds to the 2-norm distance of 2-fold quantum channels. The notion of an approximate k-design is reviewed in App. 2.A.

We will also need to compute moments of the Haar ensemble, i.e. the ability to integrate monomials of Haar random unitaries. The exact formula [58, 59] for evaluating these moments is given by

$$\int dU\, U^{j_1}_{k_1} \dots U^{j_n}_{k_n} U^{\dagger \ell_1}_{\ m_1} \dots U^{\dagger \ell_n}_{\ m_n} = \sum_{\sigma,\tau \in S_n} \delta^{j_1}_{m_{\sigma(1)}} \dots \delta^{j_n}_{m_{\sigma(n)}} \delta^{\ell_1}_{k_{\tau(1)}} \dots \delta^{\ell_n}_{k_{\tau(n)}} \mathcal{W}g(\tau\sigma^{-1})\,, \tag{2.80}$$

where, for the n-th moment, we sum over cycles of the permutation group S_n. The *Weingarten function* $\mathcal{W}g$, a function of cycles $\sigma \in S_n$, is defined in App. 2.C.3. Performing Haar integrals then simply amounts to contracting indices and computing the Weingarten functions.

2.4.2 Frame potentials for the GUE

$k = 1$ frame potential

The first frame potential for the GUE is written as

$$\mathcal{F}^{(1)}_{\text{GUE}} = \int dH_1 dH_2\, e^{-\frac{d}{2}\text{Tr}H_1^2} e^{-\frac{d}{2}\text{Tr}H_2^2} \left|\text{Tr}\left(e^{iH_1 t} e^{-iH_2 t}\right)\right|^2 . \tag{2.81}$$

Noting that the GUE measure is invariant under unitary conjugation, we find

$$\mathcal{F}^{(1)}_{\text{GUE}} = \int_{\text{Haar}} dU dV \int dH_1 dH_2\, e^{-\frac{d}{2}\text{Tr}H_1^2} e^{-\frac{d}{2}\text{Tr}H_2^2} \left|\text{Tr}\left(U^\dagger \Lambda_1^\dagger U V^\dagger \Lambda_2 V\right)\right|^2 , \tag{2.82}$$

where we define $\Lambda \equiv Ue^{-iHt}U^\dagger$, i.e. the matrix exponential of the GUE matrix in the diagonal basis. Going into the eigenvalue basis, we can express the GUE integral as

$$\mathcal{F}^{(1)}_{\text{GUE}} = \int D\lambda_1 D\lambda_2 \int dU \, \text{Tr}\big(U^\dagger \Lambda_1^\dagger U \Lambda_2\big)\text{Tr}\big(\Lambda_2^\dagger U^\dagger \Lambda_1 U\big)\,, \tag{2.83}$$

where we have used the left and right invariance of the Haar measure to write the expression as a single Haar integral. Written out explicitly with indices,

$$\mathcal{F}^{(1)}_{\text{GUE}} = \int D\lambda_1 D\lambda_2 \int dU \left(U^{j_1}_{k_1} U^{j_2}_{k_2} U^{\dagger \ell_1}_{\ m_1} U^{\dagger \ell_2}_{\ m_2} \, \Lambda_1^{\dagger m_1}{}_{j_1} \Lambda_2{}^{k_1}_{\ell_1} \Lambda_2^{\dagger k_2}{}_{\ell_2} \Lambda_1{}^{m_2}_{j_2}\right), \tag{2.84}$$

and we can do the Haar integral using the second moment

$$\begin{aligned}\int dU \, U^{j_1}_{k_1} U^{j_2}_{k_2} U^{\dagger \ell_1}_{\ m_1} U^{\dagger \ell_2}_{\ m_2} = \frac{1}{d^2-1}\Big(&\delta^{j_1}_{m_1}\delta^{j_2}_{m_2}\delta^{\ell_1}_{k_1}\delta^{\ell_2}_{k_2} + \delta^{j_1}_{m_2}\delta^{j_2}_{m_1}\delta^{\ell_1}_{k_2}\delta^{\ell_2}_{k_1} \\ &- \frac{1}{d}\delta^{j_1}_{m_1}\delta^{j_2}_{m_2}\delta^{\ell_1}_{k_2}\delta^{\ell_2}_{k_1} - \frac{1}{d}\delta^{j_1}_{m_2}\delta^{j_2}_{m_1}\delta^{\ell_1}_{k_1}\delta^{\ell_2}_{k_2}\Big)\,.\end{aligned} \tag{2.85}$$

We find

$$\mathcal{F}^{(1)}_{\text{GUE}} = \int D\lambda_1 D\lambda_2 \, \frac{1}{d^2-1}\left(\text{Tr}\Lambda_1^\dagger \text{Tr}\Lambda_1 \text{Tr}\Lambda_2^\dagger \text{Tr}\Lambda_2 + d^2 - \frac{1}{d}\Big(d\text{Tr}\Lambda_1^\dagger \text{Tr}\Lambda_1 + d\text{Tr}\Lambda_2^\dagger \text{Tr}\Lambda_2\Big)\right)$$

or equivalently

$$\mathcal{F}^{(1)}_{\text{GUE}} = \frac{1}{d^2-1}\Big(\mathcal{R}_2^2 + d^2 - 2\mathcal{R}_2\Big)\,, \tag{2.86}$$

written in terms of the 2-point form factor

$$\mathcal{R}_2 = \int D\lambda \sum_{i,j} e^{i(\lambda_i - \lambda_j)t}\,. \tag{2.87}$$

We know from the expression found in Sec. 2.2, that at early times $\mathcal{R}_2 \sim d^2$, so the early time behavior of the frame potential is dominated by the $\mathcal{R}_2^2$ term until near the dip time. At the dip time, $\mathcal{R}_2 \approx \sqrt{d}$ and $\mathcal{F}^{(1)}_{\text{GUE}} \approx 1$, achieving the Haar value and forming a 1-design. At late times $t \to \infty$, we take the late time limit of $\mathcal{R}_2$ where only the δ_{ij} terms contribute, and find $\mathcal{R}_2 \approx d$, meaning that the first frame potential $\mathcal{F}^{(1)}_{\text{GUE}} \approx 2$ or double the Haar value.

The first frame potential is plotted in Fig. 2.6.

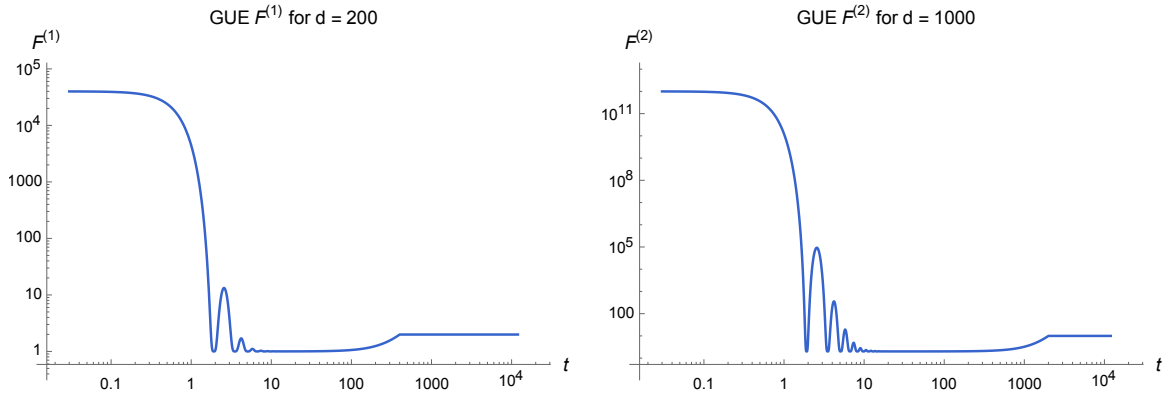


Figure 2.6: The first and second frame potentials for the GUE, using the infinite temperature 2-point and 4-point form factors computed in Sec. 2.2, plotted for $d = 200$ and $d = 1000$, respectively. We observe the decay to the Haar value at the dip time and a subsequent rise at late times.

A common intuition is that physical systems will become more and more uniformly random as time passes. Then one might expect that the frame potential, a measure of Haar randomness, would be a monotonically decreasing function with time. While it is monotonic for random local quantum circuits, we found that it is not generically monotonic for ensembles of unitaries generated by fixed Hamiltonians.[18] In Sec. 2.6, we propose an alternative quantity which may be monotonic at late times.

$k = 2$ frame potential

We can similarly compute the second frame potential using the unitary invariance of the GUE measure:

$$\begin{aligned}\mathcal{F}^{(2)}_{\text{GUE}} &= \int dH_1 dH_2\, e^{-\frac{d}{2}\text{Tr}H_1^2} e^{-\frac{d}{2}\text{Tr}H_2^2} \left|\text{Tr}\left(e^{iH_1 t} e^{-iH_2 t}\right)\right|^4 \\ &= \int D\lambda_1 D\lambda_2 \int dU\, \text{Tr}\Big(U^\dagger \Lambda_1^\dagger U \Lambda_2\Big)\text{Tr}\Big(\Lambda_2^\dagger U^\dagger \Lambda_1 U\Big)\text{Tr}\Big(U^\dagger \Lambda_1^\dagger U \Lambda_2\Big)\text{Tr}\Big(\Lambda_2^\dagger U^\dagger \Lambda_1 U\Big)\,,\end{aligned} \tag{2.88}$$

where again, Λ is the exponentiated diagonal matrix. The fourth moment of the Haar

[18] Frame potentials monotonically decrease in local random circuits and Brownian circuits [53, 3] where the time evolution is Markovian in the sense that the system samples different Hamiltonians, or infinitesimal time evolution operators, at random at each time step. In Markovian ensembles, spectral form factors are monotonically decreasing, and there is no ramp behavior. If the ensemble $\mathcal{E}$ is generated by a Markovian process and is invariant under complex transposition $\mathcal{E} = \mathcal{E}^\dagger$, then we have $\mathcal{F}^{(k)}(t) = \mathcal{R}_{2k}(2t)$.

ensemble that appears here generates $4!^2 = 576$ terms. Recalling Eq. (2.80), we can compute the fourth moment by computing the necessary Weingarten functions and summing over δ-function contractions.

We relegate the presentation of the full expression for the $k = 2$ frame potential, and the definitions of the spectral quantities on which it depends, to Appendix 2.C.2. While $\mathcal{F}^{(2)}_{\text{GUE}}$ depends on a number of spectral form factors, the dominant and interesting behavior is entirely captured by the 2-point and 4-point spectral form factors. At early times, the dominant contribution is

$$\text{Early}: \quad \mathcal{F}^{(2)}_{\text{GUE}} \approx \frac{\mathcal{R}_4^2}{d^4}\,. \tag{2.89}$$

As we approach the dip time, the spectral quantities in the second frame potential,

$$\mathcal{F}^{(2)}_{\text{GUE}} \approx 2 + \frac{\mathcal{R}_4^2}{d^4} - \frac{8\mathcal{R}_4^2}{d^6} + \frac{6\mathcal{R}_4^2}{d^8} - \frac{36\mathcal{R}_2^2}{d^4} + \frac{4\mathcal{R}_2^2}{d^2} + \frac{64\mathcal{R}_2\mathcal{R}_4}{d^6} - \frac{8\mathcal{R}_2\mathcal{R}_4}{d^4} + \dots\,, \tag{2.90}$$

are suppressed. From the calculation in Sec. 2.2, we have $\mathcal{R}_2 \approx \sqrt{d}$ and $\mathcal{R}_4 \approx d$ at the dip, meaning all terms are suppressed, with the exception of the leading constant. Thus, at the dip time, the $\mathcal{E}_t^{\text{GUE}}$ achieves the Haar value $\mathcal{F}^{(2)}_{\text{Haar}} \approx 2$ and forms an approximate unitary 2-design.

At late times, in the infinite time average, we know that $\mathcal{R}_2 \to d$, and $\mathcal{R}_4 \to 2d^2 - d$ from the two eigenvalue pairings in the sum where the exponent vanishes, i.e. $\delta_{ik}\delta_{j\ell}$ and $\delta_{i\ell}\delta_{jk}$, and accounting for the $i = j = k = \ell$ terms. This tells us that the only terms that survive at late times, and are not suppressed in d, are

$$\text{Late}: \quad \mathcal{F}^{(2)}_{\text{GUE}} \approx 2 + \frac{\mathcal{R}_4^2}{d^4} + \frac{4\mathcal{R}_2^2}{d^2}\,, \tag{2.91}$$

which gives us $\mathcal{F}^{(2)}_{\text{GUE}} \approx 10$, to leading order in $1/d$.

2.4.3 Higher k frame potentials

Let us review what we have discussed so far.

$k = 1$ Frame Potential

We computed the first frame potential for the GUE to be

$$\mathcal{F}^{(1)}_{\text{GUE}} = \frac{1}{d^2-1}\Big(\mathcal{R}_2^2 + d^2 - 2\mathcal{R}_2\Big) \approx 1 + \frac{\mathcal{R}_2^2}{d^2} - \frac{2\mathcal{R}_2}{d^2} \tag{2.92}$$

for large d. In the late time limit, where $t \to \infty$, we have that $\mathcal{R}_2 \to d$, and the late time behavior goes like $\mathcal{F}^{(1)}_{\text{GUE}} \sim 1 + \mathcal{R}_2^2/d^2$, and $\mathcal{F}^{(1)}_{\text{GUE}} \to 2$ or double the Haar value.

$$\text{Early} : \mathcal{F}^{(1)}_{\text{GUE}} \approx \frac{\mathcal{R}_2^2}{d^2}\,, \quad \text{Dip} : \mathcal{F}^{(1)}_{\text{GUE}} \approx 1\,, \quad \text{Late} : \mathcal{F}^{(1)}_{\text{GUE}} \approx 2\,. \tag{2.93}$$

$k = 2$ Frame Potential

We discussed the early and dip behaviors above. The terms unsuppressed at late times are

$$\mathcal{F}^{(2)}_{\text{GUE, late}} \approx 2 + \frac{\mathcal{R}_4^2}{d^4} + \frac{4\mathcal{R}_2^2}{d^2}\,. \tag{2.94}$$

Since $\mathcal{R}_2 \to d$ and $\mathcal{R}_4 \to 2d^2 - d$ in the late time limit, $\mathcal{F}^{(2)}_{\text{GUE}}$ approaches 10.

$$\text{Early} : \mathcal{F}^{(2)}_{\text{GUE}} \approx \frac{\mathcal{R}_4^2}{d^4}\,, \quad \text{Dip} : \mathcal{F}^{(2)}_{\text{GUE}} \approx 2\,, \quad \text{Late} : \mathcal{F}^{(2)}_{\text{GUE}} \approx 10\,. \tag{2.95}$$

$k = 3$ Frame Potential

The full expression for the third frame potential is given in App. 2.C.2. The leading order behavior at early times is $\mathcal{R}_6^2/d^6$, and at the dip time, the third frame potential approaches its Haar value. Again, the late time behavior above is better understood by looking at the dominant form factors. At late times, the terms that contribute at zeroth order in d are

$$\mathcal{F}^{(3)}_{\text{GUE, late}} \approx 6 + \frac{\mathcal{R}_6^2}{d^6} + \frac{9\mathcal{R}_4^2}{d^4} + \frac{18\mathcal{R}_2^2}{d^2} \to 96\,, \tag{2.96}$$

as $\mathcal{R}_2 \to d$, $\mathcal{R}_4 \to 2d^2$, and $\mathcal{R}_6 \to 6d^3$ to leading order in d. In summary,

$$\text{Early} : \mathcal{F}^{(3)}_{\text{GUE}} \approx \frac{\mathcal{R}_6^2}{d^6}\,, \quad \text{Dip} : \mathcal{F}^{(3)}_{\text{GUE}} \approx 6\,, \quad \text{Late} : \mathcal{F}^{(3)}_{\text{GUE}} \approx 96\,. \tag{2.97}$$

$k = 4$ Frame Potential

It is not tractable to compute the $k = 4$ frame potential, as the Haar integrals involved (the

eighth moment of the Haar ensemble), generate $(8!)^2 \sim 1.6$ billion terms. But the interesting behavior can be understood from the dominant terms at leading order in d at different time scales. Recall that the $2k$-th moment of the Haar ensemble can be written as the sum of δ-functions and the Weingarten function $\mathcal{W}g$ (defined in App. 2.C.3) over elements of the permutation group S_{2k}. At large d, the Weingarten functions go as [60, 59]

$$\mathcal{W}g(\sigma) \sim \frac{1}{d^{4k-\#\text{cycles}}}\,, \tag{2.98}$$

where '#cycles' denotes the number of cycles in the permutation σ. The Weingarten function contributing at leading order in $1/d$ is the one labeled by the partitioning of $2k$ into ones, i.e. the trivial permutation of S_{2k}, which contributes as

$$\mathcal{W}(\{1,1,\ldots\}) \sim \frac{1}{d^{2k}}\,. \tag{2.99}$$

All other Weingarten functions, labeled by the integer partitions of $2k$, contribute at subleading order at early and late times. Thus, instead of computing the full fourth frame potential, we can compute the terms of combinations of spectral functions with this Weingarten function as their coefficient. In the sum over elements of the permutation group $\sigma, \tau \in S_{2k}$, we simply need the terms where $\tau\sigma^{-1}$ is the trivial permutation, i.e. $\tau = \sigma$. Computing this we find the dominant contribution to the $k = 4$ frame potential, at leading order in $1/d$. The full expression is still too large to reproduce here, but we can comment on the relevant features. The early time behavior is

$$\mathcal{F}^{(4)}_{\text{GUE, early}} \approx \frac{\mathcal{R}_8^2}{d^8}\,. \tag{2.100}$$

At the dip, where $\mathcal{R}_n \sim d^{n/2}$, all terms are suppressed, leaving only the constant Haar value 24. Lastly, the late time behavior is

$$\mathcal{F}^{(4)}_{\text{GUE, late}} \approx 24 + \frac{\mathcal{R}_8^2}{d^8} + \frac{16\mathcal{R}_6^2}{d^8} + \frac{72\mathcal{R}_4^2}{d^4} + \frac{96\mathcal{R}_2^2}{d^2} \to 1560\,, \tag{2.101}$$

In summary,

$$\text{Early} : \mathcal{F}^{(4)}_{\text{GUE}} \approx \frac{\mathcal{R}_8^2}{d^8}\,, \quad \text{Dip} : \mathcal{F}^{(4)}_{\text{GUE}} \approx 24\,, \quad \text{Late} : \mathcal{F}^{(4)}_{\text{GUE}} \approx 1560\,. \tag{2.102}$$

k-th Frame Potential

We are now poised to discuss the general form of the k-th frame potential

$$\text{Early} : \mathcal{F}^{(k)}_{\text{GUE}} \approx \frac{(\mathcal{R}_{2k})^2}{d^{2k}}\,, \quad \text{Dip} : \mathcal{F}^{(k)}_{\text{GUE}} \approx k!\,. \tag{2.103}$$

We can also determine what the general late time value should look like. Above, we understood that the plateau value of the k-th frame potential is the sum of the Haar value and the contributions of the spectral functions. It was only the squares of the spectral functions that gave contributions which were not suppressed by $1/d$ at late times. Extrapolating from above, we expect the k-th frame potential to have

$$\mathcal{F}^{(k)}_{\text{GUE, late}} \approx \text{Haar} + \text{spectral functions} \approx k! + \frac{\mathcal{R}_{2k}^2}{d^{2k}} + c_1 \frac{\mathcal{R}_{2k-2}^2}{d^{2k-2}} + \ldots + c_{k-1}\frac{\mathcal{R}_2^2}{d^2}\,, \tag{2.104}$$

with coefficients c_ℓ. Given the way the spectral form factors are generated from Haar integration, we can understand these coefficients as the number of partial bijections of a given length. For example, for $k = 3$ there are 24 partial bijections on a 3 element set of length 2, i.e. 24 nonclosed cycles of length two, which gives us 24 ways of constructing the 2-point functions for $k = 3$. More generally, the coefficients above can be written as

$$c_\ell(k) = \binom{k}{\ell}^2 \ell!\,, \tag{2.105}$$

where for $k = 4$, we have the coefficients 1, 16, 72, 96, 24. The k-th coefficient is the Haar value $c_k(k) = k!$, i.e. the number of ways to construct 0-point functions in the Haar integration. We can then write down the general late time behavior for the k-th frame potential

$$\mathcal{F}^{(k)}_{\text{GUE, late}} \approx \sum_{\ell=0}^{k} c_\ell(k) \frac{\mathcal{R}^2_{2(k-\ell)}}{d^{2(k-\ell)}}\,. \tag{2.106}$$

Since the late time value of the $2k$-point spectral form factor is, to leading order in d, $\mathcal{R}_{2k} = k!d^k$, the late time floor value for the k-th frame potential of the GUE is

$$\mathcal{F}^{(k)}_{\text{GUE, late}} \approx \sum_{\ell=0}^{k} \binom{k}{\ell}^2 \ell!\big((k-\ell)!\big)^2 = \sum_{\ell=0}^{k} \frac{k!^2}{\ell!}\,. \tag{2.107}$$

where the first few terms of this sequence are 2, 10, 96, 1560.

We emphasize that while the purpose of this section is to understand GUE Hamiltonians, the derivations in this subsection where we relate the frame potential to spectral $2k$-point functions only used the unitary invariance of the measure to proceed in doing the calculations by Haar integration. Thus, if we are handed an ensemble whose measure is unitarily invariant, the same relations hold.

2.4.4 Frame potentials at finite temperature

We now generalize the discussion of the frame potential to ensembles at finite temperature and compute the thermal frame potential for the GUE. Again we consider the ensemble of unitary time evolutions at a fixed time t, with H drawn from an ensemble $\mathcal{E}$. One might consider generalizing the frame potential to finite temperature by defining the frame potential with respect to a thermal density matrix $\rho_\beta = e^{-\beta H}/\text{Tr}(e^{-\beta H})$, and taking thermal expectation values. With this in mind, we define the frame potential at finite temperature by taking the average over all thermal $2k$-point functions, with the operator insertions A and B spaced equidistant on the thermal circle

$$\langle AB(t)\dots AB(t)\rangle = \text{Tr}\big((e^{-\beta H/2k}Ae^{-\beta H/2k}B(t)\dots e^{-\beta H/2k}Ae^{-\beta H/2k}B(t)\big)/\text{Tr}e^{-\beta H}\,. \tag{2.108}$$

Averaging the norm-squared $2k$-point correlation function over all operators and then averaging over the ensemble, we find

$$\mathcal{F}^{(k)}_{\mathcal{E}_\beta} = \int dH_1 dH_2 \frac{\big|\text{Tr}\big(e^{-(\beta/2k-it)H_1}e^{-(\beta/2k+it)H_2}\big)\big|^{2k}}{\text{Tr}(e^{-\beta H_1})\text{Tr}(e^{-\beta H_2})/d^2}\,. \tag{2.109}$$

Note that this definition differs from the one in the Appendix of [11] by a factor of d^2. With this slight change in normalization, we reduce to the usual frame potential $\mathcal{F}_{\mathcal{E}}^{(k)}$ at infinite temperature.

$k = 1$ Frame Potential

Let us compute the first thermal frame potential for GUE Hamiltonians:

$$\mathcal{F}^{(1)}_{\text{GUE}}(t,\beta) = \int D\lambda_1 D\lambda_2 \int dU \frac{\left|\text{Tr}\left(U^\dagger e^{-(\beta/2-it)D_1} U e^{-(\beta/2+it)D_2}\right)\right|^2}{\text{Tr}(e^{-\beta H_1})\text{Tr}(e^{-\beta H_2})/d^2}\,. \tag{2.110}$$

where we use the invariance of the GUE measure under unitary conjugation, diagonalize H where D is the diagonalized Hamiltonian, and use the left and right invariance of the Haar measure to write a single Haar integral. Doing the Haar integral, we find

$$\mathcal{F}^{(1)}_{\text{GUE}}(t,\beta) = \frac{1}{d^2-1}\left(\widetilde{\mathcal{R}}_2^2(t,\beta/2) + d^2 - 2\widetilde{\mathcal{R}}_2(t,\beta/2)\right), \tag{2.111}$$

where we define

$$\widetilde{\mathcal{R}}_2(t,\beta) \equiv \left\langle \frac{Z(t,\beta)Z^*(t,\beta)}{Z(2\beta)/d} \right\rangle_{\text{GUE}} = \int D\lambda \frac{\sum_{ij} e^{it(\lambda_i-\lambda_j)} e^{-\beta(\lambda_i+\lambda_j)}}{\sum_i e^{-2\beta\lambda_i}/d}\,, \tag{2.112}$$

which is normalized such that we recover the infinite temperature form factor $\mathcal{R}_2(t)$ when $\beta \to 0$. This normalization differs from $\langle |Z(t,\beta)|^2/Z(\beta)^2\rangle$, which gives an initial value of one. Here the thermal form factor which naturally arises from the thermal frame potential has a late time value which is β-*independent*. The initial value of $\mathcal{R}_2(t,\beta)$, and thus $\mathcal{F}^{(1)}_{\text{GUE}}(t,\beta)$, depends on the β.

In stating the time scales for the thermal frame potential, we will work with the 'quenched' version of Eq. (2.112) where the numerator and denominator are averaged separately. As we mentioned in Sec. 2.2.2, the 'annealed' 2-point form factor is the correct object to consider, but we opt to work with the more analytically tractable quenched form factor. Numerically, the two functions are in close agreement with each other.

2.4.5 Time scales from GUE form factors

With an understanding of the behavior of the GUE spectral form factors from Sec. 2.2.2, we can now look at the time scales for the dip and plateau of the first frame potential

$$\mathcal{F}^{(1)}_{\text{GUE}} = \frac{1}{(d^2-1)}\left(\mathcal{R}_2^2 + d^2 - 2\mathcal{R}_2\right). \tag{2.113}$$

At $t_d \approx \sqrt{d}$, when $\mathcal{R}_2 \approx \sqrt{d}$, we reach the minimal Haar value of 1, and at the plateau time $t_p = 2d$, when $\mathcal{R}_2 = d$, we reach the late time value of 2.

There is another time scale at play here which is an artifact of working at infinite temperature. We might also ask what is the first time the form factor or frame potential reaches its minimal value. This time scale can be attributed to the first zero of the Bessel function, $J_1(2t) = 0$ at $t \approx 1.92$, and is universal for all values of d. This is the first time at which the ensemble becomes a 1-design. Something like the scrambling time, where the frame potential begins to deviate rapidly from its initial value, occurs at $\mathcal{O}(1)$ time.

Using the explicit expression for the GUE 4-point form factor, we can also verify the expected time scales in the second frame potential $\mathcal{F}^{(2)}_{\text{GUE}}$. At the dip time, $t_d \approx \sqrt{d}$, we have that all the form factors appearing in the $\mathcal{F}^{(2)}_{\text{GUE}}$ are suppressed by powers of d, and thus the leading term is the Haar value, $\mathcal{F}^{(2)}_{\text{GUE}}(t_d) \approx 2$. Further, the plateau values of the spectral form factors $\mathcal{R}_2$ and $\mathcal{R}_4$ give us the late time value of $\mathcal{F}^{(2)}_{\text{GUE}} \approx 10$.

Lastly, we can extract the time scales and values of the finite temperature frame potential from our discussion of $\mathcal{R}_2(t,\beta)$. The initial value of the first frame potential is

$$\mathcal{F}^{(1)}_{\text{GUE}}(t=0,\beta) = d^2\frac{h_1(\beta/2)^4}{h_1(\beta)^2}, \tag{2.114}$$

where $h_1(\beta) = J_1(2i\beta)/i\beta$. At the dip time, $t_d \approx h_2(\beta/2)\sqrt{d}$, the thermal form factor defined above $\widetilde{\mathcal{R}}_2(t_d,\beta/2) \approx \sqrt{d}h_3(\beta/2)/h_1(\beta)$, with the functions defined in Sec. 2.2.2. For $\beta \ll d$, we have

$$\mathcal{F}^{(1)}_{\text{GUE}}(t_d,\beta) \approx 1. \tag{2.115}$$

Finally, as we can see from time averaging Eq. (2.112), at the plateau time

$$\mathcal{F}^{(1)}_{\text{GUE}}(t_p, \beta) = 2\,, \tag{2.116}$$

for any β, as the late time value of the thermal frame potential does not depend on the temperature.

Let us briefly comment on the dip value of the k-th frame potential at infinite temperature. As we discussed, at the dip time $t_d \approx \sqrt{d}$, the frame potentials reached the Haar value and form an approximate k-design for some k. Determining the size of k requires an understanding of the corrections to the dip value. The leading order correction to the Haar value at the dip comes from $\mathcal{R}_2^2/d^2 \sim 1/d$, the coefficient of which is $c_{k-1}(k) = k!\,k$. So at the dip time

$$\mathcal{F}^{(k)}_{\text{GUE}}(t_d) \approx k!\left(1 + \frac{k}{d}\right)\,, \tag{2.117}$$

meaning we form an approximate k-design for $k \ll d$.

The claim that the GUE forms a k-design at intermediate times but then deviates from this behavior at late times might at first seem surprising, but the late time behavior makes sense if we consider the dephasing of GUE eigenvalues in the $t \to \infty$ limit. Under the exponential map $\lambda \to e^{i\lambda t}$, the GUE eigenvalues are distributed around the circle and at early times will still be correlated and logarithmically repel. However, at late times the eigenvalues will spread uniformly around the circle. Moreover, explicitly computing the level density for the GUE under the exponential map and taking the long time limit, one finds that the density becomes constant and the eigenvalues are independently and uniformly distributed. Eigenvalue statistics of Haar random unitary operators can be characterized by the following well-known relation [61][19]

$$\int_{\text{Haar}} dU\, \text{tr}(U^t)\text{tr}(U^{\dagger t}) = t \qquad k \leq d\,. \tag{2.118}$$

If we suppose that the eigenvalue distribution of U is random, then $\int dU\, \text{tr}(U^t)\text{tr}(U^{\dagger t})$ would not depend on t. Therefore, the late-time eigenvalue statistics of unitaries generated by fixed

[19] If one views t as a discrete time and U as a time evolution in a unit time with a Hamiltonian $H = i \log U$, then the above equation mimics the late-time ramp and plateau behavior.

GUE matrices is quite different from those of Haar unitaries, which have eigenvalue repulsion.

2.5 Complexity and random matrices

In recent years, the notion of quantum complexity has attracted significant attention in the study of quantum many-body systems [62, 63, 64]. By quantum complexity of a quantum state $|\psi\rangle$, we mean the minimal number of elementary local quantum gates necessary to (approximately) create $|\psi\rangle$ from a trivial product state with no entanglement. A similar characterization applies to the quantum complexity of unitary operators constructed from the identity operator. Quantum complexity provides deep insight into what kinds of physical operations are allowed (or prohibited) in a given physical system as states or operators of very large complexity cannot be prepared or implemented in a short period of time by the evolution of local Hamiltonians with finite energy density. Quantum complexity has also proven useful in condensed matter physics where topological phases of matter can be classified in terms of the quantum complexity of ground state wavefunctions [65]. More recently, it was asked whether the AMPS thought experiment can be carried out in a physically reasonable amount of time and resources by considering the computational complexity of decoding the Hawking radiation [66]. In the past few years, quantum complexity has been considered in holography as a possible CFT observable[20] to study the late-time dynamics of the AdS black holes [63, 64].

Despite all the promises of the usefulness of quantum complexity, a precise understanding of the growth of quantum complexity in quantum many-body systems, especially in AdS/CFT, continues to elude us. While it is possible to see a hint of complexity growth from entanglement dynamics at early times before the scrambling time,[21] the late-time complexity growth remains difficult to observe as the extremal surfaces do not go through the interior of the black hole and entanglement entropies get saturated at late times. From a mathematical perspective, it is extremely challenging to compute the quantum gate complexity of a given quantum state $|\psi\rangle$ as one essentially needs to consider all the possible

[20] At least with respect to some subspace of states of the boundary CFT.
[21] For example, from the level-statistics of the entanglement spectrum [67].

quantum circuits creating $|\psi\rangle$ and find the one with the minimal number of gates. Thus it would be valuable to have an analytical toy example of Hamiltonians whose dynamics indeed makes the quantum complexity of wavefunctions increase even after the scrambling time by providing a rigorous lower bound on quantum complexity.

Here, we present analysis of complexity growth of typical Hamiltonian time evolution by GUEs and show that quantum complexity indeed grows in time. A lower bound on a typical unitary operator in an ensemble $\mathcal{E}$ can be computed from a simple counting argument. Observe that short depth quantum circuits can prepare only a small number of unitary operators which occupy a tiny fraction of the whole space of unitary operators. The idea is that, if there are so many unitary operators in $\mathcal{E}$ which are sufficiently far apart and distinguishable, then most of operators in $\mathcal{E}$ cannot be created by a short depth circuit. Furthermore, it has been found that lower bounds on the number of distinguishable unitary operators in $\mathcal{E}$ can be obtained by frame potentials, a measure of randomness in $\mathcal{E}$. Although such a counting argument often gives a rather loose lower bound, it is still possible to obtain a rigorous complexity lower bound for a system of quantum many-body Hamiltonians. See [11] for a rigorous treatment and details.

To be concrete, let us consider a system of qubits where we pick a pair of qubits and apply an arbitrary two-qubit gate at each step. While the circuit complexity for generating an ensemble and the circuit complexity for generating a particular unitary in the ensemble are different, the former provides an approximate lower bound for the circuit complexity of typical unitary operators in the ensemble [11]. We define the number of quantum gates necessary to create an ensemble $\mathcal{E}$ by a quantum gate complexity $\mathcal{C}_{\text{gate}}$. The lower bound on the quantum gate complexity is then given by

$$\mathcal{C}_{\text{gate}} \geq \frac{2kn - \log_2 \mathcal{F}^{(k)}}{2\log(n)}\,, \tag{2.119}$$

up to some constant multiplicative factor. Let us consider the bound for small k. In Sec. 2.4, we found that $\mathcal{F}^{(k)}$ drops to its minimal value $\sim k!$ at $t \sim \mathcal{O}(1)$ (the first zero of the Bessel

function). We thus have

$$\mathcal{C}_{\rm gate}(t) \geq \frac{2kn - \log_2 \frac{\mathcal{R}_{2k}^2(t)}{d^{2k}}}{2\log(n)} \simeq \frac{4kn - \log_2 \mathcal{R}_{2k}^2(t)}{2\log(n)} \simeq \frac{4k(n - \log_2 \mathcal{R}_1(t))}{2\log(n)} \tag{2.120}$$

up to the first dip time $t_{\rm dip} \sim \mathcal{O}(1)$ where we have used an approximation $\mathcal{R}_{2k} \simeq (\mathcal{R}_1)^{2k}$. Thus, at $t \sim \mathcal{O}(1)$, the following lower bound on the complexity is obtained:

$$\mathcal{C}_{\rm gate}(t_{\rm dip}) \geq \mathcal{O}\left(\frac{kn}{\log(n)}\right) . \tag{2.121}$$

Converting it into a quantum circuit complexity, we obtain

$$\mathcal{C}_{\rm circuit}(t_{\rm dip}) \geq \mathcal{O}\left(\frac{k}{\log(n)}\right) . \tag{2.122}$$

This lower bound should be valid as long as $k \sim \mathcal{O}(1)$. As we have discussed in Sec. 2.2 and Sec. 2.4, the early-time oscillations of spectral form factors and frame potentials disappear at finite temperature. It would be then useful to consider the complexity lower bound based on envelope functions of form factors and frame potentials. Since the asymptotic behavior is given by $\mathcal{R}_1(t) \sim 1/t^{3/2}$, we would have

$$\mathcal{C}_{\rm gate}(\beta, t) \geq \mathcal{O}\left(\frac{k\log t}{\log(n)}\right) , \tag{2.123}$$

where β implies that we consider the asymptotic behaviors of the envelope. Thus, the quantum circuit complexity grows at least logarithmically in t up to the thermal dip time.

While the above studies are able to provide rigorous lower bounds on quantum circuit complexity, the bounds are not meaningful when k is small. To obtain a meaningful lower bound on quantum complexity, we need to evaluate the frame potential and form factor for large k. Analytically computing $\mathcal{R}_{2k}$ and $\mathcal{F}^{(k)}$ for large k seems rather challenging. Instead, we employ a certain heuristic argument to derive the decay of $\mathcal{R}_{2k}$ and $\mathcal{F}^{(k)}$. Let us begin by recalling the early-time behavior of 1-point form factor. The 1-point form factor $\mathcal{R}_1(t)$

can be analytically written via a contour integral as follows [68]

$$\mathcal{R}_1(t) = de^{-t^2/2d} \oint \frac{du}{2\pi i} \left(\frac{1}{-it}\right) \left(1 - \frac{it}{du}\right)^d e^{-itu} \,. \tag{2.124}$$

For $d \to \infty$, the integral gives the Bessel function:

$$\oint \frac{du}{2\pi i} \left(\frac{1}{-it}\right) \left(1 - \frac{it}{du}\right)^d e^{-itu} \simeq \frac{J_1(2t)}{t} \,. \tag{2.125}$$

But $J_1(2t) \simeq t$ for $t \ll 1$, so we have

$$\mathcal{R}_1(t) \simeq de^{-t^2/2d} \frac{J_1(2t)}{t} \,, \tag{2.126}$$

where the Gaussian decay is dominant for $t \ll 1$ while, for $1 \ll t \ll \sqrt{d}$, the Bessel function dominates the decay. In a similar manner, the $2k$-point form factor can be analytically written as

$$\mathcal{R}_{2k}(t) = d^{2k} e^{-kt^2/d} \oint \prod_{j=1}^{2k} \frac{du_j}{2\pi i} \left(1 + (-1)^j \frac{it}{du_j}\right)^d e^{(-1)^j i t u_j} \det\left(\frac{1}{u_j - u_k + (-1)^j it/d}\right) \,, \tag{2.127}$$

where the sign of $\pm it$ depends on the index of u_i and the integral part is equal to unity at $t = 0$. In previous sections, we have neglected the Gaussian decay because our discussions were mostly centered on small k spectral form factors. But, for large k, the Gaussian decay part is no longer negligible. Let us bound the form factor by using the Gaussian decay part only by neglecting the decay contribution from Bessel functions in the integral part:

$$\mathcal{R}_{2k}(t) \leq d^{2k} e^{-kt^2/d} \,. \tag{2.128}$$

While the validity of this inequality for large k remains unclear, we assume its validity up to the dip time $\sim \sqrt{d}$ when ramp behavior kicks in. The notion of unitary k-design and its application to complexity would be meaningful only up to $k \sim \mathcal{O}(d)$ (see [11] for instance).

By using this approximate bound for $k = cd$ with $c \sim \mathcal{O}(1)$, we will have

$$\mathcal{F}^{(cd)} \lesssim d^{2k} e^{-2ct^2} \tag{2.129}$$

up to the dip time $\sim \sqrt{d}$. This leads to the following estimate of quantum complexity growth for the GUE:

$$\mathcal{C}_{\text{gate}} \gtrsim \frac{ct^2}{\log(n)}\,, \tag{2.130}$$

which predicts a quadratic growth of quantum complexity.

Let us compare our estimate with predictions from the AdS/CFT correspondence. According to the conjecture that quantum complexity is proportional to the volume in the bulk, the early-time complexity (volume) growth is quadratic in time, and then becomes linear in time. Our analysis above suggests that the complexity growth for the GUE is (at least) quadratic in t for a long time until very close to the saturation of quantum complexity $\sim d$. One may find that t^2 complexity growth is unphysical as the system has evolved only for time t. The point is that the GUE Hamiltonian is generically non-local and is comprised of $\mathcal{O}(n)$-body terms whereas we measure quantum complexity by using two-local quantum gates as building blocks.

2.6 Characterization of Haar-invariance

From the perspective of operator delocalization, it is clear why the GUE fails to characterize information scrambling and dynamics in local quantum systems at early times. Recall that the GUE is Haar-invariant, meaning

$$\int_{U\in\text{Haar}} dU \int_{H\in\text{GUE}} dH\, f(UHU^\dagger) = \int_{H\in\text{GUE}} dH\, f(H)\,, \tag{2.131}$$

where U is integrated over the unitary group $U(d)$ and where $f(H)$ is an arbitrary function. As a consequence, a typical GUE Hamiltonian is non-local (or $\mathcal{O}(n)$-local), so local operators are delocalized essentially immediately. Indeed, the Haar-invariance of the GUE

ensemble and non-locality of its Hamiltonians resulted in unusual behaviors of OTOCs whose decay time was shorter than that of 2-point correlation functions. It thus appears that *local* chaotic Hamiltonians and a typical Hamiltonian from a Haar-invariant ensemble behave in a dramatically different way.

However, previous studies on chaotic Hamiltonians suggest that at late times, Haar-invariant Hamiltonian ensembles, such as the GUE, GOE and GSE, capture behaviors of correlation functions remarkably well. This apparent tension between early time and late time behaviors may be resolved in the following manner. Initially, any ensemble of local Hamiltonians is not Haar-invariant because Hamiltonians are made of local terms. This can be clearly seen from the fact that the OTOC, $\langle A(0)B(t)A(0)B(t)\rangle$, behaves rather differently depending on the sizes of operators A, B. Yet, after the scrambling time when local operators become delocalized by Hamiltonian evolution, it becomes harder to tell whether the original operators $A(0), B(0)$ were local or not, and we expect that the unitary ensemble becomes 'approximately' Haar-invariant.

With this observation in mind, we are naturally led to consider a fine-grained characterization of Haar-invariance which we shall call *k-invariance*. Intuitively, k-invariance refers to an ensemble of unitary operators which appear to be Haar-invariant up to k-th moments. More precisely, let $\mathcal{E}$ be an ensemble of unitary operators. We define a Haar-invariant extension $\widetilde{\mathcal{E}}$ of this ensemble by:

$$\int_{U\in\widetilde{\mathcal{E}}} dU = \int_{W\in\mathrm{Haar}} dW \int_{U\in\mathcal{E}} d(WUW^\dagger)\,. \tag{2.132}$$

From the construction, we can easily see $W\widetilde{\mathcal{E}}W^\dagger = \widetilde{\mathcal{E}}$ for any unitary operator W, and so the *Haar'ed* ensemble is independent of any basis. Let us consider the k-fold twirl superoperator:

$$\Phi_{\mathcal{E}}^{(k)}(\cdot) = \int_{U\in\mathcal{E}} dU\, U^{\otimes k}(\cdot)U^{\dagger\otimes k}\,. \tag{2.133}$$

Then, $\mathcal{E}$ is said to be k-invariant if and only if

$$\Phi_{\mathcal{E}}^{(k)}(\cdot) = \Phi_{\widetilde{\mathcal{E}}}^{(k)}(\cdot)\,. \tag{2.134}$$

An ensemble of unitaries is Haar-invariant if and only if it is k-invariant for all $k \geq 1$. Similar definitions apply to Haar-invariance with respect to orthogonal and symplectic groups.

The utility of k-invariance can be seen from an explicit relation between correlation functions and spectral statistics. Recall that we have derived the following relation in the GUE by using the Haar-invariance of the GUE measure:

$$\langle A_1(0)B_1(t)\dots A_k(0)B_k(t)\rangle_{\rm GUE} \simeq \langle A_1B_1\dots A_kB_k\rangle \frac{\mathcal{R}_{2k}(t)}{d^{2k}}\,. \tag{2.135}$$

It is clear that the same derivation applies to any ensemble which is k-invariant. The implication is that, after the *k-invariance time*, the behavior of $2k$-point OTOCs can be completely determined by the spectral statistics alone. The physical significance of the k-invariance time is that it is the time scale when OTOCs behave in a similar way regardless of the locality or non-locality of the operators A_j, B_j (as well as their time-ordering). A similar conclusion holds for k-th frame potentials which can be written only in terms of spectral form factors for k-invariant ensembles. Thus, k-invariance and its associated time scale will be a useful notion to characterize the loss of locality from the perspective of $2k$-point OTOCs and the onset of random matrix behavior.

How can one verify that some ensemble $\mathcal{E}$ is k-invariant? One formal approach is to use frame potentials. Let us define the following operator

$$S = \int_{\mathcal{E}} dU U^{\otimes k} \otimes U^{\dagger\otimes k} - \int_{\widetilde{\mathcal{E}}} dU U^{\otimes k} \otimes U^{\dagger\otimes k}\,, \tag{2.136}$$

which corresponds to the difference between tensor expanders from $\mathcal{E}$ and its Haar-invariant

extension $\widetilde{\mathcal{E}}$. Then we have

$$\begin{aligned} 0 \leq \mathrm{tr}(S^\dagger S) = &\int_{U,V\in\mathcal{E}} dU dV |\mathrm{tr}(U^\dagger V)|^{2k} \\ &- \int_{U,V\in\mathcal{E}} dU dV \int_{W\in\mathrm{Haar}} dW \, |\mathrm{tr}(U^\dagger W V W^\dagger)|^{2k} \\ &- \int_{U,V\in\mathcal{E}} dU dV \int_{W\in\mathrm{Haar}} dW \, |\mathrm{tr}(W U^\dagger W^\dagger V)|^{2k} \\ &+ \int_{U,V\in\mathcal{E}} dU dV \int_{W,Y\in\mathrm{Haar}} dW dY \, |\mathrm{tr}(W U^\dagger W^\dagger Y V Y^\dagger)|^{2k} \\ = &\mathcal{F}_{\mathcal{E}}^{(k)} - \mathcal{F}_{\widetilde{\mathcal{E}}}^{(k)} , \end{aligned} \tag{2.137}$$

where $\mathcal{F}_{\mathcal{E}}^{(k)}$ is the k-th frame potential for an ensemble $\mathcal{E}$. Here we used the fact that the Haar unitary ensemble is left and right invariant. Therefore, we arrive at the following inequality

$$\mathcal{F}_{\mathcal{E}}^{(k)} \geq \mathcal{F}_{\widetilde{\mathcal{E}}}^{(k)} \tag{2.138}$$

with equality if and only if $\mathcal{E}$ being k-invariant. The difference $\mathcal{F}_{\mathcal{E}}^{(k)} - \mathcal{F}_{\widetilde{\mathcal{E}}}^{(k)}$ measures the 2-norm distance to being k-invariant.[22] The above derivation is a straightforward generalization of a method used in [57].

Haar-invariance in a spin system

Let us examine k-invariance for the random non-local (RNL) spin system discussed in Sec. 2.3.1 where we defined the Hamiltonian in Eq. (2.48) as the sum over all 2-body operators with random Gaussian couplings $J_{ij\alpha\beta}$. The time evolution of the first frame potential for this ensemble as well as its Haar-conjugated generalization are shown in Fig. 2.7 along side the difference $\mathcal{F}_{\mathcal{E}}^{(1)} - \mathcal{F}_{\widetilde{\mathcal{E}}}^{(1)}$, measuring the distance to 1-invariance. We only report numerics for a modest spin system of $n = 6$ spins. The difficulty of performing frame potential numerics is mentioned in App. 2.D.

We find that in this chaotic spin system, at early times we quickly deviate from 1-invariance, but after evolution by the system's chaotic dynamics, we observe an approach

[22] For a more rigorous analysis, the diamond distance should be considered. While the diamond norm is difficult to compute in general, there are some examples of ensembles of realistic Hamiltonians where the diamond norm can be analytically computed. We hope to address this in a future publication.

to approximate 1-invariance at late times. For this system, we see that the frame potential approaches, but does not equal, its Haar-invariant counterpart at later times. But we found numerically that increasing the number of sites makes this late time difference smaller. Thus we expect that at large N for chaotic systems, we reach k-invariance at late times.

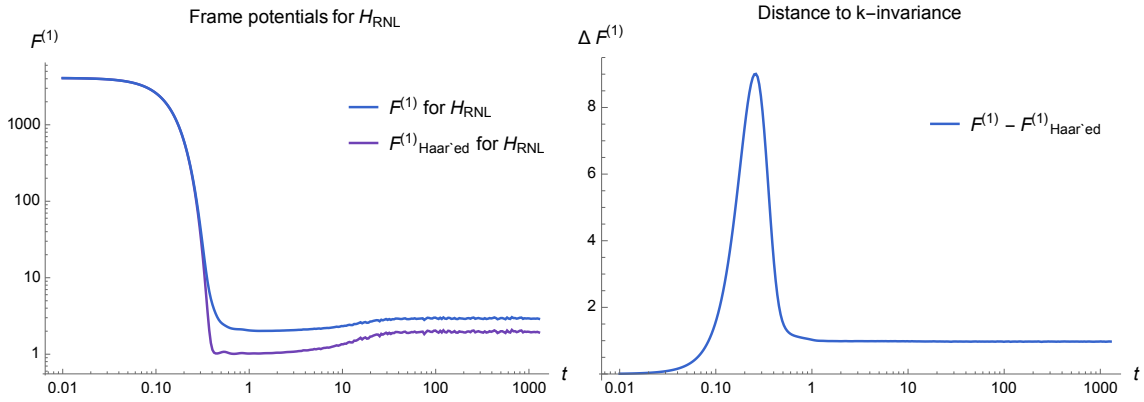


Figure 2.7: On the left we plot the first frame potential $\mathcal{F}^{(1)}_{\mathcal{E}_{\text{RNL}}}$ for H_{RNL} along side the first frame potential for its Haar-invariant extension $\mathcal{F}^{(1)}_{\widetilde{\mathcal{E}}_{\text{RNL}}}$, computed numerically using the 2-point form factor as in Eq. (2.86). On the right we plot the difference, measuring the 2-norm distance to 1-invariance and observe approximate 1-invariance at late times.

Comments on k-invariance

While frame potentials provide a quantitative way of judging if an ensemble $\mathcal{E}$ is k-invariant or not, it would be beneficial to relate it to some physical observables such as correlation functions. It is perhaps not a big surprise that k-invariance can be verified by $2k$-point OTOCs. The following statement holds:

$$\begin{aligned}&\langle A_1(0)B_1(t)\dots A_k(0)B_k(t)\rangle_{\mathcal{E}}\\ &\qquad = \langle \tilde{A}_1(0)\tilde{B}_1(t)\dots \tilde{A}_k(0)\tilde{B}_k(t)\rangle_{\mathcal{E}} \quad \forall \tilde{A}_j, \tilde{B}_j \quad \Longleftrightarrow \quad \mathcal{E} \text{ is } k\text{-invariant}\end{aligned} \tag{2.139}$$

where A_j, B_j are Pauli operators, and $\tilde{A}_j, \tilde{B}_j$ are some transformations from A_j, B_j such that

$$\tilde{A}_j = W A_j W^\dagger \qquad \tilde{B}_j = W B_j W^\dagger \tag{2.140}$$

where W is an arbitrary element of unitary $2k$-design. The proof is straightforward and thus is skipped.

Motivated by late-time random matrix universality of chaotic quantum systems, we have introduced a novel quantum information theoretic concept, k-invariance, as a possible way of bridging early-time and late-time physics. We would like to comment on a few caveats. First, consider an ensemble of unitary operators $\mathcal{E}$ generated by some Hamiltonians. Since $\mathcal{E}_{t=0} = \{\mathbb{I}\}$, the ensemble is Haar-invariant at time $t = 0$. Thus, an ensemble is initially k-invariant and is expected to immediately deviate at $t > 0$ and then eventually become approximately k-invariant. Therefore $\mathcal{F}_{\mathcal{E}}^{(k)} - \mathcal{F}_{\widetilde{\mathcal{E}}}^{(k)}$, which quantifies k-invariance, is not a monotonic quantity under time evolution. However, we expect that it is monotonically decreasing at late times. We observe these features in the non-local spin system described above. Depending on the symmetries of the system of interest, we would need to consider the Haar measure with respect to an appropriate Lie group $G \subset U(d)$.

Second, for realistic physical systems with local Hamiltonians, it is not likely that an ensemble $\mathcal{E}_t$ becomes k-invariant in an exact sense even at very late times. This can be seen from a recent work which shows that the late-time value of infinite temperature OTOCs $\langle A(0)B(t)A(0)B(t)\rangle$ of q-local Hamiltonians is $\mathcal{O}(1/N)$ if operators A, B are local and have overlaps with the Hamiltonian [69], based on an Eigenstate Thermalization Hypothesis (ETH) argument. A similar argument applies to late-time values of two-point correlators. On the other hand, the Haar average of OTOCs is $\mathcal{O}(1/d^2)$ (or $\mathcal{O}(1/d)$ for an average of absolute values). Thus, OTOCs for local operators and OTOCs for non-local operators may have significantly different late-time values. However, it should be noted that a prediction from the AdS/CFT seems to suggest that correlation functions may become exponentially small $e^{-\mathcal{O}(S)}$ even if A, B are local operators. This may suggest a subtle but important distinction between ordinary strongly interacting systems and gravitational systems which leads to a far-reaching question concerning the universality of gravity and the universality of random matrix theory, seen from the lens of k-invariance.

Let us conclude the section with a brief remark on the Eigenstate Thermalization Hypothesis (ETH). The notion of k-invariance may be viewed as a dynamical analog of Berry's conjecture about random eigenvectors, which was the motivation behind ETH [70, 71, 21].

A basic assumption of ETH is that matrix elements of a local operator O, with respect to energy eigenstates, look "random" inside some sufficiently small energy window ΔE. A system achieving k-invariance roughly tells us that energy eigenstates may be treated as random vectors after sufficiently long times for studying dynamics via OTOCs.[23] Given the prevalence of eigenstate thermalization in strongly correlated many-body systems,[24] a precise relation between k-invariance, ETH and OTOCs would provide clarity on defining what it means for a quantum system to be chaotic.

2.7 Discussion

Random matrix theory provides a powerful paradigm for studying late-time chaos. We have leveraged the technology of random matrix theory and Haar-invariance to study correlation functions like OTOCs which diagnose early-time chaos, and frame potentials which diagnose randomness and complexity. The salient feature of the GUE which gave us computational traction is its Haar-invariance, namely that the ensemble looks the same in any basis. As a result, the dynamics induced by GUE Hamiltonians is non-local ($\mathcal{O}(N)$-local) with respect to any tensor factor decomposition of the Hilbert space, and so the dynamics immediately delocalizes quantum information and other more subtle forms of correlations. Accordingly, the GUE captures features of the long-time physics of a local system that has been delocalized.

In a chaotic quantum system described by a local Hamiltonian, there are two temporal regimes of interest: times before the system scrambles and thus has mostly local correlations, and times after the system scrambles when correlations have effectively delocalized. We suggested that the transition between these two regimes may be due to the onset of approximate Haar-invariance, and we defined k-invariance as a precise characterization. A careful understanding of Haar-invariance for ensembles of local quantum systems could yield precise insights into the apparent breakdown of locality, and tell us in what time regimes we can use Haar-invariance to calculate late-time physics (i.e., correlation functions, frame potentials, complexity, etc.) A concrete way of studying delocalization of operators and

[23] The related notion of quantum ergodicity and randomness of eigenstates was recently discussed in [72].

[24] See [21] and references therein. Interestingly, evidence for ETH has also been discussed recently both in the SYK model [73] as well as in its free fermion counterpart [74].

the emergence of k-invariance would be to compare connected pieces of OTOCs with local and non-local operators and observe their eventual convergence. Of particular interest is to find the 2-invariance time when all the 4-point OTOCs, regardless of sizes of operators, start to behave in a similar manner. This time scale must be at least the scrambling time since OTOCs with local operators start to decay only around the scrambling time while OTOCs with non-local operators decay immediately. Relatedly, a recent work [75] strives to understand the onset of random matrix behavior at early times.

In this chapter, we computed correlation functions averaged over an ensemble of Hamiltonians. Chaotic systems described by disordered ensembles tend to have small variance in their correlators, and their averaged correlation functions are close to those computed for a simple instance of the ensemble. Even in regimes where replica symmetries are broken, performing time bin averaging reproduces the averaged behaviors very well. We find in App. 2.D.3 that the time bin-averaged frame potential in the large d limit for two samples agrees with averaging over the whole ensemble.

We conclude by mentioning a far reaching goal, but one that provides the conceptual pillars for these ideas, namely understanding black holes as quantum systems. While black holes are thermodynamic systems whose microscopic details remain elusive, questions about information loss can be precisely framed by late-time values of correlation functions within AdS/CFT [36], where unitary evolution can be discussed in terms of the boundary CFT. Ultimately, we would like to use random matrix theory to characterize chaos and complexity in local quantum systems and identify late-time behaviors which are universal for gravitational systems. An interesting future question is to see if gravitational systems are described by random matrices in the sense of k-invariance and pinpoint some late-time behavior which results from gravitational universality.

2.A Scrambling and 2-designs

Recently there has been growing interest in scrambling and unitary designs from the high energy and quantum information communities. Here we provide a short summary of different ways of quantifying them for infinite temperature cases.

2.A.1 Scrambling

We begin with scrambling. Consider a system of qubits and non-overlapping local ($\mathcal{O}(1)$-body) Pauli operators V, W and compute OTOC $= \langle VW(t)VW(t)\rangle$ where $W(t) = UWU^\dagger$. The initial value of OTOC at $t = 0$ is 1. Scrambling is a phenomenon where the OTOC becomes $\mathcal{O}(\epsilon)$ with $\epsilon \ll 1$ being a small but finite constant:

$$\langle VW(t)VW(t)\rangle = \mathcal{O}(\epsilon) \qquad \text{for all pairs of local operators } V, W \tag{2.141}$$

It is often the case that OTOCs with local operators are the slowest to decay. This can be seen from our analysis on 4-point spectral form factors. So, by the scrambling time, OTOCs with non-local operators are already $\mathcal{O}(\epsilon)$ or smaller. The scrambling time is lower bounded by $\mathcal{O}(\log(n))$ in the case of 0-dimensional $\mathcal{O}(1)$-local systems due to a Lieb-Robinson–like argument [3].

Scrambling has caught significant attention from the quantum gravity community since it is closely related to the Hayden-Preskill thought experiment on black hole information problems [1]. Assume that V, W act on qubits on some local regions A, D respectively, and define their complements by $B = A^c, C = D^c$. Imagine that A is an unknown quantum state $|\psi\rangle$ thrown into a "black hole" B, and the whole system evolves by some time-evolution operator $U = e^{-iHt}$. At time t, we collect the "Hawking radiation" D and attempt to reconstruct (an unknown) $|\psi\rangle$ from measurement on D. Such a thought experiment was considered by Page who argued that, if a black hole's dynamics U is approximated by a random unitary operator, then reconstructing $|\psi\rangle$ is not possible unless we collect more than $n/2$ qubits of the Hawking radiation [76]. As we shall show in Appendix 2.B, the impossibility of reconstruction of A from D is reflected in the smallness of the 2-point correlation functions:

$$|\langle VW(t)\rangle| = \mathcal{O}(\epsilon) \qquad \text{for local } V, W \quad \longrightarrow \quad \text{no reconstruction of } A \text{ from } D. \tag{2.142}$$

The famous calculations by Hawking and Unruh imply that these two-point correlators are thermal, and quickly become small.

Hayden and Preskill considered a situation where a black hole B has already emitted

half of its contents, and we have collected its early radiation and stored it in some secure quantum memory M. The quantum memory M is maximally entangled with B, and the question is whether we can reconstruct $|\psi\rangle$ by having access to M. It has been shown that scrambling, as defined above, implies that we can reconstruct $|\psi\rangle$ with some good average fidelity by collecting the Hawking radiation on D at time t:

$$\langle VW(t)VW(t)\rangle = \mathcal{O}(\epsilon) \quad \longrightarrow \quad \text{reconstruction of } A \text{ from } D \text{ and } M. \tag{2.143}$$

Therefore, scrambling implies the possibility of recovering local quantum information via local measurements on the Hawking radiation. A random unitary operator U typically gives very small OTOCs which enables reconstruction of A in the Hayden-Preskill thought experiment.

Reconstruction problems in the Hayden-Preskill setting are closely related to the problem of decoupling. A crucial difference between scrambling and decoupling is that decoupling typically considers A, D to be some finite fraction of the whole system and concerns the reconstruction of unknown many-body quantum states supported on a big region A. Since we quantify the reconstruction via fidelity for many-body quantum states, the requirement tends to be more stringent. The relation between scrambling and decoupling is discussed in [77] in the context of local random circuits.

2.A.2 Unitary designs

Next let us discuss unitary 2-designs. Consider an ensemble of time evolution operators U_j with probability distributions p_j; $\mathcal{E} = \{U_j, p_j\}$ with $\sum_j p_j = 1$. The 2-fold channels of $\mathcal{E}$ and the Haar ensemble are

$$\Phi_{\mathcal{E}}(\rho) = \sum_j p_j U_j \otimes U_j(\rho) U_j^\dagger \otimes U_j^\dagger \qquad \Phi_{\text{Haar}}(\rho) = \int_{\text{Haar}} dU \; U \otimes U(\rho) U^\dagger \otimes U^\dagger \,. \tag{2.144}$$

If $\Phi_{\mathcal{E}}(\rho) = \Phi_{\text{Haar}}(\rho)$ for all ρ, then we say $\mathcal{E}$ is 2-design. One can check if $\mathcal{E}$ is 2-design or not by looking at OTOCs. Consider the OTOC $\langle VW(t)VW(t)\rangle$ for arbitrary Pauli operators V, W which are not necessarily local operators. We will be interested in the ensemble averages

of OTOCs:

$$\langle VW(t)VW(t)\rangle_{\mathcal{E}} \equiv \sum_j p_j \langle VU_jWU_j^\dagger VU_jWU_j^\dagger\rangle\,. \tag{2.145}$$

If $\langle VW(t)VW(t)\rangle_{\mathcal{E}} = \langle VW(t)VW(t)\rangle_{\text{Haar}}$ for all pairs of Pauli operators V, W, then the ensemble forms a unitary 2-design [11].

A typical unitary operator from a 2-design achieves scrambling because

$$|\langle VW(t)VW(t)\rangle|_{\text{Haar}} \simeq \frac{1}{d} \qquad \langle VW(t)VW(t)\rangle_{\text{Haar}} \simeq \frac{1}{d^2} \tag{2.146}$$

for any (possibly non-local) Pauli operators V, W. The first equation implies that the OTOC value for a single instance from the ensemble is typically $1/d$ in absolute value while the second equation implies that the OTOC, after ensemble averaging, is $1/d^2$. Since OTOCs are small, a typical 2-design unitary operator U implies scrambling, but the converse is not always true. Recall that scrambling only requires OTOC $= \mathcal{O}(\epsilon)$. There is thus a big separation in the smallness of the OTOC, and the scrambling time may be much shorter than the 2-design time. Also, scrambling requires OTOC $= \mathcal{O}(\epsilon)$ only for local operators while a 2-design unitary makes the OTOC small for all pairs of Pauli operators. The lower bound for the exact 2-design time is $\mathcal{O}(\log(n))$, but no known protocol achieves this time scale.

One important distinction between scrambling and the 2-design time is how small the OTOCs becomes. The phenomena of scrambling concerns the deviation of OTOC values from the maximal value 1. The concept of a 2-design concerns the deviation of OTOC values from the minimal value $\mathcal{O}(1/d)$. The former is related to early-time chaos and the latter is related to late-time chaos.

2.A.3 Approximate 2-designs

Finally, let us briefly discuss the notion of approximate 2-design. When two quantum operations $\Phi_{\mathcal{E}}$ and Φ_{Haar} are close to each other, we say that $\mathcal{E}$ is an approximate 2-design. In order to be quantitative, however, we need to pick appropriate norms with which two

quantum operations can be compared. The 2-norm distance can be defined in a simple way via

$$\begin{aligned} &\text{2-norm} = \sqrt{\mathrm{Tr}(SS^\dagger)} \\ &S = \int \sum_j p_j U_j \otimes U_j \otimes U_j^\dagger \otimes U_j^\dagger - \int_{\text{Haar}} dU\, U \otimes U \otimes U^\dagger \otimes U^\dagger \,. \end{aligned} \tag{2.147}$$

If $S = 0$, then $\Phi_{\mathcal{E}}$ and Φ_{Haar} would be the same. We say that $\mathcal{E}$ is a δ-approximate 2-design in the 2-norm if $\sqrt{\mathrm{Tr}(SS^\dagger)} \leq \delta$.

Frame potentials are closely related to the 2-norm distance because $\mathrm{Tr}(SS^\dagger) = \mathcal{F}_{\mathcal{E}} - \mathcal{F}_{\text{Haar}} \geq 0$. In [11], a relation between the frame potential and OTOCs has been derived

$$\int dA dB dC dD |\langle AB(t)CD(t)\rangle_{\mathcal{E}}|^2 = \frac{\mathcal{F}_{\mathcal{E}}^{(2)}}{d^6} \,. \tag{2.148}$$

In practice, the main contribution to the left-hand side comes from OTOCs of the form $\langle AB(t)AB(t)\rangle_{\mathcal{E}}$. For simplicity of discussion, let us assume that $\langle AB(t)CD(t)\rangle_{\mathcal{E}} = 0$ when $C \neq A$ or $D \neq B$ (where A, B, C, D are non-identity Pauli operators). Then, a simple analysis leads to

$$|\langle AB(t)AB(t)\rangle_{\mathcal{E}}|^2 \simeq \delta^2 \tag{2.149}$$

for typical non-identity Pauli operators A, B. Thus, being a δ-approximate 2-design in the 2-norm implies that OTOCs are typically small. However, this does not necessarily imply scrambling because OTOCs with local operators are often the slowest to decay. In order to guarantee scrambling, we would need a $\frac{\delta}{d}$-approximate design in the 2-norm (under an assumption on $\langle AB(t)CD(t)\rangle_{\mathcal{E}} = 0$ for $C \neq A$ or $D \neq B$). For this reason, an alternative distance measure called the diamond norm is often used in quantum information literature. See [78] for relations between different norms.

2.B Information scrambling in black holes

In this Appendix, we discuss behaviors of 2-point correlators and 4-point OTOCs from the viewpoint of information scrambling in black holes. We begin by deriving a formula which relates two-point autocorrelation functions and mutual information. We will be interested in the following quantity

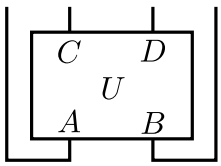
$$\left|\langle O_A O_D(t)\rangle_{\text{avg}}\right|^2 \equiv \frac{1}{d_A^2 d_D^2} \sum_{O_A\in\mathcal{P}_A} \sum_{O_D\in\mathcal{P}_D} |\langle O_A O_D(t)\rangle|^2 \,, \tag{2.150}$$

where $\langle O_A O_D(t)\rangle = \frac{1}{d}\text{Tr}(O_A U O_D U^\dagger)$ and U is the time-evolution operator of the system, and $\mathcal{P}_A$ and $\mathcal{P}_D$ are sets of Pauli operators on A and D. There are d_A^2 and d_D^2 Pauli operators.

The relation between apparent information loss and two-point correlators can be understood by using the state representation $|U\rangle$ of a unitary operator U. Given a unitary operator U acting on an n-qubit Hilbert space $\mathcal{H}$, one can view U as a pure quantum state $|U\rangle$ defined on a $2n$-qubit Hilbert space $\mathcal{H}\otimes\mathcal{H}$:

$$|U\rangle \equiv U\otimes\mathbb{I}|\text{EPR}\rangle \,, \qquad |\text{EPR}\rangle = \frac{1}{\sqrt{2^n}}\sum_{j=1}^{2^n} |j\rangle\otimes|j\rangle \,. \tag{2.151}$$

Or equivalently, $|U\rangle \equiv \frac{1}{\sqrt{2^n}}\sum_{i,j} U_{i,j}|i\rangle\otimes|j\rangle$ where $U = \sum_{i,j} U_{i,j}|i\rangle\langle j|$. One easily sees that the quantum state $|U\rangle$ is uniquely determined by a unitary operator U. The state representation allows us to view $|U\rangle_{ABCD}$ as a four-partite quantum state:

$|U\rangle = \frac{1}{\sqrt{2^n}}$ 

, (2.152)

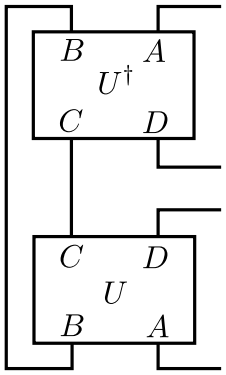
where $B = A^c$ and $D = C^c$ in the original system of qubits. Given the state representation $|U\rangle$ of a unitary operator, we can derive the following formula

$$\left|\langle O_A O_D(t)\rangle_{\text{avg}}\right|^2 = \frac{1}{d_A^2 d_D^2} 2^{I^{(2)}(A,D)} \,, \tag{2.153}$$

where $I^{(2)}(A,D)$ is the Rényi-2 mutual information between A and D for $|\Psi\rangle$, defined by $I^{(2)}(A,D) \equiv S_A^{(2)} + S_D^{(2)} - S_{AD}^{(2)}$.

To derive the formula, let ρ_{AD} be the reduced density matrix of $|U\rangle$ on AD. Its graphical representation is

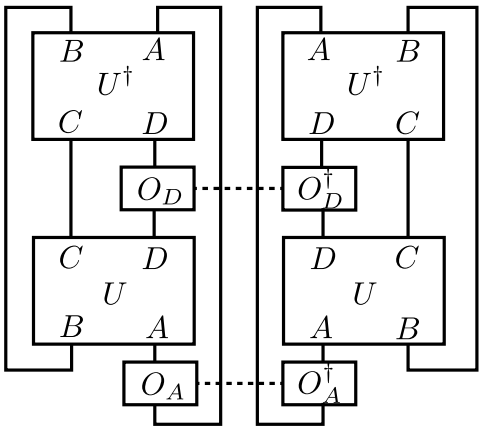
$$\rho_{AD} = \frac{1}{d}$$



(2.154)

The averaged 2-point correlator is given by

$$\left|\langle O_A O_D(t)\rangle_{\text{avg}}\right|^2 = \frac{1}{d^2}$$



(2.155)

where dotted lines represent averaging over Pauli operators. By using $\frac{1}{d}\sum_{O\in\mathcal{P}} O \otimes O^\dagger = \text{SWAP}$, we obtain

$$|\langle O_A O_D(t)\rangle_{\text{avg}}|^2 = \frac{\text{Tr}(\rho_{AD}^2)}{d_A d_D} = \frac{1}{d_A^2 d_D^2} 2^{I^{(2)}(A,D)}\,. \tag{2.156}$$

Let us think a little more about the formula we have derived. For strongly interacting

systems, it is typically the case that

$$\langle O_A O_D(t)\rangle \simeq 0 \qquad \text{if} \qquad \mathrm{Tr}(O_A O_D) = 0\,. \tag{2.157}$$

So, the following relation for the autocorrelation functions holds approximately:

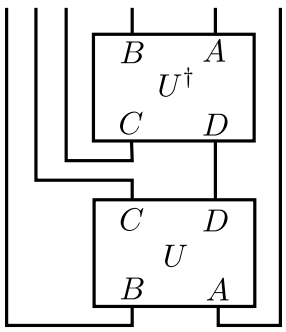
$$\sum_{O_A \in \mathcal{P}_A} |\langle O_A O_A(t)\rangle|^2 \simeq 2^{I^{(2)}(A,D)}\,, \tag{2.158}$$

where we took A and D to be the same subset of qubits.

The above formula has an interpretation as information retrieval from the early Hawking radiation. Consider scenarios where Alice throws a quantum state $|\psi\rangle$ into a black hole and Bob attempts to reconstruct it from the Hawking radiation. In accordance with such thought experiments, let A be qubits for Alice's quantum state, B be the black hole, C be the remaining black hole and D be the Hawking radiation. Then, the averaged 2-point correlation functions have an operational interpretation as Bob's strategy to retrieve Alice's quantum state. Let us assume that the initial state of the black hole is unknown to Bob and model it by a maximally mixed state $\rho_B = \frac{\mathbb{I}_B}{d_B}$. Alice prepares an EPR pair $|\mathrm{EPR}\rangle_{AR}$ on her qubits and her register qubits. Notice the difference from the Hayden-Preskill setup where Bob had access to some reference system B' which is maximally entangled with the black hole B. In this decoding problem, we do not grant such access to Bob. He just collects the Hawking radiation D and tries to reconstruct Alice's quantum state.

The most obvious strategy is to apply the inverse $U^\dagger$. However, Bob does not have access to qubits on C. So, he applies $U_{CD}^\dagger \otimes \mathbb{I}_R$ to $\rho_C \otimes \rho_{DR}$ where $\rho_C = \frac{\mathbb{I}_C}{d_C}$. Graphically, this corresponds to

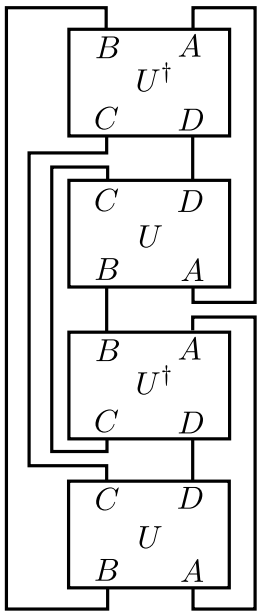
$$|\Psi\rangle = \frac{d}{\sqrt{d_A d_B d_C}}$$



(2.159)

The success of decoding is equivalent to distillation of an EPR pair between A and R. So, we compute the EPR fidelity. Namely, letting Π be a projector onto an EPR pair between A and R, we have

$$F = \langle \Psi | \Pi | \Psi \rangle = \frac{1}{d^2}$$



(2.160)

which leads to

$$F = \mathrm{Tr}(\rho_{BC}^2) = \mathrm{Tr}(\rho_{AD}^2) = d_A d_D |\langle O_A O_D(t) \rangle_{\mathrm{avg}}|^2 \,. \tag{2.161}$$

Therefore, the decay of 2-point correlation functions indeed implies that Bob cannot reconstruct Alice's quantum state.

Finally, let us summarize the known relations between correlation functions and mutual information:

$$\langle O_A O_D(t) O_A O_D(t) \rangle_{\mathrm{avg}} = 2^{-I^{(2)}(A,BD)} \tag{2.162}$$

$$d_A^2 d_D^2 |\langle O_A O_D(t) \rangle_{\mathrm{avg}}|^2 = 2^{I^{(2)}(A,D)} \,. \tag{2.163}$$

Note that the first formula proves that the decay of OTOCs leads to large $I^{(2)}(A, BD)$ which implies the possibility of Bob decoding Alice's quantum state by accessing both the early radiation B and the new Hawking radiation D. These two formulae allow us to formally

show that a black hole can be viewed as a quantum error-correcting code. Let A, D be degrees of freedom corresponding to incoming and outgoing Hawking radiation, and B, C be degrees of freedom corresponding to other exotic high energy modes at the stretched horizon. Since a black hole is thermal, we know that $|\langle O_A O_D(t)\rangle_{\text{avg}}|$ decays at $t \sim \mathcal{O}(\beta)$. Also, due to the shockwave calculation by Shenker and Stanford [4], we know that $\langle O_A O_D(t) O_A O_D(t)\rangle_{\text{avg}}$ decays at $t \sim \mathcal{O}(\beta \log N)$. These results imply that after the scrambling time:

$$I^{(2)}(A, D) \simeq 0 \qquad I^{(2)}(A, C) \simeq 0\,. \tag{2.164}$$

The implication is that quantum information injected from A gets delocalized and non-locally is hidden between C and D. The error-correction property can be seen by

$$I^{(2)}(A, BD) \simeq 2a \qquad I^{(2)}(A, BC) \simeq 2a \qquad I^{(2)}(A, CD) \simeq 2a\,, \tag{2.165}$$

where a is the number of qubits on A. Namely, if we see the black hole as a quantum code which encodes A into BCD, then the code can tolerate erasure of any single region B, C, D. In other words, accessing any two of B, C, D is enough to reconstruct Alice's quantum state. Thus, black hole dynamics, represented as a four-partite state $|U\rangle_{ABCD}$, can be interpreted as a three-party secret sharing quantum code.

2.C Spectral correlators and higher frame potentials

In this Appendix we will present formulas for form factors from random matrix theory. For GUE$(d, 0, 1/\sqrt{d})$, $d \times d$ matrices with off-diagonal complex entries and real diagonal entries chosen with variance $\sigma^2 = 1/d$, the joint probability of eigenvalues for GUE, with normalizing factors, is

$$P(\lambda_1, \ldots, \lambda_d) = \frac{d^{d^2/2}}{(2\pi)^{d/2} \prod_{p=1}^{d} p!}\, e^{-\frac{d}{2}\sum_i \lambda_i^2} \prod_{i<j} (\lambda_i - \lambda_j)^2 \tag{2.166}$$

and the joint probability distribution of n eigenvalues (i.e., the n-point spectral correlation function), defined as

$$\rho^{(n)}(\lambda_1, \ldots, \lambda_n) = \int d\lambda_{n+1} \ldots d\lambda_d P(\lambda_1, \ldots, \lambda_d) \,. \tag{2.167}$$

We can compactly express $\rho^{(n)}(\lambda_1, \ldots, \lambda_n)$ in terms of a kernel K [43, 23] as

$$\rho^{(n)}(\lambda_1, \ldots, \lambda_n) = \frac{(d-n)!}{d!} \det\left(K(\lambda_i, \lambda_j)\right)_{i,j=1}^{n} \,. \tag{2.168}$$

In the large d limit, the kernel K is approximately

$$K(\lambda_i, \lambda_j) \equiv \begin{cases} \dfrac{d}{\pi} \dfrac{\sin(d(\lambda_i - \lambda_j))}{d(\lambda_i - \lambda_j)} & \text{for} \quad i \neq j \\ \\ \dfrac{d}{2\pi} \sqrt{4 - \lambda_i^2} & \text{for} \quad i = j \,, \end{cases} \tag{2.169}$$

where the $i \neq j$ case is called the sine kernel, and the $i = j$ case is simply the Wigner semicircle. In the large d limit, the basic approach for computing spectral form factors will be expanding the determinant in Eq. (2.168) using the kernel in Eq. (2.169), and computing the Fourier transform of the resulting sums of product of kernels. Thus we will have sums of integrals of the form [23]

$$\begin{aligned} &\int \prod_{i=1}^{m} d\lambda_i \, K(\lambda_1, \lambda_2) K(\lambda_2, \lambda_3) \ldots K(\lambda_{m-1}, \lambda_m) K(\lambda_m, \lambda_1) \, e^{i \sum_{i=1}^{m} k_i \lambda_i} \\ &\quad = \frac{d}{\pi} \int d\lambda \, e^{i \sum_{i=1}^{m} k_i \lambda_i} \int dk \, g(k) g\Big(k + \frac{k_1}{2d}\Big) g\Big(k + \frac{k_2}{2d}\Big) \ldots g\Big(k + \frac{k_{m-1}}{2d}\Big) \,, \end{aligned} \tag{2.170}$$

where we define the Fourier transform of the sine kernel

$$g(k) \equiv \int dr \, e^{2\pi i k r} \, \frac{\sin(\pi r)}{\pi r} = \begin{cases} 1 & \text{for } |k| < \frac{1}{2} \\ 0 & \text{for } |k| > \frac{1}{2} \end{cases} \,. \tag{2.171}$$

The delta function singularity from the $\int d\lambda \, e^{i \sum_{i=1}^{m} k_i \lambda}$ integral in Eq. (2.170) is an artifact of our expansion around infinite d, namely that $\frac{d}{\pi} \frac{\sin(d(\lambda_i - \lambda_j))}{d(\lambda_i - \lambda_j)}$ is not regulated in the $(\lambda_i + \lambda_j)$

direction. The most direct method to soften this divergence is to impose a cutoff

$$\frac{d}{\pi}\int d\lambda\, e^{i\sum_{i=1}^{m} k_i\lambda_i} \longrightarrow \frac{d}{\pi}\int_{-\pi/2}^{\pi/2} d\lambda\, e^{i\sum_{i=1}^{m} k_i\lambda_i}\,, \tag{2.172}$$

which is fixed by the normalization condition

$$\frac{d}{\pi}\int_{-\pi/2}^{\pi/2} d\lambda\, e^{i\sum_{i=1}^{m} k_i\lambda_i}\int dk\, g(k)g\Big(k+\frac{k_1}{2d}\Big)g\Big(k+\frac{k_2}{2d}\Big)\ldots g\Big(k+\frac{k_{m-1}}{2d}\Big)\Big|_{k_1,\ldots,k_m=0} = d\,. \tag{2.173}$$

While the 'box approximation' of applying the cutoff allows us to compute higher-point spectral correlators in the large d limit, it does lead to errors relative to an exact answer whose closed form is not tractable.[25] Thus we must be careful to keep track of these errors and compare with numerics. However, we find that at infinite temperature, the box approximation of the spectral form factors is analytically controlled at early times like $\mathcal{O}(1)$ and late times greater than $\mathcal{O}(\sqrt{d})$.

To understand the errors of the box approximation, we first consider various cases heuristically: when we have $\sum_i k_i = 0$, the λ integral in Eq. (2.170) is directly fixed by normalization. When $\sum_i k_i \neq 0$, the λ integral in Eq. (2.170) dephases and so decays when $|\sum_i k_i|$ is large, and thus the induced error is unimportant at long times. At small, $\mathcal{O}(1)$ values of the $|k_i|$'s (assuming that m is $\mathcal{O}(1)$), the error induced by the box approximation is also small and the value is still close to the $\sum_i k_i = 0$ value.

For instance, carefully keeping track of factors of d tells us that in $\mathcal{R}_4$, for early times like $\mathcal{O}(1)$ the error is suppressed by $\mathcal{O}(1/d)$ relative to largest order terms, while for late times after $\mathcal{O}(\sqrt{d})$ the error is suppressed by $\mathcal{O}(1/\sqrt{d})$ relative to the largest order terms.

In this discussion, particularly for $\sum_i k_i = 0$, we assumed simple sine kernel correlations and found r_2 to be a pure linear function. However, a more delicate treatment shows some other transition time scale at early times, which likely complicates the functional form of r_2 and gives a different slope for the ramp. We briefly address this issue for our numerics in App. 2.D.

Since the dephasing of the λ integral at large $|\sum_i k_i|$ is suppressed at finite temperature,

[25] For instance, the Fourier transform of the semicircle distribution decays as $t^{-3/2}$, whereas the Fourier transform of a box decays as t^{-1}.

to better capture long-time finite temperature eigenvalue correlations we use a modified kernel $\widetilde{K}$ which is valid in the short distance limit $|\lambda_a - \lambda_b| \sim \mathcal{O}(1/d)$ [25, 68],

$$\widetilde{K}(\lambda_i, \lambda_j) = \frac{\sin\left(d\pi(\lambda_i - \lambda_j)\rho^{(1)}((\lambda_i + \lambda_j)/2)\right)}{\pi(\lambda_i - \lambda_j)}\,, \tag{2.174}$$

which naturally provides a cutoff in the $(\lambda_i + \lambda_j)$ direction. However, this approximation assumes the continued domination of the regulated integral in the short distance limit, which may not be true for large β. However, for small β the modified kernel is reliable. In the generic case, one should consider the full expression of Hermite polynomials as the sine kernel, and correctly take the limit. A complicated formula has been derived in [25, 68] from a saddle point approximation.

2.C.1 Expressions for spectral correlators

Using the analysis above, it is straightforward to compute form spectral correlation functions for the GUE. It is convenient to define

$$r_1(t) \equiv \frac{J_1(2t)}{t}\,, \quad r_2(t) \equiv \begin{cases} 1 - \frac{t}{2d} & \text{for } t < 2d \\ 0 & \text{for } t > 2d \end{cases}\,, \quad r_3(t) \equiv \frac{\sin(\pi t/2)}{\pi t/2}\,. \tag{2.175}$$

as mentioned earlier. The infinite temperature form factors which appear in the calculation of the first and second frame potentials are

$$\begin{aligned} \mathcal{R}_2(t) &= \int D\lambda \sum_{i,j=1}^{d} e^{i(\lambda_i - \lambda_j)t}\,, & \mathcal{R}_{4,1}(t) &= \int D\lambda \sum_{i,j,k=1}^{d} e^{i(\lambda_i + \lambda_j - 2\lambda_k)t}\,, \\ \mathcal{R}_4(t) &= \int D\lambda \sum_{i,j,k,\ell=1}^{d} e^{i(\lambda_i + \lambda_j - \lambda_k - \lambda_\ell)t}\,, & \mathcal{R}_{4,2}(t) &= \int D\lambda \sum_{i,j=1}^{d} e^{2i(\lambda_i - \lambda_j)t}\,. \end{aligned} \tag{2.176}$$

As $\mathcal{R}_{4,2}$ is simply $\mathcal{R}_2(2t)$, we only need to compute the first three spectral correlation functions. We will also investigate the finite temperature version of $\mathcal{R}_2$, which we defined as

$$\mathcal{R}_2(t, \beta) \equiv \int D\lambda \sum_{i,j=1}^{d} e^{i(\lambda_i - \lambda_j)t} e^{-\beta(\lambda_i + \lambda_j)}\,. \tag{2.177}$$

$\mathcal{R}_2$ at infinite temperature

We start by computing $\mathcal{R}_2$ at infinite temperature:

$$\mathcal{R}_2(t) = d + \int d\lambda_1\, d\lambda_2 \Big(K(\lambda_1,\lambda_1)K(\lambda_2,\lambda_2) - K^2(\lambda_1,\lambda_2)\Big) e^{i(\lambda_1-\lambda_2)t}\,. \tag{2.178}$$

Evaluating the first term in the integral, we find

$$\int d\lambda_1 K(\lambda_1,\lambda_1) e^{i\lambda_1 t} \int d\lambda_2 K(\lambda_2,\lambda_2) e^{-i\lambda_2 t} = d^2 r_1^2(t)\,. \tag{2.179}$$

The second term can be evaluated using Eq. (2.170), and we find

$$\int d\lambda_1 d\lambda_2 K^2(\lambda_1,\lambda_2) e^{i(\lambda_1-\lambda_2)t} = d r_2(t)\,. \tag{2.180}$$

The final result is

$$\mathcal{R}_2(t) = d + d^2 r_1^2(t) - d r_2(t)\,. \tag{2.181}$$

$\mathcal{R}_2$ at finite temperature

As explained above, to better capture long-time correlations at finite temperature we will use the short-distance-limit kernel $\widetilde{K}$. Firstly, for $i = j$, we have

$$d \int D\lambda\, e^{-2\beta\lambda_1} = d r_1(2i\beta)\,. \tag{2.182}$$

For $i \neq j$ we have

$$\begin{aligned} & d(d-1) \int D\lambda\, e^{i(\lambda_1-\lambda_2)t - \beta(\lambda_1+\lambda_2)} \\ &\quad = \int d\lambda_1 d\lambda_2 \Big(\widetilde{K}(\lambda_1,\lambda_1)\widetilde{K}(\lambda_2,\lambda_2) - \widetilde{K}^2(\lambda_1,\lambda_2)\Big) e^{i(\lambda_1-\lambda_2)t-\beta(\lambda_1+\lambda_2)} \\ &\quad = d^2 r_1(t+i\beta) r_1(-t+i\beta) - d r_1(2i\beta) r_2(t)\,. \end{aligned} \tag{2.183}$$

Putting everything together, we obtain

$$\mathcal{R}_2 = d r_1(2i\beta) + d^2 r_1(t+i\beta) r_1(-t+i\beta) - d r_1(2i\beta) r_2(t)\,. \tag{2.184}$$

$\mathcal{R}_4$ at infinite temperature

We now compute $\mathcal{R}_4(t)$, again by separately considering coincident eigenvalues, using the determinant of kernels, and Fourier transforming to find

$$\begin{aligned}\mathcal{R}_4(t) = {} & d^4 r_1^4(t) - 2d^3 r_1^2(t) r_2(t) r_3(2t) - 4d^3 r_1^2(t) r_2(t) + 2d^3 r_1(2t) r_1^2(t) + 4d^3 r_1^2(t) \\ & + 2d^2 r_2^2(t) + d^2 r_2^2(t) r_3^2(2t) + 8d^2 r_1(t) r_2(t) r_3(t) - 2d^2 r_1(2t) r_2(t) r_3(2t) \\ & - 4d^2 r_1(t) r_2(2t) r_3(t) + d^2 r_1^2(2t) - 4d^2 r_1^2(t) - 4d^2 r_2(t) + 2d^2 \\ & - 7d r_2(2t) + 4d r_2(3t) + 4d r_2(t) - d\,. \end{aligned} \tag{2.185}$$

We can simplify this formula at early times of $\mathcal{O}(1)$ and late times greater than $\mathcal{O}(\sqrt{d})$ by dropping subdominant terms and find

$$\mathcal{R}_4 \approx d^4 r_1^4(t) + 2d^2 r_2^2(t) - 4d^2 r_2(t) + 2d^2 - 7d r_2(2t) + 4d r_2(3t) + 4d r_2(t) - d\,, \tag{2.186}$$

where the $2d^2 r_2^2(t)$ term gives a quadratic rise at late times, akin to the ramp in $\mathcal{R}_2$.

$\mathcal{R}_{4,1}$ at infinite temperature

We find that

$$\begin{aligned}\mathcal{R}_{4,1}(t) = {} & d^3 r_1(2t) r_1^2(t) - d^2 r_1(2t) r_2(t) r_3(2t) - 2d^2 r_1(t) r_2(2t) r_3(t) \\ & + d^2 r_1^2(2t) + 2d^2 r_1^2(t) + 2d r_2(3t) - d r_2(2t) - 2d r_2(t) + d\,. \end{aligned} \tag{2.187}$$

Just as above, we can approximate $\mathcal{R}_{4,1}$ at early and late times by

$$\mathcal{R}_{4,1} \approx d^3 r_1(2t) r_1^2(t) + 2d r_2(3t) - d r_2(2t) - 2d r_2(t) + d\,. \tag{2.188}$$

2.C.2 Expressions for higher frame potentials

$k = 2$ frame potential

We computed the second frame potential for the GUE to be

$$\begin{aligned}\mathcal{F}^{(2)}_{\text{GUE}} =& \Big(\left(d^4 - 8d^2 + 6\right)\mathcal{R}_4^2 + 4d^2\left(d^2 - 9\right)\mathcal{R}_4 + 4\left(d^6 - 9d^4 + 4d^2 + 24\right)\mathcal{R}_2^2 \\ &- 8d^2\left(d^4 - 11d^2 + 18\right)\mathcal{R}_2 + 2\left(d^4 - 7d^2 + 12\right)\mathcal{R}_{4,1}^2 - 4d^2\left(d^2 - 9\right)\mathcal{R}_{4,2} \\ &+ \left(d^4 - 8d^2 + 6\right)\mathcal{R}_{4,2}^2 - 8\left(d^4 - 8d^2 + 6\right)\mathcal{R}_2\mathcal{R}_4 - 4d\left(d^2 - 4\right)\mathcal{R}_4\mathcal{R}_{4,1} \\ &+ 16d\left(d^2 - 4\right)\mathcal{R}_2\mathcal{R}_{4,1} - 8\left(d^2 + 6\right)\mathcal{R}_2\mathcal{R}_{4,2} + 2\left(d^2 + 6\right)\mathcal{R}_4\mathcal{R}_{4,2} \\ &- 4d\left(d^2 - 4\right)\mathcal{R}_{4,1}\mathcal{R}_{4,2} + 2d^4\left(d^4 - 12d^2 + 27\right)\Big) \\ &\Big/\Big((d-3)(d-2)(d-1)d^2(d+1)(d+2)(d+3)\Big)\,.\end{aligned}$$

with form factors as defined in Eq. (2.176). Let us try and extract the interesting behavior encoded in the expression. We know the maximal value of the spectral n-point functions defined above at early times, $\mathcal{R}_2 \sim d^2$, $\mathcal{R}_4 \sim d^4$, $\mathcal{R}_{4,1} \sim d^3$, and $\mathcal{R}_{4,2} \sim d^2$. From the expression for the frame potential above, we keep the terms that are not suppressed in $1/d$, i.e. can contribute at least at zeroth order:

$$\begin{aligned}\mathcal{F}^{(2)}_{\text{GUE}} \sim 2 &- \frac{8\mathcal{R}_2}{d^2} - \frac{36\mathcal{R}_2^2}{d^4} + \frac{4\mathcal{R}_2^2}{d^2} + \frac{4\mathcal{R}_4}{d^4} + \frac{6\mathcal{R}_4^2}{d^8} - \frac{8\mathcal{R}_4^2}{d^6} + \frac{\mathcal{R}_4^2}{d^4} + \frac{\mathcal{R}_{4,2}^2}{d^4} - \frac{14\mathcal{R}_{4,1}^2}{d^6} \\ &+ \frac{2\mathcal{R}_{4,1}^2}{d^4} + \frac{16\mathcal{R}_2\mathcal{R}_{4,1}}{d^5} + \frac{16\mathcal{R}_4\mathcal{R}_{4,1}}{d^7} - \frac{4\mathcal{R}_4\mathcal{R}_{4,1}}{d^5} + \frac{2\mathcal{R}_4\mathcal{R}_{4,2}}{d^6} - \frac{4\mathcal{R}_{4,1}\mathcal{R}_{4,2}}{d^5} \\ &+ \frac{64\mathcal{R}_2\mathcal{R}_4}{d^6} - \frac{8\mathcal{R}_2\mathcal{R}_4}{d^4}\,,\end{aligned}$$

with the Haar value appearing at the beginning. At early times, the leading order behavior is $\mathcal{F}^{(2)}_{\text{GUE}} \sim \mathcal{R}_4^2/d^4$. From our calculation of the n-point form factors, we know that at the dip time all form factor terms above are suppressed in d, meaning the frame potential goes like the Haar value. Knowing the late time value of the 2-point and 4-point form factors, the terms above that will contribute at late times are

$$\text{Late}: \quad \mathcal{F}^{(2)}_{\text{GUE}} \approx 2 + \frac{\mathcal{R}_4^2}{d^4} + \frac{4\mathcal{R}_2^2}{d^2}\,, \tag{2.189}$$

which gives ≈ 10 in the large d limit. In the strict $t \to \infty$ limit, where $\mathcal{R}_2 \to d$, $\mathcal{R}_4 \to 2d^2 - d$, and $\mathcal{R}_{4,1}, \mathcal{R}_{4,2} \to d$, we have

$$\mathcal{F}^{(2)}_{\mathrm{GUE}} = \frac{10d^2 + 22d - 20}{d^2 + 5d + 6} \quad \text{and} \quad \mathcal{F}^{(2)}_{\mathrm{GUE}} \approx 10 \text{ for } d \gg 1\,. \tag{2.190}$$

As the left-hand side expression is valid for any d at late times, in doing the numerics and taking the sample size to be large, this is the value for d we should converge to.

$k = 3$ **frame potential**

The full expression for the third frame potential of the GUE is

$$\begin{aligned}
\mathcal{F}^{(3)}_{\mathrm{GUE}} = \quad & \big(6d^{14} + 18\mathcal{R}_2^3 d^{12} - 36\mathcal{R}_2 d^{12} - 318d^{12} - 846\mathcal{R}_2^3 d^{10} + 9\mathcal{R}_4^2 d^{10} + 18\mathcal{R}_{4,1}^2 d^{10} + 9\mathcal{R}_{4,2}^2 d^{10} + 1836\mathcal{R}_2 d^{10} - 72\mathcal{R}_2\mathcal{R}_4 d^{10} + 36\mathcal{R}_4 d^{10} - 36\mathcal{R}_{4,2} d^{10} + 5550d^{10} \\
& + 144\mathcal{R}_2\mathcal{R}_{4,1} d^9 - 36\mathcal{R}_4\mathcal{R}_{4,1} d^9 - 36\mathcal{R}_{4,1}\mathcal{R}_{4,2} d^9 + 11574\mathcal{R}_2^2 d^8 - 369\mathcal{R}_4^2 d^8 + \mathcal{R}_6^2 d^8 - 828\mathcal{R}_{4,1}^2 d^8 + 9\mathcal{R}_2^2\mathcal{R}_{4,2}^2 d^8 - 18\mathcal{R}_2\mathcal{R}_{4,2}^2 d^8 - 441\mathcal{R}_{4,2}^2 d^8 + 6\mathcal{R}_{6,1}^2 d^8 \\
& + 4\mathcal{R}_{6,2}^2 d^8 + 12\mathcal{R}_{6,3}^2 d^8 + 4\mathcal{R}_{6,4}^2 d^8 - 29772\mathcal{R}_2 d^8 + 3276\mathcal{R}_2\mathcal{R}_4 d^8 - 1728\mathcal{R}_4 d^8 + 36\mathcal{R}_2\mathcal{R}_6 d^8 - 18\mathcal{R}_4\mathcal{R}_6 d^8 - 12\mathcal{R}_6 d^8 - 36\mathcal{R}_2^2\mathcal{R}_{4,2} d^8 + 18\mathcal{R}_4\mathcal{R}_{4,2} d^8 \\
& + 1800\mathcal{R}_{4,2} d^8 - 36\mathcal{R}_{4,1}\mathcal{R}_{6,1} d^8 - 24\mathcal{R}_{6,4} d^8 - 37158d^8 - 6192\mathcal{R}_2\mathcal{R}_{4,1} d^7 + 1332\mathcal{R}_4\mathcal{R}_{4,1} d^7 + 36\mathcal{R}_6\mathcal{R}_{4,1} d^7 + 108\mathcal{R}_2\mathcal{R}_{4,1}\mathcal{R}_{4,2} d^7 + 1548\mathcal{R}_{4,1}\mathcal{R}_{4,2} d^7 \\
& - 144\mathcal{R}_2\mathcal{R}_{6,1} d^7 + 108\mathcal{R}_4\mathcal{R}_{6,1} d^7 - 12\mathcal{R}_6\mathcal{R}_{6,1} d^7 - 36\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,1} d^7 + 36\mathcal{R}_{4,2}\mathcal{R}_{6,1} d^7 + 72\mathcal{R}_{4,1}\mathcal{R}_{6,2} d^7 - 24\mathcal{R}_{6,1}\mathcal{R}_{6,2} d^7 + 144\mathcal{R}_2\mathcal{R}_{6,3} d^7 - 72\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,3} d^7 \\
& + 72\mathcal{R}_{4,2}\mathcal{R}_{6,3} d^7 - 24\mathcal{R}_{6,2}\mathcal{R}_{6,3} d^7 - 48\mathcal{R}_{6,3}\mathcal{R}_{6,4} d^7 - 39978\mathcal{R}_2^2 d^6 + 3726\mathcal{R}_4^2 d^6 - 41\mathcal{R}_6^2 d^6 + 11610\mathcal{R}_{4,1}^2 d^6 - 297\mathcal{R}_2^2\mathcal{R}_{4,2}^2 d^6 + 594\mathcal{R}_2\mathcal{R}_{4,2}^2 d^6 + 6750\mathcal{R}_{4,2}^2 d^6 \\
& - 204\mathcal{R}_{6,1}^2 d^6 - 156\mathcal{R}_{6,2}^2 d^6 - 348\mathcal{R}_{6,3}^2 d^6 - 148\mathcal{R}_{6,4}^2 d^6 + 169812\mathcal{R}_2 d^6 - 42768\mathcal{R}_2\mathcal{R}_4 d^6 + 24732\mathcal{R}_4 d^6 - 1512\mathcal{R}_2\mathcal{R}_6 d^6 + 738\mathcal{R}_4\mathcal{R}_6 d^6 + 528\mathcal{R}_6 d^6 \\
& + 1512\mathcal{R}_2^2\mathcal{R}_{4,2} d^6 - 432\mathcal{R}_2\mathcal{R}_{4,2} d^6 - 162\mathcal{R}_2\mathcal{R}_4\mathcal{R}_{4,2} d^6 - 486\mathcal{R}_4\mathcal{R}_{4,2} d^6 + 18\mathcal{R}_2\mathcal{R}_6\mathcal{R}_{4,2} d^6 - 18\mathcal{R}_6\mathcal{R}_{4,2} d^6 - 27972\mathcal{R}_{4,2} d^6 + 1224\mathcal{R}_{4,1}\mathcal{R}_{6,1} d^6 + 144\mathcal{R}_2\mathcal{R}_{6,2} d^6 \\
& - 144\mathcal{R}_4\mathcal{R}_{6,2} d^6 + 16\mathcal{R}_6\mathcal{R}_{6,2} d^6 + 72\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,2} d^6 - 72\mathcal{R}_{4,2}\mathcal{R}_{6,2} d^6 - 48\mathcal{R}_{6,2} d^6 - 360\mathcal{R}_{4,1}\mathcal{R}_{6,3} d^6 + 120\mathcal{R}_{6,1}\mathcal{R}_{6,3} d^6 - 144\mathcal{R}_2\mathcal{R}_{6,4} d^6 + 72\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,4} d^6 \\
& - 72\mathcal{R}_{4,2}\mathcal{R}_{6,4} d^6 + 32\mathcal{R}_{6,2}\mathcal{R}_{6,4} d^6 + 1032\mathcal{R}_{6,4} d^6 + 89040 d^6 + 72576\mathcal{R}_2\mathcal{R}_{4,1} d^5 - 11232\mathcal{R}_4\mathcal{R}_{4,1} d^5 - 1188\mathcal{R}_6\mathcal{R}_{4,1} d^5 - 3132\mathcal{R}_2\mathcal{R}_{4,1}\mathcal{R}_{4,2} d^5 - 18792\mathcal{R}_{4,1}\mathcal{R}_{4,2} d^5 \\
& + 5040\mathcal{R}_2\mathcal{R}_{6,1} d^5 - 3564\mathcal{R}_4\mathcal{R}_{6,1} d^5 + 396\mathcal{R}_6\mathcal{R}_{6,1} d^5 + 1044\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,1} d^5 - 1044\mathcal{R}_{4,2}\mathcal{R}_{6,1} d^5 - 2232\mathcal{R}_{4,1}\mathcal{R}_{6,2} d^5 + 744\mathcal{R}_{6,1}\mathcal{R}_{6,2} d^5 - 5040\mathcal{R}_2\mathcal{R}_{6,3} d^5 \\
& + 432\mathcal{R}_4\mathcal{R}_{6,3} d^5 - 48\mathcal{R}_6\mathcal{R}_{6,3} d^5 + 2088\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,3} d^5 - 2088\mathcal{R}_{4,2}\mathcal{R}_{6,3} d^5 + 648\mathcal{R}_{6,2}\mathcal{R}_{6,3} d^5 + 288\mathcal{R}_{4,1}\mathcal{R}_{6,4} d^5 - 96\mathcal{R}_{6,1}\mathcal{R}_{6,4} d^5 + 1488\mathcal{R}_{6,3}\mathcal{R}_{6,4} d^5 - 522\mathcal{R}_4^2 d^4 \\
& - 52128\mathcal{R}_2^2 d^4 + 458\mathcal{R}_6^2 d^4 - 55692\mathcal{R}_{4,1}^2 d^4 + 2430\mathcal{R}_2^2\mathcal{R}_{4,2}^2 d^4 - 4860\mathcal{R}_2\mathcal{R}_{4,2}^2 d^4 - 35190\mathcal{R}_{4,2}^2 d^4 + 1794\mathcal{R}_{6,1}^2 d^4 + 1660\mathcal{R}_{6,2}^2 d^4 + 2388\mathcal{R}_{6,3}^2 d^4 + 1440\mathcal{R}_{6,4}^2 d^4 \\
& - 274320\mathcal{R}_2 d^4 + 146412\mathcal{R}_2\mathcal{R}_4 d^4 + 17172\mathcal{R}_2\mathcal{R}_6 d^4 - 8244\mathcal{R}_4\mathcal{R}_6 d^4 - 6276\mathcal{R}_6 d^4 - 15876\mathcal{R}_2^2\mathcal{R}_{4,2} d^4 + 18144\mathcal{R}_2\mathcal{R}_{4,2} d^4 + 3078\mathcal{R}_2\mathcal{R}_4\mathcal{R}_{4,2} d^4 + 324\mathcal{R}_4\mathcal{R}_{4,2} d^4 \\
& - 342\mathcal{R}_2\mathcal{R}_6\mathcal{R}_{4,2} d^4 + 342\mathcal{R}_6\mathcal{R}_{4,2} d^4 + 141408\mathcal{R}_{4,2} d^4 - 10764\mathcal{R}_{4,1}\mathcal{R}_{6,1} d^4 - 4608\mathcal{R}_2\mathcal{R}_{6,2} d^4 + 3672\mathcal{R}_4\mathcal{R}_{6,2} d^4 - 408\mathcal{R}_6\mathcal{R}_{6,2} d^4 - 1368\mathcal{R}_{4,2}\mathcal{R}_{6,2} d^4 \\
& + 1368\mathcal{R}_{4,2}\mathcal{R}_{6,2} d^4 + 1968\mathcal{R}_{6,2} d^4 + 7200\mathcal{R}_{4,1}\mathcal{R}_{6,3} d^4 - 2400\mathcal{R}_{6,1}\mathcal{R}_{6,3} d^4 + 3312\mathcal{R}_2\mathcal{R}_{6,4} d^4 - 288\mathcal{R}_4\mathcal{R}_{6,4} d^4 + 32\mathcal{R}_6\mathcal{R}_{6,4} d^4 - 1368\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,4} d^4 \\
& + 1368\mathcal{R}_{4,2}\mathcal{R}_{6,4} d^4 - 752\mathcal{R}_{6,2}\mathcal{R}_{6,4} d^4 - 11568\mathcal{R}_{6,4} d^4 - 96000 d^4 - 199728\mathcal{R}_2\mathcal{R}_{4,1} d^3 - 4392\mathcal{R}_4\mathcal{R}_{4,1} d^3 + 9144\mathcal{R}_6\mathcal{R}_{4,1} d^3 + 26352\mathcal{R}_2\mathcal{R}_{4,1}\mathcal{R}_{4,2} d^3 \\
& + 51552\mathcal{R}_{4,1}\mathcal{R}_{4,2} d^3 - 37296\mathcal{R}_2\mathcal{R}_{6,1} d^3 + 27432\mathcal{R}_4\mathcal{R}_{6,1} d^3 - 3048\mathcal{R}_6\mathcal{R}_{6,1} d^3 - 8784\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,1} d^3 + 8784\mathcal{R}_{4,2}\mathcal{R}_{6,1} d^3 + 17928\mathcal{R}_{4,1}\mathcal{R}_{6,2} d^3 - 5976\mathcal{R}_{6,1}\mathcal{R}_{6,2} d^3 \\
& + 37296\mathcal{R}_2\mathcal{R}_{6,3} d^3 - 1080\mathcal{R}_4\mathcal{R}_{6,3} d^3 + 120\mathcal{R}_6\mathcal{R}_{6,3} d^3 - 17568\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,3} d^3 + 17568\mathcal{R}_{4,2}\mathcal{R}_{6,3} d^3 - 190512\mathcal{R}_2\mathcal{R}_{4,2} d^2 - 100800\mathcal{R}_4 d^4 - 5736\mathcal{R}_{6,2}\mathcal{R}_{6,3} d^3 \\
& - 720\mathcal{R}_{4,1}\mathcal{R}_{6,4} d^3 + 240\mathcal{R}_{6,1}\mathcal{R}_{6,4} d^3 - 11952\mathcal{R}_{6,3}\mathcal{R}_{6,4} d^3 + 141840\mathcal{R}_2^2 d^2 - 49284\mathcal{R}_4^2 d^2 - 1258\mathcal{R}_6^2 d^2 + 111852\mathcal{R}_{4,1}^2 d^2 + 1098\mathcal{R}_2^2\mathcal{R}_{4,2}^2 d^2 - 2196\mathcal{R}_2\mathcal{R}_{4,2}^2 d^2 \\
& + 53712\mathcal{R}_{4,2}^2 d^2 - 3756\mathcal{R}_{6,1}^2 d^2 - 3188\mathcal{R}_{6,2}^2 d^2 + 108\mathcal{R}_{6,3}^2 d^2 - 2736\mathcal{R}_{6,4}^2 d^2 + 288000\mathcal{R}_2 d^2 + 5472\mathcal{R}_2\mathcal{R}_4 d^2 - 47376\mathcal{R}_2\mathcal{R}_6 d^2 + 22644\mathcal{R}_4\mathcal{R}_6 d^2 + 14400\mathcal{R}_6 d^2 \\
& + 14400\mathcal{R}_2^2\mathcal{R}_{4,2} d^2 - 9396\mathcal{R}_2\mathcal{R}_4\mathcal{R}_{4,2} d^2 + 49824\mathcal{R}_4\mathcal{R}_{4,2} d^2 + 1044\mathcal{R}_2\mathcal{R}_6\mathcal{R}_{4,2} d^2 - 1044\mathcal{R}_6\mathcal{R}_{4,2} d^2 - 115200\mathcal{R}_{4,2} d^2 + 22536\mathcal{R}_{4,1}\mathcal{R}_{6,1} d^2 + 24624\mathcal{R}_2\mathcal{R}_{6,2} d^2 \\
& - 16488\mathcal{R}_4\mathcal{R}_{6,2} d^2 + 1832\mathcal{R}_6\mathcal{R}_{6,2} d^2 + 4176\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,2} d^2 - 4176\mathcal{R}_{4,2}\mathcal{R}_{6,2} d^2 - 19200\mathcal{R}_{6,2} d^2 - 45720\mathcal{R}_{4,1}\mathcal{R}_{6,3} d^2 + 15240\mathcal{R}_{6,1}\mathcal{R}_{6,3} d^2 + 8352\mathcal{R}_2\mathcal{R}_{6,4} d^2 \\
& - 8352\mathcal{R}_4\mathcal{R}_{6,4} d^2 + 928\mathcal{R}_6\mathcal{R}_{6,4} d^2 + 4176\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,4} d^2 - 4176\mathcal{R}_{4,2}\mathcal{R}_{6,4} d^2 + 5520\mathcal{R}_{6,2}\mathcal{R}_{6,4} d^2 + 19200\mathcal{R}_{6,4} d^2 + 133200\mathcal{R}_2\mathcal{R}_{4,1} d + 53208\mathcal{R}_4\mathcal{R}_{4,1} d \\
& - 12312\mathcal{R}_6\mathcal{R}_{4,1} d - 62208\mathcal{R}_2\mathcal{R}_{4,1}\mathcal{R}_{4,2} d + 4608\mathcal{R}_{4,1}\mathcal{R}_{4,2} d + 32400\mathcal{R}_2\mathcal{R}_{6,1} d - 36936\mathcal{R}_4\mathcal{R}_{6,1} d + 4104\mathcal{R}_6\mathcal{R}_{6,1} d + 20736\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,1} d - 20736\mathcal{R}_{4,2}\mathcal{R}_{6,1} d \\
& - 33048\mathcal{R}_{4,1}\mathcal{R}_{6,2} d + 11016\mathcal{R}_{6,1}\mathcal{R}_{6,2} d - 32400\mathcal{R}_2\mathcal{R}_{6,3} d - 25272\mathcal{R}_4\mathcal{R}_{6,3} d + 2808\mathcal{R}_6\mathcal{R}_{6,3} d + 41472\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,3} d - 41472\mathcal{R}_{4,2}\mathcal{R}_{6,3} d + 16632\mathcal{R}_{6,2}\mathcal{R}_{6,3} d \\
& - 16848\mathcal{R}_{4,1}\mathcal{R}_{6,4} d + 5616\mathcal{R}_{6,1}\mathcal{R}_{6,4} d + 22032\mathcal{R}_{6,3}\mathcal{R}_{6,4} d - 216000\mathcal{R}_2^2 - 2160\mathcal{R}_4^2 + 240\mathcal{R}_6^2 - 105840\mathcal{R}_{4,1}^2 - 12960\mathcal{R}_2^2\mathcal{R}_{4,2}^2 + 25920\mathcal{R}_2\mathcal{R}_{4,2}^2 - 34560\mathcal{R}_{4,2}^2 \\
& - 2160\mathcal{R}_{6,1}^2 - 2160\mathcal{R}_{6,2}^2 - 19440\mathcal{R}_{6,3}^2 - 960\mathcal{R}_{6,4}^2 + 43200\mathcal{R}_2\mathcal{R}_4 + 14400\mathcal{R}_2\mathcal{R}_6 - 4320\mathcal{R}_4\mathcal{R}_6 + 172800\mathcal{R}_2\mathcal{R}_{4,2} + 25920\mathcal{R}_2\mathcal{R}_4\mathcal{R}_{4,2} - 69120\mathcal{R}_4\mathcal{R}_{4,2} \\
& - 2880\mathcal{R}_2\mathcal{R}_6\mathcal{R}_{4,2} + 2880\mathcal{R}_6\mathcal{R}_{4,2} + 12960\mathcal{R}_{4,1}\mathcal{R}_{6,1} + 14400\mathcal{R}_2\mathcal{R}_{6,2} + 4320\mathcal{R}_4\mathcal{R}_{6,2} - 480\mathcal{R}_6\mathcal{R}_{6,2} - 11520\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,2} + 11520\mathcal{R}_{4,2}\mathcal{R}_{6,2} + 90720\mathcal{R}_{4,1}\mathcal{R}_{6,3} \\
& - 30240\mathcal{R}_{6,1}\mathcal{R}_{6,3} - 28800\mathcal{R}_2\mathcal{R}_{6,4} - 2880\mathcal{R}_6\mathcal{R}_{6,4} + 25920\mathcal{R}_4\mathcal{R}_{6,4} - 11520\mathcal{R}_2\mathcal{R}_{4,2}\mathcal{R}_{6,4} + 11520\mathcal{R}_{4,2}\mathcal{R}_{6,4} - 6720\mathcal{R}_{6,2}\mathcal{R}_{6,4}\big) \\
& \big/\big((d-5)(d-4)(d-3)(d-2)(d-1)d^2(d+1)(d+2)(d+3)(d+4)(d+5)\big)\,.
\end{aligned}$$

The expression is best appreciated from a distance.

2.C.3 Expressions for Weingarten

Lastly, we give the definition of the unitary Weingarten function, which appeared in the integration of Haar random unitaries in Eq. (2.80). The $2k$-th moment of the Haar ensemble appeared in the k-th frame potential. For the n-th moment, the Weingarten function is a

function of an element σ of the permutation group S_n and presented as defined in [59],

$$\mathcal{W}g(\sigma) = \frac{1}{(n!)^2} \sum_{\lambda} \frac{\chi_\lambda(e)^2 \chi_\lambda(\sigma)}{s_\lambda(1)}\,, \tag{2.191}$$

where we sum over integer partitions of n (recall that the conjugacy classes of S_n are labeled by integer partitions of n). χ_λ is an irreducible character of S_n labeled by λ (as each irrep of S_n can be associated to an integer partition) and e is the identity element. $s_\lambda(1) = s_\lambda(1, \ldots, 1)$ is the Schur polynomial evaluated on d arguments and indexed by the partition λ. For instance, the Weingarten functions needed to compute the first frame potential were

$$\mathcal{W}g(\{1,1\}) = \frac{1}{d^2-1} \quad \text{and} \quad \mathcal{W}g(\{2\}) = -\frac{1}{d(d^2-1)}\,. \tag{2.192}$$

2.D Additional numerics

We conclude with a few numerical checks on the formulae we derived for the form factors and frame potentials.

2.D.1 Form factors and numerics

As we mentioned in Sec. 2.2.2 and discussed in App. 2.C.1, in order to derive expressions for the form factors for the GUE we had to make approximations which should be compared to numerics for the GUE.

We briefly remind the reader that at infinite temperature, we derived the expression

$$\mathcal{R}_2(t) = d^2 r_1^2(t) - d r_2(t) + d\,. \tag{2.193}$$

Numerical checks of this expression are shown in Fig. 2.8. We see that the approximations employed work well at $\beta = 0$, reproducing the early time oscillations, dip, plateau, and ramp features. But there is some discrepancy in the ramp behavior which merits discussion. As we take $d \to \infty$, the difference between the predicted ramp and numerical ramp is not suppressed. In Fig. 2.8, we see that the relative error between the numerics and analytic

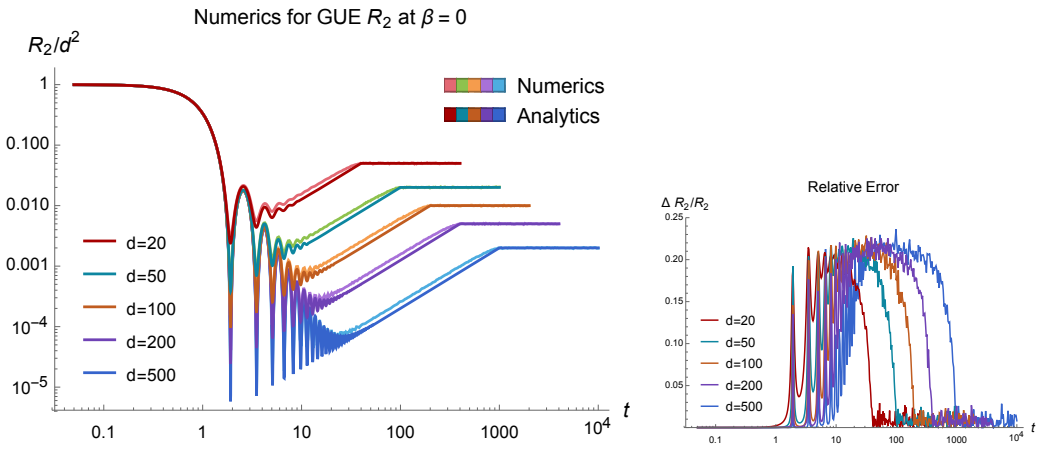


Figure 2.8: Numerical checks of the GUE 2-point spectral form factor at infinite temperature for various values of d and normalized by d^2. The analytic expressions derived in Sec. 2.2 are in the lighter shades and the numerics for GUE are in darker shades. Numerics were done 10000 samples from the GUE. On the right we plot the relative error between the numerics and analytic predictions. We observe good agreement at early and late times, and see deviations around the ramp.

prediction does not decrease as we increase d, indicating that this difference in the ramp prediction is not an artifact of finite d numerics. On a log-log plot, this shift from the numerics suggests that we capture the correct linear behavior, but with a slightly different slope for the ramp.

The $r_2(t) = 1 - t/2d$ function which controls the slope behavior comes from the Fourier transform of the square of the sine kernel. Recall that in our approximation, we integrated over the entire semicircle. A phenomenological observation is that the modified ramp function defined by $\tilde{r}_2(t) \equiv 1 - 2t/\pi d$, where we change the slope to $2/\pi$, does a much better job of capturing the ramp behavior. Working in the short-distance limit of the 2-point correlator $\rho^{(2)}(\lambda_1, \lambda_2)$ (as in [24]) and integrating the sine kernel over the entire semicircle, we obtain $\tilde{r}_2$ whose behavior we only trust near the dip.

Numerically, we find that this modified slope of $2\pi/d$ better captures the r_2 function near the dip, with error that is suppressed as we take $d \to \infty$. The same numerics are reported in Fig. 2.9, but with the modified ramp behavior. There is still some discrepancy near the plateau time when we transition to the constant plateau value, but the ramp behaviors near the dip are in much better agreement.

We understand the Bessel function contribution to $\mathcal{R}_2(t)$, which arises from 1-point func-

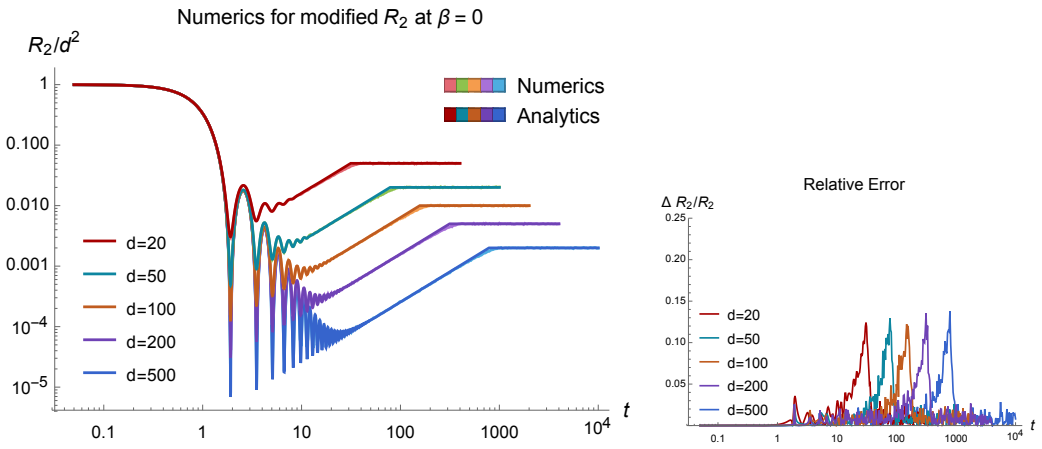


Figure 2.9: The same numerics as reported in Fig. 2.8, but now compared to the analytic expression with the modified ramp behavior $\tilde{r}_2(t)$.

tions. The subtlety above is really in the connected piece of the 2-point function

$$\mathcal{R}_2(t)_{\text{con}} \equiv \mathcal{R}_2(t) - d^2 r_1^2(t)\,. \tag{2.194}$$

Numerically, we see that the connected 2-point form factor for the GUE exhibits three different behaviors: an early time quadratic growth, an intermediate linear growth, and then a late-time constant plateau. The closed form expression we derived in Sec. 2.2 should be viewed as a coarse approximation before the plateau, approximately capturing the linear regime. The modified ramp function $\tilde{r}_2(t) = 1 - 2t/\pi d$ appears to capture the linear behavior near the dip with the correct slope. In [68], a more detailed treatment of the connected correlator is given at early times. From the integral representation of the connected 2-point form factor, they find that

$$\text{Early}: \quad \mathcal{R}_2(t)_{\text{con}} \approx t^2 - \frac{1}{2}\,t^4 + \frac{1}{3}\,t^6 + \ldots \tag{2.195}$$

to leading order in d (Eq.(2.28) in [68]). The three behaviors are compared with numerics in Fig. 2.10.

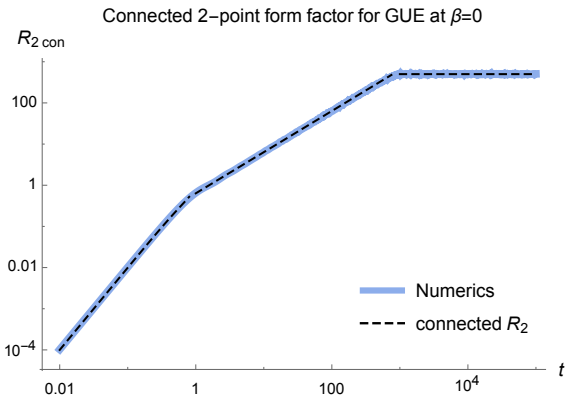


Figure 2.10: Numerics for the connected 2-point spectral form factor for GUE at infinite temperature plotted for $d = 500$ with 10000 random samples. The dashed line is the expression Eq. (2.196) approximating the three regimes of the connected form factor.

In summary, the three regimes of the connected 2-point form factor are roughly captured by

$$\mathcal{R}_2(t)_{\text{con}} = \begin{cases} \sim t^2 & \text{for} \quad t \lesssim 1\,, \\ \sim \frac{2}{\pi} t & \text{for} \quad 1 \lesssim t \lesssim 2d\,, \\ d & \text{for} \quad t \gtrsim 2d\,. \end{cases} \tag{2.196}$$

The early time quadratic behavior does not play an important role in our analysis of GUE correlation functions and frame potentials, but is of independent physical interest. This intruiging early-time behavior of the connected 2-point form factor is explored in [75].

At finite temperature we find good agreement between the expression $\mathcal{R}_2(t, \beta)$ and numerics at early and late times, but again see a deviation of the dip and ramp behaviors from the analytic prediction, as shown in Fig. 2.11. Using the modified ramp $\tilde{r}_2$ we find closer agreement at small β, but as we increase β the predicted ramp behavior again starts to deviate from the numerics, indicating that there is a β-dependence to the slope that we do not fully understand. But as we discussed in App. 2.C.1, we only trust the short-distance approximation at finite temperature, and thus $\mathcal{R}_2(t, \beta)$, for small β. We also report numerics for the $\mathcal{R}_4$ expression in Fig. 2.12.

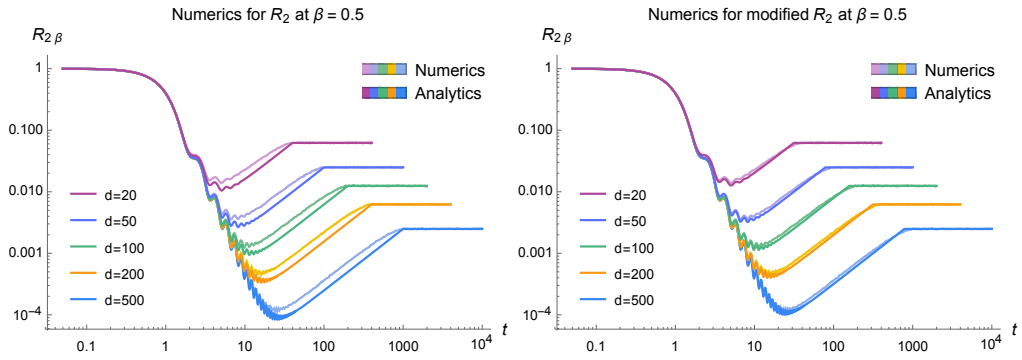


Figure 2.11: Numerical checks of the finite temperature 2-point spectral form factor for GUE at $\beta = 0.5$, plotted for various values of d and normalized by their initial values. Numerics were done with a GUE sample size of 10000. The left figure uses the expression for $\mathcal{R}_2(t, \beta)$ derived in Sec. 2.2.2 and 2.C.1, whereas the right figure uses the modified ramp $\tilde{r}_2$ discussed above.

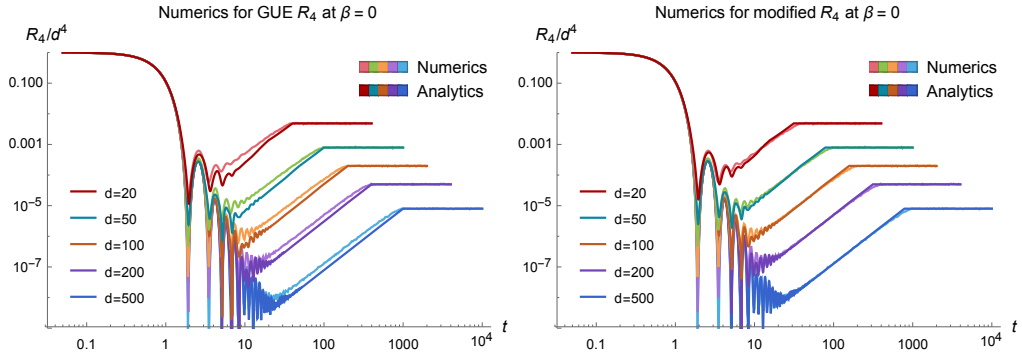


Figure 2.12: Numerical checks the infinite temperature 4-point spectral form factor for GUE with 10000 samples, plotted for various values of d and normalized by their initial values. The left figure uses the $\mathcal{R}_4$ expression derived in App. 2.C.1, and the right figure uses $\tilde{r}_2$.

2.D.2 Frame potentials and numerics

As the frame potential depends on the eigenvectors of the elements in the ensemble (and not just the eigenvalues as per the form factors) and requires a double sum over the ensemble, numerical simulation of the frame potential is harder than for the form factors. For an ensemble of $d \times d$ matrices, we need to consider sample sizes greater than d^{2k} for the k-th frame potential, which amounts to summing over many samples for fairly modest Hilbert space dimension. Instead, for a given d, we can sequentially increase the sample size and extrapolate to large $|\mathcal{E}_{\text{GUE}}|$. In Fig. 2.13 we consider the first frame potential for the GUE at $d = 32$ and, in the limit of large sample size, find good agreement with the analytic expression

computed from $\mathcal{R}_2$. Alternatively, we can numerically compute the frame potentials by ignoring the coincident contributions to the double sum in $\mathcal{F}^{(k)}$, i.e. when $U = V$. For a finite number of samples, these terms contribute $d^2/|\mathcal{E}|$ to the sum, meaning we must look at large ensembles before their contribution does not dominate entirely. Ignoring these terms, we can time average over a sliding window to compute the frame potential with only a few samples, as shown in Fig. 2.13.

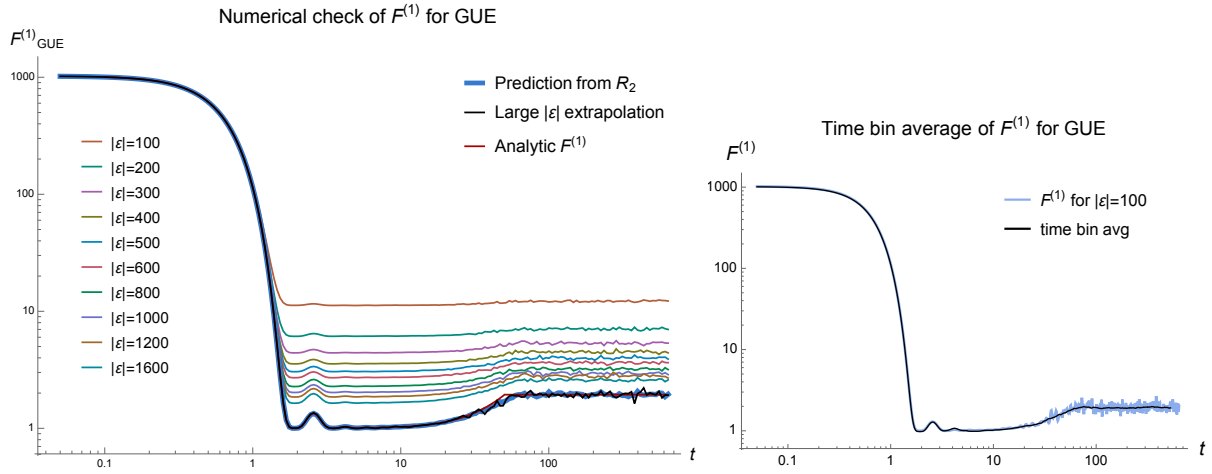


Figure 2.13: Numerical computation of first frame potential for the GUE at $d = 32$. On the left, we sequentially increase the number of samples and extrapolate to large sample size (red line), which agrees with the both the frame potential computed from $\mathcal{R}_2$ numerics as in Eq. (2.86) (blue line) and the analytic expression we derived for $\mathcal{F}^{(1)}_{\text{GUE}}$. On the right, we time bin average $\mathcal{F}^{(1)}_{\text{GUE}}$ as described above and, for $d = 32$ and 100 samples, we find good agreement with the quantities on the left.

2.D.3 Minimal realizations and time averaging

Given an ensemble of disordered systems, one can ask whether a quantity averaged over the ensemble is the same as for a single random instance of the ensemble. It is known that up until the dip time, the spectral form factor is self-averaging, meaning that single instance captures the average for large d [79]. However, the spectral form factor is not self-averaging at late times. We can try to extract the averaged behavior from a single instance in regimes dominated by large fluctuations by averaging over a moving time window. In Fig. 2.14, we see that for a single instance of the GUE, the time average of the spectral form factor at finite β gives the same result as the ensemble average for sufficiently large d. For the frame

potential, we can consider two instances, the smallest ensemble for which the frame potential makes sense. Ignoring the coincident terms in the sum, we see that the frame potential is also self-averaging at early times and that the time average at late times agrees with the ensemble average and analytic expression.

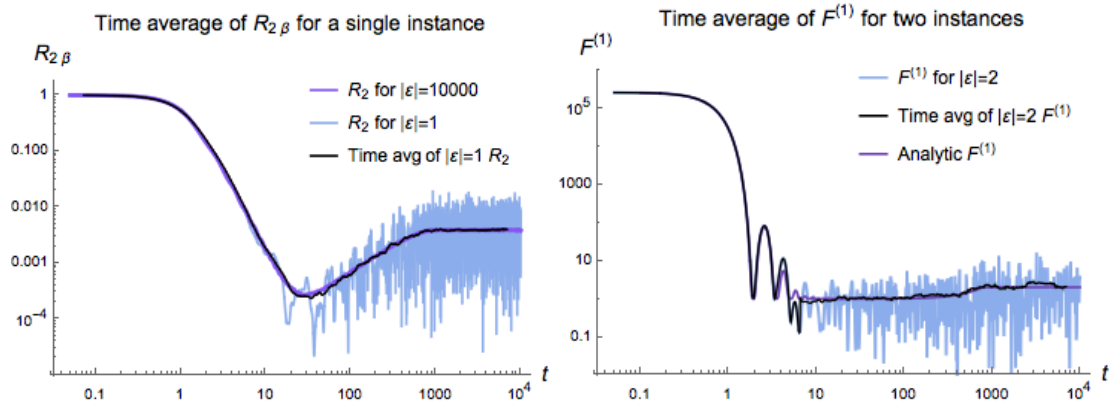


Figure 2.14: On the left: the time average of the thermal 2-point form factor at $\beta = 5$ and $d = 500$. On the right: the time average of the first frame potential for $d = 500$ computed for two instances. In both figures, the time average of the minimal number of instances agrees with the ensemble average.

Chapter 3

Late-time chaos and symmetry

This chapter is partially based on work in progress with Beni Yoshida, which will appear elsewhere in a future publication.

Abstract

We study the universal features at late-times in chaotic systems with symmetry and compute the late-time asymptotic form of the out-of-time ordered correlation function for different symmetry classes. As realizing different antiunitary symmetries picks out a quotient of the full unitary group, we use Weingarten calculus for the associated compact symmetric space and find different asymptotic forms of the OTOC in each symmetry class, where the leading terms are universal and the symmetry appears in the $1/d$ corrections. We also generalize the notion of a k-design and k-invariance to different symmetry classes and apply them to both random matrix ensembles and physical systems with symmetry.

3.1 Introduction

Computing averages over the unitary group has become an indispensable tool in quantum information. Understanding the properties of Haar random unitaries and states allows one to extract universal or generic features of quantum systems. Random unitaries have also proved useful in approximating chaotic dynamics, where we typically average over the full unitary group $U(d)$ to study universal features of chaotic systems, for instance the entanglement

in random subsystems and late-time values of correlation functions, and ask at what time-scales an ensemble of unitaries representing the time-evolution of a physical system looks Haar-random. Roughly speaking, as $e^{-iHt} \in U(d)$ is always valued in the unitary group, we average over $U(d)$ to extract universal features and ask when some ensemble $\mathcal{E}_t$ of time evolutions looks Haar-random with respect to $U(d)$. If the physical system of interest realizes some symmetry, we need to refine our measures of chaos and randomness as the unitary time-evolution might really be valued in a subset of the unitary group. Of course, given a fixed Hamiltonian, the subset of the unitary group that time-evolution will explore is determined by H, but we are interested in symmetry classes of Hamiltonians where the existence of a symmetry determines a specific subset, and in this sense is universal.

The famous insight due to Dyson [80], was that systems which realize (or break) anti-unitary symmetries fall into three different symmetry classes, labeled: A, AI, and AII. These refer to the Cartan labels for compact symmetric spaces, where here it is the unitary generated by the Hamiltonian e^{-iHt} that is valued in the subset of the unitary group. Dyson's three classes are:

A:	$U(d)$	no symmetry
AI:	$U(d)/O(d)$	time-rev inv w/ $T^2 = 1$
AII:	$U(2d)/Sp(2d)$	time-rev inv w/ $T^2 = -1$

where T is the time-reversal anti-unitary operator. Note that when equipped with an invariant measure, the ensembles on these spaces are often referred to as the circular ensembles CUE, COE, and CSE.

One of our goals in this chapter will be to consider averaging over these other spaces in order to understand the universal behavior of chaotic systems with symmetry. There is a further refinement of this classification scheme by Altland and Zirnbauer [81] for free fermion Hamiltonians. One way of interpreting this classification scheme is that the first-quantized Hamiltonian lives in the tangent space of 10 different compact symmetric spaces, depending on how that Hamiltonian realizes time-reversal and charge-conjugation symmetry (and their combination). We are interested in interacting systems acting on the physical Hilbert space. We are able to compute averages over the 10 symmetric spaces in Altland and Zirnbauer's classification, but the implications for physical systems of interacting fermions are less clear.

We will mention averages of OTOCs over all 10 classes, but we will primarily be interested in the 3-fold classification applicable to general interacting quantum systems: A, AI, and AII.

3.2 Haar-averages and symmetry classes

We start by briefly overviewing the basic machinery of taking Haar-averages over the unitary group, a formalism which has become known as Weingarten calculus. We then discuss extensions of these averages to other Lie groups and to different quotients of the full unitary group. The formalism for performing these averages is known as Weingarten calculus. Weingarten calculus allows us to integrate moments of orthogonal and symplectic groups as well as quotients of the unitary group called compact symmetric spaces. We should think of this as averaging over a symmetry class.

Although integrating over the unitary group has appeared in the literature of a number of different subfields, including lattice gauge theory, 2d QCD and matrix models, and transport in disordered systems, the methods of taking Haar integrals have become indispensable tools in quantum information theory. Diagrammatic methods for computing Haar-moments were developed in the lattice gauge theory and condensed matter literature in the 80's and 90's.[1] But exact expressions for moments of the unitary group were given by Collins [58] and Collins and Śniady [59]. These ideas are deeply rooted in Schur-Weyl duality for the unitary group. The formalism for taking Haar-averages is known as Weingarten calculus, acknowledging Weingarten's early work on the subject [60].

Collins and Śniady [59] detailed how to compute moments of $U(d)$ and the other compact Lie groups $O(d)$ and $Sp(2d)$. Weingarten calculus for the orthogonal group was further clarified and expanded in work by Collins and Matsumoto [88], as well as in [89, 90]. Building on this work, Matsumoto generalized these methods to compute moments of compact symmetric spaces [91, 92].

Before diving in, let's just sketch what these moments look like. For the compact Lie groups, integrating gives a double sum over elements of the symmetric group or a subset

[1] For example, in [82, 83, 84, 85, 86, 87].

thereof, and for the compact symmetric spaces, quotients of the Lie groups by a Lie subgroup, the moments take the form of a single sum

$$\text{Lie groups}: \quad \int dU\, U^{\otimes k} \otimes (U^{\otimes})^{\dagger} = \sum_{\sigma,\tau} \delta_{\sigma}\delta_{\tau} \mathcal{W}g^{G}(\sigma^{-1}\tau, d)\,, \tag{3.1}$$

$$\text{symmetric spaces}: \int dU\, U^{\otimes k} \otimes (U^{\otimes})^{\dagger} = \sum_{\sigma} \delta_{\sigma} \mathcal{W}g^{G/K}(\sigma, d)\,, \tag{3.2}$$

where the index structure of the product of delta functions will depend on the permutation(s) we sum over. As we will see late, if the group or subgroup is orthogonal or symplectic, the moments themselves involve transposes, symplectic conjugation, or chiral conjugation. The point of the above expressions is heuristic.

Comment on terminology:

In a slight abuse of terminology, we will refer to averages over the unitary group as 'Haar-averages,' even though the Haar-measure may be put on any compact Lie group and by involution on to quotients thereof. So when we say 'integrate over Haar,' we mean over the unitary group with respect to the Haar measure. In all other cases, we will explicitly say which space we integrate over.

3.2.1 QI review and Haar-randomness

Here we give a brief overview of some quantum information theoretic definitions and notation. The discussion here follows the discussion in [11, 12] as well as [93]. We consider a finite-dimensional Hilbert space $\mathcal{H}$ of dimension d. We are interested in an ensemble of unitaries $\mathcal{E} = \{p_i, U_i\}$ that acts on this Hilbert space, where $U_i \in U(d)$ and p_i's are probabilities. Said equivalently, we have a subset of the unitary group and some probability measure on it. The ensemble could be finite or infinite, either discrete or continuous. The Haar measure is the unique left and right invariant measure on $U(d)$,

$$\int_{\text{Haar}} dU f(U) = \int_{\text{Haar}} dU f(UV) = \int_{\text{Haar}} dU f(VU)\,, \tag{3.3}$$

for a function f and for all $V \in U(d)$. As we are often interested in moments of an ensemble, consider an operator $\mathcal{O}$ acting on the k-fold Hilbert space $\mathcal{H}^{\otimes k}$. The k-fold channel with respect to an ensemble of unitaries is

$$\Phi_{\mathcal{E}}^{(k)}(\mathcal{O}) \equiv \int_{\mathcal{E}} dU (U^{\otimes k})^{\dagger} \mathcal{O} U^{\otimes k} \,. \tag{3.4}$$

A unitary k-design is an ensemble $\mathcal{E}$ for which the k-fold channels of the ensemble and Haar are equal, meaning that for any operator $\mathcal{O} \in \mathcal{A}(\mathcal{H}^{\otimes k})$

$$\Phi_{\mathcal{E}}^{(k)}(\mathcal{O}) = \Phi_{\text{Haar}}^{(k)}(\mathcal{O}) \,. \tag{3.5}$$

This means the ensemble of unitaries $\mathcal{E}$ exactly reproduces the first k moments of the Haar ensemble. This idea captures how randomly distributed the unitaries in $\mathcal{E}$ are. The more randomly distributed around the unitary group, the more moments of the Haar ensemble we can reproduce. When we reproduce the moments exactly, $\mathcal{E}$ is a unitary k-design. Given some ensemble, one might ask how close to forming a k-design we are. One quantity which measures this distance is the k-th frame potential [57], defined as a double integral over the ensemble $\mathcal{E}$ as

$$\mathcal{F}_{\mathcal{E}}^{(k)} = \int_{U,V \in \mathcal{E}} dU dV \left|\text{Tr}(U^{\dagger} V)\right|^{2k} . \tag{3.6}$$

For some ensemble of unitaries $\mathcal{E} = \{p_i, U_i\}$, the frame potential tells us how Haar-random the ensemble is. More precisely, for a given k, it tells us how close the ensemble is to forming a unitary k-design. For any ensemble of unitaries $\mathcal{E}$, the frame potential is lower bounded by the Haar value

$$\mathcal{F}_{\mathcal{E}}^{(k)} \geq \mathcal{F}_{\text{Haar}}^{(k)} \,, \quad \text{where} \quad \mathcal{F}_{\text{Haar}}^{(k)} = k! \,, \tag{3.7}$$

as we discussed above. It will be useful to review the proof of this here as we will use similar techniques to lower bound frame potentials for different symmetry classes. Consider the operator S, defined as the difference between the k-fold channels of $\mathcal{E}$ and the Haar ensemble:

$$S = \int_{\mathcal{E}} dU \, (U^{\otimes k})^{\dagger} \otimes U^{\otimes k} - \int_{\text{Haar}} dU \, (U^{\otimes k})^{\dagger} \otimes U^{\otimes k} \,. \tag{3.8}$$

Noting that the norm of the operator is positive semi-definite $\mathrm{Tr}\, S^\dagger S \geq 0$, this implies

$$\int_{U,V\in\mathcal{E}} dU dV \left|\mathrm{Tr}(U^\dagger V)\right|^{2k} - 2\int_{U\in\mathcal{E}} dU \int_{V\in\mathrm{Haar}} dV \left|\mathrm{Tr}(U^\dagger V)\right|^{2k} + \int_{U,V\in\mathrm{Haar}} dU dV \left|\mathrm{Tr}(U^\dagger V)\right|^{2k} \geq 0\,.$$

The first and last terms are respectively the frame potentials for $\mathcal{E}$ and the Haar ensemble. Using the left-invariance of the Haar measure, the middle term is also just the frame potential for the Haar ensemble. Therefore, we find that

$$\mathcal{F}^{(k)}_{\mathcal{E}} \geq \mathcal{F}^{(k)}_{\mathrm{Haar}}\,, \tag{3.9}$$

as desired. In this sense the frame potential quantifies a distance to randomness.

Haar-random unitaries

We will quickly review integration over Haar-random unitaries. The approach that leans heavily on the representation theory of symmetric groups in described in App. 3.A, while the exposition here is much more information theoretic in spirit.

To compute averages over the unitary group explicitly, we need to integrate monomials of Haar-random unitaries. Here the convenient notation and discussion of the Weingarten function follow [94]. As we mentioned, the formalism is called Weingarten calculus and follows from Schur-Weyl duality for the unitary group. The exact expression [58, 59] for integrating over $U(d)$ is

$$\int dU\, U_{i_1 j_1}\ldots U_{i_k j_k} U^\dagger_{\ell_1 m_1}\ldots U^\dagger_{\ell_k m_k} = \sum_{\sigma,\tau\in S_k} \delta_\sigma(\vec{\imath}|\vec{m})\delta_\tau(\vec{\jmath}|\vec{\ell})\mathcal{W}g^U(\sigma^{-1}\tau, d)\,, \tag{3.10}$$

where for the k-th moment we sum over elements the permutation group S_k. Here we have defined the δ-function contraction of indices indexed by a permutation $\sigma \in S_k$ as

$$\delta_\sigma(\vec{\imath}|\vec{\jmath}) = \delta_{i_1 j_{\sigma(1)}}\ldots\delta_{i_k j_{\sigma(k)}}\,. \tag{3.11}$$

The unitary *Weingarten function* acts on elements of S_k and has a simple interpretation in terms of an inverse of a matrix which counts the cycles of a permutation, which we review

in our discussion of Schur-Weyl duality below. The expression of the unitary Weingarten function as a Fourier expansion in terms of characters of the symmetric group is given in App. 3.A

Schur-Weyl duality for the unitary group is the statement that if an operator $\mathcal{O}$ in the algebra of operators acting on the k-fold Hilbert space $\mathcal{H}^{\otimes k}$, commutes with all operators $V^{\otimes k}$ where V acts on $\mathcal{H}$, then that operator must be a linear combination of permutation operators P_σ

$$\text{Schur-Weyl} \quad \rightarrow \quad \mathcal{O} = \sum_{\sigma \in S_k} c_\sigma P_\sigma \,, \tag{3.12}$$

where the permutation operator P_σ permutes the k copies of $\mathcal{H}$. More explicitly, for a state $|\psi_1, \ldots, \psi_k\rangle$ on $\mathcal{H}^{\otimes k}$, the permutation operator acts as $P_\sigma |\psi_1, \ldots, \psi_k\rangle = |\psi_{\sigma(1)}, \ldots, \psi_{\sigma(k)}\rangle$, and equivalently by conjugation for an operator on the k-fold Hilbert space. As the k-fold channel of Haar $\Phi^{(k)}_{\text{Haar}}(\mathcal{O})$ is invariant under k-fold unitary conjugation, by the invariance of the Haar measure, this means that the channel can be written as a linear combination of permutation operators. Moreover, again by the invariance of the Haar measure, the argument of the k-fold channel $\Phi^{(k)}_{\text{Haar}}(\mathcal{O})$ is also invariant under unitary k-fold conjugation, we can write the channel as

$$\Phi^{(k)}_{\text{Haar}}(\mathcal{O}) = \sum_{\sigma,\tau \in S_k} \mathcal{W}g^U_{\sigma,\tau} P_\sigma \text{Tr}(P_\tau \mathcal{O}) \,. \tag{3.13}$$

This also simply an application of Eq. (3.10) to the k-fold channel. Here $\mathcal{W}g^U_{\sigma,\tau}$ is the Weingarten matrix, the entries of which are the unitary Weingarten functions $\mathcal{W}g^U(\sigma^{-1}\tau, d)$. Taking the k-fold channel of a permutation operator must give $\Phi^{(k)}_{\text{Haar}}(P_\sigma) = P_\sigma$, and thus the Weingarten matrix is simply the inverse of the matrix of inner products of permutation operators

$$\mathcal{W}g^U_{\sigma,\tau} = \big(\text{Tr}(P_\sigma P_\tau)\big)^{-1} \,, \tag{3.14}$$

where elements of the matrix $\text{Tr}(P_\sigma P_\tau) = d^{\#\text{cycles}(\sigma\tau)}$ and simply count the number of cycles in the product of permutations.

It is straightforward to derive the value of the Haar frame potential and will be useful when we generalize our discussion to other groups and quotients. From the invariance of the

Haar measure, the frame potential Eq. (3.6) for the Haar ensemble is

$$\mathcal{F}^{(k)}_{\text{Haar}} = \int dU \, \text{Tr}(U^{\otimes k})\text{Tr}((U^{\otimes k})^\dagger) = \sum_{\sigma,\tau \in S_k} \mathcal{W}g^U_{\sigma,\tau} \text{Tr}(P_\sigma P_\tau) = k! \,, \tag{3.15}$$

for $k < d$. When viewed as a matrix equation we just have the trace of the identity matrix on S_k, and thus the above sum just counts the number of elements in S_k.

Random orthogonal matrices

The exact formula for integrating over Haar random orthogonal matrices is [59, 88]

$$\int dO \, O_{i_1 j_1} \dots O_{i_{2k} j_{2k}} = \sum_{\sigma,\tau \in M_{2k}} \Delta_\sigma(\vec{\imath}) \Delta_\tau(\vec{\jmath}) \mathcal{W}g^O(\sigma^{-1}\tau, d) \,, \tag{3.16}$$

where we sum over elements of M_{2k}, the set of pair partitions on the set $2k$ element set $\{1, 2, \dots, 2k\}$ and we have defined a combination of δ-functions labeled by $\sigma \in S_{2k}$

$$\Delta_\sigma(\vec{\imath}) = \delta_{i_{\sigma(1)} i_{\sigma(2)}} \dots \delta_{i_{\sigma(2k-1)} i_{\sigma(2k)}} \,. \tag{3.17}$$

We discuss more details of this formula and the representation theoretic expression for the orthogonal Weingarten function in App. 3.A.

Schur-Weyl duality for the orthogonal group is given by the action of the Brauer algebra [59], which has a basis of pair partitions on a set of $2k$ elements. The set of pair partitions M_{2k} has dimension $(2k)!/(2^k k!)$ and can be realized as a subset of the permutation group S_{2k}. We describe pair partitions in more detail in the appendix, but for now, the important thing for writing the Haar-integrals over $O(d)$ is that these are just a subset of S_{2k}. With this in mind, Schur-Weyl duality for $O(d)$ implies that if an operator $\mathcal{O}$ commutes with all operators $V^{\otimes k}$ with $V \in O(d)$ on the k-fold Hilbert space $\mathcal{H}^{\otimes k}$, then it can be written as

$$\text{Schur-Weyl} \quad \rightarrow \quad \mathcal{O} = \sum_{\sigma \in M_{2k}} c_\sigma S_\sigma \,, \tag{3.18}$$

where S_σ are no longer simply permutation operators on the k copies, but instead involve permutations and projections. For example, the basis elements for $k = 2$ are the identity

and swap operator (as in the unitary case), and also a projection onto EPR pairs

$$S_I = I \otimes I\,, \quad S_{\text{SWAP}}\,|i,j\rangle = |j,i\rangle\ , \quad S_{\text{EPR}} = \sum_{j=1}^{d} |j\rangle\langle j|\ . \tag{3.19}$$

Just like in the unitary case, Schur-Weyl duality for the orthogonal group means we can write the k-fold channel over $O(d)$ as

$$\Phi_O^{(k)}(\mathcal{O}) = \sum_{\sigma,\tau\in M_{2k}} \mathcal{W}g^O_{\sigma,\tau} S_\sigma \text{Tr}(S_\tau \mathcal{O})\,, \tag{3.20}$$

where here $\mathcal{W}g^O_{\sigma,\tau}$ is the Weingarten matrix for the orthogonal Weingarten functions. Taking the channel of the operator S_σ must give S_σ, and thus the Weingarten matrix is the inverse of the inner product of the basis elements S_σ, written as a matrix equation as

$$\mathcal{W}g^O_{\sigma,\tau} = \big(\text{Tr}(S_\sigma S_\tau)\big)^{-1}\,. \tag{3.21}$$

It is now simple to compute the frame potential Eq. (3.6) for the orthogonal ensemble. From the invariance of the Haar measure on $O(d)$, we find

$$\mathcal{F}_O^{(k)} = \int dO\, \text{Tr}(O^{\otimes k})\text{Tr}((O^{\otimes k})^\dagger) = \sum_{\sigma,\tau\in M_{2k}} \mathcal{W}g^O_{\sigma,\tau}\text{Tr}(S_\sigma S_\tau) = \frac{(2k)!}{2^k k!}\,, \tag{3.22}$$

for $k < d$. Viewed as a matrix equation, the inner product on S_σ's is just the inverse of the Weingarten function, and the frame potential is just equal to the number of elements in M_{2k}.

Random symplectic matrices

The compact sympectic group is a subgroup of the unitary group defined as the intersection of symplectic matrices $Sp(2d,\mathbb{C})$ and the unitary group $U(2d)$, defined with even dimension. It is a real Lie group and consists of the unitary matrices for which $U^D U = \mathbb{I}$. The symplectic

transpose is defined as

$$\mathcal{O}^D = J\mathcal{O}^T J^T\,, \quad \text{where} \quad J = \begin{pmatrix} 0 & \mathbb{I}_d \\ -\mathbb{I}_d & 0 \end{pmatrix}\,, \tag{3.23}$$

and $\mathbb{I}_d$ is a $d\times d$ identity matrix. Another way of expressing the symplectic unitary condition is as unitaries for which $UJU^T = J$, where we can also write $J = iY \otimes \mathbb{I}_d$. Some useful identities to keep in mind are $J^T J = \mathbb{I}$ and $J^2 = -\mathbb{I}$.

The expressions for integrating over symplectic matrices were discussed in [59, 88, 95], where the moments are given explicitly as

$$\int dS\, S_{i_1 j_1} \dots S_{i_{2k} j_{2k}} = \sum_{\sigma,\tau\in M_{2k}} \Delta'_\sigma(\vec{\imath})\Delta'_\tau(\vec{\jmath})\mathcal{W}g^{Sp}(\sigma^{-1}\tau, d)\,, \tag{3.24}$$

for symplectic unitaries S as defined above. In the k-th moment we sum over pair partitions M_{2k} and contract indices according to a J-graded combination δ-functions following the notation in [92]

$$\Delta'_\sigma(\vec{\imath}) = \delta'_{i_{\sigma(1)} i_{\sigma(2)}} \dots \delta'_{i_{\sigma(2k-1)} i_{\sigma(2k)}}\,, \tag{3.25}$$

where δ' is j-graded as $\delta'_{ij} = (J)_{i,j}$. For a more details and a discussion of the symplectic Weingarten functions, refer to the discussion in App. 3.A. We note the similarity with the orthogonal Haar integrals.

Schur-Weyl duality for the symplectic group is much the same as the orthogonal case, instead given by the action of the J-graded Brauer algebra and means that if an operator $\mathcal{O}$ commutes with all symplectic matrices $S^{\otimes k}$ on $\mathcal{H}^{\otimes k}$, then it can be written as

$$\text{Schur-Weyl} \quad \to \quad \mathcal{O} = \sum_{\sigma\in M_{2k}} c_\sigma S'_\sigma\,, \tag{3.26}$$

where S'_σ are the basis elements of the graded Brauer algebra. The elements can be written as simple permutation and projection operators on the k-copies of the Hilbert space realized just as the orthogonal elements but adding J when we contract the i indices and J^T when

we contract j indices. The k-fold channel over $Sp(2d)$ is

$$\Phi_{Sp}^{(k)}(\mathcal{O}) = \sum_{\sigma,\tau \in M_{2k}} \mathcal{W}g_{\sigma,\tau}^{Sp} S'_{\sigma} \mathrm{Tr}(S'_{\tau}\mathcal{O}) \,, \tag{3.27}$$

where here $\mathcal{W}g_{\sigma,\tau}^{Sp}$ is the symplectic Weingarten matrix, again the inverse of the inner product of the basis elements S'_{σ}, written as a matrix equation as $\mathcal{W}g_{\sigma,\tau}^{Sp} = \left(\mathrm{Tr}(S'_{\sigma}S'_{\tau})\right)^{-1}$. Just as in the orthogonal case, the frame potential for the unitary symplectic group counts the number of pair partitions, $\mathcal{F}_{Sp}^{(k)} = (2k)!/2^k k!$, for $k < d$.

3.3 Late-time OTOCs for symmetry classes

We are interested in the asymptotic form of the OTOCs in a chaotic system. Consider the OTO 4-point function

$$\langle AB(t)CD(t)\rangle \,, \tag{3.28}$$

where $B(t) = U_t B U_t^{\dagger}$ with the unitary time-evolution operator $U_t = e^{-iHt}$ generated by some physical Hamiltonian. For very chaotic systems, we expect that long after the system has scrambled, the OTOC will reach a small late-time value. Although in a physical system, the OTOC will fluctuate in time around this late-time floor, we should be able to study the universal aspects of this floor value. By taking the unitary operator U to be Haar-random and averaging the OTOC over the entire unitary group, we find a universal form of the late-time asymptotic value [11, 96].

For any quantum system, we always consider unitary dynamics $U_t \in U(d)$, i.e. we time evolve with elements of the unitary group. The Haar averaged correlators are universal in this sense; for generic chaotic systems our time-evolution moves unfettered over $U(d)$. But if our chaotic system is endowed with some symmetry, say time-reversal invariance, the unitary time-evolution operator will instead live in some subset of the unitary group. Thus, we must take a more fine-grained look at the universal properties. Within a given symmetry class we should instead average correlators over the subset to explore universal properties.

Haar random unitaries

For a chaotic system with no symmetries, we expect that the late-time value of the OTOC approaches its Haar-random value. We first review the Haar-averaged 2-point function and OTO 4-point function. The expressions for $U(d)$ averaged correlators have been previously derived and discussed in [11, 96]. The Haar-averaged 2-point function is

$$\langle AB(t)\rangle_{\text{Haar}} = \frac{1}{d}\int_{\text{Haar}} dU\, \text{Tr}(AUBU^\dagger) = \frac{1}{d^2}\text{Tr}(A)\text{Tr}(B) = \langle A\rangle\langle B\rangle\,, \tag{3.29}$$

given simply by the disconnected piece. We will see that the 2-point functions differentiate between some, but not all, of the different symmetry classes.

We can also compute the Haar-averaged OTO 4-point function using the second moment of $U(d)$, and find that

$$\begin{aligned}\langle AB(t)CD(t)\rangle_{\text{Haar}} &= \frac{1}{d}\int_{\text{Haar}} dU\, \text{Tr}(AUBU^\dagger CUDU^\dagger) \\ &= \frac{d^2}{d^2-1}\Big(\langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle - \frac{1}{d^2}\langle AC\rangle\langle BD\rangle - \langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle\Big)\,,\end{aligned} \tag{3.30}$$

or equivalently, written more compactly as

$$\begin{aligned}&\langle AB(t)CD(t)\rangle_{\text{Haar}} \\ &\qquad = \langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle - \langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle - \frac{1}{d^2-1}\langle AC\rangle_c\langle BD\rangle_c\,,\end{aligned} \tag{3.31}$$

where $\langle\cdot\rangle_c$ denotes the connected piece of the 2-point function, $\langle AB\rangle_c = \langle AB\rangle - \langle A\rangle\langle B\rangle$. Although it is clear from the above expression, for later comparison against forms of the OTOC for different symmetry classes we note that for traceless operators, the leading order contribution to the Haar-averaged 4-point function is

$$\langle AB(t)CD(t)\rangle_{\text{Haar}} \approx -\frac{1}{d^2}\langle AC\rangle\langle BD\rangle\,. \tag{3.32}$$

We can also compute the 6-point OTOC for Haar random unitaries. The leading order piece

is universal, meaning it contributes asymptotically for all symmetric classes. The explicit form is given in the appendix. It will be interesting to note some aspects of higher point functions. For traceless operators, the Haar-averaged 6-point OTOC is

$$\begin{aligned}\langle AB(t)CD(t)EF(t)\rangle_{\text{Haar}} = &\frac{(d^2-2)\langle AEC\rangle\langle BFD\rangle}{(d-1)(d-1)(d+1)(d+2)} + \frac{2\langle ACE\rangle\langle BFD\rangle}{(d-1)(d-1)(d+1)(d+2)} \\ &+ \frac{2\langle AEC\rangle\langle BDF\rangle}{(d-1)(d-1)(d+1)(d+2)} + \frac{2\langle ACE\rangle\langle BDF\rangle}{(d-1)(d-1)(d+1)(d+2)}\,.\end{aligned}$$

To leading order this gives

$$\langle AB(t)CD(t)EF(t)\rangle_{\text{Haar}} \approx \frac{\langle AEC\rangle\langle BFD\rangle}{d^2}\,. \tag{3.33}$$

We can also compute the Haar-averaged 8-point OTOCs, but refrain from reproducing them here. Again, we find that the leading order contribution is universal and holds asymptotically for all classes.

AI: $U(d)/O(d)$ Time-reversal symmetric

If our system has time-reversal symmetry, with $T^2 = 1$, then the unitary time-evolution operator is valued in $U \in U(d)/O(d)$. The universal form of the late-time correlators is found by averaging over the quotient space $U(d)/O(d)$. The AI averaged 2-point function is

$$\begin{aligned}\langle AB(t)\rangle &= \frac{1}{d}\int_{\text{AI}} dU\,\text{Tr}(AUBU^\dagger) = \frac{1}{d+1}\big(\langle AB^T\rangle + d\langle A\rangle\langle B\rangle\big) \\ &= \langle A\rangle\langle B\rangle + \frac{1}{d+1}\langle AB^T\rangle_c\,,\end{aligned} \tag{3.34}$$

which is already distinct from the Haar-averaged case. The OTO 4-point function is

$$\langle AB(t)CD(t)\rangle = \frac{1}{d}\int_{\text{AI}} dU\,\text{Tr}(AUBU^\dagger CUDU^\dagger)\,, \tag{3.35}$$

which we can integrate using the second moment of AI random unitaries

$$\int dU\,U_{ij}U_{mn}U^\dagger_{k\ell}U^\dagger_{pq} = \sum_{\sigma\in S_4}\delta_\sigma(ijmn|\ell kqp)\mathcal{W}g^{\text{AI}}(\sigma;d)\,. \tag{3.36}$$

Integrating and computing the AI Weingarten functions, we find

$$\begin{aligned}
\langle AB(t)CD(t)\rangle_{\text{AI}} = \\
&\frac{d(d+2)}{(d+1)(d+3)}\Big(\langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle + \frac{1}{d}\langle ACB^T\rangle\langle D\rangle + \frac{1}{d}\langle AD^TC\rangle\langle B\rangle \\
&\quad + \frac{1}{d}\langle AB^TD^T\rangle\langle C\rangle + \frac{1}{d}\langle A\rangle\langle B^TCD^T\rangle + \frac{1}{d^2}\langle AB^TCD^T\rangle + \frac{1}{d^2}\langle AD^TCB^T\rangle\Big) \\
-\frac{d}{(d+1)(d+3)}&\Big(d\langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle + \frac{1}{d}\langle AB^TC\rangle\langle D\rangle + \frac{1}{d}\langle ACD^T\rangle\langle B\rangle + \frac{1}{d}\langle AC\rangle\langle BD\rangle \\
&\quad + \frac{1}{d^2}\langle ACD^TB^T\rangle + \frac{1}{d^2}\langle ACB^TD^T\rangle + \frac{1}{d}\langle AB^T\rangle\langle CD^T\rangle + \frac{1}{d^2}\langle AB^TD^TC\rangle \\
&\quad + \langle AB^T\rangle\langle C\rangle\langle D\rangle + \langle A\rangle\langle B^TC\rangle\langle D\rangle + \frac{1}{d}\langle A\rangle\langle B^TD^TC\rangle + \frac{1}{d}\langle AD^T\rangle\langle B^TC\rangle \\
&\quad + \langle A\rangle\langle B\rangle\langle CD^T\rangle + \langle AD^T\rangle\langle B\rangle\langle C\rangle + \frac{1}{d^2}\langle AD^TB^TC\rangle + \frac{1}{d}\langle AD^TB^T\rangle\langle C\rangle\Big). \quad (3.37)
\end{aligned}$$

The important thing to note is that to leading order at large d, we find

$$\langle AB(t)CD(t)\rangle_{\text{AI}} = \langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle + \langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle + \frac{1}{d}(\ldots) \quad (3.38)$$

The same as the Haar-integrated OTOC, but with different terms suppressed in powers of d. We have yet to find an elegant form of the suppressed terms as connected components. The leading order contribution for traceless operators is

$$\begin{aligned}
\langle AB(t)CD(t)\rangle_{\text{AI}} = -\frac{1}{d^2}\big(&\langle AC\rangle\langle BD\rangle + \langle AD^T\rangle\langle BC^T\rangle + \langle AB^T\rangle\langle CD^T\rangle \\
&- \langle AB^TCD^T\rangle + \langle AD^TCB^T\rangle\big)\ldots \quad (3.39)
\end{aligned}$$

We see that the leading order term from the Haar-averaged OTOC appears but with four additional terms. We can also compute the 6-point OTOC, and interestingly, we find that the leading order contribution for traceless operators is

$$\langle AB(t)CD(t)EF(t)\rangle_{\text{AI}} = \frac{\langle AD^T\rangle\langle BE^T\rangle\langle CF^T\rangle}{d} + \ldots, \quad (3.40)$$

which contributes at a higher order than the Haar-averaged 6-point functions.

AII: $U(2d)/Sp(2d)$ Time-reversal symmetric

If our system has time-reversal symmetry, but instead where $T^2 = -1$, then the unitary time-evolution operator is valued in $V \in U(2d)/Sp(2d)$. The AII averaged 2-point function is

$$\langle AB(t)\rangle_{\rm AII} = \frac{1}{d}\int dV\, {\rm Tr}(AVBV^\dagger) = \langle A\rangle\langle B\rangle - \frac{1}{2d-1}\langle AB^D\rangle_c\,. \tag{3.41}$$

We can further compute the Haar-averaged OTO 4-pt function $\langle AB(t)CD(t)\rangle$ for AII by integrating $\tilde{V} = JV$ as

$$\int dU\, \tilde{V}_{i_1j_2}\tilde{V}_{i_3j_4}\tilde{V}^\dagger_{j_2j_1}\tilde{V}^\dagger_{j_4j_3} = \sum_{\sigma\in S_4}\delta_\sigma(i_1i_2i_3i_4|j_1j_2j_3j_4)\mathcal{W}g^{\rm AII}(\sigma;d)\,. \tag{3.42}$$

Integrating and computing the AII Weingarten functions, we find

$$\begin{aligned}
&\langle AB(t)CD(t)\rangle_{\rm AII} = \\
&\quad\frac{(2d)^2(d-1)}{d(2d-1)(2d-3)}\Big(\langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle + \frac{1}{2d}\langle ACB^D\rangle\langle D\rangle + \frac{1}{2d}\langle AD^DC\rangle\langle B\rangle \\
&\qquad + \frac{1}{2d}\langle AB^DD^D\rangle\langle C\rangle + \frac{1}{2d}\langle A\rangle\langle B^DCD^D\rangle + \frac{1}{4d^2}\langle AB^DCD^D\rangle + \frac{1}{4d^2}\langle AD^DCB^D\rangle\Big) \\
&\quad - \frac{2d}{(2d-1)(2d-3)}\Big(2d\langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle + \frac{1}{2d}\langle AB^DC\rangle\langle D\rangle + \frac{1}{2d}\langle ACD^D\rangle\langle B\rangle + \frac{1}{2d}\langle AC\rangle\langle BD\rangle \\
&\qquad - \frac{1}{4d^2}\langle ACD^DB^D\rangle - \frac{1}{4d^2}\langle ACB^DD^D\rangle + \frac{1}{2d}\langle AB^D\rangle\langle CD^D\rangle - \frac{1}{4d^2}\langle AB^DD^DC\rangle \\
&\qquad - \langle AB^D\rangle\langle C\rangle\langle D\rangle - \langle A\rangle\langle B^DC\rangle\langle D\rangle + \frac{1}{2d}\langle A\rangle\langle B^DD^DC\rangle + \frac{1}{2d}\langle AD^D\rangle\langle B^DC\rangle \\
&\qquad - \langle A\rangle\langle B\rangle\langle CD^D\rangle - \langle AD^D\rangle\langle B\rangle\langle C\rangle - \frac{1}{4d^2}\langle AD^DB^DC\rangle + \frac{1}{2d}\langle AD^DB^D\rangle\langle C\rangle\Big)\,.
\end{aligned} \tag{3.43}$$

Note that some of the subleading terms come with different signs than in the previous case and also instead of transpose (natural for orthogonal groups) we have the symplectic transpose $\mathcal{O}^D = J\mathcal{O}^TJ^T$. Again, we find that at large d

$$\langle AB(t)CD(t)\rangle_{\rm AII} = \langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle + \langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle + \frac{1}{d}(\ldots) \tag{3.44}$$

The same as the Haar-integrated OTOC, but with different terms suppressed in powers of d. The leading order contribution to the AII OTOC for traceless operators is

$$\langle AB(t)CD(t)\rangle_{\text{AII}} = -\frac{1}{4d^2}\big(\langle AC\rangle\langle BD\rangle + \langle AD^D\rangle\langle BC^D\rangle + \langle AB^D\rangle\langle CD^D\rangle \\ - \langle AB^DCD^D\rangle + \langle AD^DCB^D\rangle\big)\dots \tag{3.45}$$

Computing the 6-point AII averaged OTOC, we find that to leading order for traceless operators

$$\langle AB(t)CD(t)EF(t)\rangle_{\text{AII}} = -\frac{\langle AD^D\rangle\langle BE^D\rangle\langle CF^D\rangle}{2d} + \dots\,, \tag{3.46}$$

which just as in the AII case, contributes at order $1/d$ instead of at order $1/d^2$ as was the case for Haar-random unitaries.

BD: $O(d)$ random orthogonal

Averaging over the orthogonal group, we have

$$\langle AB(t)\rangle_O = \frac{1}{d}\int dO\,\text{Tr}(AOBO^T) = \langle A\rangle\langle B\rangle\,. \tag{3.47}$$

So interestingly, this is the first case where we cannot see the difference at the level of the 2-point function, but to probe this symmetry class, we must look at the OTOCs. The OTO 4-point function

$$\begin{aligned}\langle AB(t)CD(t)\rangle_O &= \frac{1}{d}\int dO\,\text{Tr}(AOBO^TCODO^T)\\ &= \frac{d(d+1)}{(d+2)(d-1)}\Big(\langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle + \frac{1}{d}\langle AC^T\rangle\langle BD^T\rangle\Big)\\ &\quad - \frac{d}{(d+2)(d-1)}\Big(d\langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle + \langle AC^T\rangle\langle B\rangle\langle D\rangle + \langle BD^T\rangle\langle A\rangle\langle C\rangle\\ &\qquad + \frac{1}{d}\langle AC\rangle\langle BD^T\rangle + \frac{1}{d}\langle AC^T\rangle\langle BD\rangle + \frac{1}{d}\langle AC\rangle\langle BD\rangle\Big)\,,\end{aligned} \tag{3.48}$$

again with the same leading order terms. We can express this more compactly as

$$\begin{aligned}\langle AB(t)CD(t)\rangle_O = \\ \langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle - \langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle + \frac{d+1}{(d+2)(d-1)}\langle AC^T\rangle_c\langle BD^T\rangle_c \\ - \frac{1}{(d+2)(d-1)}\big(\langle AC\rangle_c\langle BD^T\rangle_c + \langle AC^T\rangle_c\langle BD\rangle_c + \langle AC\rangle_c\langle BD\rangle_c\big)\,. \end{aligned} \tag{3.49}$$

For traceless operators and to order $1/d^2$, the orthogonal OTOC is

$$\langle AB(t)CD(t)\rangle_O = \frac{1}{d}\langle AC^T\rangle\langle BD^T\rangle - \frac{1}{d^2}\big(\langle AC\rangle\langle BD\rangle + \langle AC^T\rangle\langle BD\rangle + \langle AC\rangle\langle BD^T\rangle\big)\ldots \tag{3.50}$$

Note that the Haar, AI, and AII averaged OTOC all had the leading order contribution appear at order $1/d^2$, whereas here the contribution appears at order $1/d$.

We can also compute the 6-point OTOC for orthogonal evolution. To leading order for traceless operators we find

$$\begin{aligned}\langle AB(t)CD(t)E(t)F(t)\rangle_O = \frac{1}{d^2}\big(\langle AEC\rangle\langle BFD\rangle + \langle AC^TE^T\rangle\langle BFD^T\rangle \\ + \langle AE^TC\rangle\langle BD^TF^T\rangle + \langle AEC^T\rangle\langle BF^TD\rangle\big) + \ldots \end{aligned} \tag{3.51}$$

C: $Sp(2d)$ random symplectic

Averaging over the symplectic group, we find the averaged 2-point functions

$$\langle AB(t)\rangle_{Sp} = \frac{1}{2d}\int dS\,\mathrm{Tr}(ASBS^\dagger) = \langle A\rangle\langle B\rangle\,. \tag{3.52}$$

Again, to probe this symmetry class, we must look at the OTOCs. The OTO 4-point function

$$\begin{aligned}\langle AB(t)CD(t)\rangle_{Sp} &= \frac{1}{2d}\int dS\,\mathrm{Tr}(ASBS^\dagger CSDS^\dagger)\\ &= \frac{2d(2d-1)}{2(d-1)(2d+1)}\Big(\langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle - \frac{1}{2d}\langle AC^D\rangle\langle BD^D\rangle\Big)\\ &\quad - \frac{2d}{2(d-1)(2d+1)}\Big(2d\langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle - \langle AC^D\rangle\langle B\rangle\langle D\rangle - \langle BD^D\rangle\langle A\rangle\langle C\rangle\\ &\qquad + \frac{1}{2d}\langle AC\rangle\langle BD^D\rangle + \frac{1}{2d}\langle AC^D\rangle\langle BD\rangle + \frac{1}{2d}\langle AC\rangle\langle BD\rangle\Big)\,, \quad (3.53)\end{aligned}$$

again with the same leading order terms. We can express the OTOC more compactly as

$$\begin{aligned}&\langle AB(t)CD(t)\rangle_{Sp} =\\ &\quad \langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle - \langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle - \frac{2d-1}{2(d-1)(2d+1)}\langle AC^D\rangle_c\langle BD^D\rangle_c\\ &\quad - \frac{1}{2(d-1)(2d+1)}\big(\langle AC\rangle_c\langle BD^D\rangle_c + \langle AC^D\rangle_c\langle BD\rangle_c + \langle AC\rangle_c\langle BD\rangle_c\big)\,. \quad (3.54)\end{aligned}$$

For traceless operators and to order $1/d^2$, the symplectic OTOC is

$$\langle AB(t)CD(t)\rangle_{Sp} = -\frac{1}{2d}\langle AC^D\rangle\langle BD^D\rangle - \frac{1}{4d^2}\big(\langle AC\rangle\langle BD\rangle + \langle AC^D\rangle\langle BD\rangle + \langle AC\rangle\langle BD^D\rangle\big)\ldots \quad (3.55)$$

where again we find that the leading order contribution appears at order $1/d$. We can also compute the 6-point OTOC for symplectic matrices. To leading order for traceless operators we find

$$\begin{aligned}\langle AB(t)CD(t)E(t)F(t)\rangle_{Sp} = \frac{1}{4d^2}\big(&\langle AEC\rangle\langle BFD\rangle + \langle AC^DE^D\rangle\langle BFD^D\rangle\\ &+ \langle AE^DC\rangle\langle BD^DF^D\rangle + \langle AEC^D\rangle\langle BF^DD\rangle\big) + \ldots \quad (3.56)\end{aligned}$$

AIII: $U(d)/(U(a)\times U(b))$

This is the first of the so-called chiral ensembles, where $d = a+b$. Most generally, elements of AIII are defined as $I_{ab}U^\dagger I_{ab}U$, with Haar-random U and $I_{ab} = \mathrm{diag}(I_a, -I_b)$. For convenience,

we define $\omega = a - b$, the difference in dimensions. The 2-point function

$$\begin{aligned}\langle AB(t)\rangle_{\text{AIII}} &= \frac{1}{d}\int_{\text{AIII}} dW\, \text{Tr}(AWBW^\dagger) \\ &= \frac{(\omega+1)(\omega-1)}{(d+1)(d-1)}\langle A\tilde{B}\rangle + \frac{4ab}{d(d-1)(d+1)}\langle A\rangle\langle B\rangle \\ &= \langle A\rangle\langle B\rangle + \frac{\omega^2-1}{(d+1)(d-1)}\langle A\tilde{B}\rangle_c\,, \end{aligned} \tag{3.57}$$

where we have defined $\tilde{B} = I_{ab}BI_{ab}$ As a sanity check we can plug in the identity for the two operators and we get one. Similar to the previous cases where we sum over elements of the cyclic group S_k, the OTO 4-point function can be integrated using the AIII integral

$$\int dW\, W_{ij}W_{k\ell}W_{mn}W_{pq} = \sum_{\sigma\in S_4}\delta_\sigma(ikmp|j\ell nq)\mathcal{W}g^{\text{AIII}}(\sigma;a,b)\,. \tag{3.58}$$

We can compute the full expression for the AIII averaged OTOC, but the expression contains 24 terms and itself is not terribly enlightening. The important aspects to note are that to leading order in $1/d$, we find

$$\langle AB(t)CD(t)\rangle_{\text{AIII}} = \langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle + \langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle + \frac{1}{d}(\ldots)\,, \tag{3.59}$$

the same leading order contribution as in the other classes. For traceless operators, to leading order in $1/d$, the AIII OTOC is

$$\begin{aligned}\langle AB(t)CD(t)\rangle_{\text{AIII}} =& \frac{1}{d^2}\big(\langle A\tilde{D}C\tilde{B}\rangle - \langle AC\rangle\langle BD\rangle + (\omega^2-1)\langle A\tilde{B}\rangle\langle C\tilde{D}\rangle \\ &+ (\omega^2-1)\langle A\tilde{D}\rangle\langle C\tilde{B}\rangle\big) + \ldots \end{aligned} \tag{3.60}$$

We can also compute the 6-point OTOC and find that for traceless operators, the leading order contribution appears at $1/d^2$.

BDI: $O(d)/(O(a)\times O(b))$

The second of the chiral ensembles, where $d=a+b$. Generally, elements of BDI are defined as $I_{ab}O^TI_{ab}O$, with Haar-random orthogonal O. The 2-point function averaged over the space is

$$\begin{aligned}\langle AB(t)\rangle_{\rm BDI} &= \frac{1}{d}\int dW\,{\rm Tr}(AWBW^\dagger)\\ &= \frac{\omega^2(d+1)-2d}{d(d+2)(d-1)}\langle A\tilde{B}\rangle + \frac{4ab}{d(d+2)(d-1)}\big(d\langle A\rangle\langle B\rangle+\langle A\tilde{B}^T\rangle\big)\\ &= \langle A\rangle\langle B\rangle + \frac{\omega^2(d+1)-2d}{d(d+2)(d-1)}\langle A\tilde{B}\rangle_c + \frac{4ab}{d(d+2)(d-1)}\langle A\tilde{B}^T\rangle_c\,, \end{aligned}\tag{3.61}$$

where $\tilde{B}=I_{ab}BI_{ab}$ and $\omega=a-b$. Again we can check that insert identity operators gives one. Using the second moment of BDI random unitaries

$$\int dW\,W_{i_1i_2}W_{i_3i_4}W_{i_5i_6}W_{i_7i_8} = \sum_{\sigma\in M_8}\Delta_\sigma(\vec{\imath})\mathcal{W}g^{\rm BDI}(\sigma;a,b)\,,\tag{3.62}$$

we can compute the averaged OTO 4-point function and find to leading order

$$\langle AB(t)CD(t)\rangle_{\rm BDI} = \langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle + \langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle + \frac{1}{d}(\ldots)\,.\tag{3.63}$$

For traceless operators and to order $1/d^2$, the BDI averaged OTOC is

$$\begin{aligned}\langle AB(t)CD(t)\rangle_{\rm BDI} =& \frac{1}{d}\langle AC^T\rangle\langle BD^T\rangle - \frac{1}{d^2}\Big(\langle AC\rangle\langle BD\rangle + \langle AC^T\rangle\langle BD\rangle + \langle AC\rangle\langle BD^T\rangle\\ &+\langle A\tilde{B}^T\rangle\langle C\tilde{D}^T\rangle + \langle A\tilde{D}^T\rangle\langle C\tilde{B}^T\rangle + 2\langle AC^T\rangle\langle BD^T\rangle + 2\langle A\tilde{D}\rangle\langle B\tilde{C}\rangle\\ &+2\langle A\tilde{B}\rangle\langle C\tilde{D}\rangle - \omega^2\langle A\tilde{D}\rangle\langle C\tilde{B}\rangle - \omega^2\langle A\tilde{B}\rangle\langle C\tilde{D}\rangle - \langle A\tilde{D}C\tilde{B}\rangle\\ &-\langle A\tilde{B}^TC\tilde{D}^T\rangle - \langle A\tilde{D}^TC\tilde{B}^T\rangle - \langle AC^T\tilde{D}^T\tilde{B}\rangle - \langle A\tilde{D}\tilde{B}^TC^T\rangle\Big)\,.\end{aligned}\tag{3.64}$$

CII: $Sp(2d)/(Sp(2a)\times Sp(2b))$

The last of the chiral ensembles, again where $d=a+b$. Elements of this space can be written as $V=\tilde{I}_{ab}S^D\tilde{I}_{ab}S$, where S is a Haar-random symplectic matrix $\tilde{I}_{ab}={\rm diag}(I_{ab},I_{ab})$.

The averaged 2-point function is

$$\langle AB(t)\rangle_{\text{CII}} = \frac{1}{2d}\int dW\,\text{Tr}(AWBW^\dagger) \tag{3.65}$$

$$= \frac{\omega^2(2d-1)-d}{d(d-1)(2d+1)}\langle A\tilde{B}\rangle + \frac{4ab}{d(d-1)(2d+1)}\big(2d\langle A\rangle\langle B\rangle - \langle A\tilde{B}^D\rangle\big) \tag{3.66}$$

$$= \langle A\rangle\langle B\rangle + \frac{\omega^2(2d-1)-d}{d(d-1)(2d+1)}\langle A\tilde{B}\rangle_c + \frac{4ab}{d(d-1)(2d+1)}\langle A\tilde{B}^D\rangle_c\,, \tag{3.67}$$

where here $\tilde{B} = \tilde{I}_{ab}B\tilde{I}_{ab}$. At leading order in the CII averaged OTOC, we again find the same universal contribution. For traceless operators and to order $1/d^2$, the CII averaged OTOC is

$$\begin{aligned}\langle AB(t)CD(t)\rangle_{\text{CII}} =& -\frac{1}{2d}\langle AC^D\rangle\langle BD^T\rangle - \frac{1}{4d^2}\Big(\langle AC\rangle\langle BD\rangle + \langle AC^D\rangle\langle BD\rangle + \langle AC\rangle\langle BD^T\rangle \\ &+ \langle A\tilde{B}^D\rangle\langle C\tilde{D}^D\rangle + \langle A\tilde{D}^D\rangle\langle C\tilde{B}^D\rangle + 2\langle AC^D\rangle\langle BD^D\rangle + 2\langle A\tilde{D}\rangle\langle B\tilde{C}\rangle \\ &+ 2\langle A\tilde{B}\rangle\langle C\tilde{D}\rangle - \omega^2\langle A\tilde{D}\rangle\langle C\tilde{B}\rangle - \omega^2\langle A\tilde{B}\rangle\langle C\tilde{D}\rangle - \langle A\tilde{D}C\tilde{B}\rangle \\ &- \langle A\tilde{B}^DC\tilde{D}^T\rangle - \langle A\tilde{D}^DC\tilde{B}^D\rangle - \langle AC^D\tilde{D}^D\tilde{B}\rangle - \langle A\tilde{D}\tilde{B}^DC^D\rangle\Big)\,.\end{aligned} \tag{3.68}$$

DIII: $O(2d)/U(d)$

One of the BdG ensembles, where elements of the space can be written as $V = O^DO$ with Haar-random orthogonal matrix O. The 2-pt function is

$$\langle AB(t)\rangle_{\text{DIII}} = \frac{1}{2d}\int dV\,\text{Tr}(AVBV^\dagger) = \langle A\rangle\langle B\rangle - \frac{1}{2d-1}\langle AB^D\rangle_c\,. \tag{3.69}$$

The DIII averaged OTOC gives the same asymptotic contribution as in the other symmetry classes. To second order in $1/d$, we find that for traceless operators, the DIII OTOC is

$$\begin{aligned}\langle AB(t)CD(t)\rangle_{\text{DIII}} =& \frac{1}{2d}\langle AC^T\rangle\langle BD^T\rangle - \frac{1}{4d^2}\Big(\langle AC\rangle\langle BD\rangle + \langle AB^D\rangle\langle CD^D\rangle + \langle AD^D\rangle\langle CB^D\rangle \\ &- 2\langle AC^T\rangle\langle BD^T\rangle + 3\langle AC^T\rangle\langle BD\rangle + 3\langle AC\rangle\langle BD^T\rangle - \langle AD^DB^DC\rangle \\ &- \langle ACD^DB^D\rangle - \langle AB^DCD^D\rangle - \langle AD^DCB^D\rangle - \langle ADCB\rangle\Big)\,.\end{aligned} \tag{3.70}$$

CI: $Sp(2d)/U(d)$

Another of the BdG ensembles, where elements of the space can be written as $V = IS^D IS$ with Haar-random symplectic matrix S. The 2-pt function is

$$\langle AB(t)\rangle_{\text{CI}} = \frac{1}{2d}\int dV\, \text{Tr}(AVBV^D) = \langle A\rangle\langle B\rangle + \frac{1}{2d+1}\langle AIB^D I\rangle_c\,. \tag{3.71}$$

The CI averaged OTOC gives the same asymptotic contribution as in the other symmetry classes. To second order in $1/d$, we find that for traceless operators, the CI OTOC is

$$\begin{aligned}\langle AB(t)CD(t)\rangle_{\text{CI}} = &-\frac{1}{2d}\langle AC^D\rangle\langle BD^D\rangle - \frac{1}{4d^2}\Big(\langle AC\rangle\langle BD\rangle + \langle A\tilde{D}^D\rangle\langle C\tilde{B}^D\rangle + \langle A\tilde{B}^D\rangle\langle C\tilde{D}^D\rangle \\ &- 2\langle AC^D\rangle\langle BD^D\rangle + 3\langle AC^D\rangle\langle BD\rangle + 3\langle AC\rangle\langle BD^D\rangle - \langle A\tilde{D}C\tilde{B}\rangle \\ &+ \langle A\tilde{D}\tilde{B}^D C^D\rangle + \langle AC^D\tilde{D}^D\tilde{B}\rangle - \langle A\tilde{B}^D C\tilde{D}^D\rangle - \langle A\tilde{D}^D C\tilde{B}^D\rangle\Big)\,.\end{aligned} \tag{3.72}$$

3.4 Symmetric k-designs and k-invariance

In Chapter 2, we introduced the notion of k-invariance to capture the onset of a random matrix description in (an ensemble of) physical systems. Given the above discussion, one might expect that for a chaotic system with symmetry (such as time-reversal invariance), we should refine the definition to account for the symmetry.

Recall that the frame potential provides a quantitative measure of randomness of an ensemble of unitaries. Considering the ensemble of unitaries generated by an ensemble of Hamiltonians

$$\mathcal{E}_t = \left\{e^{-iHt}\,, \quad \text{with} \quad H \in \mathcal{E}_H\right\}, \tag{3.73}$$

where the ensemble of Hamiltonians might be a system with quenched disorder, like SYK or a disordered spin chain, or a random matrix ensemble. By computing the frame potential for this ensemble $\mathcal{F}^{(k)}_{\mathcal{E}}(t)$, a function of time, we quantified how chaotic or random the system is at certain timescales. If the k-th frame potential is equal to the frame potential for the invariant measure on the unitary group, the Haar ensemble, then we form a k-design, meaning we reproduce the first k moments of the unitary group. The frame potential is

lower-bounded by the Haar value and thus is a well-defined distance to randomness

$$\text{distance to } k\text{-design:} \qquad \mathcal{F}^{(k)}_{\mathcal{E}_t}(t) - \mathcal{F}^{(k)}_{\text{Haar}} \geq 0\,, \tag{3.74}$$

where $\mathcal{F}^{(k)}_{\text{Haar}} = k!$. In [12], we introduced a more interesting lower bound given by the notion of k-invariance. If the k-th frame potential is instead equal to the frame potential for the Haar-conjugated ensemble $\tilde{\mathcal{E}} = U\mathcal{E}U^\dagger$, which also lower bounds it, then we say the ensemble is k-invariant

$$\text{distance to } k\text{-invariance:} \qquad \mathcal{F}^{(k)}_{\mathcal{E}_t}(t) - \mathcal{F}^{(k)}_{\tilde{\mathcal{E}}_t}(t) \geq 0\,. \tag{3.75}$$

This captures something more interesting about the time-evolution of systems. The intuition should be that we are measuring how random the dynamics look by how invariant they are under a unitary change of basis. For example, if we consider the random matrix ensemble GUE, then as the ensemble is by definition unitarily invariant, then we are always exactly k-invariant for all k. If we consider a physical system, we expect that k-invariance is a good measure of how scrambled the system is under chaotic time-evolution.

In summary, the ensemble forms a k-design or is k-invariant iff

$$k\text{-design}: \quad \mathcal{F}^{(k)}_{\mathcal{E}_t}(t) = \mathcal{F}^{(k)}_{\text{Haar}}\,, \quad \text{and} \quad k\text{-invariant}: \quad \mathcal{F}^{(k)}_{\mathcal{E}_t}(t) = \mathcal{F}^{(k)}_{\tilde{\mathcal{E}}_t}(t)\,. \tag{3.76}$$

If a system is chaotic, then we expect the late-time floor value of the difference $\mathcal{F}^{(k)}_{\mathcal{E}_t}(t) - \mathcal{F}^{(k)}_{\tilde{\mathcal{E}}_t}(t)$ to be small. But just as we have seen in previous sections, the late-time behavior depends on the symmetries of the system. Now we ask how to refine the definition of a k-design and k-invariance for systems with symmetry.

Given some ensemble of Hamiltonians with symmetry

$$\mathcal{E}_t = \left\{e^{-iHt}\,, \quad H \in \mathcal{E}_H\right\}, \quad \text{where} \quad e^{-iHt} \in U(d)/K \tag{3.77}$$

and where K is some subgroup of $U(d)$ determined by the symmetry class. For example, $K = O(d)$ for some systems with time-reversal symmetry (the AI case). For a given quotient of the unitary group, a symmetric k-design is an ensemble that reproduces the first k-

moments of the compact symmetric space

$$\text{symmetric } k\text{-design:} \qquad \mathcal{F}^{(k)}_{\mathcal{E}_t}(t) - \mathcal{F}^{(k)}_{U(d)/K} \geq 0\,. \tag{3.78}$$

For example, if we wanted to understand the time-scales at which the random matrix ensemble GOE becomes randomly distributed over (a quotient of) the unitary group, we should compute $\mathcal{F}^{(k)}_{\mathcal{E}_t}(t) - \mathcal{F}^{(k)}_{U(d)/O(d)}$. We can also generalize the notion of k-invariance. Recall that before we defined the Haar-invariant ensemble $\tilde{\mathcal{E}}_t = U\mathcal{E}_t U^\dagger$, conjugating by Haar-random unitaries. For an ensemble of Hamiltonians with some symmetry, we should define the invariant ensemble with respect to the symmetry $\text{sym}(\mathcal{E}_t) = V\mathcal{E}V^\dagger$ where $V \in K$. The distance to symmetric k-invariance is thus

$$\text{symmetric } k\text{-invariance:} \qquad \mathcal{F}^{(k)}_{\mathcal{E}_t}(t) - \mathcal{F}^{(k)}_{\text{sym}(\mathcal{E}_t)}(t) \geq 0\,. \tag{3.79}$$

For example, if we have the GOE, then the invariant ensemble $\text{sym}(\mathcal{E}_{\text{GOE}}) = O\mathcal{E}_{\text{GOE}}O^\dagger = \mathcal{E}_{\text{GOE}}$, as the GOE is orthogonally invariant. This means that the GOE is always k-invariant with respect to its symmetry class, just as for the GUE. More generally, for an ensemble of physical Hamiltonians with symmetry we should look at the late-time value of its symmetric k-invariance to observe the onset of random matrix behavior in its symmetry class.

3.4.1 Frame potentials for symmetric k-designs

The frame potential for any ensemble of unitaries $\mathcal{E}$ is lower bounded by its Haar value $k!$, and in this sense defines a distance to randomness. More precisely, it is the 2-norm distance of the k-fold channels with respect to the ensemble $\mathcal{E}$ and the Haar ensemble. Simply stated, we can compute the k-th frame potential for some ensemble and the number we get tells us how close we are to reproducing the first k moments of the unitary group. As we discussed above, if our ensemble of unitaries has some symmetry, we should expect that the distance to randomness should be refined to account for the symmetry. For instance, if our system has time-reversal symmetry, the ensemble of unitary time-evolutions should reach a minimal value corresponding to the compact symmetric space.

We will first review the computation of the frame potential for the full unitary group $U(d)$ and then compute the frame potentials for the Lie groups $O(d)$ and $Sp(2d)$. Then we compute the frame potentials for some of the compact symmetric spaces $U(d)/K$. We then review the derivation of the lower bound by the Haar frame potential using tensor expanders, and then generalize to different symmetry classes.

Haar random unitaries

As we reviewed in Sec. 3.2, Schur-Weyl duality allows one to write the k-fold channel over the Haar ensemble as a sum over permutations with Weingarten coefficients. When writing out the frame potential we find Weingarten coefficents and their inverses, which count the cycles in an element of S_k, cancel such that the expression is simply a sum over elements of S_k, which gives

$$\mathcal{F}^{(k)}_{\text{Haar}} = k! \,. \tag{3.80}$$

Symmetric k-designs

We now generalize the notion of a k-design to different symmetry classes. Consider some ensemble of unitaries valued in a quotient of the unitary group

$$\mathcal{E} = \{p_i, U_i\} \quad \text{where} \quad U_i \in U(d)/K \,, \tag{3.81}$$

and where K is a closed subgroup of the full unitary group. As we discussed previously, the involution for which K is a fixed point set defines a map which allows us to realize the quotient space as a subset of the unitary group, this is essentially the Cartan embedding. An invariant probability measure on the compact symmetric space is induced from this map.

Analogous to the proof before that the frame potential for any ensemble of unitaries is lower-bounded by the Haar-value, we now show that any ensemble of unitaries valued in $U(d)/K$ is lowered bounded by the frame potential for the compact symmetric space. Consider the operator S, the difference between the k-fold channels of $\mathcal{E}$ and the invariant

ensemble for the compact symmetric space $U(d)/K$:

$$S = \int_{\mathcal{E}} dU \, (U^{\otimes k})^\dagger \otimes U^{\otimes k} - \int_{U(d)/K} dU \, (U^{\otimes k})^\dagger \otimes U^{\otimes k} \,. \tag{3.82}$$

As the operator $S^\dagger S$ is positive semi-definite $\mathrm{Tr}\, S^\dagger S \geq 0$, this implies that

$$\mathcal{F}^{(k)}_{\mathcal{E}} + \mathcal{F}^{(k)}_{U(d)/K} - 2 \int_{U \in \mathcal{E}} dU \int_{V \in U(d)/K} dV \left| \mathrm{Tr}(U^\dagger V) \right|^{2k} \geq 0 \,. \tag{3.83}$$

Now we note that as the elements of the ensemble $\mathcal{E}$ are in the compact symmetric space $U(d)/K$, then we can use the left/right invariance of the probability measure on the quotient to absorb the U's into the V's, and the last term simply becomes the frame potential for $U(d)/K$

$$\mathcal{F}^{(k)}_{\mathcal{E}} \geq \mathcal{F}^{(k)}_{U(d)/K} \quad \text{for} \quad \mathcal{E} = \{p_i, U_i \in U(d)/K\} \,. \tag{3.84}$$

Note that if the ensemble $\mathcal{E}$ were simply some ensemble of unitaries, not necessarily restricted to live within a quotient of the unitary group, then the above derivation would not hold. This is not surprising; all we have shown here is that the frame potential for the invariant measure on the quotient space lower bounds the frame potential for any ensemble consisting of unitaries from that quotient.

Frame potentials for AI

Recall that the frame potential is lower bounded by its Haar random unitary value $\mathcal{F}^{(k)}_{\mathrm{Haar}} = k!$. This was how we measured randomness, as a distance to reproducing the first k moments of the Haar ensemble. But if we are given some ensemble of unitary time evolutions respecting an additional symmetry, like time-reversal invariance, we should instead look at the ensembles ability to reproduce moments of a subset of the full unitary group, with respect to that symmetry. Namely, if the measure is orthogonally invariant, as with certain Z_2^T invariant systems, then we should quantify randomness with respect to the frame potential of quotient space $U(d)/O(d)$.

Consider the frame potential for the compact symmetric space AI

$$\mathcal{F}^{(k)}_{\mathrm{AI}} = \int_{U(d)/O(d)} dU dV \left|\mathrm{Tr}(U^\dagger V)\right|^{2k} = \int_{U(d)/O(d)} dU \left|\mathrm{Tr}(U)\right|^{2k}. \tag{3.85}$$

To compute first frame potential we sum over elements of S_2, for which we need the Weingarten function $\mathcal{W}g^{\mathrm{AI}}(\sigma \in S_2, d) = 1/(d+1)$, and thus find

$$\mathcal{F}^{(1)}_{\mathrm{AI}} = \frac{2d}{d+1}, \quad \text{and} \quad \mathcal{F}^{(1)}_{\mathrm{AI}} \approx 2 \quad \text{for} \quad d \gg 1. \tag{3.86}$$

The first thing to note is that unlike for the compact Lie groups, the frame potentials for the compact symmetric spaces depend on the dimension d. Schur-Weyl duality for the compact Lie groups allowed us to write the k-fold channels as linear combinations of a set of basis elements (which depended on the group). This meant that the Weingarten function arose as the inverse of the inner products of these basis elements. For the compact symmetric space, this is no longer the case. For instance, the AI Weingarten function is a product of the unitary Weingarten matrix and the inner product of orthogonal basis elements [91, 92]. Said in the language of these papers convolution on S_{2k} of the unitary Weingarten function and the class function for the orthogonal group (of which the orthogonal Weingarten function is the inverse). So the AI frame potentials will not have an exact closed form in terms of k, but what we are really interested in is the asymptotic value.

To compute the second frame potential for AI, we sum over elements of S_4 for which we need the Weingarten functions

$$\mathcal{W}g^{\mathrm{AI}}(\{1,1\}, d) = \frac{d+2}{d(d+1)(d+3)}, \qquad \mathcal{W}g^{\mathrm{AI}}(\{2\}, d) = \frac{-1}{d(d+1)(d+3)}, \tag{3.87}$$

where the argument of the Weingarten function is the coset type of the permutation, as reviewed in App. 3.A. Computing the second frame potential we find

$$\mathcal{F}^{(2)}_{\mathrm{AI}} = \frac{8(d^2+2d-2)}{(d+1)(d+3)}, \quad \text{and} \quad \mathcal{F}^{(2)}_{\mathrm{AI}} \approx 8 \quad \text{for} \quad d \gg 1. \tag{3.88}$$

We can now discuss the general form of the k-th frame potential. As we explain in App. 3.A,

just like for the orthogonal Weingarten functions, the AI Weingarten functions are constant on the double coset of S_{2k}, which means they are labeled by the coset-type of the permutation. The coset-type of a permutation $\sigma \in S_{2k}$ is an integer partition of k. The leading term in the large d limit will be given by the Weingarten function of coset-type $\{1, 1, \ldots\}$. Given the simple relation between orthogonal and AI functions, this arises from the asymptotic form of $\mathcal{W}g^O$ [59]

$$\mathcal{W}g^O(\sigma, d) \sim \frac{1}{d^{2k-\ell(\mu)}}\,, \tag{3.89}$$

where $\sigma \in S_{2k}$, μ is the coset-type of σ, and $\ell(\mu)$ is the length of the integer partition μ. As we sum over elements of S_{2k} in the AI Haar-integral (not pair partitions as in the orthogonal case), the asymptotic form of the frame potential is given by counting the number of elements of S_{2k} of coset-type $\{1, 1, \ldots\}$, which gives

$$\mathcal{F}^{(k)}_{\text{AI}} \approx 2^k k! \quad \text{for} \quad d \gg 1\,. \tag{3.90}$$

Thus, given that the frame potential of a compact symmetric space is lower bounded by its Haar-value Eq. (3.84), any ensemble of unitaries in $U(d)/O(d)$ forms an AI symmetric k-design if and only if its frame potential equals $2^k k!$.

Frame potentials for AII

The discussion here is essentially the same as the previous subsection. We now want to consider the frame potential for the compact symmetric space AII

$$\mathcal{F}^{(k)}_{\text{AII}} = \int_{U(2d)/Sp(2d)} dU dV \left|\text{Tr}(U^\dagger V)\right|^{2k} = \int_{U(2d)/Sp(2d)} dU \left|\text{Tr}(U)\right|^{2k}. \tag{3.91}$$

We will just compute the first frame potential before turning to the more general discussion. For the first frame potential we need the Weingarten function $\mathcal{W}g^{\text{AII}}(\sigma \in S_2, d) = \epsilon(\sigma)/(2d-1)$, which depends on the sign of the permutation. Proceeding we find

$$\mathcal{F}^{(1)}_{\text{AII}} = \frac{4d}{2d-1}\,, \quad \text{and} \quad \mathcal{F}^{(1)}_{\text{AII}} \approx 2 \quad \text{for} \quad d \gg 1\,, \tag{3.92}$$

again depending on the dimension d.

More generally, although the exact values of the AII frame potentials are different from AI, the asymptotic form is the same. Just like for the symplectic Weingarten functions, the AII Weingarten functions are also constant on the double coset of S_{2k} and are labeled by the coset-type of σ. The asymptotic form of $\mathcal{W}g^{Sp}$ is the same as the orthogonal Weingarten functions [59], meaning the leading order contribution will arise from permutations of coset-type $\{1, 1, \ldots\}$. Summing over elements of S_{2k} in computing the AII frame potential we simply count the number of permutations of this coset-type, which gives

$$\mathcal{F}^{(k)}_{\text{AII}} \approx 2^k k! \quad \text{for} \quad d \gg 1\,. \tag{3.93}$$

As Eq. (3.84) shows that frame potential of a compact symmetric space is lower bounded by its Haar-value, any ensemble of unitaries in $U(2d)/Sp(2d)$ forms a AII symmetric k-design if and only if its frame potential equals $2^k k!$.

3.4.2 Random matrices and symmetry

In Sec. 3.4, we generalized the notion of a k-design to ensembles with symmetry. Let's now turn to random matrix ensembles with symmetry and determine the time scales at which the evolution forms a k-design. In [12], we computed the frame potentials for the Gaussian unitary ensemble and found that at an intermediate time scale, the dip time $t_d \sim \sqrt{d}$, the GUE forms a k-design and reproduces $k \ll d$ moments of the Haar ensemble, but at late-times the frame potential increases and the GUE is no longer Haar-random. We refer the reader to [12] for details. For comparison to other random matrix ensembles, we quickly review the result for the $k = 1$ frame potential of the GUE. The k-th frame potential for the GUE can be computed by going to the unitarily rotating to the eigenvalue basis and then Haar-integrating, resulting in an expression in terms of the $2k$-spectral form factors. For $k = 1$ we found

$$\mathcal{F}^{(1)}_{\text{GUE}} = \frac{1}{d^2-1}\left(\mathcal{R}_2^2 + d^2 - 2\mathcal{R}_2\right), \quad \text{where} \quad \mathcal{R}_2 = \int D\lambda \sum_{i,j} e^{i(\lambda_i - \lambda_j)t} \tag{3.94}$$

is the 2-point form factor for the GUE. At very early times, $\mathcal{R}_2 \sim d^2$ so the behavior of the frame potential is approximately $\sim \mathcal{R}_2^2/d^2$ until near the dip time. At the dip time t_d, $\mathcal{R}_2 \approx \sqrt{d}$ and the frame potential achieves its Haar-random value $\mathcal{F}^{(1)}_{\text{GUE}} \approx 1$. But at late times $\mathcal{R}_2 \approx d$ and the frame potential becomes $\mathcal{F}^{(1)}_{\text{GUE}} \approx 2$, which means the GUE no longer forms a 1-design. More generally, for the k-th frame potential [12] found that

$$\text{Early} : \mathcal{F}^{(k)}_{\text{GUE}} \approx \frac{\mathcal{R}_{2k}^2}{d^{2k}}\,, \qquad \text{Dip} : \mathcal{F}^{(k)}_{\text{GUE}} \approx k!\,, \qquad \text{Late} : \mathcal{F}^{(k)}_{\text{GUE}} \approx \sum_{\ell=0}^{k} \frac{k!^2}{\ell!}\,. \tag{3.95}$$

Frame potentials for the GOE

Now we consider the Gaussian orthogonal ensemble and the ensemble of time-evolutions generated by GOE Hamiltonians

$$\mathcal{E}_t = \left\{ e^{-iHt}\,,\ \text{where } H \in \text{GOE} \right\}. \tag{3.96}$$

The GOE is the random matrix ensemble for time-reversal invariant Hamiltonians with $T^2 = 1$. For electronic systems, these are Hamiltonians with time-reversal invariance and preserving $SU(2)$ spin. In the language of this chapter, these are systems in the symmetry class AI; H is a real Hermitian matrix and $e^{-iHt} \in U(d)/O(d)$. The symmetric k-design condition for this symmetry class is

$$\text{AI } k\text{-design} : \quad \mathcal{F}^{(k)}_{\mathcal{E}_t}(t) \geq \mathcal{F}^{(k)}_{U(d)/O(d)}\,. \tag{3.97}$$

Recall that we also showed that for the compact symmetric space $U(d)/O(d)$, the frame potential is

$$\mathcal{F}^{(k)}_{U(d)/O(d)} \approx 2^k k! \quad \text{for} \quad d \gg 1 \quad \text{and} \quad d > k\,, \tag{3.98}$$

where the first few terms of this sequence are $2, 8, 48$, and 384.

We now want to compute the frame potentials for the GOE and show that at the dip time the GOE forms an AI k-design. We give an explicit derivation of the first GOE frame potential and then quote our results for higher k.

$k = 1$ **frame potential**

The first frame potential for the GOE

$$\mathcal{F}^{(1)}_{\text{GOE}} = \int dH_1 dH_2 \, e^{-\frac{d}{2}\text{Tr}H_1^2} e^{-\frac{d}{2}\text{Tr}H_2^2} \left| \text{Tr}\left(e^{iH_1 t} e^{-iH_2 t}\right) \right|^2 , \tag{3.99}$$

can be computed by noting that the GOE measure is invariant under conjugation by orthogonal matrices, meaning we can diagonalize the Hermitian matrices integrating over the invariant measure on the orthogonal group. For convenience, we define $\Lambda \equiv U e^{-iHt} U^\dagger$, i.e. the diagonal matrix exponential of the GOE Hamiltonian. Thus in the eigenvalue basis, we can express the frame potential as

$$\mathcal{F}^{(1)}_{\text{GOE}} = \int D\lambda_1 D\lambda_2 \int dO \, \text{Tr}\left(O^T \Lambda_1^\dagger O \Lambda_2\right) \text{Tr}\left(\Lambda_2^\dagger O^T \Lambda_1 O\right) , \tag{3.100}$$

where we have used the left/right invariance of the Haar measure. To integrate this we need the second moment of the orthogonal group $O(d)$, as we reviewed in Sec. 3.2

$$\int dO \, O_{j_1 k_1} O_{j_2 k_2} O_{j_3 k_3} O_{j_4 k_4} = \sum_{\sigma,\tau \in M_4} \delta_{j_{\sigma(1)} j_{\sigma(2)}} \delta_{j_{\sigma(3)} j_{\sigma(4)}} \delta_{k_{\tau(1)} k_{\tau(2)}} \delta_{k_{\tau(3)} k_{\tau(4)}} \mathcal{W}g^O(\sigma^{-1}\tau) , \tag{3.101}$$

where we sum over the pair partitions M_4. The three elements of M_4 are $\{\{1,2\},\{3,4\}\}$, $\{\{1,4\},\{2,3\}\}$, and $\{\{1,3\},\{2,4\}\}$. The two Weingarten functions for the orthogonal group we will need are

$$\mathcal{W}g^O(\{1,1\}) = \frac{d+1}{d(d+2)(d-1)} , \quad \text{and} \quad \mathcal{W}g^O(\{2\}) = \frac{-1}{d(d+2)(d-1)} , \tag{3.102}$$

labeled by the coset type of the permutation. Recall that $\mathcal{W}g^O(\sigma)$ is a function on $\sigma \in S_{2k}$ and is constant on the double coset of S_{2k}, meaning that the Weingarten function only depends on the coset type of σ. In the above sum, the two cases correspond to whether the pair partitions in the sum are equal or not, i.e. we use $\mathcal{W}g^O(\{1,1\})$ when $\sigma = \tau$ in the sum, and $\mathcal{W}g^O(\{2\})$ when $\sigma \neq \tau$. Summing over the pair partitions and doing the δ-function

contractions, we find

$$\mathcal{F}^{(1)}_{\text{GOE}} = \frac{1}{d(d+2)(d-1)}\big((d+1)\mathcal{R}_2^2 + 2d^3 - 4d\mathcal{R}_2\big)\,, \tag{3.103}$$

where again $\mathcal{R}_2$ is the 2-point form factor. As a sanity check, we can evaluate at $t = 0$, where $\mathcal{R}_2 = d^2$, and find $\mathcal{F}^{(1)}_{\text{GOE}} = d^2$. Just like in the GUE, we see that the early time behavior of the first frame potential is dominated by the 2-point spectral form factor for the GOE

$$\text{Early}: \quad \mathcal{F}^{(1)}_{\text{GOE}} \approx \frac{\mathcal{R}_2^2}{d^2}\,. \tag{3.104}$$

At the dip time, when $\mathcal{R}_2 \approx \sqrt{d}$, we find that

$$\text{Dip}: \quad \mathcal{F}^{(1)}_{\text{GOE}} \approx 2\,. \tag{3.105}$$

At late times, in the limit $t \to \infty$ where $\mathcal{R}_2 = d$, we find

$$\text{Late}: \quad \mathcal{F}^{(1)}_{\text{GOE}} = \frac{3d^2}{d(d+2)} \approx 3 \quad \text{for} \quad d \gg 1\,. \tag{3.106}$$

$k = 2$ frame potential

We can compute the second frame potential for the GOE using the fourth moment of the orthogonal group, integrating and computing the Weingarten functions we find an expression in terms of the spectral form factors for the GOE. We give the first few terms of the expression:

$$\mathcal{F}^{(1)}_{\text{GOE}} = 8 + \frac{\mathcal{R}_4^2}{d^4} + \frac{12\mathcal{R}_4^2}{d^6} + \frac{16\mathcal{R}_4}{d^4} + \frac{8\mathcal{R}_2^2}{d^2} - \frac{8\mathcal{R}_2^2}{d^3} - \frac{32\mathcal{R}_2}{d^2} - \frac{16\mathcal{R}_4\mathcal{R}_2}{d^4} + \dots \tag{3.107}$$

Again we note that at early times, the second frame potential is dominated by the 4-point form factor, but at the dip time we find that the frame potential reaches a value of 8 and forms an AI symmetric 2-design, but not a unitary 2-design. At late times, we deviate away from this and get the asymptotic value 20 in the $t \to \infty$ limit and for $d \gg 1$. In summary,

for the second frame potential for the GOE, we find

$$\text{Early} : \mathcal{F}^{(2)}_{\text{GOE}} \approx \frac{\mathcal{R}_4^2}{d^4}\,, \qquad \text{Dip} : \mathcal{F}^{(2)}_{\text{GOE}} \approx 8\,, \qquad \text{Late} : \mathcal{F}^{(2)}_{\text{GOE}} \approx 20\,. \tag{3.108}$$

k-th frame potential

We can now comment on the general features for the k-th frame potential for the GOE. Just as for the unitary Weingarten functions, the orthogonal Weingarten functions have an asymptotic form at large d

$$\mathcal{W}g^O(\sigma) \sim \frac{1}{d^{4k-\#\text{cycles}}}\,, \tag{3.109}$$

in terms of the number of cycles in the permutation σ. Note that the k-th frame potential involves the $2k$-th moment of the orthogonal group, hence the $4k$ factor in the denominator. This tells us that the leading order terms will be for Weingarten function evaluated on the identity permutation, i.e. when $\sigma = \tau$ in the Haar-integration. Recall that we sum over the subset of S_{4k} consisting of pair partitions. For the k-th frame potential, which requires the $2k$-th moment, the number of pair partitions is $(4k)!/(4^k(2k)!)$. It turns out that the subset of these elements on the double coset labeled by the integer partition $2k$, which generate the index contractions and give the leading order contribution, contains $2^k k!$ elements, which are the terms that survive at the dip of k-th frame potential for the GOE. The early time decay and the dip value for the GOE frame potentials are

$$\text{Early} : \mathcal{F}^{(k)}_{\text{GOE}} \approx \frac{\mathcal{R}_{2k}^2}{d^{2k}}\,, \qquad \text{Dip} : \mathcal{F}^{(k)}_{\text{GOE}} \approx 2^k k!\,, \tag{3.110}$$

for $d \gg 1$. This means that at the dip time, the GOE forms an AI symmetric k-design, just as the GUE forms a unitary k-design at the dip. Moreover, as was also the case for the GUE, at late times the random matrix eigenvalues dephase and we are no longer Haar-random with respect to the symmetry class.

Frame potentials for the GSE

The discussion here is essentially the same as for the GOE, so we will keep the discussion somewhat brief. Consider the ensemble of unitaries generated by the time evolution of GSE

Hamiltonians, i.e. where H is drawn from GSE

$$\mathcal{E} = \left\{ e^{-iHt}, \quad \text{where} \quad H \in \text{GSE} \right\}. \tag{3.111}$$

Recall that we define the GSE to consist of $2d \times 2d$ Hamiltonians, and thus $2d$ is the dimension of the Hilbert space in the following discussion.

We can compute the first frame potential noting that the GSE measure is invariant under conjugation by symplectic matrices, meaning we can diagonalize the Hermitian matrices integrating over the invariant measure on the symplectic group. In the eigenvalue basis, we can express the frame potential as

$$\mathcal{F}^{(1)}_{\text{GSE}} = \int D\lambda_1 D\lambda_2 \int dS \, \text{Tr}\big(U^D \Lambda_1^\dagger U \Lambda_2\big) \text{Tr}\big(\Lambda_2^\dagger U^D \Lambda_1 U\big) \,, \tag{3.112}$$

where U^D is the symplectic adjoint of U, and we've used the left/right invariance of the Haar measure. To integrate, we use the moments of the Haar measure on the symplectic group. The formulae are essentially the same as in the last subsection. To integrate we need the second moment of $Sp(2d)$

$$\int dU \, U^D_{j_1 k_1} U_{j_2 k_2} U^D_{j_3 k_3} U_{j_4 k_4} = \sum_{\sigma,\tau \in P_4} \delta_{j_{\sigma(1)} j_{\sigma(2)}} \delta_{j_{\sigma(3)} j_{\sigma(4)}} \delta_{k_{\tau(1)} k_{\tau(2)}} \delta_{k_{\tau(3)} k_{\tau(4)}} \mathcal{W}g^{Sp}(\sigma^{-1}\tau) \,. \tag{3.113}$$

The two symplectic Weingarten functions we will need are

$$\mathcal{W}g^{Sp}(\{1,1\}) = \epsilon(\sigma) \frac{2d-1}{2d(2d-2)(2d+1)} \,, \qquad \mathcal{W}g^{Sp}(\{2\}) = \epsilon(\sigma) \frac{1}{2d(2d-2)(2d+1)} \,. \tag{3.114}$$

Summing over the pair partitions and doing the δ-function contractions, we find

$$\mathcal{F}^{(1)}_{\text{GSE}} = \frac{1}{2d(2d-2)(2d+1)} \big((2d-1)\mathcal{R}_2^2 + 2(2d)^3 - 8d\mathcal{R}_2 \big) \,, \tag{3.115}$$

where here $\mathcal{R}_2$ is the spectral form factor for the GSE. Again at the dip $\mathcal{F}^{(1)}_{\text{GSE}} \approx 2$ and we form a symmetric AII 1-design, but not a unitary 1-design. At late times $t \to \infty$, we find $\mathcal{F}^{(1)}_{\text{GSE}} \approx 3$, and deviate away from Haar-randomness. While the exact expressions and

subleading corrections are different from the GOE, and the actual expressions for the GOE and GSE spectral form factors are different, the essential aspects of the discussion are the same. At the dip time, which is the same time-scale for GSE form factors, we find the asymptotic value of $2^k k!$, which is also the symmetric k-design value for the symmetry class AII. In summary, just as for the GOE, the early time decay and the dip value for the GSE frame potentials are

$$\text{Early} : \mathcal{F}^{(k)}_{\text{GSE}} \approx \frac{\mathcal{R}^2_{2k}}{(2d)^{2k}}\,, \qquad \text{Dip} : \mathcal{F}^{(k)}_{\text{GSE}} \approx 2^k k!\,, \tag{3.116}$$

meaning we form a AII symmetric k-design and then at late-times become less Haar-random with respect to the AII compact symmetric space.

3.5 Random subsystems and symmetry

In this section we study the reduced density matrix of a random state with respect to different symmetry classes. We are interested in the trace distance of $\rho_A(U)$, the reduced density matrix of a subsystem in a random state, to the maximally mixed state. Simply stated, we want to understand Page's theorem for different symmetry classes. This section differs slightly from the goals of the previous sections, where we studied the time-evolution and late-time features of chaotic systems with symmetry. Studying random subsystems for different symmetry classes is a nice application of the tools developed in this chapter, but will also be physically interesting.

Haar-random unitary states

Consider a bipartite system $\mathcal{H} = \mathcal{H}_A \otimes \mathcal{H}_B$, a Haar-random state $U|\psi_0\rangle$, and its reduced density matrix on A:

$$\rho_A(U) = \text{Tr}_B U\rho_0 U^\dagger\,, \quad \text{where} \quad \rho_0 = |\psi_0\rangle\langle\psi_0|\,. \tag{3.117}$$

Denote the dimensions $d = \dim \mathcal{H}$, $d_A = \dim \mathcal{H}_A$ and $d_B = \dim \mathcal{H}_B$, and define the difference with the maximally mixed state as $\Delta\rho_A \equiv \rho_A(U) - \rho_A^\infty$, where $\rho_A^\infty = \mathbb{I}_A/d_A$ is the maximally

mixed state on the A subsystem. Recall that we can upper bound the 1-norm distance $\int dU \, \|\Delta\rho_A\|_1$ by computing the Haar-average of the squared 2-norm distance

$$\int dU \, \|\Delta\rho_A\|_2^2 = \int dU \, \mathrm{Tr}\Big(\rho_A^2(U) - \frac{2}{d_A}\rho_A(U) + \frac{\mathbb{I}_A}{d_A^2}\Big) \,. \tag{3.118}$$

The last two terms are always trivially Haar-averaged, even if we are averaging over some symmetry class and not the full unitary group; the non-trivial piece is always the $\mathrm{Tr}\rho_A^2(U)$ term. Note, this is the same computation done to show decoupling $\rho_{RE} = \rho_R \otimes \rho_E$ in [1].

Recalling that for any operator $\mathcal{O}$ acting on a Hilbert space $\mathcal{H}_A$, the 2-norm bounds the 1-norm as $\|\mathcal{O}\|_1^2 \leq d_A \|M\|_2^2$, we can bound the trace distance as

$$\int dU \, \big\|\Delta\rho_A\big\|_1^2 \leq d_A \int dU \, \mathrm{Tr}\rho_A^2(U) - 1 \,. \tag{3.119}$$

For Haar-random unitaries, it is simple to compute

$$\int_{\mathrm{Haar}} dU \, \mathrm{Tr}\rho_A^2(U) = \frac{d_A + d_B}{d+1} \,, \tag{3.120}$$

and thus we find that

$$\int dU \, \big\|\Delta\rho_A\big\|_1^2 \leq \frac{d_A^2 - 1}{d+1} \approx \frac{d_A}{d_B} \,, \tag{3.121}$$

meaning that if the A subsystem is smaller than half the total system then typical states on that subsystem look very close to maximally mixed.

Random orthogonal states

It is straight-forward to extend this calculation to other symmetry classes and ask if the same is true for random states with respect to some quotient of the full unitary group. Obviously, we expect it to be true and that typical states should look maximally mixed on subsystems, but we want to know what the corrections are.

We now consider a Haar-random state $O\,|\psi_0\rangle$, with respect to the orthogonal group $O(d)$.

All we need to compute is the average of $\mathrm{Tr}\rho_A^2(O)$

$$\int_{O(d)} dO\, \mathrm{Tr}\rho_A^2(O) = \frac{d_A + d_B + 1}{d+2}\,, \tag{3.122}$$

which is almost identical to the Haar-random case. We find

$$\int dO\, \left\|\Delta\rho_A\right\|_1^2 \leq \frac{d_A^2 + d_A - 2}{d+2} \approx \frac{d_A}{d_B}\,, \tag{3.123}$$

with leading corrections that appear distinct in the orthogonal case and go like $1/d_B$ and $-2/d$.

Random AI states

We now consider a Haar-random state $U\left|\psi_0\right\rangle$, where U is drawn randomly from the quotient $U(d)/O(d)$. Again, we compute the average of $\mathrm{Tr}\rho_A^2(U)$

$$\int_{\mathrm{AI}} dU\, \mathrm{Tr}\rho_A^2(U) = \frac{(d^2+3d-2)(d_A+d_B)}{d(d+1)(d+3)} + \frac{2(d+2)\mathrm{Tr}(\rho_A^2) - 2\mathrm{Tr}(\rho_B^2)}{d(d+1)(d+3)}\,, \tag{3.124}$$

which is a little more non-trivial than the Haar case. Note that the purity of the reduced density matrix on the A and B subsystems in the initial state $\left|\psi_0\right\rangle$ appears in the $1/d_B^2$ corrections. Again the leading order term gives $\int dU\, \|\Delta\rho_A\|_1 \lesssim d_A/d_B$, and leading corrections which go like $1/d$.

Note that in this case, we have assumed that the initial state is pure and so it transposes to itself, but if we assume otherwise, the expression of the averaged density matrix on the subsystem involves transposed density matrices.

Random symplectic states

We now consider a Haar-random state $S\left|\psi_0\right\rangle$, where S is drawn Haar-randomly from the symplectic group $Sp(2d)$. Again, we compute the average of $\mathrm{Tr}\rho_A^2(S)$

$$\int_{Sp(2d)} dS\, \mathrm{Tr}\rho_A^2(S) = \frac{d_A + d_B - 1)(2d-1) + 2}{2(d-1)(2d+1)}\,. \tag{3.125}$$

Again the leading order term gives $\int dU \, \|\Delta\rho_A\|_1 \lesssim d_A/d_B$, with leading corrections that go like $1/b$ and $a/2d$, differing from the above cases at order $1/d_B^2$. We have assumed that the initial state is pure, but if we start with a mixed state, the expression of the averaged density matrix on the subsystem involves ρ^D.

Random AII states

We now consider a Haar-random state $U|\psi_0\rangle$, where U is drawn randomly from the quotient $U(2d)/Sp(2d)$. Again, we compute the average of $\mathrm{Tr}\rho_A^2(U)$

$$\int_{\mathrm{AII}} dU \, \mathrm{Tr}\rho_A^2(U) = \frac{(d-1)(d_A+d_B)}{d(2d-1)} + \frac{2(d-1)\mathrm{Tr}(\rho_A^2) - \mathrm{Tr}(\rho_B^2)}{d(2d-1)(2d-3)} \,. \tag{3.126}$$

The leading order term gives $\int dU \, \|\Delta\rho_A\|_1 \lesssim d_A/d_B$, with $1/d$ corrections and the purity of the subsystems in the initial state appearing at order $1/d_B^2$. Again, we have assumed that the initial state is pure, but if we start with a mixed state, the expression of the averaged density matrix on the subsystem involves ρ^D.

3.A An overview of Weingarten calculus

In this appendix, we review Weingarten calculus and the formalism for averaging over the unitary group, as well as orthogonal and symplectic groups. Then discuss averaging over compact symmetric spaces, quotients of Lie groups.

3.A.1 Integrating over compact Lie groups

Integration over the unitary group

As was given in Sec. 3.2, the expression for integrating over Haar-random unitaries is [58, 59]

$$\int dU \, U_{i_1 j_1} \dots U_{i_k j_k} U^\dagger_{\ell_1 m_1} \dots U^\dagger_{\ell_k m_k} = \sum_{\sigma,\tau \in S_k} \delta_\sigma(\vec{\imath}|\vec{m}) \delta_\tau(\vec{\jmath}|\vec{\ell}) \mathcal{W}g^U(\sigma^{-1}\tau, d) \,, \tag{3.127}$$

where we sum over elements of S_k and denote a δ-function contraction of indices indexed by a permutation $\sigma \in S_k$ as

$$\delta_\sigma(\vec{i}|\vec{j}) \equiv \prod_{n=1}^{k} \delta_{i_n, j_{\sigma(n)}} = \delta_{i_1, j_{\sigma(1)}} \dots \delta_{i_k, j_{\sigma(k)}} \,. \tag{3.128}$$

As we discussed in Ch. 3.2, the unitary Weingarten functions can be computed noting that $\Phi^{(k)}_{\text{Haar}}(P_\sigma) = P_\sigma$, for the k-fold channel of Haar given in Eq. (3.13), and where P_σ are permutation operators on the k copies of $\mathcal{H}^{\otimes k}$. The unitary Weinarten matrix is the inverse of the matrix of inner products $\mathcal{W}g^U_{\sigma,\tau} = (\text{Tr}(P_\sigma P_\tau))^{-1}$, the entries of which are simply given by counting the closed loops in the contracted permutations, e.g. $\text{Tr}(P_\sigma P_\tau) = d^{\text{cycles}(\sigma\tau)}$. The elements of the Weingarten matrix are the Weingarten function $\mathcal{W}g^U(\sigma^{-1}\tau, d)$.

The unitary Weingarten function acts on elements of S_k and has a Fourier expansion in terms of characters of the symmetric group [59] as

$$\mathcal{W}g^U(\sigma, d) = \frac{1}{k!} \sum_{\lambda \vdash k} \frac{\chi^\lambda(\mathbb{I})\chi^\lambda(\sigma)}{c_\lambda(d)} \,, \tag{3.129}$$

where we sum over integer partitions λ of k, as conjugacy classes of S_k are labeled by integer partitions. $\chi^\lambda(\sigma)$ is an irreducible character of S_k labeled by λ and here $\mathbb{I}$ is the identity element of the symmetric group, and $\chi^\lambda(\mathbb{I})$ is simply the dimension of the irreducible representation λ. The factor in the denominator, from which the d dependency of the Weingarten function arises, is a polynomial

$$c_\lambda(d) = \prod_{(i,j)\in\lambda} (d + j - 1) \,, \tag{3.130}$$

where we take the product over (i, j) the coordinates of the blocks in the Young diagram of λ. Said equivalently, λ is some integer partition of k, with elements λ_i. The product is taken over i from 1 to the length of the partition, and j from 1 to λ_i. While this expression of the unitary Weingarten function relies heavily on machinery from the representation theory of symmetric groups, and is less intuitive than the more quantum information theoretic derivation above, this form is far more tractable if we are interested in explicitly computing

large moments of $U(d)$ as we do not need to invert a $k! \times k!$ matrix.

Moreover, the in expansion in terms of characters makes it clear that the Weingarten functions only depend on the cycle-type of the permutation σ. Recall that the cycle-type of $\sigma \in S_k$ is an integer partition λ of k denoting the disjoint cycles in σ, with lengths λ_i written as an ordered list $(\lambda_1, \lambda_2, \ldots)$, where $\lambda_1 \geq \lambda_2 \geq \ldots$ and $\sum_i \lambda_i = k$. For instance, the permutation $\{2, 1, 3, 4\}$ has cycle-type $(2, 1, 1)$. We will always denote the cycle type of a permutation with parentheses (curly brackets being used to denote coset-type). The Weingarten functions used to compute second moments, labeled by their cycle-type, are

$$\mathcal{W}g^U((1,1), d) = \frac{1}{(d-1)(d+1)}\,, \qquad \mathcal{W}g^U((2), d) = \frac{-1}{d(d-1)(d+1)}\,. \tag{3.131}$$

The expression in Eq. (3.129) allows us to implement the calculation of $\mathcal{W}g^U$ in Mathematica and compute, for instance, Weingarten functions for the sixteenth moments with ease.

Integration over the orthogonal group

The formula for integrating over monomials of orthogonal matrices is [59, 88]

$$\int dO\, O_{i_1 j_1} \ldots O_{i_{2k} j_{2k}} = \sum_{\sigma,\tau \in M_{2k}} \Delta_\sigma(\vec{\imath})\Delta_\tau(\vec{\jmath})\mathcal{W}g^O(\sigma^{-1}\tau, d)\,, \tag{3.132}$$

where we sum over elements of M_{2k}, the set of all pair partitions of the set $\{1, 2, \ldots, 2k\}$ and we have defined a combination of δ-functions indexed by a pair partition σ

$$\Delta_\sigma(\vec{\imath}) \equiv \prod_{n=1}^{k} \delta_{i_{\sigma(2n-1)} i_{\sigma(2n)}} = \delta_{i_{\sigma(1)}} \delta_{i_{\sigma(2)}} \ldots \delta_{i_{\sigma(2k-1)}} \delta_{i_{\sigma(2k)}}\,. \tag{3.133}$$

A pair partition $\sigma \in M_{2k}$ is a partition of a set of $2k$ elements into pairs, written as $\{\{\sigma(1), \sigma(2)\}, \ldots, \{\sigma(2k-1), \sigma(2k)\}$, where $\sigma(2n-1) < \sigma(2n)$ and $\sigma(1) < \sigma(3) < \ldots < \sigma(2k-1)$. Let's consider a simple example: the set of pair partitions of 4 elements, M_4, contains three elements

$$\big\{\{1,2\},\{3,4\}\big\}\,, \quad \big\{\{1,4\},\{2,3\}\big\}\,, \quad \big\{\{1,3\},\{2,4\}\big\}\,. \tag{3.134}$$

Realized as a subset of S_{2k}, these are simply the three permutations $\{1,2,3,4\}$, $\{1,4,2,3\}$, and $\{1,3,2,4\}$. Generally, M_{2k} contains $(2k)!/(2^k k!)$ elements.

Recall that Schur-Weyl duality for the unitary group is the statement that irreducible representations of the unitary group can be decomposed into tensors irreps of the symmetric group and the unitary group summed over partitions, which allows us to write down a general formula for integrating monomials of unitaries over the Haar measure. For the orthogonal group the action of the Brauer algebra gives the analog of Schur-Weyl duality; the Brauer algebra has a natural action on the space of polynomials commuting with the action of $O(d)$. The Brauer algebra has a basis formed by pair partitions, which can be realized as a subset of the symmetric group.

As we mentioned in Ch. 3.2, the k-fold channel of the orthogonal group is

$$\Phi_O^{(k)}(\mathcal{O}) = \sum_{\sigma,\tau \in M_{2k}} \mathcal{W}g^O_{\sigma,\tau} S_\sigma \mathrm{Tr}(S_\tau \mathcal{O})\,, \tag{3.135}$$

where S_σ are the basis elements of the Brauer algebra. As $\Phi_O^{(k)}(S_\sigma) = S_\sigma$, we simply compute the matrix of inner products $\mathrm{Tr}(S_\sigma S_\tau)$, the entries given by counting the number of loops in the graph they define $d^{\mathrm{cycles}(\sigma\tau)}$. The orthogonal Weingarten matrix must be the inverse $\mathcal{W}g^O_{\sigma,\tau} = (\mathrm{Tr}(S_\sigma S_\tau))^{-1}$, which is unique.

As the above construction of $\mathcal{W}g^O$ requires inverting a factorially large matrix, if we are interested in computing higher moments of $O(d)$ we should instead use a convenient Fourier expansion. The orthogonal Weingarten function has been discussed in [88, 90, 89]. Just as the unitary Weingarten function admits an expansion in characters of the symmetric group, the orthogonal Weingarten function has an expansion in zonal spherical functions

$$\mathcal{W}g^O(\sigma) = \frac{2^k k!}{(2k)!} \sum_{\lambda \vdash k} \frac{\omega^\lambda(\sigma)\chi^\lambda(\mathbb{I})}{z_\lambda(d)}\,, \quad \text{where} \quad z_\lambda(d) = \prod_{(i,j)\in\lambda} (d+2j-i-1) \tag{3.136}$$

and we sum over integer partitions of k. Here $\omega^\lambda(\sigma)$ is the zonal spherical function, which can be defined in terms of symmetric characters

$$\omega^\lambda(\sigma) = \frac{1}{2^k k!} \sum_{\xi \in H_k} \chi^{2\lambda}(\sigma\xi)\,, \tag{3.137}$$

where H_k is the hyperoctahedral group, a subgroup of S_{2k} of order $2^k k!$, and is the centralizer of the permutation build out of k disjoint transpositions. Pair partitions are representatives of the left-cosets of the hyperoctahedral group in S_{2k}. For more on zonal spherical functions and H_k, see [97]. A quick aside: the denominator in the unitary Weingarten functions $c_\lambda(d)$ in Eq. (3.129), arises from the Schur polynomial, a symmetric function naturally associated with the symmetric group. The denominator for the orthogonal Weingarten function in Eq. (3.136) arises from the zonal polynomial, the symmetric polynomial naturally associated with the pair (S_{2k}, H_k).

The orthogonal Weingarten functions depend only on the coset-type of the permutation, with the coset-type defined in [97] as follows: Compose the pair partition corresponding to the identity permutation $\{\{1,2\},\{3,4\},\ldots\}$, with the pair partition of $\sigma \in S_{2k}$ and count the length of the cycles, which are always even $\{2\lambda_1, 2\lambda_2, \ldots\}$. This gives an integer partition of k as $\lambda = \{\lambda_1, \lambda_2, \ldots\}$. The partition $\lambda \vdash k$ is said to be the coset-type of the permutation $\sigma \in S_{2k}$ and is constant on the the double cosets of S_{2k}. Note that we denote the cycle-type of a permutation with parentheses and the coset-type with curly brackets. For example, the permutation $\{1,3,2,4\}$ has cycle-type $(2,1,1)$ and coset-type $\{2\}$. The orthogonal Weingarten functions used to compute second moments, labeled by their coset-type, are

$$\mathcal{W}g^O(\{1,1\},d) = \frac{d+1}{d(d-1)(d+2)}\,, \qquad \mathcal{W}g^O(\{2\},d) = \frac{-1}{d(d-1)(d+2)}\,. \tag{3.138}$$

3.A.2 Integrating over compact symmetric spaces

We now review Weingarten calculus for compact symmetric spaces—quotients of Lie groups by Lie subgroups.

AI: $U(d)/O(d)$

The first of the circular ensembles, AI denotes the compact symmetric space $U(d)/O(d)$. The group $O(d)$ is not normal in $U(d)$, but the elements of the coset can be realized by the Cartan embedding into $U(d)$. Said equivalently, the involution $\Omega(U) = U^*$ on $U(d)$; $O(d)$ is the closed subgroup fixed by the involution (as $O^* = O$). The quotient $U(d)/O(d)$ can

then be identified with a subset of $U(d)$ defined by $U \mapsto (U^*)^{-1}U = U^T U$. Simply stated, this is the subset of $U(d)$ of symmetric unitaries. This is the symmetry class generated by time-reversal symmetric Hamiltonians (where the time-reversal operator squares to unity).

If V is a Haar random unitary then a random element of $U(d)/O(d)$ is simply $V^T V$. The map above induces an invariant probability measure on the space from the Haar measure. The set and the induced measure are often referred to as the circular orthogonal ensemble (COE). Exact expressions for averages over random elements of this space were worked out in [91, 92]. From the expressions derived there, we can take averages over AI as

$$\int_{\mathrm{AI}} dU\, U_{i_1 i_2} \dots U_{i_{2k-1} i_{2k}} U^\dagger_{j_1 j_2} \dots U^\dagger_{j_{2k-1} j_{2k}} = \sum_{\sigma \in S_{2k}} \delta_\sigma(\vec{\imath}|\vec{\jmath}) \mathcal{W}g^{\mathrm{AI}}(\sigma, d)\,, \tag{3.139}$$

where the AI Weingarten functions are simply related to the Weingarten functions for the orthogonal group as

$$\mathcal{W}g^{\mathrm{AI}}(\sigma, d) = \mathcal{W}g^{\mathrm{O}}(\sigma, d+1)\,. \tag{3.140}$$

This implies that the Weingarten functions for AI only depend on the coset type of σ. We give the first few Weingarten functions labeled by the coset type of $\sigma \in S_{2k}$

$$\mathcal{W}g(\{1\}, d) = \frac{1}{d+1}\,, \quad \mathcal{W}g(\{1,1\}, d) = \frac{d+2}{d(d+1)(d+3)}\,, \quad \mathcal{W}g(\{2\}, d) = -\frac{1}{d(d+1)(d+3)}\,.$$

More generally, the higher Weingarten functions can be computed using the Fourier expansion of the Orthogonal Weingarten function.

AII: $U(2d)/Sp(2d)$

The next of the circular ensembles, AII denotes the compact symmetric space $U(2d)/Sp(2d)$. Again, the elements of the coset can be realized as a subset of $U(2d)$ by the Cartan embedding. Consider the involution $\Omega(U) = (U^D)^{-1}$ on $U(2d)$, where now $Sp(2d)$ is the subgroup fixed Ω. The quotient space $U(2d)/Sp(2d)$ is identified with a subset of $U(2d)$ defined by $U \mapsto U^D U$, which this is the subset of symplectically symmetric unitaries. This is the symmetry class generated by time-reversal symmetric Hamiltonians where the time-reversal operator squares to -1.

If V is a Haar random unitary then a random element of $U(2d)/Sp(2d)$ is simply $V^D V$, and again the map induces an invariant probability measure on the space. In this case the set and the induced measure are often referred to as the circular symplectic ensemble (CSE). Averaging over random instances of $U(2d)/Sp(2d)$ was worked out in [92]. From the expressions derived there, we can take averages over AII as

$$\int_{\text{AII}} dU\, J^T U_{i_1 i_2} \dots J^T U_{i_{2k-1} i_{2k}} J U^\dagger_{j_1 j_2} \dots J U^\dagger_{j_{2k-1} j_{2k}} = \sum_{\sigma \in S_{2k}} \delta_\sigma(\vec{\imath}|\vec{\jmath}) \mathcal{W}g^{\text{AI}}(\sigma, d)\,, \tag{3.141}$$

where the AII Weingarten functions are related to the Weingarten functions for the symplectic group as

$$\mathcal{W}g^{\text{AII}}(\sigma, d) = \mathcal{W}g^{\text{Sp}}(\sigma, d - 1/2)\,. \tag{3.142}$$

This implies that the Weingarten functions for AII also only depend on the coset type and signature of σ. For example, a few of the Weingarten functions labeled by the coset type of $\sigma \in S_{2k}$ are

$$\mathcal{W}g(\{1,1\}, d) = \frac{\epsilon(\sigma)(d-1)}{d(2d-1)(2d-3)}\,, \quad \mathcal{W}g(\{2\}, d) = -\frac{\epsilon(\sigma)}{2d(2d-1)(2d-3)}\,.$$

Weingarten calculus can also be extended to other symmetric spaces, including the chiral and BdG ensembles. We refer the reader to [92] for further details.

3.B Higher-point correlation functions

As we mentioned in the text, the leading terms in the asymptotic form of the correlation functions is the same irrespective of the symmetry class. For the 2-point, OTO 4-point, and OTO 6-point function, we list the leading order terms in the $d \to \infty$ limit, which appear to be universal.

$$\langle AB(t)\rangle \approx \langle A\rangle\langle B\rangle \tag{3.143}$$

$$\langle AB(t)CD(t)\rangle \approx \langle AC\rangle\langle B\rangle\langle D\rangle + \langle A\rangle\langle C\rangle\langle BD\rangle + \langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle \tag{3.144}$$

$$\begin{aligned}\langle AB(t)CD(t)EF(t)\rangle \approx \langle A\rangle\langle C\rangle\langle E\rangle\langle BDF\rangle + \langle B\rangle\langle D\rangle\langle F\rangle\langle ACE\rangle + \langle A\rangle\langle D\rangle\langle BF\rangle\langle CE\rangle \\ +\langle B\rangle\langle E\rangle\langle AC\rangle\langle DF\rangle + \langle C\rangle\langle F\rangle\langle AE\rangle\langle BD\rangle - \langle A\rangle\langle B\rangle\langle C\rangle\langle E\rangle\langle DF\rangle \\ -\langle A\rangle\langle B\rangle\langle D\rangle\langle F\rangle\langle CE\rangle - \langle A\rangle\langle C\rangle\langle D\rangle\langle E\rangle\langle BF\rangle - \langle A\rangle\langle C\rangle\langle E\rangle\langle F\rangle\langle BD\rangle \\ -\langle B\rangle\langle C\rangle\langle D\rangle\langle F\rangle\langle AE\rangle - \langle B\rangle\langle D\rangle\langle E\rangle\langle F\rangle\langle AC\rangle + 2\langle A\rangle\langle B\rangle\langle C\rangle\langle D\rangle\langle E\rangle\langle F\rangle\end{aligned} \tag{3.145}$$

We also compute the asymptotic form of the OTO 8-point function and checked that it was universal, i.e. the leading contribution for all symmetry classes, but as the leading term contains 55 terms, we refrain from reproducing it here.

3.C k-invariance in spin-systems

In this appendix we present a numerical investigation of k-invariance in spin systems and SYK models. The quantity we investigate here is the 1-invariance, given as

$$\text{1-inv}: \quad \mathcal{F}^{(1)}_{\mathcal{E}}(t) - \mathcal{F}^{(1)}_{\tilde{\mathcal{E}}}(t)\,, \tag{3.146}$$

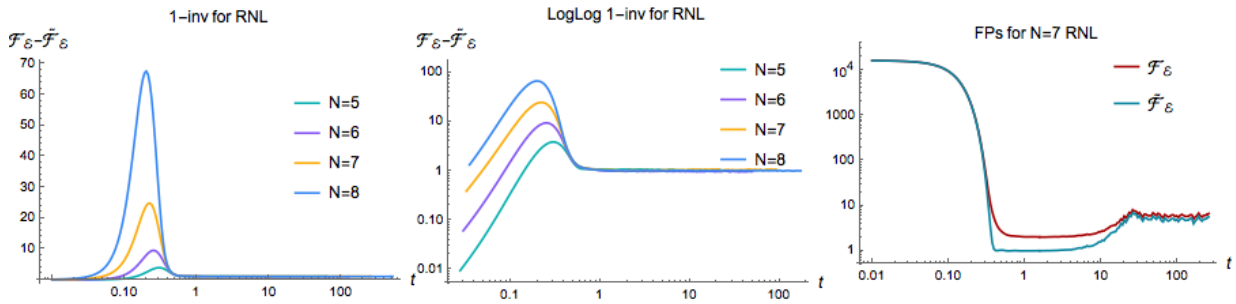
and where $\tilde{\mathcal{E}} = U\mathcal{E}U^\dagger$ for Haar-random U.

Random Nonlocal 2-body

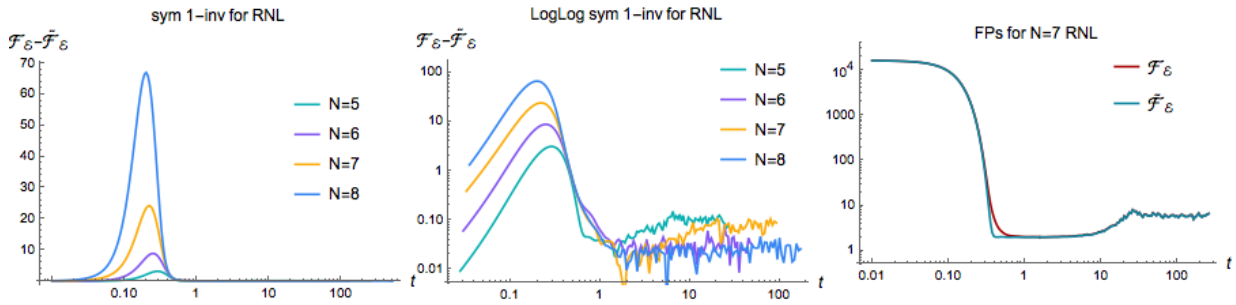
Consider systems of N spins all-to-all interacting with random couplings and summed over all possible 2-body interactions

$$\text{RNL}: \qquad H = \sum_{ij\alpha\beta} J_{\alpha\beta ij} S^\alpha_i S^\beta_j\,, \tag{3.147}$$

where each J is Gaussian random with variance $1/N$, and we sum over all possible 2-body interactions between all pairs of sites i, j, and α, β run over the local Paulis at the site. We plot the 1-invariance for 5, 6, 7, and 8 spins



and observe an increase in the distance to 1-inv at intermediate times, but a constant value at late times, which does not change as we increase N. On the right we plot the difference between the frame potential and the Haar'ed frame potential for $N = 7$ spins. Note that the system is time-reversal invariant, where even and odd N alternate between $T^2 = \pm 1$. Computing the symmetric 1-invariance to account for the symmetry of the system, we find that the late-time floor value drops to zero

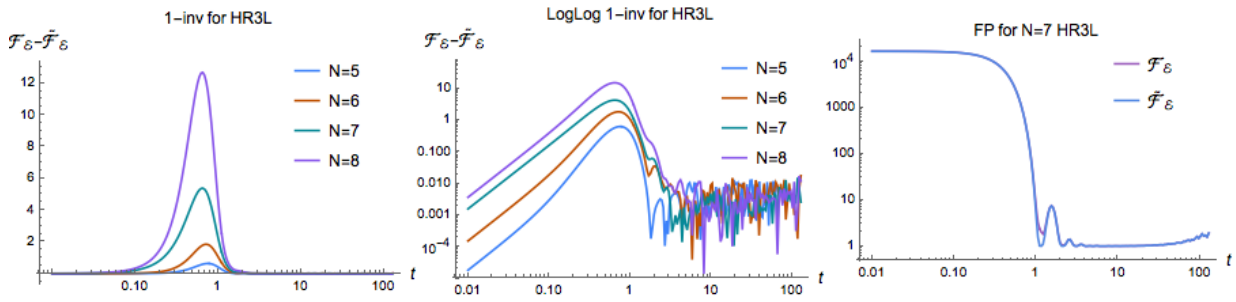


Random Nonlocal 3-body

Consider systems of N spins all-to-all interacting with random couplings and summed over all possible 3-body interactions

$$\text{HR3L}: \qquad H = \sum_{ijk\alpha\beta\gamma} J_{\alpha\beta\gamma ijk} S_i^\alpha S_j^\beta S_k^\gamma \,, \tag{3.148}$$

where each J is Gaussian random with variance $1/N^2$, we sum over all possible 3-body interactions between all triplets of sites i, j, k, and α, β, γ run over the local Paulis at the site. We plot the 1-invariance for the 3-local Hamiltonian



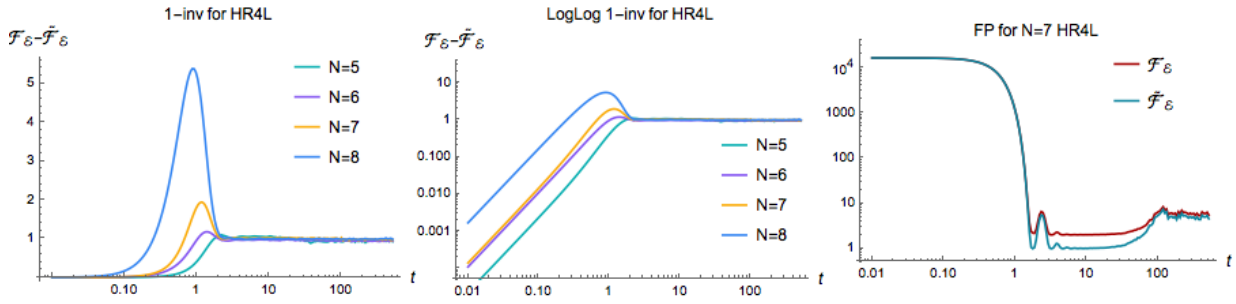
and observe an N-dependent distance to 1-inv at intermediate times, but independence at late times. The floor value here fluctuates around zero and thus the 3-local system appears to achieve 1-variance at late times. On the right we plot the difference between the frame potential and the Haar'ed FP for $N = 7$ spins.

Random Nonlocal 4-body

Consider systems of N spins all-to-all interacting with random couplings and summed over all possible 4-body interactions

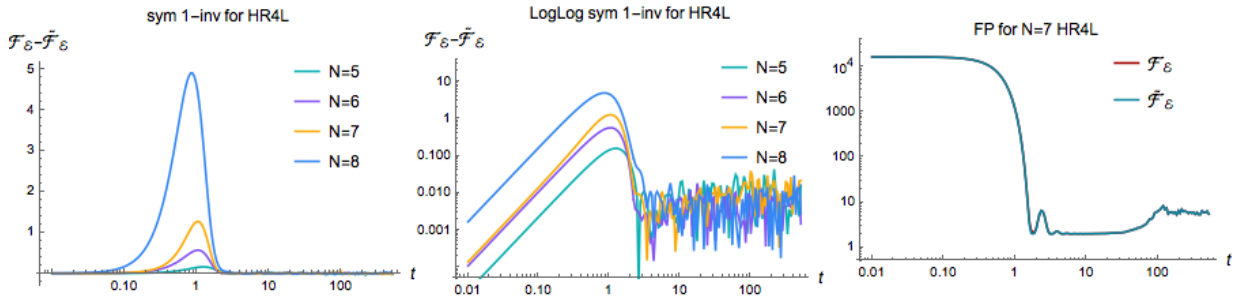
$$\text{HR4L}: \qquad H = \sum_{ijk\alpha\beta\gamma\delta} J_{ijk\ell}^{\alpha\beta\gamma\delta} S_i^\alpha S_j^\beta S_k^\gamma S_\ell^\delta \,, \tag{3.149}$$

where each J is Gaussian random with variance $1/N^2$, and we sum over all possible 3-body interactions between all triplets of sites i, j, k, and α, β, γ run over the local Paulis at the site. We plot the 1-invariance for the 4-local model



and, just as in the 2-local model, observe an N-dependence at intermediate times and a late time floor value that seems to be robust in the large N limit. On the right we plot the difference between the frame potential and the Haar'ed FP for $N = 7$ spins. Just like

the 2-local model, the 4-local model is time-reversal invariant. Computing the symmetric 1-invariance we find that late-time floor value drops to zero

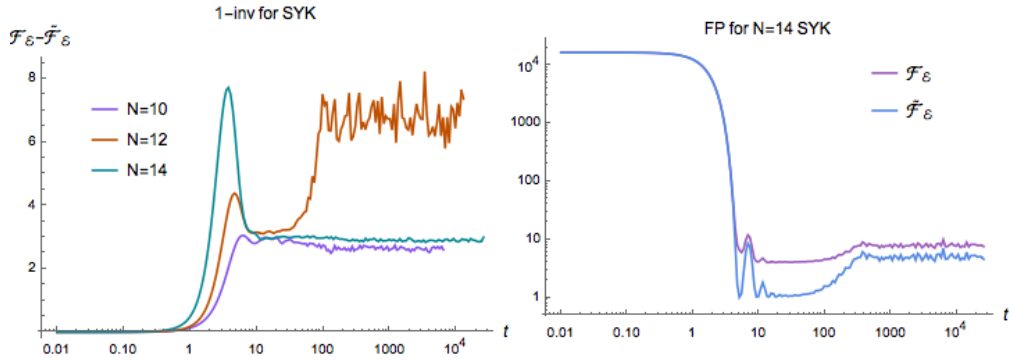


Sachdev-Ye-Kitaev Model

Consider a system of N all-to-all interacting Majoranas with random couplings

$$\text{SYK}: \qquad H = \sum_{ijk\ell} J_{ijk\ell}\chi_i\chi_j\chi_k\chi_\ell\,, \tag{3.150}$$

where each J is Gaussian random with variance $6/N^3$, and we sum over all 4-local interactions. We plot the 1-invariance for 10, 12, and 14 Majoranas



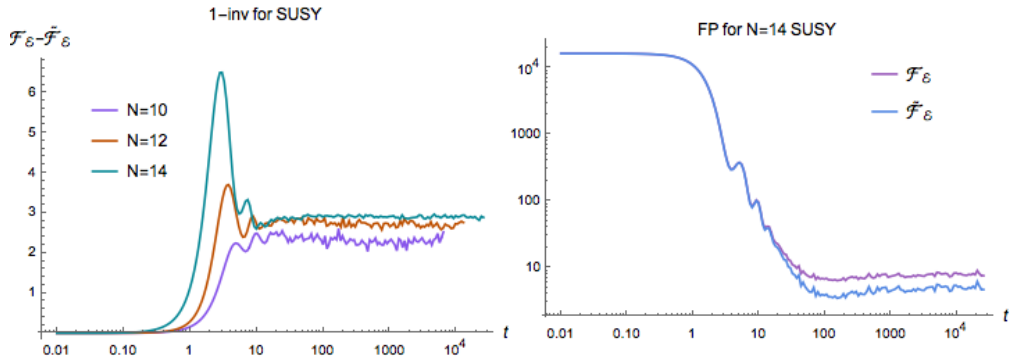
and observe the N-dependence of the distance to 1-inv at intermediate times, but also a dependence at late times. This shouldn't be surprising as 10 and 14 correspond to GUE, but 12 corresponds to GSE, so we expect a difference measure of k-invariance. To correctly account for the late-time behavior we should extract the contribution from a single charge sector.

Supersymmetric SYK Model

Consider a system of N all-to-all interacting Majoranas with random couplings

$$\text{SUSY}: \qquad H = Q^2 \quad \text{where} \quad Q = i \sum_{ijk\ell} C_{jk\ell} \chi_j \chi_k \chi_\ell \,, \tag{3.151}$$

and where each C is Gaussian random with variance $2/N$, and we sum over all 3-local interactions in the supercharge. We plot the 1-invariance for SUSY SYK and find

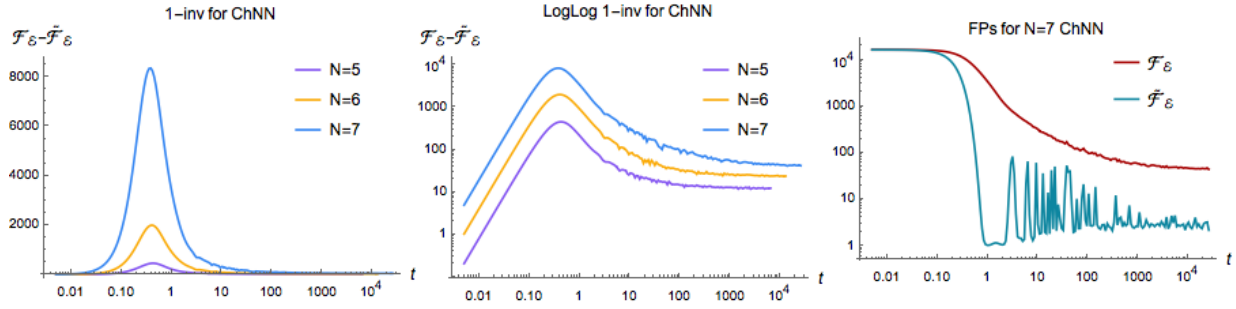


Chaotic Disordered 1D Spin Chain

Consider a system of N spins interacting with nearest-neighbor couplings and disordered transverse and parallel fields

$$\text{ChNN}: \qquad H = -\sum_i Z_i Z_{i+1} - \sum_i g_i Z_i - \sum_i h_i X_i \,, \tag{3.152}$$

where both h and g are Gaussian random with variance $1/N$. We compute the 1-invariance for the chaotic spin-chain



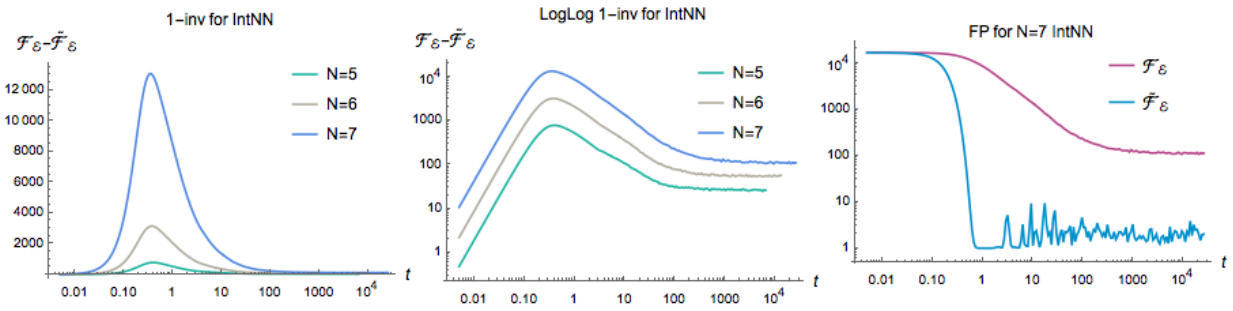
and observe an N-dependence of the distance at intermediate times, but also a dependence at late times. Both the intermediate values and late time values are huge compared to the random systems above. On the right we plot the difference between the frame potential and its Haar'ed equivalent for $N = 7$ spins.

Integrable Disordered 1D Spin Chain

Consider a system of N spins interacting with nearest-neighbor couplings and disordered transverse field

$$\text{IntNN}: \qquad H = -\sum_i Z_i Z_{i+1} - \sum_i h_i X_i \,, \tag{3.153}$$

where both h are Gaussian random with mean 1 and variance $1/N$. Note, we also tried mean zero couplings and the result is essentially the same. We plot the 1-invariance for the integrable spin-chain



and observe an N-dependence of the distance to 1-inv both at intermediate times and late times. Again, both the intermediate values and late time values are huge compared to the random systems above and even larger than the chaotic local spin chain. On the right we plot the difference between the frame potential and the Haar'ed version for $N = 7$ spins. This is evidence that the late-times floor captures the chaotic and nonlocal nature of the system.

Chapter 4

Chaos and random matrices in supersymmetric SYK

This chapter is essentially the same as

- N. Hunter-Jones, J. Liu, "Chaos and random matrices in supersymmetric SYK," *JHEP* **05** (2018) 202, arXiv:1710.08184 [hep-th].

Abstract

We use random matrix theory to explore late-time chaos in supersymmetric quantum mechanical systems. Motivated by the recent study of supersymmetric SYK models and their random matrix classification, we consider the Wishart-Laguerre unitary ensemble and compute the spectral form factors and frame potentials to quantify chaos and randomness. Compared to the Gaussian ensembles, we observe the absence of a dip regime in the form factor and a slower approach to Haar-random dynamics. We find agreement between our random matrix analysis and predictions from the supersymmetric SYK model, and discuss the implications for supersymmetric chaotic systems.

4.1 Introduction

A recent surge of interest in quantum chaos has revolved around a strongly-interacting quantum system called the Sachdev-Ye-Kitaev (SYK) model [6, 35]. This model of N all-

to-all randomly interacting Majorana fermions is solvable at strong-coupling and appears to be in the same universality class as black holes, exhibiting an emergent reparametrization invariance and an extensive ground-state entropy. More compellingly, the out-of-time order correlation function (OTOC) of the theory [6, 7] saturates a universal bound on chaotic growth [27], a seemingly unique feature of gravity [4, 5] and conformal field theories with a holographic dual [29]. The low-energy description of the theory in terms of a Schwarzian effective action also encapsulates dilaton gravity in AdS_2 [98, 99]. This model should be seen as a valuable resource for understanding both black holes and quantum chaos.

There have already been a myriad of generalizations of the SYK model, including an extension by Fu, Gaiotto, Maldacena, and Sachdev, to a supersymmetric model of strongly interacting Majoranas [100], which has been further explored in [101, 102, 103, 104, 105]. The supersymmetric version of the model also displays many of the same holographic properties. Notably, at strong-coupling the theory has an emergent superconformal symmetry which renders it solvable and allows one to compute correlation functions. At low-energies the symmetry is broken, giving a Schwarzian-like effective action which mimics supergravity in AdS_2 [106]. Like its non-supersymmetric counterpart, the model has random matrix universality in its spectral statistics [41, 107] and appears to exhibit thermalization in its eigenstates [73, 108], both hinting at underlying chaotic dynamics.

Although we lack a precise definition of quantum chaos, there are still universal features one expects of quantum chaotic systems: most notably, having the spectral statistics of a random matrix [8]. Information scrambling [1, 2] and chaotic correlation functions [4] have also been extolled as symptoms of chaos. Ideas from quantum information have helped make these notions more precise, quantifying how scrambling [10] and randomness [11] are encoded in OTOCs. Similarly, [12] explored the connection to random matrix dynamics, quantifying randomness and scrambling in the time evolution by random matrix Hamiltonians and computing a quantity called the frame potential. The onset of random matrix behavior can also be seen in the spectral form factor, which has been studied in the SYK model [9].

Motivated by this, we may ask the question: what are the universal features of supersymmetric SYK models, or more generally, of all supersymmetric quantum chaotic systems? And how do we quantify them from an information-theoretic standpoint?

To address this, we consider the Wishart-Laguerre ensembles, also termed random covariance matrices [43], which appeared in the random matrix classification of the supersymmetric SYK models [107]. Recall that the Hamiltonian in supersymmetric quantum mechanics is constructed as the square of a supercharge. Loosely speaking, the intution is that this random matrix ensemble arises from squaring the Gaussian random matrices, just as we might think of a chaotic supersymmetric system defined by a disordered supercharge. In this chapter we consider the simplest Wishart-Laguerre ensemble,[1] the Wishart-Laguerre unitary ensemble (LUE), corresponding to supersymmetric quantum systems without additional discrete symmetries. In the following, we will quantitatively derive predictions for the spectral form factors, frame potential, and the out-of-time-ordered correlators, where a central distinction from the non-supersymmetric models arises in the spectral 1-point functions, which modifies the early time decay of the spectral form factor. A slower decay in the LUE frame potential indicates less efficient information scrambling and the failure of the ensemble to become Haar-random. Our predictions for the LUE match those from the 1-loop partition function of the supersymmetric SYK model.

The chapter is organized as follows: In Section 4.2, we review the supersymmetric model and spectral form factor, discussing its universal features and behavior in SYK models. In Section 4.3, we review the basic tools in random matrix theory and then compute spectral form factors for the Wishart-Laguerre ensemble. In Section 4.4, we explore chaos in this random matrix ensemble by computing the frame potentials and correlation functions, and comment on its complexity growth. In Section 4.5, we discuss chaos in supersymmetric SYK and compare with the random matrix predictions, concluding in Section 4.6. In Appendix 4.A we present some numerical checks of our expressions.

[1] Interestingly, Wishart ensembles have appeared in studying the reduced density matrix in systems evolved with random matrix Hamiltonians [109]. Wishart ensembles have also appeared in random matrix contructions of supergravity to explore the space of AdS vacua [110].

4.2 Setup and overview

4.2.1 Supersymmetric SYK model

We first briefly review the supersymmetric extension of the SYK model. For an in-depth discussion of the original model, see [7]. Consider N all-to-all interacting Majorana fermions ψ_i with random couplings, which anticommute as $\{\psi_i, \psi_j\} = \delta_{ij}$. The $(2q-2)$-point $\mathcal{N}=1$ supersymmetric model is constructed from the supercharge Q, a q-body Majorana interaction with odd q. The Hamiltonian is then given by the square of the supercharge as

$$H = Q^2\,, \quad \text{where} \quad Q = i^{(q-1)/2} \sum_{i_1<\ldots<i_q} C_{i_1\ldots i_q}\psi_{i_1}\ldots\psi_{i_q}\,, \tag{4.1}$$

with Gaussian random couplings $C_{i_1\ldots i_q}$ of mean and variance

$$\left\langle C_{i_1\ldots i_q}\right\rangle = 0\,, \qquad \left\langle C^2_{i_1\ldots i_q}\right\rangle = \frac{J^2(q-1)!}{N^{q-1}}\,, \tag{4.2}$$

and where J is a positive constant. We also define $\mathcal{J}$ as $J^2 = 2^{q-1}\mathcal{J}^2/q$, with a slightly more convenient scaling in q.

In the large N limit, this model shares many of the same appealing holographic features as the SYK model, such as chaotic correlation functions, a zero-temperature entropy, and an emergent superconformal symmetry which is broken at low-energies, admitting a Schwarzian-like desciption [100]. We can compute the free energy at large N by evaluating at the saddle point, and at low temperatures find

$$\log Z = -\beta E_0 + N s_0 + \frac{cN}{2\beta} + \ldots\,, \tag{4.3}$$

where s_0 is the zero-temperature entropy density and c is the specific heat. In the supersymmetric theory we have $c = \alpha\pi^2/\mathcal{J}$ with a constant α, which becomes $c = \pi^2/4q^2\mathcal{J}$ in the large q limit. The ground-state entropy density is computed to be $s_0 = \frac{1}{2}\log(2\cos\frac{\pi}{2q})$ and the ground state energy E_0 can be subtracted off.

The SYK model with N Majoranas enjoys a random matrix classification, where the symmetry class of the theory is dictated by a particle-hole symmetry [41, 9]. Depending on N,

the spectrum will display level statistics of one of the three Gaussian ensembles: GUE, GOE, or GSE. For the supersymmetric extension of SYK, we can similarly classify the random matrix behavior for a given number of Majoranas N, going beyond Dyson's classification to the extended 10-fold symmetry classification of Altland-Zirnbauer [81]. Understanding how anti-unitary symmetries act on the supercharge Q, we can identify the appropriate symmetry class [107]. The Hamiltonian, given as the square of the supercharge, then has random matrix description in terms of the Wishart-Laguerre ensembles. The level statistics are still those of the Gaussian ensembles, but the spectral correlations are different. Roughly, we can think of the supersymmetric SYK behaving like the square of Gaussian random matrices, which are the Wishart ensembles. For more details, see [107] as well as an extension of the classification to the $\mathcal{N} = 2$ supersymmetric models [105].

Speaking generally, there a number of reasons one might wish to consider supersymmetric generalizations of SYK. For instance, much is understood about the low-energy physics in nearly AdS_2 spacetimes purportedly dual to the low-energy dynamics in SYK, but the exact holographic dual of the theory is not known. As many of the best understood examples of AdS/CFT are supersymmetric, one might hope that this particular construction might provide guidance on the correct UV completion of the SYK model. Less ambitiously, considering the supersymmetric models might be useful in contructing higher dimension analogs [102].

4.2.2 Spectral form factor

Quantum chaotic systems are often defined to have the spectral statistics of a random matrix. An object familiar in random matrix theory which exhibits these universal properties is the spectral form factor. We will introduce this object more precisely in our review of random matrix theory in Sec. 4.3.1, but the 2-point spectral form factor $\mathcal{R}_2(t, \beta)$ can be given simply in terms of the analytically continued partition function

$$\mathcal{R}_2(t, \beta) \equiv \left\langle Z(\beta, t) Z^*(\beta, t) \right\rangle , \quad \text{where} \quad Z(\beta, t) \equiv \mathrm{Tr}\left(e^{-\beta H - itH}\right) , \tag{4.4}$$

and where the average $\langle \cdot \rangle$ is taken over an ensemble of Hamiltonians (e.g. SYK, or some disordered spin system, or a random matrix ensemble). This object was discussed more recently in [9], where they studied the form factor in SYK and found that the theory revealed random matrix behavior at late times. From the bulk point of view, one motivation for studying this object was a simple version of black hole information loss [36]: 2-point functions appear to decay exponentially in terms of local bulk variables, whereas a discrete spectrum implies a finite late-time value. The same inconsistency is apparent in the spectral form factor.

Some characteristic features of the time-evolved form factor $\mathcal{R}_2(t)$, exhibited in both the SYK model and in random matrix theories, are: an early time decay from an initial value called the slope, a crossover at intermediate times called the dip, a steady linear rise called the ramp, and a late-time floor called the plateau. In Fig. 4.1 we observe these features in SYK. While the early time decay depends on the specific system, the ramp and plateau should be universal features of quantum chaotic systems. The ramp is characteristic of spectral rigidity: the long-range logarithmic repulsion of eigenvalues. The anticorrleation of eigenvalues causes the linear increase in the form factor. At late times, or at energy scales smaller than the mean spacing, the form factor reaches a plateau as degeneracies are rare and neighboring eigenvalues repel in chaotic systems.

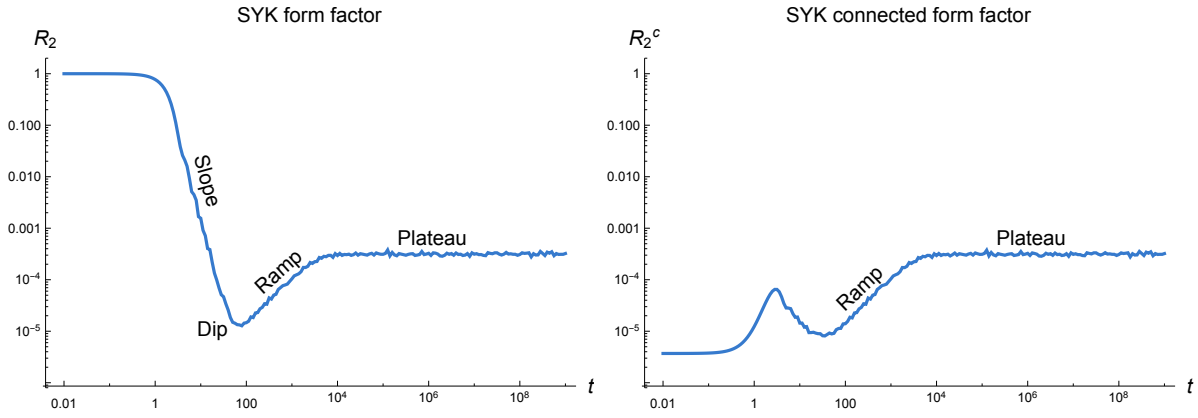


Figure 4.1: The 2-point spectral form factor and its connected component for SYK with $N = 24$ Majoranas at inverse temperature $\beta = 1$, computed for 800 realizations of disorder. We observe the slope, dip, ramp, and plateau behaviors.

SYK form factor and GUE

Recently, [9] studied the form factor in SYK and found agreement with random matrix theory, showing analytically and numerically the aspects of the dip, ramp, and plateau of SYK agree with those of the Gaussian unitary ensemble (GUE), an ensemble of $d \times d$ random Hermitian matrices. We will avoid explicitly introducing and defining the original Majorana, instead simply mentioning a few details to better frame the discussion of the model's supersymmetric extension.

The emergent reparamentrization invariance of SYK at strongly-coupled is broken spontaneously and explicitly at low-energies, yielding an effective description in terms of the Schwarzian derivative [6, 7]. The 1-loop partition function of the Schwarzian theory $Z^{\text{Sch}}_{\text{1-loop}} \sim e^{cN/2\beta}/\beta^{3/2}$, can be analytically continued to $\beta + it$ to study the form factor of SYK. At early times, $\mathcal{R}_2(t,\beta)$ is dominated by the disconnected piece which gives a $1/t^3$ power law decay, normalized by its initial value we have

$$\frac{\langle Z(\beta,t)Z^*(\beta,t)\rangle}{\langle Z(\beta)\rangle^2} \simeq \frac{\beta^3 e^{-cN/\beta}}{t^3}\,, \tag{4.5}$$

for times greater than $t \gtrsim \sqrt{N}$ when the time dependence in the exponent disappears and where c is the specific heat of the theory. To isolate this contribution, [9] considered a special limit (a 'triple scaled' limit) where only the Schwarzian contributes. Moreover, [111] showed that the Schwarzian theory is 1-loop exact and recieves no higher-order corrections, indicating that the power-law decay predicted by the Schwarzian should dominate the disconnected form factor for long times.[2] This power law decay is simply the Laplace transform of the statement that the spectrum has a square-root edge[3]

$$\rho(E) \sim \sinh\sqrt{2cEN}\,. \tag{4.6}$$

Knowing the free energy in the large N limit, we can also show that the form factor of SYK transitions to a ramp at a dip time $t_d \sim e^{Ns_0/2}$, growing linearly until a plateau time of $t_p \sim e^{Ns_0 + cN/2\beta}$, where s_0 is the zero-temperature entropy density.

[2] For more on solving the Schwarzian theory, see [112, 113].

[3] As discussed in [7]. The spectral density of SYK has been further studied in [9, 42, 114].

Many of these features of the SYK form factor agree with the universal predictions from GUE. The form factor for GUE has been studied extensively in the random matrix literature [23, 24, 68] and references therein, and revisited more recently in the context of SYK and black holes in [9, 45, 12]. Simply stating the results, the early-time decay of the GUE form factor transitions to a linear ramp at a dip time of $t_d \sim \sqrt{d}$, growing linearly until the plateau time $t_p \sim d$. We note that around the plateau time the ramp is not quite linear as nonperturbative effects become important as we transition to the plateau [115]. The non-universal early time decay also has the same power law $1/t^3$, due to the fact the Wigner semicircle law for Gaussian random matrices $\rho(\lambda) = \frac{1}{2\pi}\sqrt{4-\lambda^2}$, also exhibits a square-root edge.

Supersymmetric SYK form factor

From the large N partition function of the supersymmetric theory, we can also make predictions as to the behavior of the spectral form factor. We will present a more explicit treatment of this in Sec. 4.5. At low-energies, the fluctuations around the large N saddle point of the supersymmetric theory break superconformal symmetry; the action for these reparametrizations is a super-Schwarzian [100], where the action integrates over τ and a superspace coordinate θ and the super-Schwazian acts just like the standard Schwarzian derivative except as a super-derivative, respecting a similar chain rule. The action gives a 1-loop partition function

$$Z^{\text{sSch}}_{\text{1-loop}}(\beta) \sim \frac{1}{\sqrt{\beta \mathcal{J}}} e^{N s_0 + cN/2\beta}\,, \tag{4.7}$$

which differs in the 1-loop determinant from the SYK model. The super-Schwarzian theory is also 1-loop exact [111], ensuring its validity away from very early times. Analytically continuing the partition function $\beta \to \beta + it$, disconnected piece of the form factor which dominates at early times, is

$$\frac{\langle Z(\beta,t) Z^*(\beta,t)\rangle}{\langle Z(\beta)\rangle^2} \simeq \frac{\beta e^{-cN/\beta}}{t}\,, \tag{4.8}$$

exhibiting a $1/t$ decay in the slope, slower than the decay in SYK. This can also be understood as the contribution from the edge of the spectrum, where the Laplace transform of the 1-loop partition function gives

$$\rho(E) \sim \frac{1}{\sqrt{\mathcal{J}E}} \cosh\left(\sqrt{2cNE}\right), \tag{4.9}$$

observing a square-root growth at the edges of the spectrum.

As we discuss later, computing the ramp function for supersymmetric SYK, we find the ramp and slope intersect at a dip time $t_d \sim e^{Ns_0}$, which is the same time scale as the ramp's transition to the plateau $t_p \sim e^{Ns_0}$. The slow decay at early times means that the slope transitions to ramp behavior at the same time-scale as the plateau time, i.e. the ramp is hidden beneath the slope. We plot the 2-point form factor for the model in Fig. 4.2. Subtracting the disconnected contribution reveals the ramp in the connected form factor, also plotted. The lack of a dip in the supersymmetric model will have implications for our discussion of the frame potential and randomness.

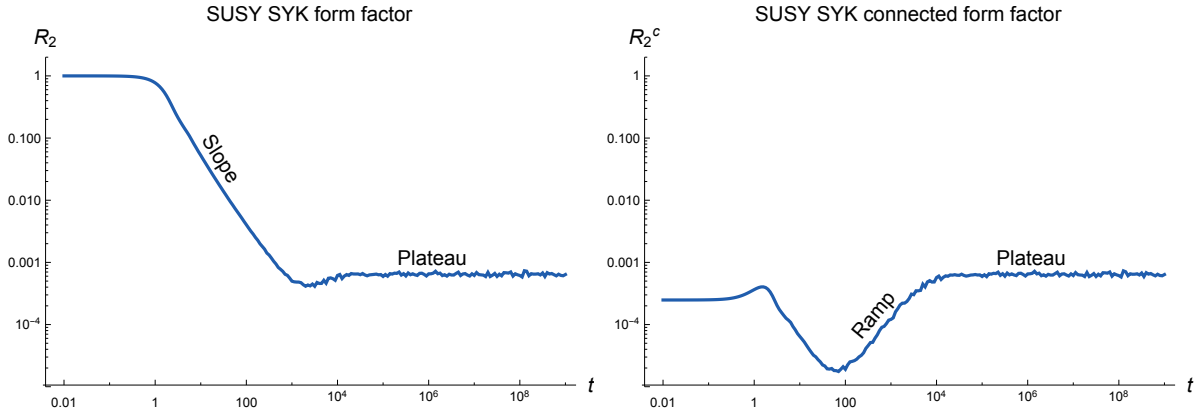


Figure 4.2: The 2-point spectral form factor and its connected piece for the supersymmetric SYK model with $N = 24$ Majoranas at inverse temperature $\beta = 1$, computed for 800 realizations of disorder. We observe the slope and plateau behaviors, while the ramp is obscured by the slow early-time decay of the 1-point function.

Notation

A brief comment on notation. In recent work studying the spectral form factor, the normalized 2-point form factor is often denoted as $g(t, \beta)$, and its connected component as

$g_c(t,\beta)$:

$$g(t,\beta) \equiv \frac{\langle Z(\beta,t)Z^*(\beta,t)\rangle}{\langle Z(\beta)\rangle^2}\,, \quad \text{and} \quad g_c(t,\beta) \equiv g(t,\beta) - \frac{\langle Z(\beta,t)\rangle\langle Z^*(\beta,t)\rangle}{\langle Z(\beta)\rangle^2}\,. \tag{4.10}$$

While in [12], we denoted the 2-point form factor as $\mathcal{R}_2(t,\beta)$, and more generally the $2k$-th form factor as $\mathcal{R}_{2k}(t,\beta)$. Just to be clear

$$g(t,\beta) = \frac{\mathcal{R}_2(t,\beta)}{\langle Z(\beta)\rangle^2}\,, \quad \text{or at } \beta = 0: \quad g(t,0) = \frac{\mathcal{R}_2(t)}{d^2}\,. \tag{4.11}$$

For us, working directly with the numerator turns out to be more convenient when discussing the frame potential and correlation functions, and avoids subtleties regarding the appropriate or tractable normalization, i.e. 'quenched' vs 'annealed'.

4.3 Form factors for Wishart matrices

4.3.1 Basic setup in random matrix theory

In this chapter, we consider the Wishart-Laguerre Unitary Ensemble (LUE), an ensemble of $d \times d$ random matrices which can be generated as $H^\dagger H$, where H is a complex Gaussian random matrix with normally distributed complex entries drawn with mean 0 and variance $\sigma^2 = 1/d$. This is the 'physics normalization', where the spectrum does not scale with system size.[4] The joint probability distribution of LUE eigenvalues is given by

$$P(\lambda)d\lambda = C\,|\Delta(\lambda)|^2 \prod_{k=1}^{d} e^{-\frac{d}{2}\lambda_k} d\lambda_k\,, \tag{4.12}$$

where $\Delta(\lambda)$ is the Vandermonde determinant and the constant factor is defined such that the distribution integrates to unity. One can think of LUE matrices as square of a Gaussian random matrix. More generally, we could define $d \times d$ Wishart matrices generated by $d' \times d$ Gaussian matrices, where $d' \geq d$, which gives a slightly more general eigenvalue distribution. But given the supersymmetric Hamiltonians we consider defined as the square of the

[4]Note that it is common in the random matrix literature to instead work with unit variance $\sigma^2 = 1$.

supercharge, we just consider Wishart matrices generated by square matrices with $d = d'$. We average over the random matrix ensemble as

$$\langle \mathcal{O} \rangle \equiv \int D\lambda \, \mathcal{O} \quad \text{where} \quad \int D\lambda = C \int \prod_k d\lambda_k |\Delta(\lambda)|^2 e^{-\frac{d}{2}\sum_k \lambda_k} \,. \tag{4.13}$$

The spectral density is given by integrating the joint probability $P(\lambda)$ over $d-1$ variables,

$$\rho(\lambda) = \int d\lambda_1 d\lambda_2 \dots d\lambda_{d-1} \, P(\lambda_1, \lambda_2, \dots \lambda_{d-1}, \lambda) \,. \tag{4.14}$$

More generally, we can define the k-point spectral correlation function by integrating over all but k arguments

$$\rho^{(k)}(\lambda_1, \lambda_2, \dots, \lambda_k) = \int d\lambda_{k+1} d\lambda_{k+2} \dots d\lambda_d \, P(\lambda_1, \lambda_2, \dots \lambda_k, \lambda_{k+1}, \dots, \lambda_d) \,. \tag{4.15}$$

Recall that for the Gaussian ensembles, we may take the large d limit famously recover Wigner's semicircle law for the distribution of eigenvalues. Instead in the LUE, we take the large d limit and find [116]

$$\rho(\lambda) = \frac{1}{2\pi\lambda} \sqrt{\lambda(4-\lambda)} \,, \tag{4.16}$$

which is referred to as the Marčenko-Pastur distribution.

Just as in the GUE, the LUE is a determinantal point process, which means the k-point spectral correlators are given by a kernel K as

$$\rho^{(k)}(\lambda_1, \dots, \lambda_k) = \frac{(d-k)!}{d!} \det \big(K(\lambda_i, \lambda_j) \big)_{i,j=1}^k \,. \tag{4.17}$$

Demonstrating the universality of Dyson's sine kernel [22], the Wishart ensemble has sine

kernel statistics in the large d limit [117, 43], meaning

$$K(\lambda_i, \lambda_j) = \begin{cases} \dfrac{\sin\left(d\rho(u)\pi(\lambda_i - \lambda_j)\right)}{\pi(\lambda_i - \lambda_j)} & \text{for} \quad i \neq j \\ \dfrac{d}{2\pi\lambda_i}\sqrt{\lambda_i(4-\lambda_i)} & \text{for} \quad i = j\,, \end{cases} \tag{4.18}$$

where u is an arbitrary constant valued in $[0, 4]$. We will fix the value of u numerically.[5]

The spectral form factor, defined as the Fourier transform of the spectral correlation functions, is a standard quantity to consider in random matrix theory; see [23] for an overview. We define the 2-point spectral form factor in terms of the analytically continued partition function $Z(\beta, t)$ as[6]

$$\mathcal{R}_2(t,\beta) \equiv \langle Z(\beta,t)Z^*(\beta,t)\rangle = \int D\lambda \sum_{i,j} e^{i(\lambda_i-\lambda_j)t} e^{-\beta(\lambda_i+\lambda_j)}\,, \tag{4.19}$$

where the continued partition function $Z(\beta, t)$ is

$$Z(\beta,t) = \mathrm{Tr}\left(e^{-\beta H - iHt}\right)\,. \tag{4.20}$$

More generally, we consider k-point spectral form factors which we define as

$$\mathcal{R}_{2k}(t,\beta) \equiv \left\langle \left(Z(\beta,t)Z^*(\beta,t)\right)^{2k} \right\rangle \tag{4.21}$$

$$= \int D\lambda \sum_{i,j} e^{i(\lambda_{i_1}+\ldots+\lambda_{i_k}-\lambda_{j_1}-\ldots\lambda_{j_k})t} e^{-\beta(\lambda_{i_1}+\ldots+\lambda_{i_k}+\lambda_{j_1}+\ldots+\lambda_{j_k})}\,. \tag{4.22}$$

In the following subsections, we will compute the LUE spectral form factors and compare analytical results with numerical observations.

At large d, we compute the spectral form factors by Fourier transforming the determinant

[5] The analogous constant in considering the GUE would be fixed to $u = 0$, given the symmetry of the spectrum. However, for the LUE $u = 0$ it is divergent. The value of u specifies the center of the two eigenvalues λ_i and λ_j.

[6] This is slightly different than the standard presentation in the RMT literature, where the form factor is usually given as the Fourier transform of a connected form factor, called the cluster function. Here we work with both connected and disconnected pieces.

of kernels in Eq. (4.17). We integrate the products of K as [23]

$$\int \left(\prod_{j=1}^{n} d\lambda_j \, e^{ik_j\lambda_j}\right) K(\lambda_1,\lambda_2)K(\lambda_2,\lambda_3)\dots K(\lambda_{n-1},\lambda_n)K(\lambda_n,\lambda_1)$$
$$= \alpha_d \int d\lambda \, e^{i\sum_{j=1}^{n} k_j\lambda} \int dk \, g(k) g\Big(k+\frac{k_1}{2\pi\alpha_d}\Big) g\Big(k+\frac{k_2}{2\pi\alpha_d}\Big)\dots g\Big(k+\frac{k_{n-1}}{2\pi\alpha_d}\Big)\,, \quad (4.23)$$

where the Fourier transform of the sine kernel is

$$g(k) \equiv \int dr \, e^{2\pi i k r} \frac{\sin(\pi r)}{\pi r} = \begin{cases} 1 \text{ for } |k| < 1/2 \\ 0 \text{ for } |k| > 1/2\,, \end{cases} \quad (4.24)$$

and where $\alpha_d \equiv d\rho(u)$. The integral over the sine kernel is unbounded and can be treated by imposing a cutoff. We use the *box approximation* [12]

$$\alpha_d \int d\lambda \, e^{i\sum_{j=1}^{n} k_j\lambda} \to \alpha_d \int_{-d/2\alpha_d}^{d/2\alpha_d} d\lambda \, e^{i\sum_{j=1}^{n} k_j\lambda} = d\frac{\sin\big(\sum_{j=1}^{n} k_j/2\rho(u)\big)}{\sum_{j=1}^{n} k_j/2\rho(u)}\,, \quad (4.25)$$

fixed such that Eq. (4.23) over the truncated range with $k_i = 0$ integrates to d. This will be helpful in computing the higher-point spectral form factors, for instance, $\mathcal{R}_4$.

It will also be convenient to define the following functions which will appear in computing the LUE form factors

$$r_1(t) \equiv e^{2it}\big(J_0(2t) - iJ_1(2t)\big)$$
$$r_2(t) \equiv \begin{cases} 1 - \dfrac{t}{2\pi d\rho(u)} & \text{for} \quad 0 < t < 2\pi d\rho(u) \\ 0 & \text{for} \quad t > 2\pi d\rho(u) \end{cases}$$
$$r_3(t) \equiv \frac{\sin\big(t/2\rho(u)\big)}{t/2\rho(u)}\,. \quad (4.26)$$

4.3.2 Two-point form factor at infinite temperature

Let us start with the simplest case, the two point spectral form factor at infinite temperature $\beta = 0$. Pulling out coincident eigenvalues, we have

$$\mathcal{R}_2(t) = \int D\lambda \sum_{i,j} e^{i(\lambda_i - \lambda_j)t} = d + d(d-1) \int d\lambda_1 d\lambda_2 \, \rho^{(2)}(\lambda_1, \lambda_2) e^{i(\lambda_1 - \lambda_2)t} \,. \tag{4.27}$$

The determinant of kernels in Eq. (4.17) gives a squared 1-point function and 2-point function contribution. Using the integration formula in Eq. (4.23), we obtain

$$\mathcal{R}_2(t) = d + d^2 |r_1(t)|^2 - d r_2(t) \tag{4.28}$$

in terms of the functions defined above, and where

$$|r_1(t)|^2 = J_0^2(2t) + J_1^2(2t) \,. \tag{4.29}$$

In Fig. 4.3, we plot the infinite temperature LUE 2-point form factor as derived in Eq. (4.28) along side the GUE form factor (see [12]). Note that unlike in the GUE case there is no dip or ramp. The lack of an intermediate time scale at which the initial slope decay transitions at the dip to a linear growth to a plateau, is due to the slow decay of the 1-point functions which gives the slope.

Subtracting off the contribution from the 1-point functions defines the connected piece of the 2-point form factor

$$\mathcal{R}_2^c(t) \equiv \left\langle |Z(\beta, t)|^2 \right\rangle - \left\langle Z(\beta, t) \right\rangle^2 = d - d r_2(t) \,, \tag{4.30}$$

which exposes the linear growth before the plateau. The connected components are also plotted in Fig. 4.3.

The transition point in the function of r_2 is defined as the plateau time $t_p = 2\pi\alpha_d$, where $\alpha_d = d\rho(u)$. The value of $2\pi\alpha_d$ is not straightforwardly fixed given the unbounded support when integrating over kernels. The constant also determines the linear slope of the ramp function r_2 prior to the plateau. As we discuss in App. 4.A, the constant u is fixed by

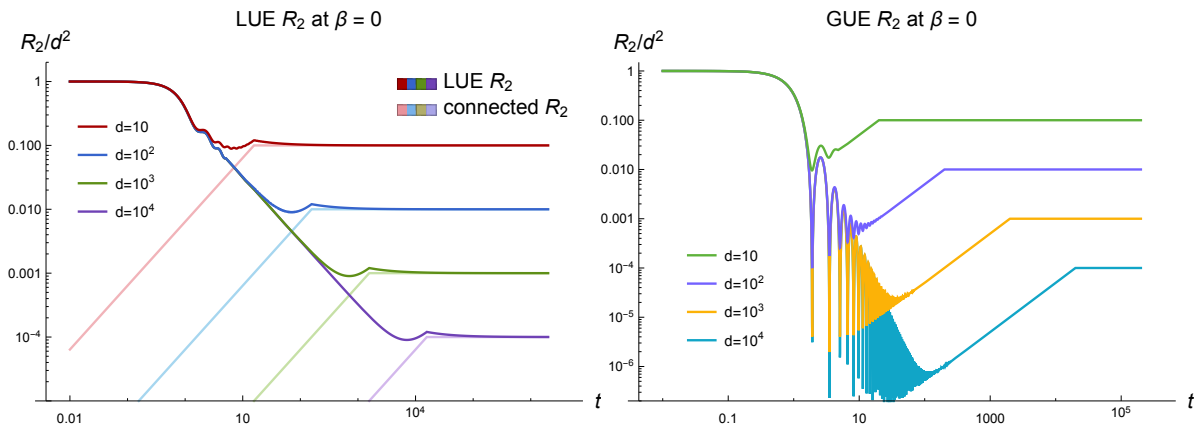


Figure 4.3: On the left: the 2-point spectral form factor and its connected component for the LUE at infinite temperature, as given in Eq. (4.28), plotted for different values of d and normalized by the initial value d^2. We observe the slow $1/t$ decay down to the plateau value, hiding the linear ramp in the connected piece. On the right: the 2-point spectral form factor for the GUE at infinite temperature, with a faster early-time decay exposing the ramp.

numerically fitting to the ramp. We find a plateau time of $t_p \sim \pi d/2$ for the LUE 2-point form factor.

Using the asymptotic form of the Bessel function,

$$J_k(z) \sim \sqrt{\frac{2}{\pi z}} \cos\left(z - \frac{k\pi}{2} - \frac{\pi}{4}\right), \tag{4.31}$$

we conclude that the disconnected piece decays at early times (for t much smaller than d but larger than $\mathcal{O}(1)$) as

$$r_1(t) r_1^*(t) = J_0^2(2t) + J_1^2(2t) \sim \frac{1}{\pi t}\left(\cos^2(2t - \pi/4) + \sin^2(2t - \pi/4)\right) = \frac{1}{\pi t}. \tag{4.32}$$

This $\mathcal{O}(1/t)$ decay of the LUE form factor is to be contrasted with the slower $\mathcal{O}(1/t^3)$ decay in both the GUE and the SYK model [12, 9]. However, the connected piece, dominated by the universal sine kernel in the large d limit, still sees the steady linear rise $\mathcal{O}(t)$ at intermediate time scales. This fact reaffirms the expectation that the decay in the disconnected piece, the Fourier transformed one-point functions, is model dependent. However, the ramp in the connected 2-point function is a universal feature of quantum chaotic systems.

In addition to a hidden dip, another difference with the GUE result is the lack of an

oscillating decay in the LUE at infinite temperature. In the GUE, the Bessel function decay at $\beta = 0$ gives a true dip time $\mathcal{O}(1)$. The envelope of this decay was what we considered as the decay to a dip given that a finite β smoothed out the oscillations.

4.3.3 Two-point form factor at finite temperature

Now let us consider the two point form factor at finite temperature. For small β, one may effectively insert the one point distribution in the integration formula. We walk through the computation in some detail as it will mimic the calculation of the supersymmetric SYK form factor in Sec. 4.5. To be concrete, we write

$$\begin{aligned}
\mathcal{R}_2(t,\beta) &= \int D\lambda \sum_{i,j} e^{i(\lambda_i-\lambda_j)t} e^{-\beta(\lambda_i+\lambda_j)} \\
&= d\int d\lambda\, \rho(\lambda) e^{-2\beta\lambda} + d(d-1)\int d\lambda_1 d\lambda_2\, \rho^{(2)}(\lambda_1,\lambda_2) e^{i(\lambda_1-\lambda_2)t} e^{-\beta(\lambda_1+\lambda_2)} \\
&= d\int d\lambda\, \rho(\lambda) e^{-2\beta\lambda} + \int d\lambda_1 d\lambda_2 \Big(K(\lambda_1,\lambda_1)K(\lambda_2,\lambda_2) - K^2(\lambda_1,\lambda_2)\Big) e^{i(\lambda_1-\lambda_2)t} e^{-\beta(\lambda_1+\lambda_2)} \\
&= dr_1(2i\beta) + d^2 r_1(t+i\beta) r_1(-t+i\beta) - \int d\lambda_1 d\lambda_2 K^2(\lambda_1,\lambda_2) e^{i(\lambda_1-\lambda_2)t} e^{-\beta(\lambda_1+\lambda_2)}\,, \qquad (4.33)
\end{aligned}$$

simply integrating the kernels as specified above. For the final integral, we make the change of variables

$$u_1 = \frac{1}{2}(\lambda_1+\lambda_2)\,, \quad u_2 = \lambda_1 - \lambda_2\,. \qquad (4.34)$$

which allows us to compute

$$\begin{aligned}
&\int d\lambda_1 d\lambda_2\, K^2(\lambda_1,\lambda_2) e^{i(\lambda_1-\lambda_2)t} e^{-\beta(\lambda_1+\lambda_2)} = \int du_1 du_2 \left(\frac{\sin(d\pi u_2)}{\pi u_2}\right)^2 e^{iu_2 t - 2\beta u_1} \\
&\approx \int du_1\, e^{-2\beta u_1}\rho(u_1) \int du_2 \left(\frac{\sin(d\pi u_2)}{\pi u_2}\right)^2 e^{iu_2 t} = dr_1(2i\beta) r_2(t)\,, \qquad (4.35)
\end{aligned}$$

where we regulate the unbounded integral with the insertion of $\rho(u_1)$. The 2-point spectral form factor at finite temperature is

$$\mathcal{R}_2(t,\beta) = dr_1(2i\beta) + d^2 r_1(t+i\beta) r_1(-t+i\beta) - dr_1(2i\beta) r_2(t)\,. \qquad (4.36)$$

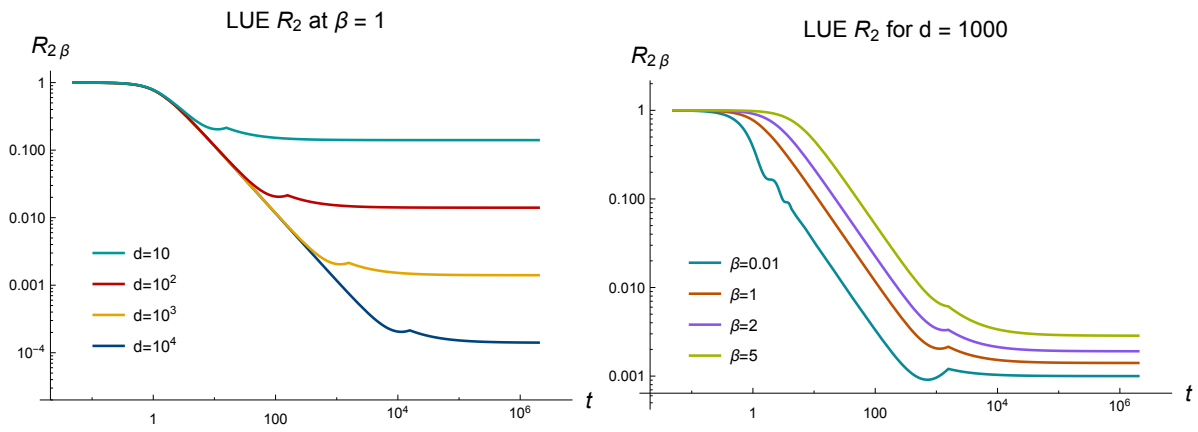


Figure 4.4: The 2-point spectral form factor for LUE at finite temperature, as given in Eq. (4.36), plotted for different values of d and at different temperatures, normalized by the initial value. The plateau value depends on both d and β, while the plateau time is just d dependent.

We plot the analytic result in Fig. 4.4 and observe that at finite temperature there is still no clear dip time in LUE, unlike for the GUE, and that the plateau time t_p does not depend on β. For the LUE, we define $h_1(\beta) \equiv r_1(2i\beta)$, a purely real function of the inverse temperature, with the plateau value

$$\mathcal{R}_2(t_p, \beta) = h_1(2\beta)d\,. \tag{4.37}$$

At small but finite β we have

$$h_1(2\beta) = 1 - 2\beta + 4\beta^2 + \mathcal{O}(\beta^3)\,, \tag{4.38}$$

compared to the GUE result $1 + 2\beta^2 + \mathcal{O}(\beta^4)$ [12], one can see that the LUE plateau value is smaller than GUE, which is also observed in numerics.

4.3.4 Four-point form factor at infinite temperature

As an example of a higher point form factor, we compute the 4-point $\mathcal{R}_4$ at infinite temperature. By definition we have

$$\mathcal{R}_4(t) \equiv \left\langle Z(t)Z(t)Z(t)^*Z(t)^* \right\rangle_{\text{LUE}} = \int D\lambda \sum_{i,j,k,\ell} e^{i(\lambda_i+\lambda_j-\lambda_k-\lambda_\ell)t}\,. \tag{4.39}$$

To evaluate the expression we must consider all possible ways in which the eigenvalues can collide in the sum, i.e. all equal, $\lambda_i = \lambda_j$, $\lambda_k = \lambda_\ell$, etc, and treat them separately. Making use of the 2-point form factors we derived above, and computing the 3 and 4-point function contributions by expanding the determinant and integrating products of kernels as Eq. (4.23), we obtain

$$\begin{aligned}\mathcal{R}_4(t) = d^4|r_1(t)|^4 - 2d^3\mathrm{Re}(r_1^2(t))r_2(t)r_3(2t) - 4d^3|r_1(t)|^2 r_2(t) + 2d^3\mathrm{Re}(r_1(2t)r_1^{*2}(t)) \\ + 4d^3|r_1(t)|^2 + 2d^2 r_2^2(t) + d^2 r_2^2(t)r_3^2(2t) + 8d^2\mathrm{Re}(r_1(t))r_2(t)r_3(t) \\ - 2d^2\mathrm{Re}(r_1(2t))r_3(2t)r_2(t) - 4d^2\mathrm{Re}(r_1^*(t))r_3(t)r_2(2t) + d^2|r_1(2t)|^2 \\ - 4d^2|r_1(t)|^2 - 4d^2 r_2(t) + 2d^2 - 7dr_2(2t) + 4dr_2(3t) + 4dr_2(t) - d\,. \end{aligned} \tag{4.40}$$

In the large d limit, some of the terms above are subdominant or suppressed in d at all times, allowing us to simplify the expression as

$$\mathcal{R}_4(t) \approx d^4|r_1(t)|^4 + 2d^2 r_2^2(t) - 4d^2 r_2(t) + 2d^2 - 7dr_2(2t) + 4dr_2(3t) + 4dr_2(t) - d\,, \tag{4.41}$$

similar to the result we derived for the GUE [12]. At times much earlier than the plateau time, we have

$$\mathcal{R}_4 \approx d^4|r_1(t)|^4 + \frac{t(t-2\pi\rho(u))}{2\pi^2\rho(u)^2} \sim \frac{d^4}{\pi^2 t^2} + \frac{t(t-2\pi\rho(u))}{2\pi^2\rho(u)^2}\,. \tag{4.42}$$

Again, we find a slow decay of $\mathcal{O}(1/t^2)$ and thus no visible dip at large d. The plateau time is still $2\pi\alpha_d$, with a plateau value $\mathcal{R}_4(t_p) = 2d^2 - d \sim 2d^2$.

4.4 Chaos and Wishart matrices

We want to study the chaotic nature of time-evolution by LUE Hamiltonians. Consider the ensemble of unitary time-evolutions generated by LUE random matrices

$$\mathcal{E}_t = \left\{e^{-iHt}\,,\ \text{with } H \in \mathrm{LUE}\right\}. \tag{4.43}$$

We want to understand how random LUE time-evolution is by asking when the ensemble forms a k-design. Computing the frame potential for the ensemble will quantify a distance to Haar-randomness. We also compute correlation functions of operators evolved by the LUE to look at early-time chaos in the chaotic decay of $2k$-point functions.

4.4.1 QI overview

Before discussing the frame potential and measures of chaos for the random matrix ensemble, we will briefly overview the quantum information theoretic concepts and tools we use, namely the notion of a unitary k-design and the frame potential. For a more in-depth review of these in the context of information scrambling in chaotic systems, see [11, 12].

For a finite dimensional quantum mechanical system, with Hilbert space $\mathcal{H}$ of dimension d, the unitary group $U(d)$ can be equipped with the Haar measure, the unique left/right invariant measure on $U(d)$. Given some ensemble of unitary operators $\mathcal{E}$, we say that the ensemble forms a unitary k-design if it reproduces the first k-moments of Haar

$$\int_{\text{Haar}} dU\, (U^{\otimes k})^\dagger(\cdot)U^{\otimes k} = \int_{V\in\mathcal{E}} dV\, (V^{\otimes k})^\dagger(\cdot)V^{\otimes k}\,, \tag{4.44}$$

for any operator. More intuitively, we should think of this as capturing how random the ensemble is, in that the ensemble is sufficiently spread out over the unitary group to reproduce its statistics. A precise measure of Haar-randomness is the frame potential [57], which measures the 2-norm distance between the k-th moments of an ensemble $\mathcal{E}$ and Haar. The k-th frame potential is defined with respect to an ensemble $\mathcal{E}$ as

$$\mathcal{F}^{(k)}_{\mathcal{E}} \equiv \int_{U,V\in\mathcal{E}} dU dV \left|\text{Tr}(U^\dagger V)\right|^{2k}. \tag{4.45}$$

The frame potential for any ensemble $\mathcal{E}$ is lower bounded by the Haar value

$$\mathcal{F}^{(k)}_{\mathcal{E}} \geq \mathcal{F}^{(k)}_{\text{Haar}}\,, \tag{4.46}$$

with equality iff $\mathcal{E}$ forms a k-design. The k-frame potential for the Haar ensemble is simply $\mathcal{F}^{(k)}_{\text{Haar}} = k!$ for $k \leq d$.

The frame potential appeared in the context of information scrambling and black holes as the average of all out-of-time ordered correlators [11]

$$\frac{1}{d^{4k}}\sum_{A\text{'s},B\text{'s}}\left|\left\langle A_1 B_1(t)\dots A_k B_k(t)\right\rangle_{\mathcal{E}}\right|^{2k}=\frac{1}{d^{2(k+1)}}\mathcal{F}_{\mathcal{E}}^{(k)}\,, \tag{4.47}$$

where "$B(t)$" $= UBU^\dagger$ and $U\in\mathcal{E}$, averaged over any ensemble of unitaries $\mathcal{E}$, with each A_i and B_i summed over all Pauli operators. This makes precise an approach to randomness, where the chaotic decay of correlators at late-times means the frame potential becomes small and the ensemble forms a k-design.

4.4.2 Frame potentials

First frame potential at $\beta = 0$

We start by computing the first frame potential at infinite temperature $\mathcal{F}_{\mathcal{E}}^{(k)}$ for the ensemble of LUE time-evolutions. Following [12], we have

$$\mathcal{F}_{\text{LUE}}^{(k)}=\int dH_1 dH_2\, e^{-\frac{d}{2}\text{Tr}H_1^2}e^{-\frac{d}{2}\text{Tr}H_2^2}\left|\text{Tr}\left(e^{iH_1 t}e^{-iH_2 t}\right)\right|^2. \tag{4.48}$$

Using the unitary invariance of the ensemble and integrating using the second moment of the Haar ensemble, we find

$$\mathcal{F}_{\text{LUE}}^{(k)}=\frac{1}{d^2-1}\left(\mathcal{R}_2^2+d^2-2\mathcal{R}_2\right), \tag{4.49}$$

with the same dependence on the form factors as in the GUE case.

In Fig. 4.5 we plot our analytic form of the first frame potential of the LUE at infinite temperature. We can see that there are significant differences between the supersymmetric and non-supersymmetric cases. The slow decay of the LUE means there the ensemble does not form a k-design at the dip. At late-times after the plateau, we find the frame potential approaches a value of 2.

First frame potential at finite β

We can also generalize the frame potential to finite temperature by averaging over all thermal $2k$-point functions with operators spaced equidistant on the thermal circle (i.e. inserting $\rho^{1/2k}$ between operators in the $2k$-OTOC). Averaging over operators, we find [11]

$$\mathcal{F}^{(k)}_{\mathcal{E}_\beta} = \int_{\mathcal{E}} dH_1 dH_2 \, \frac{\left|\mathrm{Tr}\left(e^{-(\beta/2k-it)H_1}e^{-(\beta/2k+it)H_2}\right)\right|^{2k}}{\mathrm{Tr}(e^{-\beta H_1})\mathrm{Tr}(e^{-\beta H_2})/d^2}\,, \tag{4.50}$$

with the normalization that gives the standard frame potential as $\beta \to 0$. For the LUE, we compute the finite temperature frame potential just as above, Haar integrating to find

$$\mathcal{F}^{(1)}_{\mathrm{LUE}}(t,\beta) = \frac{1}{d^2-1}\left(\widetilde{\mathcal{R}}_2^2(\beta/2) + d^2 - 2\widetilde{\mathcal{R}}_2(\beta/2)\right), \tag{4.51}$$

where we define a slightly more conveniently normalized form factor

$$\widetilde{\mathcal{R}}_2(t,\beta) = \int D\lambda \, \frac{\sum_{ij} e^{it(\lambda_i-\lambda_j)}e^{-\beta(\lambda_i+\lambda_j)}}{\sum_i e^{-2\beta\lambda_i}/d}\,. \tag{4.52}$$

As it is more analytically tractable, we opt to separately average the numerator and denominator (the 'quenched' version), and checked numerically that the results are in good agreement. We see that at early times, near $t = 0$, we have the β-dependent value

$$\mathcal{F}^{(1)}_{\mathrm{LUE}} \approx d^2 \frac{h_1(\beta/2)^4}{h_1(\beta)^2}\,, \tag{4.53}$$

while at late times, after the plateau time, we have $\mathcal{F}^{(1)}_{\mathrm{LUE}}(t_p,\beta) = 2$.

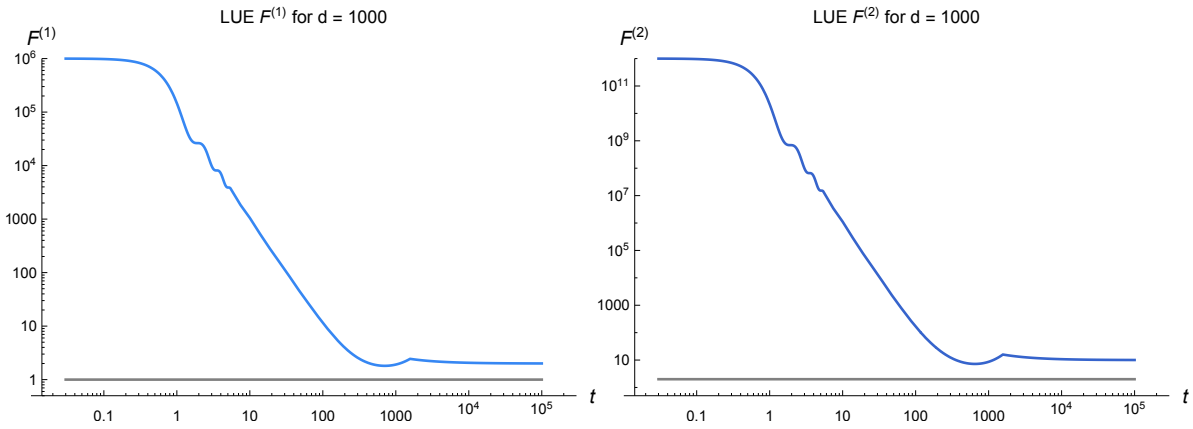


Figure 4.5: We show the first and second frame potentials for the LUE at infinite temperature at $d = 1000$. The slow decay means we do not form a k-design at the dip time. For comparison, the Haar value is plotted in grey.

Second frame potential at $\beta = 0$

The second frame potential for the LUE at infinite temperature is expressed in terms of the spectral form factors as [12]

$$\begin{aligned}\mathcal{F}^{(2)}_{\text{LUE}} = \frac{1}{(d^2-9)(d^2-4)(d^2-1)d^2}\Big(&\left(d^4-8d^2+6\right)\mathcal{R}_4^2 + 4d^2\left(d^2-9\right)\mathcal{R}_4 \\ &+ 4\left(d^6-9d^4+4d^2+24\right)\mathcal{R}_2^2 - 8d^2\left(d^4-11d^2+18\right)\mathcal{R}_2 - 4d^2\left(d^2-9\right)\mathcal{R}_{4,2} \\ &+ \left(d^4-8d^2+6\right)\mathcal{R}_{4,2}^2 + 2\left(d^4-7d^2+12\right)\mathcal{R}_{4,1}^2 - 8\left(d^4-8d^2+6\right)\mathcal{R}_2\mathcal{R}_4 \\ &- 4d\left(d^2-4\right)\mathcal{R}_4\mathcal{R}_{4,1} + 16d\left(d^2-4\right)\mathcal{R}_2\mathcal{R}_{4,1} - 8\left(d^2+6\right)\mathcal{R}_2\mathcal{R}_{4,2} \\ &+ 2\left(d^2+6\right)\mathcal{R}_4\mathcal{R}_{4,2} - 4d\left(d^2-4\right)\mathcal{R}_{4,1}\mathcal{R}_{4,2} + 2d^4\left(d^4-12d^2+27\right)\Big)\,, \end{aligned} \tag{4.54}$$

where we have defined

$$\mathcal{R}_{4,1}(t) \equiv \int D\lambda \sum_{i,j,k=1}^{d} e^{i(\lambda_i+\lambda_j-2\lambda_k)t}\,, \qquad \mathcal{R}_{4,2}(t) \equiv \int D\lambda \sum_{i,j=1}^{d} e^{2i(\lambda_i-\lambda_j)t}\,. \tag{4.55}$$

The 4-point form factor with two coincident eigenvalues, $\mathcal{R}_{4,2}(t)$, is simply $\mathcal{R}_2(2t)$. The 3-point form factor $\mathcal{R}_{4,1}(t)$ for the LUE can be computed just as in Sec. 4.3, where we find

$$\begin{aligned}\mathcal{R}_{4,1}(t) = d^3\mathrm{Re}\big(r_1(2t)r_1^{*2}(t)\big) - d^2\mathrm{Re}(r_1(2t))r_3(2t)r_2(t) - 2d^2\mathrm{Re}(r_1^*(t))r_3(t)r_2(2t) \\ + d^2|r_1(2t)|^2 + 2d^2|r_1(t)|^2 + 2dr_2(3t) - dr_2(2t) - 2dr_2(t) + d\,. \end{aligned} \tag{4.56}$$

We plot the second frame potential for LUE alongside the first frame potential in Fig. 4.5. The second frame potential has an initial value of d^4 and late-time value of 10, just as for the GUE. But again the difference arises at intermediate time scales, where the LUE fails to form a k-design.

4.4.3 Correlation functions

As we discussed before, the recent interest in quantum chaos has involved extensive discussion of out-of-time order correlation functions (OTOCs). Namely, the following 4-point functions of pairs of operators in thermal states

$$\langle AB(t)AB(t)\rangle_\beta \quad \text{where} \quad B(t) = e^{-iHt}Be^{iHt}\,. \tag{4.57}$$

We consider OTOCs with operators evolved by LUE Hamiltonians and averaged over the random matrix ensemble. In [12], we studied $2k$-OTOCs and related them to spectral quantities, both by averaging over the operators in the correlation function or over an ensemble of Hamiltonians. In that work, we averaged $2k$-OTOCs over the GUE and related the correlators to spectral quantities using the unitary invariance of the measure. As the LUE is similarly invariant, the relation between correlation functions averaged over the random matrix ensemble and the form factors will be the same as thus parts of the discussion here will closely follow [12]; the differentiating aspects of LUE time-evolution thus lie in the spectral form factors themselves.

First we look at the 2-point function and integrate over Hamiltonians drawn from the LUE, using the unitary invariance of the measure and Haar integrating in the eigenvalue

basis

$$\langle AB(t)\rangle_{\text{LUE}} = \int dH \langle AB(t)\rangle = \frac{\mathcal{R}_2(t)-1}{d^2-1}\langle AB\rangle_c + \langle A\rangle\langle B\rangle\,, \tag{4.58}$$

where $\langle AB\rangle_c$ denotes the connected correlator. For non-identity Paulis, the expression is nonzero for $B = A^\dagger$, and thus

$$\text{LUE average}: \quad \langle AA^\dagger(t)\rangle_{\text{LUE}} \approx \frac{\mathcal{R}_2(t)}{d^2}\,, \tag{4.59}$$

for $\mathcal{R}_2(t) \gg 1$. We note that, just as is the case for GUE, if we instead average the same 2-point function over all operators A, we arrive at the same expression

$$\text{Operator average}: \quad \int dA\, \langle AA^\dagger(t)\rangle = \frac{\mathcal{R}_2(t)}{d^2}\,, \tag{4.60}$$

which is true regardless of the Hamiltonian. The fact that the LUE averaged 2-point function equals the operator averaged correlator means that LUE does not care about the size or locality of the operator A, given that we made no assumptions about about A in computing Eq. (4.59), and thus is blind to phenomena relevant for early-time chaos such as operator growth.

We next compute the 4-point OTOC averaged over the LUE, using the fourth moment of Haar and looking at the leading order behavior

$$\langle AB(t)AB(t)\rangle_{\text{LUE}} = \int dH\, \langle AB(t)AB(t)\rangle \approx \frac{\mathcal{R}_4(t)}{d^4}\,, \tag{4.61}$$

for non-identity Pauli operators A and B. Note that the OTOCs of the form $\langle AB(t)CD(t)\rangle$ are all almost zero unless $ABCD = \mathbb{I}$.

We can now comment on the time scales that LUE describes as seen from the averaged correlation functions. The time scale of 2-point function decay corresponds to the time scales for which the system thermalizes. Using the early time piece of the 2-point form factor we derived in Sec. 4.3, where the contribution from the 1-point function gives the decay

$$\langle AA^\dagger(t)\rangle_{\text{LUE}} \approx J_0^2(2t) + J_1^2(2t) \sim \frac{1}{\pi t}\,, \tag{4.62}$$

contrasted to the $1/t^3$ decay for GUE. Similarly, we can comment on scrambling in the LUE by looking at the early time decay of the LUE averaged 4-point OTOCs. The early time behavior of the 4-point form factor means the OTOC decays like

$$\langle AB(t)AB(t)\rangle_{\text{LUE}} \approx \left(J_0^2(2t) + J_1^2(2t)\right)^2 \sim \frac{1}{\pi^2 t^2}\,. \tag{4.63}$$

The characteristic time-scale for decay of LUE 2-point functions is $t_2 \sim \mathcal{O}(1)$, or for systems at finite temperature $\mathcal{O}(\beta)$. The time-scale for 4-point function decay is also order 1, but faster than the decay of 2-point functions $t_4 \sim t_2/2$. Although the decay is slower than for GUE, unsurprisingly, the conclusion about the LUE's perception of early-time chaos is the same: the LUE 4-point OTOCs decay faster than the LUE 2-point functions, which means the random matrix ensemble fails to describe scrambling at early times.

4.4.4 Complexity

Lastly, we briefly comment on the complexity growth under time-evolution of LUE Hamiltonians. Here we simply discuss the results; details and definitions of ensemble complexity and its relation to the frame potential are given in [11, 12]. The gate complexity of an ensemble $\mathcal{E}$, i.e. the number of gates needed to generate $\mathcal{E}$, is lower bounded by the frame potential as

$$\mathcal{C}(t) \geq \frac{2kn - \log \mathcal{F}_{\mathcal{E}}^{(k)}(t)}{2\log n}\,. \tag{4.64}$$

At early times before the dip time $t \ll t_d$, the dominant contribution to the k-th frame potential is $\mathcal{F}_{\mathcal{E}}^{(k)} \simeq \mathcal{R}_{2k}^2(t)/d^{2k}$ [12]. For $k \ll d$, the $2k$-th form factor goes as $\mathcal{R}_{2k} \sim r_1^{2k}$, the function defined in Eq. (4.26) in terms of Bessel functions. The decay $r_1^2 \sim 1/t$, gives a lower bound on the growth of the circuit complexity

$$\mathcal{C}(t) \geq \mathcal{O}\left(\frac{k\log t}{\log n}\right)\,, \tag{4.65}$$

where the slower decay for LUE still gives the same logarithmic lower bound as GUE. Interestingly, in GUE the 1-point function contribution to the form factor at early times is an

oscillating Bessel function decay $J_0^2(2t)/t^2$, which formally gives a dip time $\mathcal{O}(1)$. As these oscillations are not present in the LUE, we can bound the complexity up to the dip time even at infinite temperature. But for large k, we recover the quadratic growth of complexity: $\mathcal{C} \geq t^2/\log n$, hinting again at the unphysical nature of LUE evolution at early times.

4.5 Chaos in supersymmetric SYK

The supersymmetric SYK model admits a classification by Wishart-Laguerre random matrix ensembles and has a density of states which closely follows a Marčenko-Pastur distribution [107]. Having discussed the properties of LUE random matrices, we turn to the supersymmetric SYK model and check that the form factor acts similarly. From the frame potential, we then discuss the Haar-randomness of the model's time evolution.

Assuming that the spectral statistics of the theory are Gaussian, as both SYK and the Wishart matrices are, allows us to use the sine kernel to compute the spectral n-point functions. We note that if the statistics are GUE/GOE/GSE, the sine kernel is slightly modified and the ramp function differs as we approach $t \sim d$, but the universal growth of the ramp is still present. Knowing that the supersymmetric SYK model has Gaussian spectral statistics [107], we can compute the finite temperature form factor for the theory just as in Eq. (4.33), and find

$$\begin{aligned}\mathcal{R}_2(t,\beta) = \big\langle Z(\beta+it)Z(\beta-it)\big\rangle &= \int D\lambda \sum_{i,j} e^{i(\lambda_i-\lambda_j)t} e^{-\beta(\lambda_i+\lambda_j)} \\ &\approx d\int dE\,\rho(E)e^{-2\beta E} + \big|\langle Z(\beta+it)\rangle\big|^2 - d\int dE\, e^{-2\beta E}\rho(E) r_2(t)\,, \end{aligned} \tag{4.66}$$

where $r_2(t)$ is the ramp function from the LUE and we define $E = \frac{1}{2}(\lambda_1+\lambda_2)$. Continuing, we find the finite temperature form factor

$$\mathcal{R}_2(t,\beta) \approx |\langle Z(\beta+it)\rangle|^2 + Z(2\beta)\big(1-r_2(t)\big)\,. \tag{4.67}$$

As a sanity check, the late-time value $Z(2\beta)$ here matches the infinite-time average of the spectral form factor. As we discussed in Sec. 4.2, the 1-loop partition function from the

super-Schwarzian theory is

$$Z^{\text{sSch}}_{\text{1-loop}}(\beta) \sim \frac{1}{\sqrt{\beta \mathcal{J}}} e^{Ns_0 + cN/2\beta}\,, \tag{4.68}$$

where s_0 is the ground-state entropy density and c is the specific heat. At early times, the form factor is dominated by its disconnected component, decaying as $1/t$

$$\text{Early}: \quad \mathcal{R}_2(t,\beta) \sim \frac{e^{2Ns_0}}{\mathcal{J}t} \tag{4.69}$$

for times greater than $t \sim \sqrt{N} = \log d/2$, but shorter than $t \sim \sqrt{d}$. Computing the connected form factor, we find

$$\begin{aligned} \mathcal{R}_2^c(t,\beta) &\equiv \big\langle Z(\beta+it)Z(\beta-it)\big\rangle - |\langle Z(\beta+it)\rangle|^2 \\ &= Z(2\beta)\big(1-r_2(t)\big) = \frac{1}{\sqrt{2\beta\mathcal{J}}} e^{Ns_0 + cN/4\beta}\big(1-r_2(t)\big)\,. \end{aligned} \tag{4.70}$$

Equating the $1/t$ decay with the ramp gives a dip time $t_d \sim e^{Ns_0}$, the same order as the plateau time t_p. Even in light of the exactness of the super-Schwarzian theory, we should be cautious in extrapolating to very late times. It is possible that in the large N theory the slope is not well-described by the effective theory at late times and, in turn, decays faster at an intermediate time scale.

Lastly, to get a hint at the nature of scrambling and an approach to randomness in SYK and its supersymmetric extension, we numerically plot the first frame potential for each in Fig. 4.6 at infinite temperature and for $N = 16$ Majoranas. The faster decay and dip that appears for SYK means the frame potential decays quickly, forming an approximate k-design at the dip time. Although the dip value of the SYK frame potential for $N = 16$ is larger than the Haar value, we checked that as we increase N the dip value decreases and expect that SYK forms an approximate k-design in the large N limit. The frame potential for the supersymmetric model exhibits a much more gradual approach to its minimal value which is larger than in SYK, indicating less effective information scrambling and a greater distance of the ensemble to forming a k-design. It would be interesting to see, either numerically or analytically, if these behaviors persist at large N. Both theories, like their random matrix

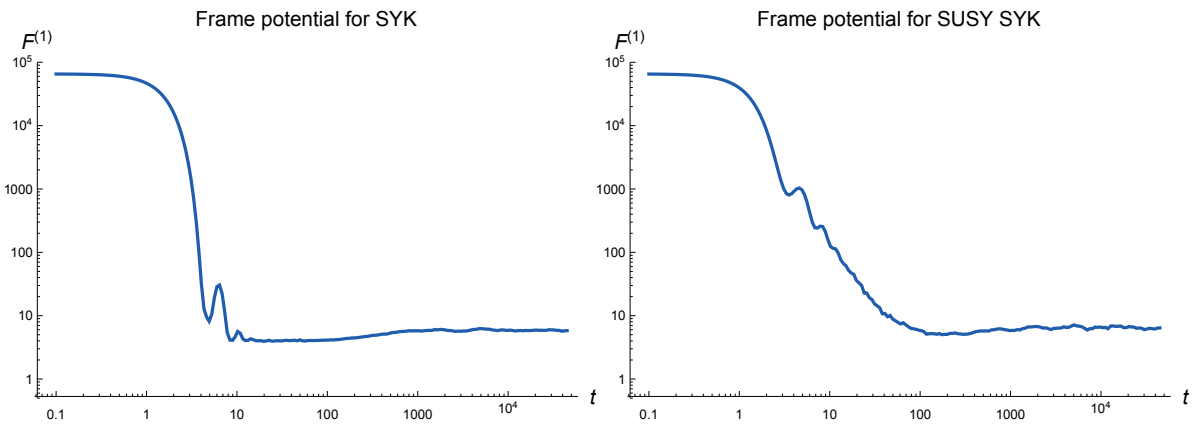


Figure 4.6: Numerics for the first frame potential of SYK and supersymmetric SYK at $\beta = 0$ for $N = 16$ Majoranas and 200 samples. The decay and dip of SYK indicates faster scrambling and an approximate k-design behavior not as readily apparent in the supersymmetric model.

counterparts, become less random and increase after the dip, deviating further from an approximate design, which suggests that k-invariance [12] might provide a better insight in how information scrambles in SYK models.

There are a few comments worth making relating the discussion here with the behavior of the form factor in similar models.[7] In the complex SYK model, the spectral form factor appears to have a $1/t^4$ power-law decay at early times [118, 73], in contrast to the Majorana and SUSY SYK models. As we discussed, the respective power-law decays in these models arise from the Schwarzian and super-Schwarzian modes governing the low-energy physics, and persist for a long time as a result of the 1-loop exactness of the effective actions. In the complex SYK model, where we have a conserved $U(1)$, there is an additional contribution to the effective action from the phase fluctuations of the reparametrization mode, as was discussed in [118]. Combined with the contribution from the Schwarzian mode, the partition function has a $Z(\beta) \sim 1/(\beta\mathcal{J})^2$ dependence. Continuing to real-time, the early-time contribution to the 2-point form factor gives a power-law decay $\mathcal{R}_2(t) \sim |Z(\beta, t)|^2 \sim 1/t^4$. As the low-energy description is likely also 1-loop exact, one expects this behavior to persist for a long time. It is further interesting to note that while the power-law indicates a more rapid onset of late-time chaos as seen by the frame potential, the additional $U(1)$-mode does

[7] We thank an anonymous JHEP referee for raising these points.

not contribute to the Lyapunov exponent of the theory [119]. Thus, like Majorana SYK and SUSY SYK models, the complex SYK model is maximally chaotic at early times, but in the above sense scrambles quicker.

We should also comment on the behavior of spectral quantities more generally in chaotic systems with gravitational duals. In 2d CFTs, an analysis of the contribution from different saddles indicates a persisting $1/t^3$ decay in the form factor for holographic CFTs, and a $1/t$ decay for rational CFTs [39].[8] A slow decay of spectral quantities also appears in the D1-D5 theory at the orbifold point, in line with the fact that the theory does not have chaotically decaying correlation functions [121] and appears to exhibit a logarithmic ramp [40], in contrast to the universal linear ramp we expect in chaotic systems. Although [9] argued for the rapid decay of spectral functions and the late-time appearance of a ramp in super Yang-Mills at strong-coupling, better analytic control of spectral quantities is needed to understand quantum chaos in holographic theories.

On a slightly less related note, a recent work [122] also considered the infinite temperature 2-point spectral form factor for Wishart matrices in a different context. Namely, they studied the statistical properties of the reduced density matrix on spatial regions in quantum many-body systems. They also comment on universal features of Wishart matrices in Floquet systems. As there is a sense in which Floquet systems may be thought of as supersymmetric quantum mechanics [123], where the Floquet unitary is built from two 'supercharges', it would be interesting to explore further connections with our work.

4.6 Conclusion and outlook

In this chapter, we considered the Wishart-Laguerre unitary ensemble in order to understand universal features of supersymmetric quantum mechanical systems. We computed the 2-point spectral form factor for the LUE and found the one-point function contribution gives a $1/t$ power law decay at early times, hiding the dip and transitioning directly into the plateau. This is relatively slow compared to the $\sim 1/t^3$ decay seen in both SYK and the GUE. The

[8] Relatedly, [120] discussed a distinction between entanglement scrambling in rational and holographic CFTs.

universal ramp behavior from the sine kernel can be seen in the connected LUE 2-point form factor. These results agree with the prediction from the 1-loop partition function in supersymmetric SYK. This slow decay implies the onset of a random matrix description occurs at much later times. This can best be seen from the frame potential, where we find a more gradual decay to Haar-random dynamics. Moreover, the frame potential for the LUE, unlike that of the GUE, does not reach the Haar value and does not form an approximate k-design. This is also what we predict and observe numerically in the supersymmetric SYK model, where the slower decay and larger dip value imply less effective information scrambling.

The supersymmetric model, while maximally chaotic, sees a slower onset of random matrix behavior—made evident by the lack of a dip in the form factor and by the slow approach to Haar-randomness in the frame potential. The apparent distinction here between early-time chaos, in terms of chaotic correlation functions, and late-time chaos, in terms of scrambling and Haar-randomness, demands a deeper understanding.

4.A Numerics

In this appendix we discuss numerics to fix an analytic form of the form factors for LUE and to further provide checks on the expressions we derived for the form factors and frame potentials. As we mentioned in Sec. 4.3, there was a free parameter u in the expressions we derived for the k-point form factors. This dependence appears in the ramp function $r_2(t)$, defined in Eq. (4.26), and determines both the slope of the linear ramp in $\mathcal{R}_2^c(t)$ and the plateau time. Numerically computing the connected 2-point form factor for $d = 500$, we fix u by fitting the ramp between times $\sim$1 and $\sqrt{d}/2$. We know that the early time behavior of the ramp is quadratic before $t \sim 1$ and expect a loss of analytic control as we approach the plateau time. We thus linearly fit points in this intermediate regime and find $u = 1.156$. We hope to derive this result more rigorously in the future.

We also present some numerical checks of our expressions for the LUE 2-point form factor in Fig. 4.7, where we find good agreement in the slope, ramp, and plateau. Our results were derived for LUE at large d and thus should capture the perturbative behavior.

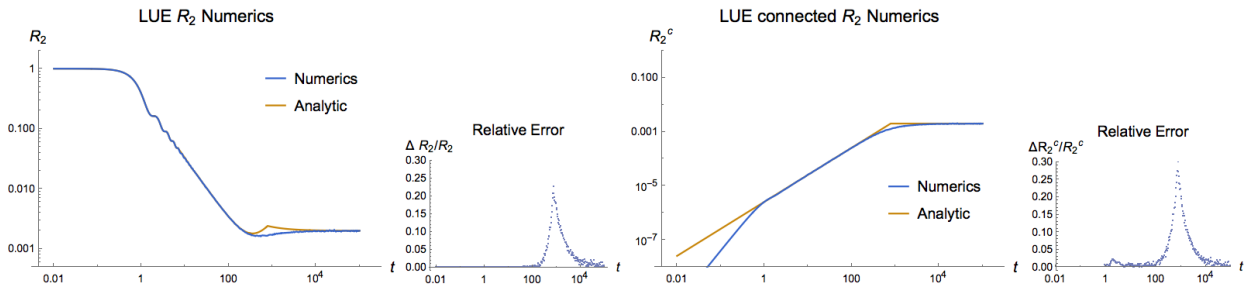


Figure 4.7: Numerics for both the LUE 2-point form factor and its connected component, compared to the analytic expressions derived in Sec. 4.3, for $d = 500$ and with 10000 samples. We find good agreement in the slope and plateau, with expected deviations around the plateau time. The very early time behavior of the connected form factor can also be understood analytically.

But in the transition to the plateau, nonperturbative effects [115] become important and our results deviate from numerics in this regime. After the plateau time, we return to contributions from the 1-point function. At very early times, before $t \sim \mathcal{O}(1)$, the connected component grows as $\mathcal{R}_2^c(t) \sim t^2$. This quadratic growth can be derived from an impressive integral representation of the connected 2-point form factor [68]. We have also checked our expressions of the finite temperature and higher point LUE spectral functions and found good agreement with numerics.

Chapter 5

Operator growth in random quantum circuits

This chapter is partially based on work in progress, which will appear elsewhere in a future publication.

Abstract

We study random quantum circuits with symmetry, where the local 2-site unitaries are drawn from a quotient of the full unitary group $U(d)$. Random quantum circuits are minimal models of unitary local chaotic dynamics and can be used to study operator growth and the emergence of dissipative hydrodynamics. We derive the transition probabilities of the local Markov process on Pauli strings in the operator growth for five classes of symmetric random circuits. We then compute the butterfly velocities and diffusion constants for a spreading operator by solving a simple random walk in each class of circuits.

5.1 Introduction

In this chapter we study operator growth in local random quantum circuits with symmetry, where instead of Haar random 2-site unitaries, we construct the circuits using unitaries drawn from a quotient of the unitary group. This is a simple application of the tools developed in Ch. 3, making use of Weingarten calculus for compact Lie groups and compact symmetric

spaces. Here we consider 5 different symmetry classes of random circuits: $O(d)$, $Sp(d)$, AI: $U(d)/O(d)$, and AII: $U(d)/Sp(d)$. In each of the five cases we build a one-dimensional random circuit with random gates from the group or symmetric space. Each of the symmetric random circuits defines a different Markov process of local updates on the evolving Pauli strings, which gives rise to operator growth under random circuit evolution.

Random quantum circuits have been studied extensively in the quantum information community, largely focused on understanding the convergence properties, for instance [124, 125, 126, 54]. More specifically, [126] showed that random quantum circuits form approximate unitary 2-designs in polynomial depth. [54] expanded on these results using the spectral gap of the moment operator. [77] studied decoupling in random quantum circuits and showed that decoupling occurs in circuits of polylogarithmic depth. More recently, in a nice series of papers random quantum circuits have been used to study entanglement growth [127] and operator spreading under random unitary dynamics [128, 129] and with conservation laws [130, 131]. The study of ballistic operator growth in random circuits contrasts interestingly with analytic results for weakly interacting metals [132, 133] and numerical results for spin-chains [134]. Moreover, random quantum circuits also give an emergent picture for the evolution of entanglement under unitary dynamics [135, 136].

The random quantum circuit models considered in this chapter are just a simple generalization of the constructions in [128, 129]. Moreover, while our construction of symmetric circuits builds in symmetry to the local random unitary, the random circuit as a whole does not obey any conservation law. Therefore, almost by construction, we do not see the beautiful picture that emerges in [130, 131]. There the random quantum circuit models obey a local conservation law, where the block diagonal random unitary preserves local z-spin, i.e. is Haar-random within fixed charge sectors of the z-basis. In these models, they find an emergent coupled diffusion process, where non-conserved operators propagate ballistically, with fronts that spread diffusively, and where conserved charge dissipates diffusively as conserved operators decay to non-conserved operators which again propagate ballistically. This coupled process gives rise to long hydrodynamic power-law tails in the weight of the evolving operator.

Instead, the models we consider here instead give rise just to the ballistic growth of an

operator and the diffusive spreading of the ballistic front. Whereas operator growth in the unitary random circuits can be understood as a simple biased random walk, in the symmetric circuits we find that the endpoint dynamics are governed by different types of random walks. The endpoint of evolving strings in the AI and AII random circuits are governed by a biased random walk with correlation between steps. The edge of an evolving Pauli string in the $O(d)$ and $Sp(d)$ random circuits is governed by a biased correlated random walk with an internal state. In each case we derive the butterfly velocity and diffusion constant of the evolving operator.

We will start by reviewing basic properties of unitary random quantum circuits, discussing the local Markov process and operator growth. We then discuss the Markov process in the symmetric random circuits, deriving the respective transition probabilities of the Markov process updating evolving Pauli strings. Then we discuss operator growth in the symmetric random circuits and derive the butterfly velocity and diffusion constant by solving a random walk governing the endpoint dynamics of operator growth.

In this chapter we will not explicitly review the formalism for taking Haar-averages. Weingarten calculus for Lie groups and compact symmetric spaces was reviewed and discussed extensively in Ch. 3. The methods for performing Haar averages over compact Lie groups were worked out in [58, 59, 88], and extended to compact symmetric spaces in [91, 92]. In computing the averages over compact symmetric spaces, we have relied heavily on the methods developed there.

Random quantum circuits

Consider a one-dimensional chain of n qudits, with local dimension q, evolved by a random circuit built from layers of 2-site unitaries. Each layer or time-step of the random circuit alternates between acting on all even links between qudits at even time-steps, and odd links at odd time steps. Explicitly, the t-th layer of the circuit for even t, $U_{t,\text{even}}$, is given by the tensor product of 2-site unitaries $U_{i,i+1}$ for even i, and the t-th layer for odd t by tensoring $U_{i,i+1}$ for odd i. We denote the unitary implementing the t-th layer as U_t, and the evolution to time t is simply the product of t layers of the circuit. The architecture of the random quantum circuit is shown in Fig. 5.1. Each 2-site unitary $U_{i,i+1}$ is drawn at random from the

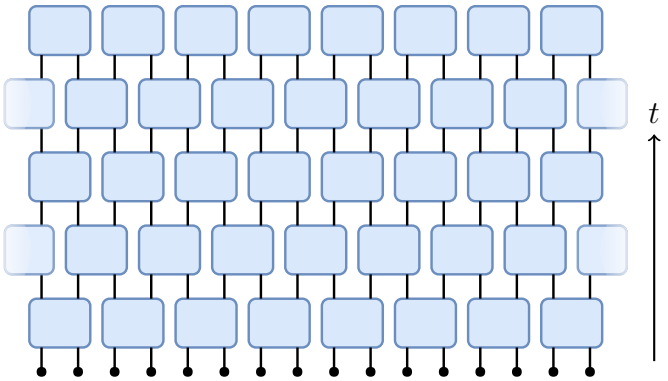


Figure 5.1: Random quantum circuits built from staggered layers of 2-site unitaries and where each unitary is drawn at random from the unitary group, or a subgroup or quotient thereof.

unitary group $U(q^2)$ or from some subgroup or quotient.

We'll go over the basic structure of operator growth in random circuits to set the stage for the rest of the discussion in this chapter. The discussion here is general and applies to all classes of random circuits in this chapter. Recall that we can expand any operator in a basis of operators (like Pauli strings for qubit systems) as

$$\mathcal{O}(t) = \sum_p \gamma_p(t)\mathcal{O}_p\,. \tag{5.1}$$

We will refer to the elements in the basis as Pauli strings regardless of local dimension; the only difference is that generalized Paulis are no longer Hermitian, but this does not affect the discussion. We should think about the coefficients $\gamma_p(t)$ as probabilities of finding the operator $\mathcal{O}_p$ in the growing operator, or equivalently the weight of the growing operator on a given Pauli string. Unitary evolution $U_t\mathcal{O}_pU_t^\dagger$ and the orthonormality of Paulis $\frac{1}{q^{2n}}\mathrm{Tr}(\mathcal{O}_a\mathcal{O}_b) = \delta_{ab}$ means the operator norm is conserved under time-evolution, and thus the probabilities are conserved $\sum_p |\gamma_p(t)|^2 = 1$.

Consider the evolution of a local operator $\mathcal{O}_0$ in the random circuit. The local 2-site unitaries will act on the operator and the range of its support will grow ballistically, spreading outwards by one site at each side every time step, as shown in Fig. 5.2. While the operator will grow to a linear combination of all Pauli strings with support on the $2t+1$ sites at time t in the light-cone of the operator, we want to know the distribution on those Pauli strings,

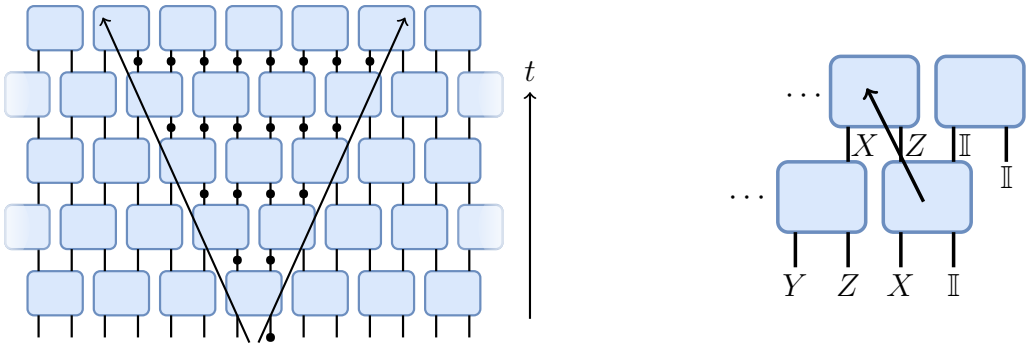


Figure 5.2: Operator growth in random quantum circuits. On the left: the ballistic growth of the support of an initially local operator. On the right: an example of the right edge of a Pauli string moving back after evolution by a layer of the circuit.

i.e. the shape of the support of that operator on Pauli strings. We can define the weight of the evolving operator on its left/right edge as [128, 129]

$$\rho_{\mathrm{L/R}}(i,t) = \sum_{p \in \mathcal{O}_{\mathrm{L/R}}(i)} |\gamma_p(t)|^2 \tag{5.2}$$

where we sum over all Pauli strings with the left or right-most non-identity Pauli operator at site i. We will see that the edges of the growing operator grow ballistically but also spread diffusively, such that $\rho_{\mathrm{L/R}}(i,t)$ obeys a simple biased diffusion equation.

As we will discuss, we can solve for this distribution of the support of the operator by studying the random walking of the edges of the Pauli strings. The random circuits define a Markov process governing the internal dynamics of the growing operator, where each Pauli string in the growing operator is stochastically updated to another Pauli string. As we will discuss in the next section, each local 2-site unitary gives a Markov process on two site pairs of the Pauli string, taking identities to themselves and non-identity Paulis uniformly to all other non-identity Paulis. For instance, two sites in a Pauli string might get updated like $(XY) \to (ZX)$. An update like this occurs for each two sites, i.e. at every link, alternating between even and odd links in the circuit.

We can understand the distribution of the each of the evolving operator by thinking about the edges of the evolving Pauli strings. In the random circuit evolution, each Pauli string is random walking through the space of Pauli strings. Moreover, the end of each Pauli

string has some probability of moving back. For instance, say the two Pauli operators at the right-most gate, which doesn't simply act on identities, is $(X\,\mathbb{I})$. There is a probability that this gets updated to $(Z\,\mathbb{I})$ and the gate to the right in the following time-step has only identities going in. Thus in the coordinate system on the links, the operator moves back, as shown in Fig. 5.2. This is a simple biased random walk on the edge of a Pauli string.

A note on notation and terminology:

To prevent any confusion, we make very clear the notation used to denote local dimensions. We will consider one-dimensional chain of qubits, where the local dimension is 2, but more generally, we denote the local dimension in qudit chains as q. In this chapter, we define the dimension of a 2-site unitary as $d \equiv q^2$, so that the local unitaries are drawn from $U(d)$ (and $U(4)$ in the case of qubits). We will also refer to the basis of operators as Pauli strings, regardless of the local dimension. Lastly, in previous chapters we defined the symplectic group with even dimension $2d$, i.e. defined the symplectic group as the intersection of symplectic matrices $Sp(2d, \mathbb{C})$ and the unitary group. Here, for consistency of the discussion and ease in comparing formulae, we denote the symplectic group $Sp(d)$ and the compact symmetric space AII: $U(d)/Sp(d)$, keeping in mind that d must be taken to be even in these two cases.

5.2 Operator growth in unitary random circuits

Let's start by reviewing the story for unitary random circuits, where each local 2-site unitary in Fig. 5.1 is drawn Haar-randomly from $U(d)$. Consider the evolution of some local operator $\mathcal{O}(t)$, at time zero acting on a single site, by the unitary random circuit. We can understand the growth of this operator in the random circuit by solving for the coefficients $|\gamma_p(t)|^2$.

$$|\gamma_p(t)|^2 = \frac{1}{q^{2n}}\left|\mathrm{Tr}\big(\mathcal{O}(t)\mathcal{O}_p\big)\right|^2 = \frac{1}{q^{2n}}\mathrm{Tr}\big(U_t\mathcal{O}(t-1)U_t^\dagger\mathcal{O}_p\big)\mathrm{Tr}\big(U_t\mathcal{O}(t-1)U_t^\dagger\mathcal{O}_p\big)\,, \tag{5.3}$$

where we consider the layer implementing the t-th time step acting on the operator at $t-1$. Expanding in the operator basis

$$|\gamma_p(t)|^2 = \frac{1}{q^{2n}} \sum_{a,b} \gamma_a(t-1)\gamma_b(t-1) \mathrm{Tr}\big(U_t \mathcal{O}_a U_t^\dagger \mathcal{O}_p\big) \mathrm{Tr}\big(U_t \mathcal{O}_b U_t^\dagger \mathcal{O}_p\big)\,, \tag{5.4}$$

we now Haar average the expression, decomposing the expression into a product over 2-site operators and 2-site random unitaries. For a single 2-site operator the expression we find averaging using the 2nd moment is

$$\int_{U(d)} dU\, \mathrm{Tr}\big(U \mathcal{O}_a U^\dagger \mathcal{O}_p\big) \mathrm{Tr}\big(U \mathcal{O}_b U^\dagger \mathcal{O}_p\big) = \frac{d^2}{d^2-1} \delta_{a,b} \big(d^2 \delta_{a,1}\delta_{p,1} + 1 - \delta_{a,1} - \delta_{p,1}\big)\,. \tag{5.5}$$

where we are just looking at the operators $\mathcal{O}_a$ and unitaries at a single 2-site gate. Taking the product over all pairs of sites on which the 2-site random unitaries act, we can write Eq. (5.4) as

$$|\gamma_p(t)|^2 = \sum_a S_{pa} |\gamma_a(t-1)|^2\,, \quad \text{where} \quad S_{pa} = \prod_s \Big(\delta_{a,1}^{(s)}\delta_{p,1}^{(s)} + \frac{1}{d^2-1}(1-\delta_{a,1}^{(s)})(1-\delta_{p,1}^{(s)})\Big)\,, \tag{5.6}$$

where the symmetric matrix determines the growth of the operator at each time step. The matrix S_{pa} is the transition matrix of a local Markov process on Pauli strings, where the Pauli string is updated stocastically at pairs of sites with transition probabilities determined by S_{pa}. The matrix S_{pa} essentially tells us that at each pair of sites we take $\mathbb{I}\mathbb{I} \to \mathbb{I}\mathbb{I}$ or a non-identity 2-site Pauli operator to any of the non-identity Paulis $\mathcal{O}_p \to \mathcal{O}_a$ each with prob $1/(d^2-1)$, which makes sense as there are d^2-1 non-identity 2-site Paulis (or rather the qudit generalization thereof).

Unitary Markov process:

The transition probabilities for the Markov process defined by a Haar-random 2-site unitary acting on the evolving Pauli string are

$$\begin{aligned} \mathbb{I}\mathbb{I} &\to \mathbb{I}\mathbb{I} \qquad \text{with prob} \quad 1 \\ \mathcal{O}_p &\to \mathcal{O}_a \qquad \text{with prob} \quad 1/(d^2-1)\,. \end{aligned}$$

These are local update rules 2-site pairs of operators in the Pauli string, which gives rise to a biased diffusion process. Looking at the endpoints, the update rule says that the left/right-most non-identity Pauli on the i-th site is mapped to a non-identity Pauli on the i and $i+1$ sites. As there are q^4-1 non-identity 2-site Paulis and q^2-1 of them have an identity at the $i+1$ site, there is a probability of the edge of the operator moving back $p=(q^2-1)/(q^4-1)$. The biased random walk governing the edge of an evolving operator gives rise to a biased diffusion process at long times. The weight of the operator on the left/right edges obeys a biased diffusion equation [128, 129]

$$\rho_{\mathrm{L/R}}(i,t) = \sum_{p\in\mathcal{O}_{\mathrm{L/R}}(i)} |\gamma_p(t)|^2 \quad \to \quad \partial_t\rho(x,t) = v_B\partial_x\rho(x,t) + D\partial_x^2\rho(x,t)\,, \tag{5.7}$$

where we sum over all Pauli strings with the left or right-most non-identity Pauli operator at site i. The edges of the growing operator grow ballistically but also spread diffusively, with butterfly velocity and diffusion constant

$$v_B = \frac{q^2-1}{q^2+1} \quad \text{and} \quad D = \frac{2q^2}{(q^2+1)^2}\,, \tag{5.8}$$

which all arises the the local update rule on the edges of the growing operator. We should think about this as a random walk on the Pauli strings where the Pauli strings want to grow larger, but there is some probability that they move back, i.e. biased diffusion.

It will be useful for later comparison with the unitary random circuits to give the $1/q$

expansion of the butterfly velocity and diffusion constant in Eq. (5.8)

$$v_B \approx 1 - \frac{2}{q^2} + \frac{2}{q^4} - \frac{2}{q^6} \qquad \text{and} \qquad D \approx \frac{2}{q^2} - \frac{4}{q^4} + \frac{6}{q^6}\,. \tag{5.9}$$

Note that it is convenient to define our coordinate system on the links, i.e. with respect to the 2-site gates, instead of the qudits. This way, the edge of an evolving Pauli either moves forwards or backwards at each time step.

Quick review of biased random walks

We can understand the derivation of the butterfly velocity and diffusion constant in the unitary random circuit as a simple biased random walk. Let's consider the right endpoint of the growing operator, so the bias refers to the preference of the operator to move right. At each time step we have a random variable x_i which takes values ± 1: we get -1 with probability p, meaning the random walker moves left, and $+1$ with probability $1-p$, meaning the operator moves right. The position of the operator at a time t is then given by

$$X(t) = \sum_{i=1}^{t} x_i\,. \tag{5.10}$$

As the endpoint dynamics are Markovian, meaning the steps at successive times are uncorrelated and the random variables x_i are iid, we simply compute the mean

$$\langle X(t)\rangle = \sum_{i=1}^{t} \langle x_i\rangle = t\langle x_i\rangle = t(1-2p) \quad \to \quad v_B = 1-2p\,, \tag{5.11}$$

which gives us the butterfly velocity v_B, and the second moment

$$\langle X(t)^2\rangle = \sum_{i,j=1}^{t} \langle x_i x_j\rangle = t\langle x_i^2\rangle = t\,, \tag{5.12}$$

which gives the variance and the diffusion constant

$$\langle X(t)^2\rangle_c = 4pt(1-p)\,, \quad \to \quad D = 2p(1-p)\,. \tag{5.13}$$

Therefore, computing the probability of moving back $p = 1/(q^2+1)$ in the random circuits gives the above velocity and diffusion constants.

We can arrive at the same conclusion noting by counting the left and right moves. Again, consider the motion of the right end of the Pauli string, and let ℓ denote the number of left steps and r the number of right moves. Clearly, $t = \ell + r$ and the position of the random walker is $x = r - \ell$. The distribution on the number of left and right moves is

$$f(\ell, r) = \binom{t}{r}(1-p)^r p^\ell \,. \tag{5.14}$$

From this we can compute the mean and variance of $x = r - \ell$ after t timesteps get v_B and D as above. Said equivalently, the binomial distribution on the endpoint of the string is simply the sum of Bernoulli distributed random variables making up the steps in the random walk.

5.3 Operator growth in symmetric random circuits

Now let's generalize the story to random circuits constructed out of local gates randomly drawn from quotients of the unitary group. We will consider 5 different symmetry classes, local orthogonal and symplectic matrices, and the two quotients of the unitary group $U(d)/O(d)$ and $U(d)/Sp(d)$. In each case the analysis is more complicated than the analysis for the random unitary circuits, but we can in fact solve the operator growth analytically. For each case we derive the Markov process that corresponds to the growth of a Pauli string in the random circuit, where the transition probabilities encoding the local update rules on a Pauli string are simply given by Haar-integrating over the quotient space. We then solve for the dynamics of the ends of the evolving operator. In the unitary random circuits, the growth of a Pauli string is a Markov process and the dynamics of the end of the operator is itself an autonomous Markov process. In the symmetric circuits the evolution of a Pauli string is also a Markov process, in each case with different Markov rules/transition probabilities. But the dynamics of the endpoints are no longer Markovian. This is because the biased random walk describing the successive time steps of the edge of the growing operator now has correlations between different time steps. In the language of random walks and diffusion the edge of the

growing operator in the symmetric circuits is a persistent biased random walk. The persistence refers to the operator having some 'inertia,' i.e. wanting to move in the same direction as the previous time step. In our symmetric random circuits, we find that the successive time steps are anticorrelated, thus the random walks are biased and anti-persistent.

In each of the four symmetric random circuits we consider, we compute the probability of an operator moving back, which gives the butterfly velocity. The analysis of the diffusion constant is a little more subtle. The anticorrelation reduces the diffusion constant from its uncorrelated value, i.e. the one that would be computed by p the probability of moving back.

Before discussing the four random circuit models, we summarize our findings in the tables below

Random circuit model	v_B	v_B for $q=2$		v_B at large q
Unitary $U(d)$	$\frac{q^2-1}{q^2+1}$	$\frac{3}{5}$	$(=0.6)$	$1-\frac{2}{q^2}+\frac{2}{q^4}-\frac{2}{q^6}$
Orthogonal $O(d)$	(5.31)	$\frac{23}{39}$	(≈ 0.5897)	$1-\frac{2}{q^2}+\frac{6}{q^5}-\frac{4}{q^6}$
Symplectic $Sp(d)$	(5.37)	$\frac{7}{15}$	(≈ 0.467)	$1-\frac{2}{q^2}-\frac{2}{q^5}+\frac{4}{q^7}$
AI $U(d)/O(d)$	(5.43)	$\frac{1}{2}$	$(=0.5)$	$1-\frac{2}{q^2}-\frac{2}{q^4}+\frac{14}{q^6}$
AII $U(d)/Sp(d)$	(5.59)	$\frac{1}{4}$	$(=0.25)$	$1-\frac{2}{q^2}-\frac{2}{q^4}-\frac{2}{q^6}$

5.3.1 Operator growth in random orthogonal circuits

The random unitary circuit picture is nice, we see the emergence of a diffusive phenomena from unitary dynamics. But as the random circuit breaks all symmetries and obeys no conservation laws, we also want to know what happens when we start building in symmetry in different ways. We could construct a unitary random circuit which obeys a conservation law explicitly in terms of some local operator. Alternatively, we could ask what happens when we consider a random circuit with 2-site gates restricted to obey some symmetry. Consider a circuit of 2-site Haar random orthogonal operators. For a single 2-site operator,

consider

$$\int_{O(d)} dU\, \mathrm{Tr}\big(U\mathcal{O}_a U^\dagger \mathcal{O}_p\big)\mathrm{Tr}\big(U\mathcal{O}_b U^\dagger \mathcal{O}_p\big) \tag{5.15}$$

$$= \frac{1}{(d-1)(d+2)}\Big((d+1)d^3\mathrm{Tr}(\mathcal{O}_a)\mathrm{Tr}(\mathcal{O}_b)\mathrm{Tr}(\mathcal{O}_p)^2 - d^2\mathrm{Tr}(\mathcal{O}_p)^2\mathrm{Tr}(\mathcal{O}_a\mathcal{O}_b^T) + \dots \tag{5.16}$$

The second moment of the orthogonal group gives an expression with nine terms, and after some reworking we find

$$= d^2\delta_{a,b}\Big(\delta_{a,1}\delta_{p,1} + \frac{1}{(d-1)(d+2)}\big((\delta_{a,1}-1)(\delta_{p,1}-1) + (\delta_{a,1}-(-1)^{Y_a})(\delta_{p,1}-(-1)^{Y_p}) + \frac{1}{d}(1-(-1)^{Y_a})(1-(-1)^{Y_p})\big)\Big)\,, \tag{5.17}$$

where Y_a and Y_p count the number of Y operators in the $\mathcal{O}_p$ and $\mathcal{O}_a$ operators at the sites we are acting on, i.e. simply accounting for whether the Pauli is even or odd under transposition. This expression means that

$$|\gamma_p(t)|^2 = \sum_a S_{pa}|\gamma_a(t-1)|^2\,, \tag{5.18}$$

with the S_{pa} defined in the expression above as

$$S_{pa} = \prod_s \Big(\delta^{(s)}_{a,1}\delta^{(s)}_{p,1} + \frac{1}{(d-1)(d+2)}\big((\delta^{(s)}_{a,1}-1)(\delta^{(s)}_{p,1}-1) + (\delta^{(s)}_{a,1}-(-1)^{Y_a})(\delta^{(s)}_{p,1}-(-1)^{Y_p}) + \frac{1}{d}(1-(-1)^{Y_a})(1-(-1)^{Y_p})\big)\Big)\,, \tag{5.19}$$

which defines the transition matrix of the orthogonal Markov process.

Orthogonal Markov process:

We should again interpret this as a Markov process with local update rules defined by S_{pa}. In the orthogonal case, the matrix S_{pa} tells us that at each site we take identities to identites,

even non-identity Paulis to even and odd non-identity Paulis to odd:

$$\begin{aligned} \mathbb{I}\mathbb{I} &\to \mathbb{I}\mathbb{I} && \text{with prob} \quad 1 \\ \text{even } \mathcal{O}_p &\to \text{even } \mathcal{O}_a && \text{with prob} \quad p_e = 2/((d-1)(d+2)) \\ \text{odd } \mathcal{O}_p &\to \text{odd } \mathcal{O}_a && \text{with prob} \quad p_o = 2/(d(d-1))\,. \end{aligned}$$

So far this is not too surprising given what we understand about $O(d)$ and the action of Paulis under the group. But we see an interesting conservation law encoded into the growing wavefront, the wavefront is growing diffusively, but there are two coupled diffusion processes, where even Paulis are mapped to even and odd Paulis are mapped to odd.

We can now solve for the operator growth in the random orthogonal circuits. First we want to compute the probability of an operator moving back. Unfortunately, this is no longer as simple as the unitary random circuit, where the endpoint undergoes autonomous Markovian dynamics. The probability of an operator moving back depends on whether the operator at the farthest most gate is even or odd, and further depends on whether the action at the previous timestep was a move forward or backward. The structure of the alternating gates at each time step means that even operators output from a gate can input odd operators at the next gate. Consider YY, an even 2-site operator, if this operator appears at the edge of the growing operator, then the input to the gate of the farthest gate at the next time step is $Y\mathbb{I}$, an odd operator. So we need to treat the dynamics of even and odd operators carefully.

We start by discussing the case for qubits and then derive formulae for general local dimension q. The probability of an even Pauli moving back is $2/9$, i.e. if we get a $X\mathbb{I}$ or $Z\mathbb{I}$. For odd Paulis we have $1/6$, i.e. if we generate a $Y\mathbb{I}$. Denote the probabilities of even and odd operators moving forwards or backwards as

$$p_{\leftarrow e} = \frac{2}{9}\,, \quad p_{e\rightarrow} = \frac{7}{9}\,, \quad p_{\leftarrow o} = \frac{1}{6}\,, \quad p_{o\rightarrow} = \frac{5}{6}\,. \tag{5.20}$$

Now we need to find the probability that the operator at the end of the string, at the rightmost gate, is even or odd by relating it to the probability it was even or odd at the previous time step. We must take into account whether the operator has moved forward or

backwards, as the probabilities of an even or odd operator depend on which action occurs. An even operator moves forward to an even operator with probability 6/7, but moves back to an even operator with probability 3/4. When an even operator moves forward to an even operator, the rightmost operator will be either $X\mathbb{I}$ or $Z\mathbb{I}$, which occurs 6/7 times, the 1/7 occurs when an even operator moves forward to an odd operator, which only happens when YY is generated and the operator at the rightmost gate is $Y\mathbb{I}$. An even operator moves back to an even operator with probability 3/4, when $X\mathbb{I}$ or $Z\mathbb{I}$ is generated and the operator at the rightmost gate is $\{\mathbb{I}, X, Z\} \otimes \{X, Z\}$.[1] Meaning the $i+1$-th entry of the gate will be an even operator, but the i-th entry has probability 3/4 of being even, and thus the operator has probability 3/4 of being even. A similar analysis holds for the odd operators, all together we find

$$\begin{aligned} p_{e\to e} &= \frac{6}{9}\,, \quad p_{e\to o} = \frac{1}{9}\,, \quad p_{o\to e} = \frac{2}{6}\,, \quad p_{o\to o} = \frac{3}{6}\,, \\ p_{e\leftarrow e} &= \frac{1}{6}\,, \quad p_{e\leftarrow o} = \frac{1}{18}\,, \quad p_{o\leftarrow e} = \frac{1}{8}\,, \quad p_{o\leftarrow o} = \frac{1}{24}\,. \end{aligned} \tag{5.21}$$

The probability of the rightmost operator being even or odd at time t is thus

$$\begin{aligned} p_e(t) &= (p_{e\to e} + p_{e\leftarrow e})p_e(t-1) + (p_{o\to e} + p_{o\leftarrow e})p_o(t-1) \\ p_o(t) &= (p_{o\to o} + p_{o\leftarrow o})p_o(t-1) + (p_{e\to o} + p_{e\leftarrow o})p_e(t-1)\,, \end{aligned} \tag{5.22}$$

and for qubits we get the equations governing the evolving probabilities:

$$p_e(t) = \frac{5}{6}p_e(t-1) + \frac{3}{8}p_o(t-1)\,, \qquad p_o(t) = \frac{5}{8}p_o(t-1) + \frac{1}{6}p_e(t-1)\,. \tag{5.23}$$

This itself is a Markov process, where the probability of finding an even or an odd operator at the end of the evolving string updates at each time step. We solve for the stationary distribution (i.e. where the probabilities are the same at each time step) and find $p_e = 9/13$ and $p_o = 4/13$. Thus we can compute the probability of an operator moving back in terms

[1] Here we have assumed that the 1-site operators coming from the random circuit behind the evolving end of the string appear with uniform probability, which quickly becomes the case after a few time steps and the interior of the evolving Pauli string rapidly mixes.

of the probability the operator at the end of the evolving string is even or odd:

$$p = p_{\leftarrow e} p_e + p_{\leftarrow o} p_o \,, \tag{5.24}$$

and find $p = 8/39$ for qubits. This gives a butterfly velocity of $v_B = 23/39 \approx 0.5897$. Very close to the unitary butterfly velocity for qubits $v_B = 3/5$. We should mention, in case the reader is suspect of how close the result is to the unitary case, that we have confirmed this value of the orthogonal butterfly velocity by computationally simulating the evolving operators in the random orthogonal circuits.

We can now derive the butterfly velocity for general local dimension. First we define the number of even or odd operators for a local dimension of d

$$n_e(d) = \frac{(d-1)(d+2)}{2} \,, \qquad n_o(d) = \frac{d(d-1)}{2} \,. \tag{5.25}$$

Thus the number of even/odd operators at a single site is $n_e(q)$ and $n_o(q)$. The probabilities of an even and odd operator moving backwards are a simple generalization of what we discussed above in the case of qubits

$$p_{\leftarrow e} = \frac{n_e(q)}{n_e(d)} \quad \text{and} \quad p_{\leftarrow o} = \frac{n_o(q)}{n_o(d)} \,. \tag{5.26}$$

The probabilities of even/odd operators moving forwards or backwards to even/odd operators are

$$\begin{aligned} p_{e\to e} &= \frac{(n_e(q)+1)n_e(q)}{n_e(d)} \,, \quad p_{e\to o} = \frac{n_o(q)^2}{n_e(d)} \,, \quad p_{o\to e} = \frac{(n_e(q)+1)n_o(q)}{n_o(d)} \,, \quad p_{o\to o} = \frac{n_e(q)n_o(q)}{n_o(d)} \,, \\ p_{e\leftarrow e} &= \frac{(n_e(q)+1)n_e(q)}{n_e(d)q^2} \,, \quad p_{e\leftarrow o} = \frac{n_o(q)^2}{n_o(d)q^2} \,, \quad p_{o\leftarrow e} = \frac{n_e(q)n_o(q)}{n_e(d)q^2} \,, \quad p_{o\leftarrow o} = \frac{(n_e(q)+1)n_o(q)}{n_o(d)q^2} \,. \end{aligned} \tag{5.27}$$

The equations determining the probability of the rightmost operator at a time t:

$$\begin{aligned} p_e(t) &= (p_{e\to e} + p_{e\leftarrow e})p_e(t-1) + (p_{o\to e} + p_{o\leftarrow e})p_o(t-1) \\ p_o(t) &= (p_{o\to o} + p_{o\leftarrow o})p_o(t-1) + (p_{e\to o} + p_{e\leftarrow o})p_e(t-1)\,, \end{aligned} \tag{5.28}$$

for general local dimension q become

$$\begin{aligned} p_e(t) &= \frac{(q+2)(q^2+1)}{2q(q^2+2)}p_e(t-1) + \frac{q^2-1}{2q^2}p_o(t-1) \\ p_o(t) &= \frac{q^2+1}{2q^2}p_o(t-1) + \frac{(q-1)(q^2-q+2)}{2q(q^2+2)}p_e(t-1)\,. \end{aligned} \tag{5.29}$$

As a sanity check, the probabilities of $p_e(t-1)$ and $p_o(t-1)$ add to unity. Solving for the stationary distribution, we find p_e and p_o and then compute the probability of an operator moving back to be

$$p = \frac{q^2+q+2}{(q+1)(q^3+2q+1)}\,. \tag{5.30}$$

This gives a butterfly velocity for orthogonal random circuits

$$v_B = \frac{q^2(q^2+q)+q-3}{(q+1)(q^3+2q+1)}\,, \tag{5.31}$$

which has a series expansion in q^2 as

$$v_B = 1 - \frac{2}{q^2} + \frac{6}{q^5} + \dots\,, \tag{5.32}$$

the same as the unitary case up to second order in $1/q$ in Eq. (5.9). Note the absence of contribution at $1/q^4$. For qubits where $q = 2$, we find the butterfly velocity $v_B = 23/39$.

The moral of the story seems to be that operator growth in orthogonal random circuits proceeds very similarly to that in unitary random circuits; the main difference is that the analysis is somewhat more tedious.

5.3.2 Operator growth in random symplectic circuits

Consider now a circuit of 2-site Haar random symplectic gates. To understand the operator growth, first consider the action of a single random gate

$$\int_{Sp(d)} dU \, \mathrm{Tr}\big(U\mathcal{O}_a U^\dagger \mathcal{O}_p\big)\mathrm{Tr}\big(U\mathcal{O}_b U^\dagger \mathcal{O}_p\big)\,. \tag{5.33}$$

Computing the second moment we find the transition matrix for the symplectic circuits

$$S_{pa} = \prod_s \Big(\delta^{(s)}_{a,1}\delta^{(s)}_{p,1} + \frac{1}{(d+1)(d-2)}\big((\delta^{(s)}_{a,1}-1)(\delta^{(s)}_{p,1}-1) + (\delta^{(s)}_{a,1}-(-1)^{S_a})(\delta^{(s)}_{p,1}-(-1)^{S_p}) \\ - \frac{1}{d}(1-(-1)^{S_a})(1-(-1)^{S_p})\big)\Big)\,, \tag{5.34}$$

which is unsurprisingly similar to the orthogonal case. Noting that under symplectic conjugation Paulis are either even or odd: $\mathcal{O}_p^D \equiv J\mathcal{O}_p^T J^T = \pm\mathcal{O}_p$. Here we have defined $(-1)^{S_a}$ to be $+1$ for symplectically even Paulis and -1 for odd Paulis. The above transition matrix tells us that Paulis which are even under symplectic conjugation are taken to symplectically even Paulis and odd to odd.

Symplectic Markov process:

In the symplectic case, we find that at each gate we take identities to identites, even non-identity Paulis to even and odd non-identity Paulis to odd:

$$\begin{aligned} \mathbb{I}\mathbb{I} &\to \mathbb{I}\mathbb{I} && \text{with prob} \quad 1 \\ \text{even } \mathcal{O}_p &\to \text{even } \mathcal{O}_a && \text{with prob} \quad p_e = 1/((2d-1)(2d+1)) \\ \text{odd } \mathcal{O}_p &\to \text{odd } \mathcal{O}_a && \text{with prob} \quad p_o = 1/(d(2d+1)) \end{aligned}$$

where again we mean the even or odd action under symplectic conjugation $J\mathcal{O}_p^T J^T = \pm\mathcal{O}_p$. Note that this is not as conceptually simple as the transpose above, where we just simple count the number of Y operators in the Pauli string. The matrix J depends on the size of the gate, so for 2-qubit operators $J = iY \otimes \mathbb{I}$. Even 2-qubit Paulis are $XY, \mathbb{I}X, ZY \ldots$ and odd are $XZ, X\mathbb{I}, \mathbb{I}Y, \ldots$, which we see are not symmetric.

The analysis for operator growth in symplectic random circuits is essentially the same as in the orthogonal case, we have that even and odd operators are mapped to themselves with uniform probability. The only difference is that the set of Paulis invariant under symplectic transpose is different that in the orthogonal case, thus giving slightly different probabilities. The equations determining the probability of the rightmost operator at a time t:

$$\begin{aligned} p_e(t) &= (p_{e\to e} + p_{e\leftarrow e})p_e(t-1) + (p_{o\to e} + p_{o\leftarrow e})p_o(t-1) \\ p_o(t) &= (p_{o\to o} + p_{o\leftarrow o})p_o(t-1) + (p_{e\to o} + p_{e\leftarrow o})p_e(t-1)\,. \end{aligned} \tag{5.35}$$

Repeating the same analysis but with the symplectic version of n_e and n_o, we find the probability of the end of an evolving Pauli string moving back to be

$$p = \frac{q^2 - q + 2}{(q^2+1)(q^2-q+1)}\,, \tag{5.36}$$

which gives a butterfly velocity $v_B = 1 - 2p$

$$v_B = \frac{q^4 - q^3 + q - 3}{(q^2+1)(q^2-q+1)} \quad \text{for qubits:} \quad v_B = \frac{7}{15}\,. \tag{5.37}$$

5.3.3 Operator growth in random AI circuits

Consider now a circuit of 2-site Haar random AI gates. To understand the operator growth, we again consider the action of a single random gate

$$\int_{\text{AI}} dU\, \text{Tr}\big(U\mathcal{O}_a U^\dagger \mathcal{O}_p\big)\text{Tr}\big(U\mathcal{O}_b U^\dagger \mathcal{O}_p\big)\,. \tag{5.38}$$

We can compute the average using the second moment of the $U(d)/O(d)$ and find the transition matrix

$$\begin{aligned} S_{pa} = \prod_s \Big(\delta^{(s)}_{a,1}\delta^{(s)}_{p,1} + \frac{1}{d(d+1)(d+3)}\big((d+4)(\delta^{(s)}_{a,1}-1)(\delta^{(s)}_{p,1}-1) \\ + 2(d+2)\delta^{(s)}_{a,p}(1-\delta^{(s)}_{p,1}) + \frac{2}{d}(1-(-1)^{(a,p)})\big)\Big)\,, \end{aligned} \tag{5.39}$$

where $(-1)^{(a,p)}$ is $+1$ if the Paulis commute $[\mathcal{O}_p, \mathcal{O}_a] = 0$, and -1 if they anticommute $\{\mathcal{O}_p, \mathcal{O}_a\} = 0$. This is a little more interesting than the cases above. Here we see that the identity is mapped to the identity, and non-id Paulis are mapped to a linear combination of all other non-identity Paulis, but with a higher probability of being mapped to itself, and different probabilities for Paulis they commute and anticommute with.

AI Markov process:

From the above transition matrix, we can list the transition probabilities for the AI Markov process

$$\begin{aligned}
&\mathbb{I}\mathbb{I} \to \mathbb{I}\mathbb{I} && && \text{with prob} \quad 1 \\
&\mathcal{O}_p \to \mathcal{O}_{a=p} && && \text{with prob} \quad p_s = (3d+8)/d(d+1)(d+3) \\
&\mathcal{O}_p \to \mathcal{O}_{a\neq p} \quad \text{and} && [\mathcal{O}_a, \mathcal{O}_p] = 0 && \text{with prob} \quad p_c = (d+4)/d(d+1)(d+3) \\
&\mathcal{O}_p \to \mathcal{O}_{a\neq p} \quad \text{and} && \{\mathcal{O}_a, \mathcal{O}_p\} = 0 && \text{with prob} \quad p_a = (d+2)^2/d^2(d+1)(d+3)\,.
\end{aligned}$$

We should now think about the butterfly velocity and the diffusion constant in the AI circuits and determine how they are affected with the above updates. As a sanity check, recall that exactly half of Pauli operators commute with a Pauli operator $\mathcal{O}_p$ (including itself and the identity) and half anti-commute. So we see that the probabilities for a non-identity Pauli operator add to one: $p_s + p_c(d^2/2 - 2) + p_a d^2/2 = 1$.

AI operator growth

Now we study operator growth in the AI random circuits. The fact that an operator is more likely to update to itself already indicates the the butterfly velocity will be slower. The AI Markov update rules make analyzing the operator growth a little more intricate than in the unitary circuits as the movement of the endpoint of the string is no longer autonomous. The probability that an operator moves back depends on what happened in the previous time-step. Nevertheless, we can still exactly solve for the operator growth. For an operator that has moved forward at the previous timestep, e.g. $X\mathbb{I}$, the probability the string moves back is $p_1 \equiv p_s + (d/2 - 2)p_c + (d/2)p_a$ (the probability it goes to itself plus the other $d - 2$

commuting or anticommuting operators that also move back). If the operator has moved back at a previous timestep, then the operator at the edge of the string has a non-identity Pauli at the rightmost site, e.g. XZ, and will move forward with higher probability (as it going to itself moves the operator forward). In this case the probability the operator moves back is $p_2 \equiv (d/2-1)p_c + (d/2)p_a$.

Now we can describe the probability that an operator moves back at time t, $p(t)$, in terms of the probability of the operator having moved back at the previous timestep

$$p(t) = (p_2 - p_1)p(t-1) + p_1 \,. \tag{5.40}$$

Either solving for the stationary probability or by summing the series we can solve for $p(t)$ and find

$$p(t) = \frac{p_1}{1 + p_1 - p_2} \,. \tag{5.41}$$

More explicitly, we have the probability of a Pauli string moving back in the AI circuit as

$$p = \frac{d^2 + 5d + 2}{d^3 + 4d^2 + 5d + 4} \quad \text{and for qubits:} \quad p = \frac{1}{4} \,. \tag{5.42}$$

This gives a butterfly velocity for AI random circuits $v_B = 1 - 2p$ in terms of the local dimension q

$$v_B = \frac{q^4 + 5q^2 + 2}{q^6 + 4q^4 + 5q^2 + 4} \quad \text{and for qubits:} \quad v_B = \frac{1}{2} \,. \tag{5.43}$$

Note that this is quite a bit slower than the unitary random circuits where $v_B = 3/5$. The above expression is exact for any local dimension q, but we can expand in powers of q to find

$$v_B = 1 - \frac{2}{q^2} - \frac{2}{q^4} + \frac{14}{q^6} + \ldots \tag{5.44}$$

for the AI circuits.

Diffusion in random AI circuits

Recall that for the growing operator in the random AI circuits, we derived the probability of an operator moving back to be

$$p(t) = p_1 + p(t-1)(p_2 - p_1)\,, \tag{5.45}$$

with the probabilities p_1 and p_2 defined above. The dependence on the probability of moving back introduces correlation between the time steps. The evolution of the end of the operator in the random circuit is a persistent biased random walk, where there is a direct correlation with the motion at a previous timestep. As the prefactor $p_2 - p_1$ is negative, there is an anticorrelation; the random walker is more inclined to move in the opposite direction from its previous motion.

The correlation between timesteps does not affect the mean $\langle X(t)\rangle$, which is why deriving the probability of moving back p directly gives us the butterfly velocity v_B. Correlations between timesteps affect the variance $\langle X\rangle_c$ and thus the diffusion constant. Generally, we can compute the variance of $X(t) = \sum_i^t x_i$

$$\langle X(t)^2\rangle_c = \sum_{i,j}\langle x_i x_j\rangle_c = 4tp(1-p) + \sum_{i\neq j}\langle x_i x_j\rangle_c\,, \tag{5.46}$$

where the first term is simply the diffusion constant. Define $c(|i-j|) \equiv \langle x_i x_j\rangle$ as the correlator between time steps, where we have translational invariance in time. Clearly, $c(0) = 4tp(1-p) = 2D_0 t$. To compute the correlations for the AI random walker, we need to look closer at the direct correlation between time steps.

The probability of moving back in an AI circuit, Eq. (5.45), means that the random variable x_i at a time step i, is a sum of two random variables

$$x_i = \alpha x_{i-1} + p_1 x_i'\,, \quad \text{where} \quad \alpha = (p_2 - p_1) \tag{5.47}$$

and where x_{i-1} is the step taken at the previous timestep and x_i' is an iid random variable, with mean and variance to be determined. There is now a direct correlation between

timesteps given by the coefficient of the x_{i-1} variable $\alpha = p_2 - p_1$. We already know the mean and variance of x_i to be

$$\langle x_i \rangle = (1-2p)\,, \quad \text{and} \quad \langle x_i^2 \rangle_c = 4p(1-p)\,. \tag{5.48}$$

We can fix the mean and variance of the random variable x_i' from Eq. (5.47). Taking the mean we find

$$\langle x_i \rangle = \alpha \langle x_{i-1} \rangle + p_1 \langle x_i' \rangle \quad \rightarrow \quad \langle x_i' \rangle = \frac{1-2p}{p}\,, \tag{5.49}$$

and taking the variance we find

$$\langle x_i^2 \rangle_c = \alpha^2 \langle x_{i-1}^2 \rangle_c + p_1^2 \langle x_i'^2 \rangle_c \quad \rightarrow \quad \langle x_i'^2 \rangle_c = 4(1-p)\frac{1-p_1+p_2}{p_1}\,. \tag{5.50}$$

First we note that if $p_2 = p_1$, the correlation between time-steps vanishes and the endpoint dynamics are uncorrelated and Markovian. In this case the probability of moving back is $p = p_1$. We see that in terms of x_i' the mean and variance of x_i from Eqs. (5.49) and (5.50) become $\langle x_i \rangle = (1-2p_1)$ and $\langle x_i^2 \rangle_c = 4p_1(1-p_1)$, as expected. As a second sanity check, if $p_1 = 0$, the probability of moving is zero and the endpoint deterministically moves forward.

We can now compute the correlator between random steps at successive times and find that for a difference

$$c(t) = \langle x_i x_j \rangle_c = 2D_0 \alpha^t\,, \quad \text{where} \quad t = |i-j|\,. \tag{5.51}$$

Therefore, rewriting the sum in Eq. (5.46) by making a change of variables and then summing the resulting series, we find

$$\langle X(t)^2 \rangle_c = 2D_0 t + 2\sum_{t'=1}^{t}(t-t')c(t') = 2D_0 t \frac{1+\alpha}{1-\alpha} - \frac{\alpha(1-\alpha^t)}{(1-\alpha)^2}\,. \tag{5.52}$$

For times $t \gg 1$, we have

$$\langle X(t)^2 \rangle_c \approx 2D_{\rm AI} t\,, \quad \text{where} \quad D_{\rm AI} = \frac{1+\alpha}{1-\alpha} D_0\,, \tag{5.53}$$

and $D_0 = 2p(1-p)$ is the uncorrelated diffusion constant.

5.3.4 Operator growth in random AII circuits

Consider now a circuit of 2-site Haar random AII gates. To understand the operator growth, we again consider the action of a single random gate

$$\int_{AII} dU \, \mathrm{Tr}\big(U\mathcal{O}_a U^\dagger \mathcal{O}_p\big)\mathrm{Tr}\big(U\mathcal{O}_b U^\dagger \mathcal{O}_p\big) \tag{5.54}$$

$$= (2d)^2\delta_{a,b}\Big(\delta_{a,1}\delta_{p,1} + \frac{1}{d(2d-3)(2d-1)}\big((d-2)(\delta_{a,1}-1)(\delta_{p,1}-1) + 2(d-1)\delta_{a,p}(1-\delta_{p,1})\big)\Big)\,. \tag{5.55}$$

Again, like in the AI circuit, we see that the identity is mapped to the identity, and non-id Paulis are mapped to a linear combination of all other non-identity Paulis, but with a higher probability of being mapped to itself.

AII Markov rules:

$$\begin{aligned}
&\mathbb{I}\mathbb{I} \to \mathbb{I}\mathbb{I} && && \text{with prob} \quad 1 \\
&\mathcal{O}_p \to \mathcal{O}_{a=p} && && \text{with prob} \quad p_s = (3d-8)/d(d-3)(d-1) \\
&\mathcal{O}_p \to \mathcal{O}_{a\neq p} && [\mathcal{O}_a, \mathcal{O}_p] = 0 && \text{with prob} \quad p_c = (d-4)/d(d-3)(d-1) \\
&\mathcal{O}_p \to \mathcal{O}_{a\neq p} && \{\mathcal{O}_a, \mathcal{O}_p\} = 0 && \text{with prob} \quad p_a = (d-2)^2/d^2(d-3)(d-1)
\end{aligned} \tag{5.56}$$

Again the analysis is much the same as the AI random circuits, where there is a higher probability of an operator going to itself, and separate probabilities of an operator updating to operators which commute or anticommute with that operator. One interesting difference here is that for the case of $q = 2$, local qubits, the updated probabilities become

$$\text{For qubits:} \qquad p_s = 1/3\,, \qquad p_a = 1/12\,, \qquad p_c = 0\,. \tag{5.57}$$

So the probability an operator updates to itself is fairly high, and no 2-qubit gate updates to a gate with which it commutes. This means that operator growth will happen fairly

slowly as the operators at the end of the Pauli string, always of the form $X\mathbb{I}$, have a much higher chance of updating to themselves, and thus moving back, than in the other symmetric circuits.

Just as in the AI random circuit, we can exactly solve for general local dimension q the probability for an evolving operator to move back

$$p = \frac{(q^2-2)(q^2-1)}{q^6-4q^4+5q^2-4}\,. \tag{5.58}$$

For $q = 2$, this gives $p = 3/8$. The butterfly velocity for random AII circuits is then given as

$$v_B = \frac{q^6-6q^4+11q^2-8}{q^6-4q^4+5q^2-4}\,. \tag{5.59}$$

For $q = 2$, qubit chains this gives a butterfly velocity of $v_B = 1/4$, substantially slower than in the other random circuits. Expanding this at large q we find

$$1-\frac{2}{q^2}-\frac{2}{q^4}-\frac{2}{q^6}\,, \tag{5.60}$$

giving the same leading order term in $1/q$, but different subleading corrections.

The diffusion constant is also given just as in the AI case, where the correlation between steps is $\alpha = p_2 - p_1$ in terms of the p_s, p_c, and p_a defined for the AII random circuits. Meaning the endpoint dynamics of the AII random walker are also a biased random walk with anticorrelated steps. The diffusion constant is $D_{\text{AII}} = \frac{1+\alpha}{1-\alpha}D_0$, where $D_0 = 2p(1-p)$ is the uncorrelated diffusion constant in terms of the stationary probability of moving back in the AII circuits p.

5.4 Discussion

In this chapter, we have studied operator growth in different classes of symmetric random circuits where the local 2-site unitaries are drawn from quotients or subgroups of the unitary group. Just as in the unitary case, we were able to understand the growth of operators in the random circuit as a Markov process on the evolving Pauli strings. Studying the operator

edge in the unitary circuits, one can solve for the growth as a biased random walk, where evolving Pauli strings have some probability of moving back. This defines a butterfly velocity and diffusion constant for the support of the operator on its edge, where the endpoints of $\mathcal{O}(t)$ grow ballistically with v_B but spread diffusively. The weight of the operator on its left and right ends $\rho_{\mathrm{L/R}}$ satisfies a biased diffusion equation, a dissipative process emerging from unitary dynamics. We expanded on this picture by considering random quantum circuits built out of random gates drawn from the orthogonal $O(d)$ and symplectic $Sp(d)$ groups, as well as the quotients of the unitary group $U(d)/O(d)$ and $U(d)/Sp(d)$. Each class of random circuits gave rise to a different Markov process on evolving Pauli strings, and operator growth in the random circuits could be solved by understanding the endpoint dynamics as a random walk. In the AI and AII circuits, we found a persistent biased random walk with anticorrelation between steps. In the orthogonal and symplectic circuits, we found a similar random walk with bias and correlation, but where the random walker also carried an internal state. In all four of the classes of random circuits, we solved for the butterfly velocities and diffusion constants by solving the simple random walk.

We can also comment more generally on properties of the random circuit models considered here. It is known that unitary circuits of polynomial depth form approximate k-designs [126, 54] and achieve optimal decoupling [77]. As we discussed in Ch. 3, Haar-random $O(d)$ and $Sp(d)$ and random unitaries from $U(d)/O(d)$ and $U(d)/Sp(d)$ do not, in general, form k-designs. The frame potential for $O(d)$ and $Sp(d)$ is $\frac{(2k)!}{2^k k!}$, and only coincides with the frame potential of $U(d)$ for $k = 1$, meaning Haar-random orthogonal and symplectic matrices form unitary 1-designs, but do not reproduce any higher moments of the unitary group. The frame potential for AI and AII is $2^k k!$ for large d, and thus they do not form k-designs for any k. For the random circuits considered in this paper, we expect orthogonal and symplectic circuits of polynomial depth to form approximate 1-designs, but will not capture any higher moments. More interestingly, in Ch. 3 we also defined the notion of a symmetric k-design, with respect to subsets of $U(d)$. We expect combining our results above and employing similar methods to [126, 54], would show that symmetric RQCs of polynomial depth form approximate symmetric k-designs.

Moreover, in Ch. 3 we looked at the reduced density matrix on subsystems of random

symmetric states, finding essentially the same result as Page's theorem: random symmetric states locally look maximally mixed. A similar calculation also shows that random symmetric unitaries also achieve decoupling with only subleading corrections to the Haar-random calculation. The expectation would be that random symmetric circuits also achieve decoupling in polylogarithmic depth, but precisely computing the depth and understanding how suboptimal random symmetric matrices are would be an interesting avenue for future work.

5.A Operator growth in random matrix theory

Here we compute analytically compute the coefficients $\gamma_p(t)$ for the Gaussian unitary ensemble (GUE) in order to study operator growth in random matrix theory. Recall that we can expand any operator in a basis of operators as

$$\mathcal{O}(t) = \sum_p \gamma_p(t)\mathcal{O}_p\,, \tag{5.61}$$

where the coefficients $\gamma_p(t)$ are probabilities of finding the operator $\mathcal{O}_p$ in the growing operator, or equivalently, the weight of the growing operator on a given Pauli string. The fact that the operator norm is conserved means that the probabilities are conserved $\sum_p |\gamma_p(t)|^2 = 1$.

We want to consider evolving by a random matrix Hamiltonian, where $U = e^{-iHt}$ where $H \in \text{GUE}$. To study operator growth we need to compute the coefficients $\gamma_p(t)$ averaged over GUE, more precisely we need the first and second moments of the coefficients. We will not review the necessary random matrix machinery here, but refer to [12] as well as definitions given in Ch. 2. Consider the growth of an operator $\mathcal{O}_0(t) = e^{-iHt}\mathcal{O}_0 e^{iHt}$.

GUE averaged $\gamma_p(t)$

We can compute the first moment

$$\gamma_p(t) = \frac{1}{d}\text{Tr}\big(\mathcal{O}_0(t)\mathcal{O}_p\big) = \frac{1}{d}\text{Tr}\big(e^{-iHt}\mathcal{O}_0 e^{iHt}\mathcal{O}_p\big)\,, \tag{5.62}$$

using the unitary invariance of the GUE measure dH, then integrating using the 2nd Haar moment, and find

$$|\gamma_p(t)| = \frac{\mathcal{R}_2 - 1}{d^2 - 1}\delta_{0,p} \quad \rightarrow \quad |\gamma_{\mathcal{O}_0}(t)| \approx \frac{\mathcal{R}_2(t)}{d^2}\,, \tag{5.63}$$

with all other coefficients vanishing.

GUE averaged $|\gamma_p(t)|^2$

More interesting is the average of $|\gamma_p(t)|^2$, which is the probability of finding a given operator $\mathcal{O}_p$ in the evolving operator. We can compute the second moment,

$$|\gamma_p(t)|^2 = \frac{1}{d^2}\mathrm{Tr}\big(e^{-iHt}\mathcal{O}_0 e^{iHt}\mathcal{O}_p\big)\mathrm{Tr}\big(e^{-iHt}\mathcal{O}_0 e^{iHt}\mathcal{O}_p\big)\,, \tag{5.64}$$

using the unitary invariance of the GUE measure dH, and then integrating using the 4th Haar moment. We find an expression involving $(4!)^2$ terms, which we will not reproduce here. First, we give the leading order behavior in d which captures the early time piece, i.e. the decay of the support on the initial operator. For the late-time behavior we need the $1/d^2$ terms, as the coefficients all decay order $1/d^2$ terms.

At leading order we find

$$|\gamma_{\mathcal{O}_0}(t)|^2 \approx \frac{\mathcal{R}_4(t)}{d^4} \quad \text{and} \quad |\gamma_{p\neq\mathcal{O}_0}(t)|^2 \approx \frac{1}{d^2}\,. \tag{5.65}$$

This isn't surprising; the support on the initial operator decays in time as $1/t^4$ and all other coefficients are around $1/d^2$ at early times.

Looking at the $1/d^2$ terms we can then discuss the late-time behavior of GUE operator growth. The coefficient of the initial operator $\mathcal{O}_0$ is

$$|\gamma_{\mathcal{O}_0}(t)|^2 \approx \frac{\mathcal{R}_4}{d^4} + \frac{1}{d^2}\Big(1 - \frac{4\mathcal{R}_2}{d^2} - \frac{4\mathcal{R}_4}{d^4} - \frac{2\mathcal{R}_{4,1}}{d^3} + \frac{\mathcal{R}_{4,2}}{d^2}\Big)\,. \tag{5.66}$$

Furthermore, we find that the probabilities of $\mathcal{O}_p$ depend on whether they commute or

anticommute with the initial operator $\mathcal{O}_0$. The second moments of the coefficients are

$$|\gamma_{p\neq\mathcal{O}_0}(t)|^2 \approx \frac{1}{d^2} - \frac{3\mathcal{R}_4}{d^6} + \frac{2\mathcal{R}_{4,1}}{d^5} \quad \text{if} \quad [\mathcal{O}_p, \mathcal{O}_0] = 0\,, \tag{5.67}$$

$$|\gamma_{p\neq\mathcal{O}_0}(t)|^2 \approx \frac{1}{d^2} + \frac{\mathcal{R}_4}{d^6} - \frac{2\mathcal{R}_{4,1}}{d^5} \quad \text{if} \quad \{\mathcal{O}_p, \mathcal{O}_0\} = 0\,. \tag{5.68}$$

This means that at early times the $\mathcal{O}_0$ coefficient decays from unity and all other coefficients are $\approx 1/d^2$. We note as a sanity check the non $\mathcal{O}_0$ coefficients vanish at $t = 0$ as $\mathcal{R}_4 = d^4$ and $\mathcal{R}_{4,1} = d^3$. Around the dip time, when all the form factors are ≈ 1, the coefficients are uniformly equal to $1/d^2$. At late-times there are $1/d^4$ fluctuations around the values $1/d^2$. The probabilities of the anticommuting operators increase with the ramp as the probabilities of the commuting operators decreases. The most interesting thing is that the initial operator becomes more likely again, with a probability twice that of the other operators. In summary,

$$\text{Dip}: \quad |\gamma_p(t)|^2 \approx \frac{1}{d^2}\,, \qquad \text{Late}: \quad |\gamma_{\mathcal{O}_0}(t)|^2 \approx \frac{2}{d^2} \quad \text{and} \quad |\gamma_{p\neq\mathcal{O}_0}(t)|^2 \approx \frac{1}{d^2}\,. \tag{5.69}$$

At the dip time, after the support on the initial operator has decayed, all operators are equally likely. But at late-times, in the plateau regime, the weight on the initial operator is twice that of all other operators.

Epilogue

I know that all beneath the moon decays,
And what by mortals in this world is brought,
In Time's great periods shall return to nought;
That fairest states have fatal nights and days;

-William Drummond of Hawthornden[1]

[1]Likely referring to the universality of chaotic dynamics

Bibliography

[1] P. Hayden and J. Preskill, "Black holes as mirrors: Quantum information in random subsystems," *JHEP* **09** (2007) 120, arXiv:0708.4025 [hep-th].

[2] Y. Sekino and L. Susskind, "Fast Scramblers," *JHEP* **10** (2008) 065, arXiv:0808.2096 [hep-th].

[3] N. Lashkari, D. Stanford, M. Hastings, T. Osborne, and P. Hayden, "Towards the Fast Scrambling Conjecture," *JHEP* **04** (2013) 022, arXiv:1111.6580 [hep-th].

[4] S. H. Shenker and D. Stanford, "Black holes and the butterfly effect," *JHEP* **03** (2014) 067, arXiv:1306.0622 [hep-th].

[5] S. H. Shenker and D. Stanford, "Stringy effects in scrambling," *JHEP* **05** (2015) 132, arXiv:1412.6087 [hep-th].

[6] A. Kitaev, "A simple model of quantum holography." Talks given at the KITP, Apr. 7, 2015 and May 27, 2015.

[7] J. Maldacena and D. Stanford, "Remarks on the Sachdev-Ye-Kitaev model," *Phys. Rev.* **D94** (2016) 106002, arXiv:1604.07818 [hep-th].

[8] O. Bohigas, M. J. Giannoni, and C. Schmit, "Characterization of Chaotic Quantum Spectra and Universality of Level Fluctuation Laws," *Phys. Rev. Lett.* **52** (1984) 1.

[9] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, "Black Holes and Random Matrices," arXiv:1611.04650 [hep-th].

[10] P. Hosur, X.-L. Qi, D. A. Roberts, and B. Yoshida, "Chaos in quantum channels," *JHEP* **02** (2016) 004, arXiv:1511.04021 [hep-th].

[11] D. A. Roberts and B. Yoshida, "Chaos and complexity by design," *JHEP* **04** (2017) 121, arXiv:1610.04903 [quant-ph].

[12] J. Cotler, N. Hunter-Jones, J. Liu, and B. Yoshida, "Chaos, Complexity, and Random Matrices," *JHEP* **11** (2017) 048, arXiv:1706.05400 [hep-th].

[13] M. Berry, "Quantum chaology, not quantum chaos," *Phys. Scr.* **40** (1989) 335.

[14] Lieb, Elliott H. and Robinson, Derek W., "The finite group velocity of quantum spin systems," *Commun. Math. Phys.* **28** (1972) 251.

[15] M. B. Hastings, "Locality in Quantum Systems," in *Quantum Theory from Small to Large Scales: Lecture Notes of the Les Houches Summer School*, vol. 95. 2010. arXiv:1008.5137 [math-ph].

[16] D. A. Roberts, D. Stanford, and L. Susskind, "Localized shocks," *JHEP* **03** (2015) 051, arXiv:1409.8180 [hep-th].

[17] E. P. Wigner, "Characteristic vectors of bordered matrices with infinite dimensions," *Ann. Math.* **62** (1955) 548.

[18] F. J. Dyson, "Statistical Theory of the Energy Levels of Complex Systems. I," *J. Math. Phys.* **3** (1962) 140.

[19] T. Guhr, A. Müller-Groeling, and H. A. Weidenmüller, "Random matrix theories in quantum physics: Common concepts," *Phys. Rept.* **299** (1998) 189, arXiv:cond-mat/9707301.

[20] F. Haake, *Quantum Signatures of Chaos*. Springer, 2010.

[21] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, "From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics," *Adv. Phys.* **65** (2016) 239, arXiv:1509.06411 [cond-mat.stat-mech].

[22] F. J. Dyson, "Statistical Theory of the Energy Levels of Complex Systems. III," *J. Math. Phys.* **3** (1962) 166.

[23] M. Mehta, *Random Matrices*. Pure and Applied Mathematics. Elsevier Science, 2004.

[24] E. Brézin and S. Hikami, "Spectral form factor in a random matrix theory," *Phys. Rev.* **E55** (1997) 4067, arXiv:cond-mat/9608116.

[25] E. Brézin and S. Hikami, "Extension of level-spacing universality," *Phys. Rev.* **E56** (1997) 264, arXiv:cond-mat/9702213.

[26] E. Brézin and A. Zee, "Universality of the correlations between eigenvalues of large random matrices," *Nucl. Phys. B* **402** (1993) 613.

[27] J. Maldacena, S. H. Shenker, and D. Stanford, "A bound on chaos," *JHEP* **08** (2016) 106, arXiv:1503.01409 [hep-th].

[28] A. I. Larkin and Y. N. Ovchinnikov, "Quasiclassical Method in the Theory of Superconductivity," *JETP* **28** (1969) 1200.

[29] D. A. Roberts and D. Stanford, "Two-dimensional conformal field theory and the butterfly effect," *Phys. Rev. Lett.* **115** (2015) 131603, arXiv:1412.5123 [hep-th].

[30] A. Kitaev and S. J. Suh, "The soft mode in the Sachdev-Ye-Kitaev model and its gravity dual," arXiv:1711.08467 [hep-th].

[31] M.-D. Choi, "Completely positive linear maps on complex matrices," *Linear Algebra Appl.* **10** (1975) 285.

[32] A. Jamiołkowski, "Linear transformations which preserve trace and positive semidefiniteness of operators," *Rep. Math. Phys.* **3** (1972) 275.

[33] D. N. Page, "Average entropy of a subsystem," *Phys. Rev. Lett.* **71** (1993) 1291, arXiv:gr-qc/9305007.

[34] A. Kitaev, "Hidden correlations in the hawking radiation and thermal noise." Talks given at the Fundamental Physics Prize Symposium, Nov. 10, 2014, and at the KITP, Feb. 12, 2015.

[35] S. Sachdev and J. Ye, "Gapless spin-fluid ground state in a random quantum heisenberg magnet," *Phys. Rev. Lett.* **70** (1993) 3339.

[36] J. M. Maldacena, "Eternal black holes in anti-de Sitter," *JHEP* **04** (2003) 021, arXiv:hep-th/0106112.

[37] A. L. Fitzpatrick, J. Kaplan, D. Li, and J. Wang, "On information loss in AdS_3/CFT_2," *JHEP* **05** (2016) 109, arXiv:1603.08925 [hep-th].

[38] A. L. Fitzpatrick and J. Kaplan, "On the Late-Time Behavior of Virasoro Blocks and a Classification of Semiclassical Saddles," *JHEP* **04** (2017) 072, arXiv:1609.07153 [hep-th].

[39] E. Dyer and G. Gur-Ari, "2D CFT Partition Functions at Late Times," arXiv:1611.04592 [hep-th].

[40] V. Balasubramanian, B. Craps, B. Czech, and G. Sárosi, "Echoes of chaos from string theory black holes," *JHEP* **03** (2017) 154, arXiv:1612.04334 [hep-th].

[41] Y.-Z. You, A. W. W. Ludwig, and C. Xu, "Sachdev-Ye-Kitaev model and thermalization on the boundary of many-body localized fermionic symmetry-protected topological states," *Phys. Rev.* **B95** (2017) 115150, arXiv:1602.06964 [cond-mat.str-el].

[42] A. M. García-García and J. J. M. Verbaarschot, "Spectral and thermodynamic properties of the Sachdev-Ye-Kitaev model," *Phys. Rev.* **D94** (2016) 126010, arXiv:1610.03816 [hep-th].

[43] T. Tao, *Topics in Random Matrix Theory*. Graduate studies in mathematics. American Mathematical Society, 2012.

[44] A. R. Brown and L. Susskind, "The Second Law of Quantum Complexity," arXiv:1701.01107 [hep-th].

[45] A. del Campo, J. Molina-Vilaplana, and J. Sonner, "Scrambling the spectral form factor: unitarity constraints and exact results," *Phys. Rev.* **D95** (2017) 126008, arXiv:1702.04350 [hep-th].

[46] L. Erdős and D. Schröder, "Phase Transition in the Density of States of Quantum Spin Glasses," *Math. Phys. Anal. Geom.* **17** (2014) 9164, arXiv:1407.1552 [math-ph].

[47] J. S. Cotler, G. R. Penington, and D. H. Ranard, "Locality from the Spectrum," arXiv:1702.06142 [quant-ph].

[48] D. Bagrets, A. Altland, and A. Kamenev, "Power-law out of time order correlation functions in the SYK model," *Nucl. Phys.* **B921** (2017) 727, arXiv:1702.08902 [cond-mat.str-el].

[49] F. G. S. L. Brandão, P. Ćwikliński, M. Horodecki, P. Horodecki, J. K. Korbicz, and M. Mozrzymas, "Convergence to equilibrium under a random Hamiltonian," *Phys. Rev.* **E86** (2012) 031101, arXiv:1108.2985 [quant-ph].

[50] S. W. Hawking, "Particle Creation by Black Holes," *Commun. Math. Phys.* **43** (1975) 199.

[51] W. G. Unruh, "Notes on black hole evaporation," *Phys. Rev.* **D14** (1976) 870.

[52] Z.-W. Liu, S. Lloyd, E. Y. Zhu, and H. Zhu, "Entanglement, quantum randomness, and complexity beyond scrambling," arXiv:1703.08104 [quant-ph].

[53] C. Dankert, R. Cleve, J. Emerson, and E. Livine, "Exact and approximate unitary 2-designs and their application to fidelity estimation," *Phys. Rev.* **A80** (2009) 012304, arXiv:quant-ph/0606161.

[54] F. G. S. L. Brandão, A. W. Harrow, and M. Horodecki, "Local Random Quantum Circuits are Approximate Polynomial-Designs," *Commun. Math. Phys.* **346** (2016) 397, arXiv:1208.0692 [quant-ph].

[55] F. Pastawski, B. Yoshida, D. Harlow, and J. Preskill, "Holographic quantum error-correcting codes: Toy models for the bulk/boundary correspondence," *JHEP* **06** (2015) 149, arXiv:1503.06237 [hep-th].

[56] P. Hayden, S. Nezami, X.-L. Qi, N. Thomas, M. Walter, and Z. Yang, "Holographic duality from random tensor networks," *JHEP* **11** (2016) 009, arXiv:1601.01694 [hep-th].

[57] A. J. Scott, "Optimizing quantum process tomography with unitary 2-designs," *J. Phys. A: Math. Theor.* **41** (2008) 055308, arXiv:0711.1017 [quant-ph].

[58] B. Collins, "Moments and cumulants of polynomial random variables on unitary groups, the Itzykson-Zuber integral, and free probability," *Int. Math. Res. Not.* **2003** (2003) 953, arXiv:math-ph/0205010.

[59] B. Collins and P. Śniady, "Integration with Respect to the Haar Measure on Unitary, Orthogonal and Symplectic Group," *Commun. Math. Phys.* **264** (2006) 773, arXiv:math-ph/0402073.

[60] D. Weingarten, "Asymptotic Behavior of Group Integrals in the Limit of Infinite Rank," *J. Math. Phys.* **19** (1978) 999.

[61] P. Diaconis and M. Shahshahani, "On the eigenvalues of random matrices," *J. Appl. Prob.* **31** (1994) 49.

[62] S. Bravyi, M. B. Hastings, and F. Verstraete, "Lieb-Robinson Bounds and the Generation of Correlations and Topological Quantum Order," *Phys. Rev. Lett.* **97** (2006) 050401, arXiv:quant-ph/0603121.

[63] L. Susskind, "Computational Complexity and Black Hole Horizons," *Fortsch. Phys.* **64** (2016) 24, arXiv:1403.5695 [hep-th].

[64] A. R. Brown, D. A. Roberts, L. Susskind, B. Swingle, and Y. Zhao, "Complexity, action, and black holes," *Phys. Rev.* **D93** (2016) 086006, arXiv:1512.04993 [hep-th].

[65] X. Chen, Z. C. Gu, and X. G. Wen, "Local unitary transformation, long-range quantum entanglement, wave function renormalization, and topological order," *Phys. Rev.* **B82** (2010) 155138, arXiv:1004.3835 [cond-mat.str-el].

[66] D. Harlow and P. Hayden, "Quantum Computation vs. Firewalls," *JHEP* **06** (2013) 085, arXiv:1301.4504 [hep-th].

[67] Z.-C. Yang, A. Hamma, S. M. Giampaolo, E. R. Mucciolo, and C. Chamon, "Entanglement complexity in quantum many-body dynamics, thermalization, and localization," *Phys. Rev.* **B96** (2017) 020408, arXiv:1703.03420 [cond-mat.str-el].

[68] E. Brézin and S. Hikami, *Random Matrix Theory with an External Source*. Springer Briefs in Mathematical Physics. Springer Singapore, 2017.

[69] Y. Huang, F. G. S. L. Brandão, and Y.-L. Zhang, "Finite-size scaling of out-of-time-ordered correlators at late times," arXiv:1705.07597 [quant-ph].

[70] M. V. Berry, "Regular and irregular semiclassical wavefunctions," *J. Phys. A: Math. Gen.* **10** (1977) 2083.

[71] M. Srednicki, "Chaos and quantum thermalization," *Phys. Rev.* **E50** (1994) 888, arXiv:cond-mat/9403051.

[72] W. W. Ho and D. Radičević, "The Ergodicity Landscape of Quantum Theories," *Int. J. Mod. Phys.* **A33** (2018) 1830004, arXiv:1701.08777 [quant-ph].

[73] J. Sonner and M. Vielma, "Eigenstate thermalization in the Sachdev-Ye-Kitaev model," arXiv:1707.08013 [hep-th].

[74] J. M. Magan, "Random free fermions: An analytical example of eigenstate thermalization," *Phys. Rev. Lett.* **116** (2016) 030401, arXiv:1508.05339 [quant-ph].

[75] H. Gharibyan, M. Hanada, S. H. Shenker, and M. Tezuka, "Onset of Random Matrix Behavior in Scrambling Systems," arXiv:1803.08050 [hep-th].

[76] D. N. Page, "Information in black hole radiation," *Phys. Rev. Lett.* **71** (1993) 3743, arXiv:hep-th/9306083.

[77] W. Brown and O. Fawzi, "Decoupling with random quantum circuits," *Commun. Math. Phys.* **340** (2015) 867, arXiv:1307.0632 [quant-ph].

[78] R. A. Low, "Pseudo-randomness and Learning in Quantum Computation," arXiv:1006.5227 [quant-ph]. PhD Thesis, 2010.

[79] R. E. Prange, "The Spectral Form Factor Is Not Self-Averaging," *Phys. Rev. Lett.* **78** (1997) 2280, arXiv:chao-dyn/9606010.

[80] F. J. Dyson, "The Threefold Way. Algebraic Structure of Symmetry Groups and Ensembles in Quantum Mechanics," *J. Math. Phys.* **3** (1962) 1199.

[81] A. Altland and M. R. Zirnbauer, "Nonstandard symmetry classes in mesoscopic normal-superconducting hybrid structures," *Phys. Rev. B* **55** (1997) 1142, arXiv:cond-mat/9602137.

[82] M. Creutz, "On invariant integration over SU(N)," *J. Math. Phys.* **19** (1978) 2043.

[83] S. Samuel, "U(N) Integrals, $1/N$, and the De Wit-'t Hooft anomalies," *J. Math. Phys.* **21** (1980) 2695.

[84] P. Mello and T. Seligman, "On the entropy approach to statistical nuclear reactions," *Nucl. Phys.* **A344** (1980) 489.

[85] P. A. Mello, "Averages on the unitary group and applications to the problem of disordered conductors," *J. Phys. A: Math. Gen.* **23** (1990) 4061.

[86] P. W. Brouwer and C. W. J. Beenakker, "Diagrammatic method of integration over the unitary group, with applications to quantum transport in mesoscopic systems," *J. Math. Phys.* **37** (1996) 4904, arXiv:cond-mat/9604059.

[87] N. Argaman and A. Zee, "Diagrammatic theory of random scattering matrices for normal-metal–superconducting mesoscopic junctions," *Phys. Rev.* **B54** (1996) 7406, arXiv:cond-mat/9603136.

[88] B. Collins and S. Matsumoto, "On some properties of orthogonal weingarten functions," *J. Math. Phys.* **50** (2009) 113516, arXiv:0903.5143 [math-ph].

[89] P. Zinn-Justin, "Jucys-Murphy Elements and Weingarten Matrices," *Lett. Math. Phys.* **91** (2010) 119, arXiv:0907.2719 [math.CO].

[90] T. Banica, "The Orthogonal Weingarten Formula in Compact Form," *Lett. Math. Phys.* **91** (2010) 105, arXiv:0906.4694 [math.CA].

[91] S. Matsumoto, "General moments of matrix elements from circular orthogonal ensembles," *Random Matrices: Theory Appl.* **01** (2012) 1250005, arXiv:1109.2409 [math.PR].

[92] S. Matsumoto, "Weingarten calculus for matrix ensembles associated with compact symmetric spaces," *Random Matrices: Theory Appl.* **02** (2013) 1350001, arXiv:1301.5401 [math.PR].

[93] Z. Webb, "The Clifford group forms a unitary 3-design," *Quantum Information & Computation* **16** (2016) 1379, arXiv:1510.02769 [quant-ph].

[94] B. Collins, S. Matsumoto, and N. Saad, "Integration of invariant matrices and moments of inverses of Ginibre and Wishart matrices," *J. Multivar. Anal.* **126** (2014) 1, arXiv:1205.0956 [math.ST].

[95] B. Collins and M. Stolz, "Borel theorems for random matrices from the classical compact symmetric spaces," *Ann. Probab.* **36** (2008) 876–895, arXiv:math/0611708 [math.PR].

[96] B. Yoshida and A. Kitaev, "Efficient decoding for the Hayden-Preskill protocol," arXiv:1710.03363 [hep-th].

[97] I. G. Macdonald, *Symmetric Functions and Hall Polynomials*. Oxford Mathematical Monographs. Oxford University Press, 1999.

[98] K. Jensen, "Chaos in AdS_2 Holography," *Phys. Rev. Lett.* **117** (2016) 111601, arXiv:1605.06098 [hep-th].

[99] J. Maldacena, D. Stanford, and Z. Yang, "Conformal symmetry and its breaking in two dimensional Nearly Anti-de-Sitter space," *PTEP* **2016** (2016) 12C104, arXiv:1606.01857 [hep-th].

[100] W. Fu, D. Gaiotto, J. Maldacena, and S. Sachdev, "Supersymmetric Sachdev-Ye-Kitaev models," *Phys. Rev.* **D95** (2017) 026009, arXiv:1610.08917 [hep-th]. [Addendum: Phys. Rev. D95 (2017) 069904].

[101] C. Peng, M. Spradlin, and A. Volovich, "A Supersymmetric SYK-like Tensor Model," *JHEP* **05** (2017) 062, arXiv:1612.03851 [hep-th].

[102] J. Murugan, D. Stanford, and E. Witten, "More on Supersymmetric and 2d Analogs of the SYK Model," *JHEP* **08** (2017) 146, arXiv:1706.05362 [hep-th].

[103] J. Yoon, "Supersymmetric SYK Model: Bi-local Collective Superfield/Supermatrix Formulation," *JHEP* **10** (2017) 172, arXiv:1706.05914 [hep-th].

[104] C. Peng, M. Spradlin, and A. Volovich, "Correlators in the $\mathcal{N} = 2$ Supersymmetric SYK Model," *JHEP* **10** (2017) 202, arXiv:1706.06078 [hep-th].

[105] T. Kanazawa and T. Wettig, "Complete random matrix classification of SYK models with $\mathcal{N} = 0$, 1 and 2 supersymmetry," *JHEP* **09** (2017) 050, arXiv:1706.03044 [hep-th].

[106] S. Forste and I. Golla, "Nearly AdS_2 sugra and the super-Schwarzian," *Phys. Lett.* **B771** (2017) 157, arXiv:1703.10969 [hep-th].

[107] T. Li, J. Liu, Y. Xin, and Y. Zhou, "Supersymmetric SYK model and random matrix theory," *JHEP* **06** (2017) 111, arXiv:1702.01738 [hep-th].

[108] N. Hunter-Jones, J. Liu, and Y. Zhou, "On thermalization in the SYK and supersymmetric SYK models," *JHEP* **02** (2018) 142, arXiv:1710.03012 [hep-th].

[109] Vinayak and M. Žnidarič, "Subsystem dynamics under random Hamiltonian evolution," *J. Phys. A: Math. Theor.* **45** (2012) 125204, arXiv:1107.6035 [quant-ph].

[110] T. C. Bachlechner, D. Marsh, L. McAllister, and T. Wrase, "Supersymmetric Vacua in Random Supergravity," *JHEP* **01** (2013) 136, arXiv:1207.2763 [hep-th].

[111] D. Stanford and E. Witten, "Fermionic Localization of the Schwarzian Theory," *JHEP* **10** (2017) 008, arXiv:1703.04612 [hep-th].

[112] D. Bagrets, A. Altland, and A. Kamenev, "Sachdev-Ye-Kitaev model as Liouville quantum mechanics," *Nucl. Phys.* **B911** (2016) 191, arXiv:1607.00694 [cond-mat.str-el].

[113] T. G. Mertens, G. J. Turiaci, and H. L. Verlinde, "Solving the Schwarzian via the Conformal Bootstrap," *JHEP* **08** (2017) 136, arXiv:1705.08408 [hep-th].

[114] A. M. García-García and J. J. M. Verbaarschot, "Analytical Spectral Density of the Sachdev-Ye-Kitaev Model at finite N," *Phys. Rev.* **D96** (2017) 066012, arXiv:1701.06593 [hep-th].

[115] A. V. Andreev and B. L. Altshuler, "Spectral statistics beyond random matrix theory," *Phys. Rev. Lett.* **75** (1995) 902, arXiv:cond-mat/9503141.

[116] V. A. Marčenko and L. A. Pastur, "Distribution of eigenvalues for some sets of random matrices," *Math. USSR-Sb.* **1** (1967) 457. Orig. in Mat. Sb. **72** (1967) 507.

[117] T. Nagao and M. Wadati, "Correlation functions of random matrix ensembles related to classical orthogonal polynomials," *J. Phys. Soc. Jpn.* **60** (1991) 3298.

[118] R. A. Davison, W. Fu, A. Georges, Y. Gu, K. Jensen, and S. Sachdev, "Thermoelectric transport in disordered metals without quasiparticles: The Sachdev-Ye-Kitaev models and holography," *Phys. Rev.* **B95** (2017) 155131, arXiv:1612.00849 [cond-mat.str-el].

[119] K. Bulycheva, "A note on the SYK model with complex fermions," *JHEP* **12** (2017) 069, arXiv:1706.07411 [hep-th].

[120] C. T. Asplund, A. Bernamonti, F. Galli, and T. Hartman, "Entanglement Scrambling in 2d Conformal Field Theory," *JHEP* **09** (2015) 110, arXiv:1506.03772 [hep-th].

[121] E. Perlmutter, "Bounding the Space of Holographic CFTs with Chaos," *JHEP* **10** (2016) 069, arXiv:1602.08272 [hep-th].

[122] X. Chen and A. W. W. Ludwig, "Universal Spectral Correlations in the Chaotic Wave Function, and the Development of Quantum Chaos," arXiv:1710.02686 [cond-mat.str-el].

[123] T. Iadecola and T. H. Hsieh, "Floquet Supersymmetry," arXiv:1710.05927 [cond-mat.str-el].

[124] J. Emerson, E. Livine, and S. Lloyd, "Convergence conditions for random quantum circuits," *Phys. Rev. A* **72** (2005) 060302, arXiv:quant-ph/0503210.

[125] M. Žnidarič, "Exact convergence times for generation of random bipartite entanglement," *Phys. Rev. A* **78** (2008) 032324, arXiv:0809.0554.

[126] A. W. Harrow and R. A. Low, "Random Quantum Circuits are Approximate 2-designs," *Commun. Math. Phys.* **291** (2009) 257, arXiv:0802.1919 [quant-ph].

[127] A. Nahum, J. Ruhman, S. Vijay, and J. Haah, "Quantum Entanglement Growth Under Random Unitary Dynamics," *Phys. Rev.* **X7** (2017) 031016, arXiv:1608.06950 [cond-mat.stat-mech].

[128] A. Nahum, S. Vijay, and J. Haah, "Operator Spreading in Random Unitary Circuits," *Phys. Rev.* **X8** (2018) 021014, arXiv:1705.08975 [cond-mat.str-el].

[129] C. von Keyserlingk, T. Rakovszky, F. Pollmann, and S. Sondhi, "Operator hydrodynamics, OTOCs, and entanglement growth in systems without conservation laws," *Phys. Rev.* **X8** (2018) 021013, arXiv:1705.08910 [cond-mat.str-el].

[130] T. Rakovszky, F. Pollmann, and C. W. von Keyserlingk, "Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation," arXiv:1710.09827 [cond-mat.stat-mech].

[131] V. Khemani, A. Vishwanath, and D. A. Huse, "Operator spreading and the emergence of dissipation in unitary dynamics with conservation laws," arXiv:1710.09835 [cond-mat.stat-mech].

[132] I. L. Aleiner, L. Faoro, and L. B. Ioffe, "Microscopic model of quantum butterfly effect: out-of-time-order correlators and traveling combustion waves," *Annals Phys.* **375** (2016) 378, arXiv:1609.01251 [cond-mat.stat-mech].

[133] A. A. Patel, D. Chowdhury, S. Sachdev, and B. Swingle, "Quantum butterfly effect in weakly interacting diffusive metals," *Phys. Rev.* **X7** (2017) 031047, arXiv:1703.07353 [cond-mat.str-el].

[134] S. Xu and B. Swingle, "Accessing scrambling using matrix product operators," arXiv:1802.00801 [quant-ph].

[135] C. Jonay, D. A. Huse, and A. Nahum, "Coarse-grained dynamics of operator and state entanglement," arXiv:1803.00089 [cond-mat.stat-mech].

[136] T. Zhou and A. Nahum, "Emergent statistical mechanics of entanglement in random unitary circuits," arXiv:1804.09737 [cond-mat.stat-mech].