Onset of random matrix behavior in scrambling systems

Hrant Gharibyan,[a] **Masanori Hanada,**[a,b,c,d] **Stephen H. Shenker**[a] **and Masaki Tezuka**[e]

[a]*Stanford Institute for Theoretical Physics, Stanford University,*
Stanford, CA 94305, U.S.A.

[b]*Yukawa Institute for Theoretical Physics, Kyoto University,*
Kyoto 606-8502, Japan

[c]*The Hakubi Center for Advanced Research, Kyoto University,*
Kyoto 606-8501, Japan

[d]*Department of Physics, University of Colorado,*
Boulder, Colorado 80309, U.S.A.

[e]*Department of Physics, Kyoto University,*
Kyoto 606-8502, Japan

E-mail: hrant@stanford.edu, hanada@yukawa.kyoto-u.ac.jp,
sshenker@stanford.edu, tezuka@scphys.kyoto-u.ac.jp

ABSTRACT: The fine grained energy spectrum of quantum chaotic systems is widely believed to be described by random matrix statistics. A basic scale in such a system is the energy range over which this behavior persists. We define the corresponding time scale by the time at which the linearly growing ramp region in the spectral form factor begins. We call this time t_{ramp}. The purpose of this paper is to study this scale in many-body quantum systems that display strong chaos, sometimes called scrambling systems. We focus on randomly coupled qubit systems, both local and k-local (all-to-all interactions) and the Sachdev-Ye-Kitaev (SYK) model. Using numerical results, analytic estimates for random quantum circuits, and a heuristic analysis of Hamiltonian systems we find the following results. For geometrically local systems with a conservation law we find t_{ramp} is determined by the *diffusion time* across the system, order N^2 for a 1D chain of N qubits. This is analogous to the behavior found for local one-body chaotic systems. For a k-local system like SYK the time is order $\log N$ but with a different prefactor and a different mechanism than the scrambling time. In the absence of any conservation laws, as in a generic random quantum circuit, we find $t_{\text{ramp}} \sim \log N$, independent of connectivity.

Contents

1 Introduction and summary

Dynamical aspects of many-body quantum chaos are of fundamental interest in various fields of physics, including condensed matter physics, quantum information theory and quantum gravity. Random Matrix Theory (RMT) provides us with a simple characterization of quantum chaos: if the level statistics at small energy separation agrees with RMT, one concludes the system under consideration is chaotic. There is a rich theory of this for few-body semiclassical quantum systems (for a general reference see [1]). For many-body systems there is a substantial amount of numerical evidence, but not yet a fully developed theory (for recent analytic approaches see [2–4]).

This work was motivated by recent numerical work we and others [6–8] performed on the Sachdev-Ye-Kitaev (SYK) model [9, 10, 12], which has become a reference model for the study of many-body quantum chaos.

Another diagnostic of strong chaos is sensitive dependence to initial conditions, classically diagnosed by the presence of a positive Lyapunov exponent. The semiclassical quantum analog of this in few-body chaos was discussed in the 1960s [13], and diagnosed by out-of-time-order correlation functions (OTOC). Motivated by a desire to understand the quantum chaotic nature of black holes [14, 15], these diagnostics were rediscovered in the general many-body context [16] and have been studied extensively in recent years [10, 17–23]. A positive value of the analogous "chaos exponent" is a signature of strong many-body chaos. Systems with such behavior are sometimes referred to as "fast scramblers" [15]. There is a characteristic time scale for such systems, the scrambling time t_{scr}, that measures when an initial perturbation affecting only a few degrees of freedom has chaotically spread to all the degrees of freedom in the system.

Such systems also display a simpler signature of thermalization. Two point functions of simple operators exponentially decay. We call the decay time of such quantities the "dissipation time" t_{diss}. In general this time is different from the inverse of the chaos exponent.

There is also a natural time scale associated to RMT level statistics. In the energy domain one can ask how far in energy space do random matrix statistics persist, that is how many energy levels apart can one go before the characteristic correlations of RMT are lost. The inverse of this energy defines a time, which in the many-body context we will call the "ramp time", t_{ramp} for reasons that will be made clear below.

The RMT behavior of the fine-grained energy spectrum is *universal* in quantum chaotic systems.[1] The structure of the spectrum at time scales shorter than t_{ramp} contains nonuniversal information about the specific system.

In simpler, one-body, systems this time scale has been studied extensively [24, 25]. For a single quantum particle propagating in a random medium, or equivalently for a banded random matrix, this time scale is the time for a particle to diffuse across the system t_{diff}. The corresponding energy is called the "Thouless energy" and we will sometimes refer to this time scale as the "Thouless time".

---

[1]The type of random matrix ensemble just depends on the symmetry class of the Hamiltonian.

The goal of this paper is to study the relationship between these various time scales in strongly chaotic many-body quantum systems.[2] These time scales are connected, via gauge/gravity duality, to phenomena in quantum black holes. This connection provides an important motivation for our work.

To disentangle the various phenomena involved we study two different types of system: k-local systems, with degrees of freedom nonlocally coupled, but only k degrees of freedom interacting at a time; and geometrically local systems, with only nearest neighbor couplings. The k-local systems we study are the SYK model and a qubit model of pairs of qubits interacting with Gaussian random couplings which we call the "2-local Randomly Coupled Qubits (RCQ)" model. The geometrically local systems we study are a one dimensional nearest neighbor qubit model with Gaussian random couplings (the "Local RCQ") and, for reference, banded matrices.

Robust analytic techniques to study RMT phenomena in such many-body systems have not yet been developed. This is a major goal for future research. So of necessity one part of our work involves numerical investigation. But powerful techniques have been developed to study related systems, ones in which the random couplings vary at each time step. For discrete time steps these are referred to as random quantum circuits, in the continuous time limit these are called Brownian circuits.[3]

As we will discuss various indicators of chaos are well captured by such systems. But a major difference with time independent Hamiltonian systems is the absence of conservation of energy. The fluctuating couplings pump energy into the system. To study the effects of conservation laws we use a recently introduced random quantum circuit that conserves spin [43, 48].

We are not able to give a definitive analysis of $t_{\rm ramp}$ in time independent Hamiltonian systems, but are able to give a sketch of an argument using a construction discussed in refs. [49, 50] that predicts results consistent with those obtained from random quantum circuits.

1.1 Summary of results

In section 2 we review the basics of RMT, including the density of states and the eigenvalue pair correlation function. In section 3 we introduce diagnostics of spectral rigidity, the hallmark of the spectrum of random matrices and quantum chaotic systems. We first discuss the "number variance" and point out artifacts that can affect this quantity in many-body systems. We then define the spectral form factor which is the Fourier transform of

---

[2]The first discussion of the Thouless (or ramp) time in the SYK model was presented in [26]. We discuss the limitations of their analysis in section 4.1.1.

[3]Random quantum circuits and Brownian circuits have played an important role in a number of recent developments. An early application in the black hole context was [27]. More recently these systems have been used in the quantum information community to provide rapid approximations to random unitary operators [28]. An important advance was the use of Markov chain techniques to estimate convergence times [29]. In particular these ideas have been applied to the study of approximate unitary k-designs [30–40]. These techniques were used to give the first logarithmic estimate of scrambling [18]. More recently they have been used to calculate OTOC for k-local [22] and geometrically local systems [41–43]. These models are currently being used as testbeds to establish "quantum supremacy" [44–47].

the pair correlation function. We discuss the robustness of this diagnostic in the presence of a varying density of states, and then point out the contamination of this signal due to sharp edges in the spectrum. We then introduce a new quantity YY^* that incorporates a Gaussian filter into the spectral form factor that removes the effect of sharp edges. We then discuss analogous quantities for random unitary matrices.

In section 4.1 we study the ramp time t_{ramp} in various k-local scrambling systems numerically. In the SYK model we use YY^* to eliminate the "slope" and show that t_{ramp} is very short, consistent with a $\log N$ dependence (although a slow power law cannot be ruled out). A ramp time $t_{\text{ramp}} \sim \log N$ is equivalent to spectral rigidity extending over $2^{N/2}/(\sqrt{N} \log N)$ eigenvalues, out of a total of $2^{N/2}$ eigenvalues.[4]

We then study a 2-local qubit system, the Randomly Coupled Qubit model (RCQ). Here again we find a short ramp time, consistent with $t_{\text{ramp}} \sim \log N$.

In order to understand the relation between this $\log N$ time and the scrambling time we turn to geometrically local systems in section 4.2. We mainly focus on a one dimensional nearest neighbor version of the RCQ. For such geometrically local systems the scrambling time is $t_{\text{scr}} \sim N$. Our numerical results indicate that $t_{\text{ramp}} \sim N^2$. Energy is conserved here and the diffusion time is $t_{\text{diff}} \sim N^2$, indicating that one must wait for all conserved quantities to diffuse across the system before random matrix behavior is established. This is one of our central findings and we give analytic arguments for it later.

In section 5 we turn to analytic methods. In particular we study analogs of the spectral form factor for random and Brownian quantum circuits where the couplings are varied at each time step. Powerful techniques using Markov chains exist to study such systems. We study k-local and geometrically local RCQ random circuits and find that the analog of the ramp time is $t_{\text{ramp}} \sim \log N$, independent of connectivity. The scrambling time in the one dimensional geometrically local random circuit is $t_{\text{scr}} \sim N$ so the mechanisms here are clearly different. In fact we find the source of the logarithm is the exponential decay of N two-point functions of single qubit operators. So the coefficient of $\log N$ is set by the dissipation time, not the Lyapunov time.

Energy is not conserved in random circuits and hence its diffusion cannot be studied there. So in order to study diffusion we examine in section 5.2 a random circuit that conserves spin, a version of the XXZ model. Here we find that the ramp time depends on connectivity. For the geometrically local case $t_{\text{ramp}} \sim N^2$. We are able to relate this to (many-body) diffusive effects. For the 2-local case we find that the diffusion time is order one, and so unimportant. Here $t_{\text{ramp}} \sim \log N$. Again we are able to trace this to the decay of N two point functions. In the geometrically local case these decay diffusively $\sim 1/t^{1/2}$ and in the 2-local case they decay exponentially with a scale given by the dissipation time.

Finally in section 6 we give a sketch of an argument to determine t_{ramp} in time independent Hamiltonian systems. One can rewrite the spectral form factor as a sum over two point functions of a complete basis of operators. Building on the random circuit intuition we argue that the result is dominated by the simplest operators that decay the slowest. This gives an answer analogous to the random circuit answer, with the time scale deter-

---

[4]Recall that the spacing between near-neighbor eigenvalues is $\sim \sqrt{N} 2^{-N/2}$.

mined by the dissipation time or the diffusive power law. There are important subtleties here involving factorization of multipoint functions that will need to be controlled to make this argument precise.

Other relevant work

There has been other recent work aimed at computing t_{ramp} in the SYK model [3, 7, 26]. García-García and Verbaarschot give a numerical study of the number variance. They find that spectral rigidity extends over of order N^2 eigenvalues in the spectrum. Translating their result to the time domain they predict $t_{\text{ramp}} \sim 2^{N/2}/N^{5/2}$ [7, 26]. This differs dramatically from our results. We discuss a possible explanation for this discrepancy in section 3.1.

Altland and Bagrets [3] use interesting techniques from few-body quantum chaos to argue that in the SYK model $t_{\text{ramp}} \sim \sqrt{N} \log N$. This disagrees with our N scaling. In addition the $N = 34$ numerical value for t_{ramp} they calculate disagrees with our numerical results presented below by roughly an order of magnitude. We do not understand this discrepancy, but make a few remarks about it in section 7. Their techniques may well serve to demonstrate the ramp at later times.

As this paper was in the final stages of editing the paper appeared. These authors [4] introduce a Floquet model without local conservation law with large site dimension q that can be treated analytically in the large q limit. This allows them to compute all moments of their unitary (which in this case corresponds to time evolution) analytically, a significant advance. For a one dimensional chain they find a ramp time of $\log N$ consistent with our results. For higher dimensions they find $t_{\text{ramp}} \sim 1$. For our systems we would expect a higher dimensional model with N sites to again have $t_{\text{ramp}} \sim \log N$. Perhaps the difference is due to the large q limit.

After submitting the first version of this paper we learned of a very interesting paper by Kos, Ljubotina, and Prosen [5] that we had missed. In this paper the authors discuss a clever model that injects just enough randomness to compute all moments of their Floquet unitary even with fixed site dimension. They see the ramp as the sum of cyclically permuted ladders that give the result in RMT (as do the authors of [4]) and make the connection to periodic orbit theory. They find a ramp time of $\log N$ in a way that seems quite generic, in accordance with our results. This technique looks quite powerful and seems amenable to generalization.

2 Random matrix theory

In this section we review a few basic properties of Random Matrix Theory (RMT) which are relevant for the following sections. For simplicity we restrict ourselves to Gaussian random Hermitian matrices, the Gaussian Unitary Ensemble (GUE) [51]. A standard reference is [52]. The partition function for the GUE is given by

$$\mathcal{Z}_{\text{GUE}} = \int \left(\prod_{i,j} dM_{ij} \right) e^{-\frac{L}{2}\text{Tr}M^2}, \tag{2.1}$$

where M is an $L \times L$ Hermitian matrix. We denote its eigenvalues E_i, because M is often taken to be the Hamiltonian of a quantum system. We introduce the normalized eigenvalue density

$$\hat{\rho}(E) = \frac{1}{L} \sum_i^L \delta(E - E_i) \,. \tag{2.2}$$

At large L the average density in this ensemble is given by the Wigner semicircle law

$$\langle \hat{\rho}(E) \rangle_{\text{GUE}} = \rho_{\text{sc}}(E) = \frac{1}{2\pi} \sqrt{4 - E^2}. \tag{2.3}$$

We see that with these normalizations the width of the entire band is order one and so the typical spacing between nearest neighbor eigenvalues is $\sim 1/L$.

The pair correlation function $R(E_1, E_2) = \langle \rho(E_1)\rho(E_2) \rangle$ is the probability of finding two eigenvalues at energies E_1 and E_2. If their separation $|E_1 - E_2| \ll 1$ and their average $\bar{E} = (E_1 + E_2)/2$ is not near the edge of the band $R(E_1, E_2)$ is given by the sine kernel formula [52], most simply stated by introducing a scaled variable

$$x = \pi L(E_1 - E_2)\rho(\bar{E}) \,. \tag{2.4}$$

Then we have

$$R(E_1, E_2) = \frac{\rho_{\text{sc}}(E_1)}{\pi L}\delta(E_1 - E_2) + \rho_{\text{sc}}(E_1)\rho_{\text{sc}}(E_2)\left(1 - \frac{\sin^2 x}{x^2}\right). \tag{2.5}$$

A sketch of this formula is given in figure 1. Note that $1/(L\rho_{\text{sc}}(\bar{E}))$ is the spacing between nearest neighbor energy levels in the vicinity of $\bar{E}$. In regions of small ρ_{sc} the eigenvalues are farther apart. The scale of the oscillations in (2.5) are set by this spacing. RMT universality means that formula (2.5) holds after scaling by the local eigenvalue density. When comparing different parts of the spectrum one must scale different regions differently. This process is called "unfolding", and becomes problematic when the fluctuations in the average density are large. We will see some effects of this later.

Note that when $|E_1 - E_2| \gg 1/L$, the nonconstant part of the pair correlation behaves like $\frac{1}{x^2}$. This shows that long range fluctuations in the eigenvalue "gas" are suppressed, a phenomenon referred to as "spectral rigidity". We can now give a more precise definition of the Thouless energy: it is the energy difference where the RMT form of spectral rigidity ceases to be valid.

In the following sections we will often use ensembles of unitary matrices. The Lie group of $L \times L$ unitary matrices has a unique measure that is invariant under left and right multiplication, Haar measure. In random matrix theory, the unitary group $\text{U}(L)$ with Haar probability measure is called the Circular Unitary Ensemble (CUE). The eigenvalues of each matrix are of the form $\lambda_k = e^{i\theta_k}$, where the eigenphases are defined in the range $0 \le \theta_k \le 2\pi$, and the probability density is [52]

$$\rho(\theta_1, \theta_2, \ldots, \theta_L) = \frac{1}{\mathcal{Z}_{\text{CUE}}} \prod_{1 \le k < j \le L} |e^{i\theta_k} - e^{i\theta_j}|^2 \,, \tag{2.6}$$
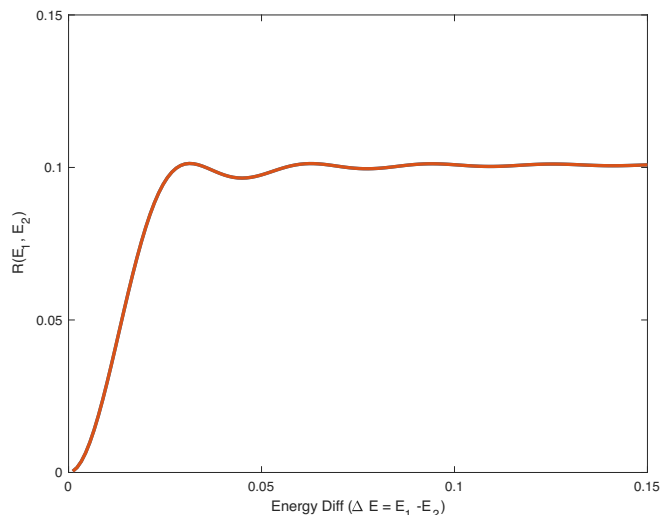
Figure 1. Pair correlation function $R(E_1, E_2)$ plotted when $E_2 = 0$ and $\Delta E = E_1$ is varied. For simplicity, we have used $L = 100$ and have suppressed the delta function in equation (2.5). The oscillations here are at the scale of the nearest neighbor eigenvalue spacing.

where $\mathcal{Z}_{\text{CUE}} = L! \cdot (2\pi)^L$ denotes the partition function. Note that the single eigenvalue density $\rho(\theta)$ is uniform on the unit circle. There are no edges, unlike the semicircle. Again, we see that with these normalizations the width of the entire band is order one and so the typical spacing between nearest neighbor eigenphases $(\theta_k) \sim 1/L$.

The pair correlation function $R(\theta_1, \theta_2) = \langle \rho(\theta_1)\rho(\theta_2) \rangle$ of eigenvalues is given by the sine kernel formula [52],

$$R(\theta_1, \theta_2) = \frac{1}{2\pi L}\delta(\theta_1 - \theta_2) + \frac{1}{(2\pi)^2}\left(1 - \frac{\sin^2 L(\theta_1 - \theta_2)/2}{L^2 \sin^2((\theta_1 - \theta_2)/2)}\right). \tag{2.7}$$

Note that, unlike the GUE ensemble, this formula is exact for all values of eigenvalue differences and not just for a small window. When the eigenvalues are close, expanding the denominator yields the same answer as GUE, as expected, because the same logarithmic repulsion acts in both cases.

3 Diagnostics of spectral rigidity

3.1 Number variance

A standard diagnostic of spectral rigidity in random matrix theory and quantum chaos is the number variance. This was first applied to the SYK model in [7]. The number variance is defined as

$$\Sigma^2(K) = \langle n^2(E, K)\rangle - \langle n(E, K)\rangle^2 \tag{3.1}$$

where $n(E, K)$ denotes the count of levels in the interval $[E, E + K\bar{\Delta}]$ where $\bar{\Delta}$ is the average nearest neighbor level spacing and the computation is done after unfolding the spectrum. If RMT universality is valid we expect this quantity to be essentially independent of E.

By the definition of $\bar{\Delta}$ we have $\langle n(E, K) \rangle = K$ and in an interval of length K the range of the expected number of energy levels is of order

$$K \pm \sqrt{\Sigma^2(K)}. \tag{3.2}$$

The number variance can be expressed in terms of the pair correlation function (see for example [53]). The result for GUE is

$$\Sigma^2(K) = \frac{1}{\pi^2} \left(\log(2\pi K) + \gamma + 1 \right) + O\left(\frac{1}{L^2} \right), \tag{3.3}$$

where $\gamma = 0.57721\ldots$ is Euler's constant. The crucial thing here is the very slow $\log K$ behavior which indicates a very rigid spectrum. Poisson distributed eigenvalues would give $\Sigma^2(K) \sim K$. This quantity has proved a very useful numerical diagnostic in various studies of few-body quantum chaotic systems [53]. But in many-body quantum chaotic systems, like SYK, there is an additional contribution to $\Sigma^2(K)$ that obscures the RMT signal. This complication was first pointed out in the study of embedded Gaussian ensembles [54]. The basic issue is that the fluctuations in the coarse grained eigenvalue density in these systems are only suppressed by powers of $1/N$. In contrast these fluctuations in random matrix theory and few-body chaos are suppressed by powers of $1/L$. Here L is the dimension of the Hilbert space and in many-body systems $L \sim e^N$. So the density fluctuations are much larger in many-body systems.

The main effect of this on $\Sigma^2(K)$ is easy to understand. Assume that the result of each sample is described by random matrix statistics rescaled by the eigenvalue density realized in that sample, *not* the mean eigenvalue density. Because the typical eigenvalue spacing Δ in each sample depends on this density it varies a lot from sample to sample. This means that $\Sigma^2(K)$ is actually of the following form (see formula 1 in [54])

$$\Sigma^2(K) \sim K^2 \frac{\langle (\Delta - \bar{\Delta})^2 \rangle}{\bar{\Delta}^2} + \mathcal{O}(\log K) + \ldots \tag{3.4}$$

For a system with zero mean energy ($\langle \text{Tr} H \rangle = 0$) and small variance a rough estimate is given by[5]

$$\frac{\langle (\Delta - \bar{\Delta})^2 \rangle}{\bar{\Delta}^2} \sim \frac{\langle (\text{Tr} H^2)^2 \rangle - (\langle \text{Tr} H^2 \rangle)^2}{(\langle \text{Tr} H^2 \rangle)^2} \tag{3.5}$$

For GUE (3.5) is of order $1/L^2$ so the correction in (3.4) is of order K^2/L^2 which is small until K covers a finite fraction of the whole energy band. In contrast in SYK ($q = 4$) the correction is of order K^2/N^4 (see for example [8] appendix F), which becomes large at $K \sim N^2$. This is the same answer found by [55] and according to [26] is consistent with numerics. In SYK the spectrum consists of $2^{N/2}$ eigenvalues and so this result would correspond to an exponentially small fraction of the spectrum. Note that this K^2/N^4 behavior has nothing to do with spectral rigidity and in fact will swamp the logarithmic signal for larger energy scales. We will argue later that spectral rigidity in this system extends over $2^{N/2}/(\sqrt{N} \log N)$ eigenvalues, corresponding to a Thouless time of $t \sim \log N$.

---

[5]This does not account for more complicated fluctuations in the shape of the energy spectrum.

We have made a preliminary check that this artifact is responsible for the results in [7, 26] by rescaling each sample to have the same variance. This increases the range of logarithmic behavior by roughly a factor of 2 for the largest N studied. These results are discussed in detail in appendix A.1. It is not clear to us whether there is even an in principle way to fully correct for these large fluctuations in this observable.[6]

3.2 Spectral form factor

We now discuss another diagnostic that is less sensitive to the kind of errors discussed above, the spectral form factor. This diagnostic also has a long history in random matrix theory and quantum chaos; see e.g. [52].[7] It was first used to study the SYK model in [8], motivated by a finite temperature extension discussed in the context of black hole physics in [59]. Here we only consider problems with finite dimensional Hilbert spaces and focus on generic behavior in the band so we restrict ourselves to infinite temperature.

The spectral form factor is defined as follows:

$$|Z(t)|^2 = \left| \mathrm{Tr} e^{-i\hat{H}t} \right|^2 = \left| \sum_j e^{-itE_j} \right|^2 = \sum_{ij} e^{i(E_i - E_j)t} \tag{3.6}$$

Using the eigenvalue correlator we can write

$$\langle |Z(t)|^2 \rangle = \int dE_1 dE_2 L^2 R(E_1, E_2) e^{i(E_1 - E_2)t}. \tag{3.7}$$

We will often discuss the normalized spectral form factor

$$g(t) = \langle |Z(t)|^2 \rangle / L^2 \tag{3.8}$$

defined so that $g(0) = 1$. Here again L is the dimension of the Hilbert space.

To get oriented let us look at this quantity for the GUE ensemble. As a first approximation we can replace $\rho_{\mathrm{sc}}(E)$ in (2.3) by 1, so the eigenvalue spacing is uniform and of order $1/L$ all across the band.[8] Then we can compute the Fourier transform of (2.5) giving the result

$$\langle |Z(t)|^2 \rangle \sim \begin{cases} t/(2\pi), & t < 2L \\ L/\pi, & t \geq 2L \end{cases} \tag{3.9}$$

This is plotted in the right panel of figure 2 for two different energy slices with different values of ρ. We called the linear t behavior at early time the ramp. It is the time domain reflection of spectral rigidity. The Thouless energy determines the earliest time for which the linear ramp behavior holds — the ramp time, t_{ramp}. In GUE this time is order one, set by the width of the band.

---

[6]Note that the variance in (3.5) goes like $1/N^q$ for general q SYK. So studying large q will improve this statistic. But we must take $q^2 < N$ to preserve k-locality (see [8, 56]). This is not sufficient to make the error in (3.4) smaller than the desired signal.

[7]For elegant analytic formulas in RMT see [57] and [58].

[8]We approximate further by extending the integration ranges to $(-\infty, \infty)$.
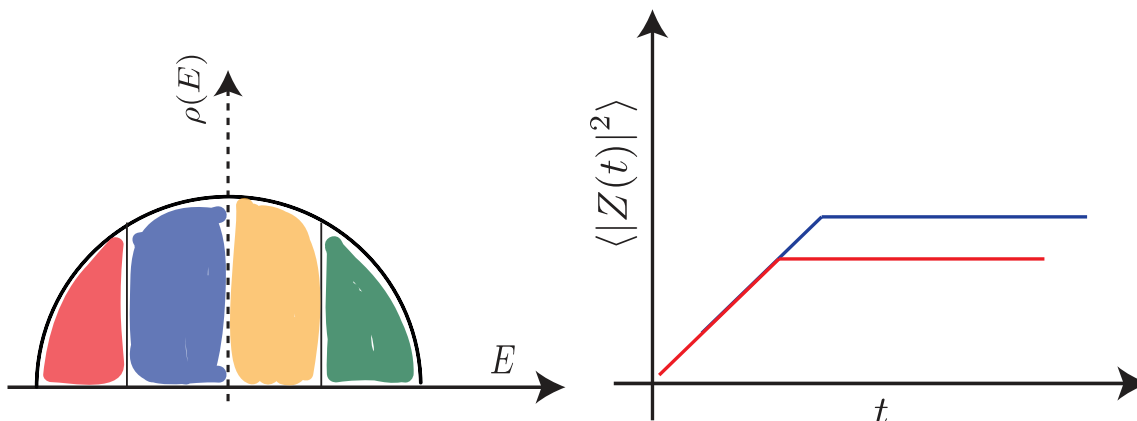
Figure 2. [Left] Equal width slices of the energy spectrum marked with different colors. [Right] Ramp and plateau coming from two slices marked with the corresponding colors.

We call the constant region the "plateau" and the time at which it begins the "plateau time", which is of order L.[9]

We now discuss the consequence of having an eigenvalue density $\rho(\bar{E})$ that varies across the band. We divide the entire energy band up into a large number of regions of equal energy width narrow enough so that ρ is roughly constant in each. Changing the constant value $\rho(\bar{E})$ in (2.5) does not change the overall form of (3.9). It only changes the plateau time and the height of the plateau (see figure 2). This means that summing up the contributions from each region to compute the total spectral form factor is very simple. Along the ramp one just adds up linear functions with equal coefficients. The final plateau height is just L, the total number of eigenvalues. The only significant change is a rounding of the sharp transition to the plateau, because different regions of the spectrum have different plateau times. The size of this effect depends on the shape of $\rho(E)$ and depends on the system being studied.

This insensitivity of the ramp means that fluctuations from sample to sample of the eigenvalue density do not affect the ramp structure, making it a robust diagnostic of the Thouless time even without unfolding. This is in contrast to the number variance, discussed above.

There is a problem, though, with using the spectral form factor to diagnose spectral rigidity when it is very long ranged because the ramp time will be very short and the height of the ramp will be very small compared to the plateau. So small effects can hide the ramp. The GUE ensemble offers a simple example of this effect when the variation of the spectral density is included. The leading contribution at relatively early times to the spectral form factor is just given by the Fourier transform of $\rho_{\rm sc}(E)$,

$$\langle |Z(t)|^2 \rangle \sim L^2 \left| \int dE \rho_{\rm sc}(E) e^{-iEt} \right|^2 = L^2 \left(\frac{J_1(2t)}{t} \right)^2 \sim L^2 \frac{\cos^2(2t)}{t^3} \tag{3.10}$$

---

[9]Note that from the definition (3.6) the presence of a plateau of height exactly L at sufficiently large t is manifest. The factor of $1/\pi$ here is an artifact of the approximations involving ρ.

The $\frac{L^2}{t^3}$ behavior for $t \gg 1$ is due to the sharp square root edge of the semicircle distribution and dominates until the "dip time" t_{dip} when it becomes smaller than the ramp, i.e., at a time when $L^2/t_{\text{dip}}^3 \sim t_{\text{dip}}$, $t_{\text{dip}} \sim L^{1/2}$.[10] This early time "slope" obscures the ramp until after t_{dip}, but the ramp time is typically much earlier than this. The necessity of following the ramp underneath the slope in SYK was already emphasized in [8].

There are strategies to deal with this. In GUE one can just subtract off the disconnected part of the spectral form factor, computing

$$\langle |Z(t)|^2 \rangle_c = \langle |Z(t)|^2 \rangle - |\langle Z(t) \rangle|^2 . \tag{3.11}$$

But this does not work in the SYK model, as explained in [8]. This model has a square root edge in its eigenvalue density [8, 61, 62] but fluctuations in the edge position from sample to sample are only suppressed by a power of N, the number of fermions, rather than L (equal to $2^{N/2}$ in SYK) as is the case in GUE. This fluctuation converts the power law slope into a Gaussian falloff in $|\langle Z(t) \rangle|^2$. But fluctuations in the edge position cancel out in $\langle |Z(t)|^2 \rangle$, leaving the $1/t^3$ slope.

In this paper we employ the alternate strategy, suggested to us by Douglas Stanford, of using a type of microcanonical ensemble with a Gaussian window to suppress the contribution of the edge. To see the onset of the ramp we introduce a variant of the spectral form factor

$$|Y(\alpha, t)|^2 = \left| \sum_i e^{-\alpha E_i^2 - itE_i} \right|^2 . \tag{3.12}$$

In other words, we replace the energy spectrum $\rho(E)$ with $e^{-\alpha E^2} \rho(E)$. This is in the spirit of a microcanonical density of states centered around $E = 0$ but with smooth edges to the energy window that will not create spurious signals in the time domain.[11] Although this changes the global shape of the energy spectrum, locally it is just a multiplication by a constant factor, and hence it does not spoil the RMT universality of the energy correlations.

When α is sufficiently large, any sharp edge of the spectrum is replaced with a Gaussian. Then Y, which is essentially the Fourier transform of this 'Gaussian-filtered energy spectrum', decays as a Gaussian at early time. Hence the non-universal part can be strongly suppressed, while preserving the universal part.

We also introduce the normalized quantity $h(\alpha, t) \equiv |Y(\alpha, t)|^2 / |Y(\alpha, t = 0)|^2$. Note $h(0, t) = g(t)$.

In the GUE case at early time

$$|Y(\alpha, t)|^2 \sim L^2 e^{-t^2/\alpha^2} . \tag{3.13}$$

This becomes of order one at $t \sim \sqrt{\log L}$. So spectral rigidity can be probed out to this scale. In fact $t_{\text{ramp}} \sim 1$ here, so we cannot fully diagnose this system using this technique.

---

[10]Many plots illustrating these phenomena are displayed in [8]. Another approach was discussed in [60], where authors use a non-Gaussian potential for RMT ensemble to change the edge.

[11]Of course we could choose to center this window around any energy we shoose, but to maximize the number of eigenvalues we choose to center it around $E = 0$. We expect the same universal features to be present as long as the center of the window is not extremely close to the edge of the spectrum.

In the many-body systems we will study we expect $t_{\rm ramp}$ to be large enough that, in principle at least, it can be fully revealed by this measure.

In our analysis we will also discuss the analog of the spectral form factor for the Haar random CUE ensemble. Here the relevant quantity is the (k-th moment) of the trace

$$\langle |{\rm Tr}(U^k)|^2 \rangle_{\rm CUE} = \Big\langle \sum_{ij} e^{i(\theta_i - \theta_j)k} \Big\rangle_{\rm CUE} = \int d\theta_1 d\theta_2 L^2 R(\theta_1, \theta_2) e^{i(\theta_1 - \theta_2)k} \tag{3.14}$$

Using the exact formula for the pair correlation function (2.7) and doing the integral we get [63, 64]

$$\langle |{\rm Tr}(U^k)|^2 \rangle_{\rm CUE} = \begin{cases} k, & k < L \\ L, & k \geq L \end{cases} \tag{3.15}$$

Note that the ramp and plateau are exact for the CUE, even at finite L. There is no slope because there is no eigenvalue edge. No unfolding is necessary because the eigenvalue density is uniform.

With these preliminaries out of the way we now turn to a systematic analysis of the models. We first discuss numerical data on time independent Hamiltonian systems, beginning with k-local systems.

4 Time independent Hamiltonian systems

4.1 k-local Hamiltonian systems

k-local systems have been studied as a good model for scrambling in black holes [14, 15]. These models thermalize very quickly, with a dissipation time of order one and a scrambling time of order $\log N$. Diffusive effects will also be very rapid. Here we will see numerically that $t_{\rm ramp}$ is also short, with a weak N dependence. In section 6 we will give an analytic argument that $t_{\rm ramp}$ also scales as $\log N$ but the physical mechanism is different from scrambling.

4.1.1 SYK model

Let us start with the SYK model [9–12]. This is a quantum mechanical model of N Majorana fermions satisfying the anticommutation relation

$$\{\hat{\psi}_a, \hat{\psi}_b\} = \delta_{ab}. \tag{4.1}$$

The dimension of the Hilbert space is $L = 2^{N/2}$. The Hamiltonian is given by

$$H = \frac{1}{4!} \sum_{a,b,c,d=1}^{N} J_{abcd} \hat{\psi}_a \hat{\psi}_b \hat{\psi}_c \hat{\psi}_d, \tag{4.2}$$

where J_{abcd} is a Gaussian random number with a mean zero and width $\sqrt{\frac{6}{N^3}} J$, and totally antisymmetric with respect to a, b, c, d. The disorder average (the average with respect to J_{abcd}) is taken after calculating observables.
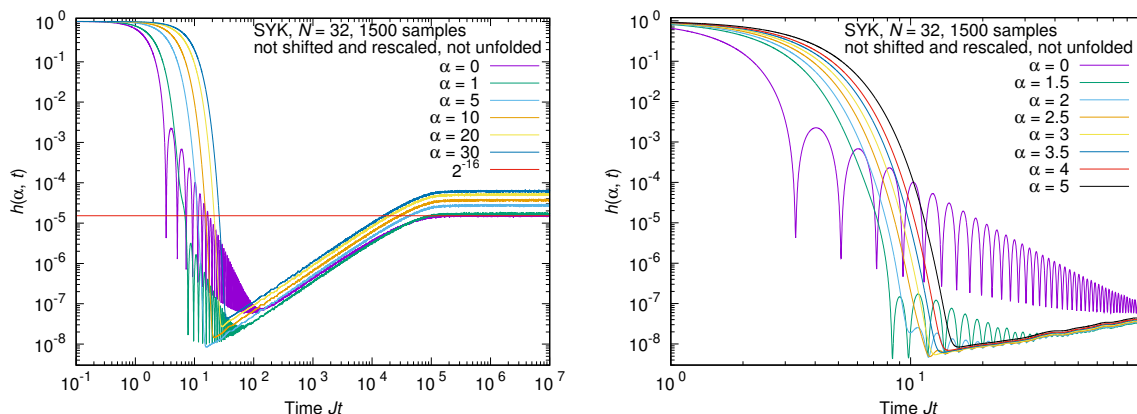
Figure 3. $h(\alpha, t) \equiv |Y(\alpha, t)|^2 / |Y(\alpha, t = 0)|^2$ at $N = 32$, with various values of α. Note that $g(t) = h(\alpha = 0, t)$.
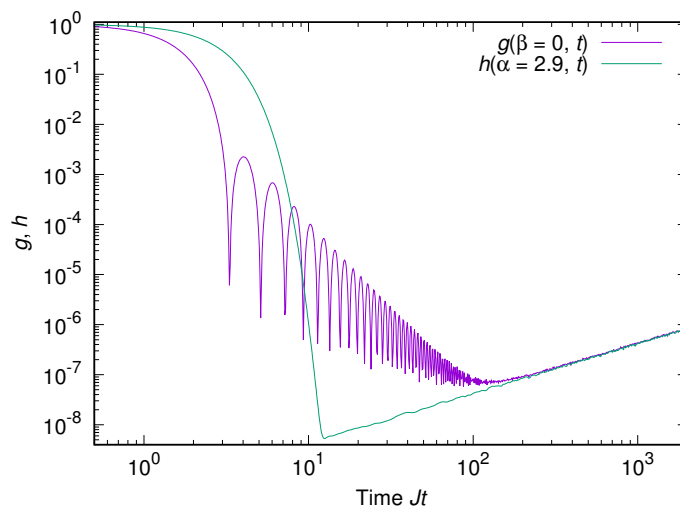
Figure 4. $g(t) = |Z(t)|^2 / |Z(0)|^2$ and $h(\alpha, t) = |Y(\alpha, t)|^2 / |Y(\alpha, t = 0)|^2$ with $\alpha = 2.9$ for $N = 32$, 1500 samples.

In the first and second panels of figure 3, h with $N = 32$ has been plotted for various values of α. As mentioned above the spectral form factor of the SYK model has a $\frac{1}{t^3}$ 'slope' which hides the onset of the ramp. This is visible in the $\alpha = 0$ curves where $h(t) = g(t)$ in figure 3. In figure 4 (using the optimal value of α) we can see that the ramp continues all the way down to the Gaussian envelope, and may well continue past it. This intersection time, which we call $t_{\min}$, gives an upper bound to the ramp time t_{ramp}. In the data displayed the ramp extends down to a time $t_{\min}$ of order 10. The plateau time is of order 10^5 here, so this represents a ratio of 10^4 in time or energy scales. Note that there are only $2^{16} \sim 66\,000$ energy levels[12] in the spectrum so rigidity extends across an appreciable

---

[12]After removing the two fold fermion parity degeneracy which is the only degeneracy present for $N = 32$ [50].

fraction of the entire spectrum![13] We should emphasize that as expected $t_{\min}$ is far less, over an order of magnitude here, than t_{dip}.

It seems from the graph that the ramp probably extends somewhat further but is just being masked by the Gaussian envelope. For this finite value of N it is impossible to do better because an envelope that decays faster in time corresponds to a Gaussian filter that is broader in energy and extends closer to the edge of the spectrum, allowing more contamination from the sharp edge. But as N grows this effect goes away because the edges of the spectrum are at energies $\pm cN$ (even though the standard deviation of the energy is of order $\sqrt{N}$). The Gaussian filter lets in an edge signal in $\langle |Y(\alpha, t)|^2 \rangle$ of order $e^{-2\alpha(cN)^2 + 2s_0 N}/t^3$ where $s_0 N$ is the zero temperature entropy. The early time ramp signal is of order one. So choosing α as small as $\frac{s_0}{c^2 N}$ is sufficient to mask the slope contribution from the edge. The Fourier transform of the Gaussian envelope decays $\sim e^{2s_\infty N - \frac{t^2}{4\alpha}}$ where $s_\infty N$ is the infinite temperature entropy. So an $\alpha \sim 1/N$ is small enough for the Gaussian envelope to become order one in a time of order one, allowing the study of t_{ramp} as short as order one. Thus this quantity provides a reliable definition of the Thouless time in SYK and other many-body systems with similar properties. The required N values may well be computationally prohibitive though.

We discuss our detailed algorithm for determining $t_{\min}$ in appendix A.2. The oscillations in the data make this determination quite noisy. The values listed in table 1 in the appendix provide upper bounds for t_{ramp}. The best we can say about the N dependence is that it is weak. Power laws faster than N^1 and exponentials, with order one coefficients, are disfavored. A $\log N$ behavior would certainly be consistent.

To gain more experience we now examine a k- local model without a sharp edge in its density of states.

4.1.2 2-local randomly coupled qubits (RCQ)

We now consider the 2-local-RCQ model.[14] This is a model of N spin-$(1/2)$ particles on a complete graph that interact with couplings of random strength. It has the same kind of k-local connectivity as the SYK model and hence the same pattern of sparseness in its Hamiltonian matrix. We therefore expect similar scaling properties of its random matrix statistics. The Hamiltonian is of the form

$$H = \frac{1}{\sqrt{8(N-1)}} \sum_{i<j}^{N} \sum_{\alpha,\beta=0}^{3} J_{ij}^{\alpha\beta} \sigma_i^\alpha \otimes \sigma_j^\beta \tag{4.3}$$

where σ_i^α represents an element of the set $\{\sigma^0, \sigma^1, \sigma^2, \sigma^3\} = \{I, X, Y, Z\}$ acting at the i-th site and couplings $J_{ij}^{\alpha\beta}$ are randomly drawn from a normal distribution. This choice of normalization is analogous to the typical SYK normalization discussed in section 4.1.1, such that

$$\langle \overline{\text{Tr}} H^2 \rangle_J = \frac{1}{L} \langle \text{Tr} H^2 \rangle_J = N \tag{4.4}$$

---

[13]Of course a precise comparison would require evaluating various numerical factors of order one.

[14]This model is often referred to as the "Quantum Spin Glass" [56]. We see no evidence of a spin glass phase in our system so we prefer a different name.
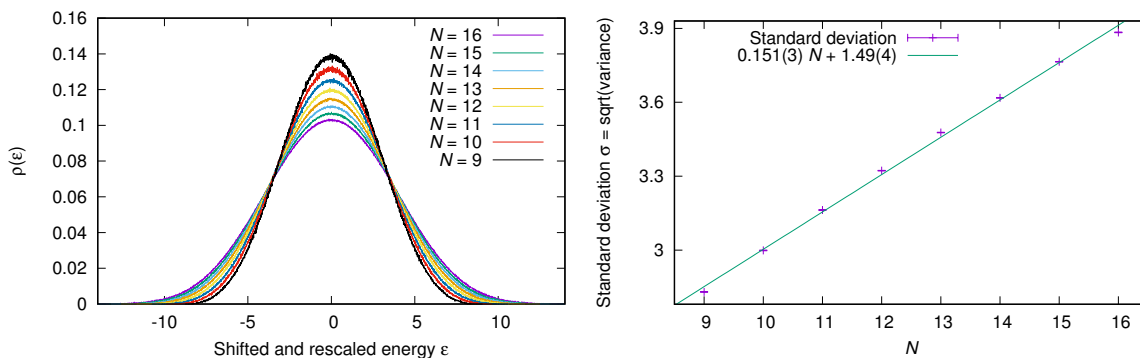
Figure 5. The density of shifted and renormalized eigenvalues for the 2-local RCQ model for $N = 9, 10, \ldots, 16$, and the averaged variance of eigenvalues with a straight line fit. $(2^{16-N} \times 100)$ samples have been used for each N.

where $L = 2^N$ is the dimension of the Hilbert space, $\langle \cdot \rangle_J$ denotes ensemble averaging over Gaussian couplings $J_{ij}^{\alpha\beta}$ and $\overline{\mathrm{Tr}}$ denotes the normalized trace.

We plot the eigenvalue density in figure 5.[15] We see that the energy distribution function $\rho(E)$ has a Gaussian tail for sufficiently large N ($N \gg 4$) (see [56] for discussion of the spectral density). Therefore the slope behaves like e^{-Nt^2} which decays very quickly, unlike SYK.

Spectral form factor

The spectral form factor for the 2-local RCQ system with N from 8 to 16 is shown in figure 6. There are several noteworthy features of this graph. First, as just discussed, the slope decays very quickly. Then there is a bump. This is followed by a short intermediate region. Then a ramp begins, at quite early times. To study the bump we have experimented with using YY^*. We were unable to isolate a sizable transition region to the ramp while excluding the bump. We do not understand its origin, unfortunately.

Nonetheless it is clear that the ramp begins early, at times between 2 and 5, increasing slowly with N. For $N = 16$ (see figure 7) the plateau time is about 10 000 times greater than the ramp time. The entire spectrum contains $2^{16} \sim 66\,000$ eigenvalues. So as in the SYK model we see rigidity extending across an appreciable part of the entire spectrum.

We now attempt a more quantitative study of the N dependence of t_{ramp}. In the absence of a detailed theory of the early time behavior we proceed phenomenologically. We define the time scale where the bump ends and the ramp starts as the ramp time, t_{ramp}. In order to define the ramp time systematically, we quantify the distance of the spectral form factor from the linear fit to the ramp, by using the fractional error

$$\epsilon(t) = \frac{g(t) - g_{\mathrm{ramp}}(t)}{g_{\mathrm{ramp}}(t)}. \tag{4.5}$$

Here we use $g_{\mathrm{ramp}}(t)$ to denote the linear fit to the ramp. We consider four different estimates for the ramp time, picking the points where $\epsilon(t)$ equal to 5 %, 10 %, 20 % and

---

[15]We give details about this figure and present other data about this model, including nearest neighbor eigenvalue histograms and the number variance in appendix A.3.
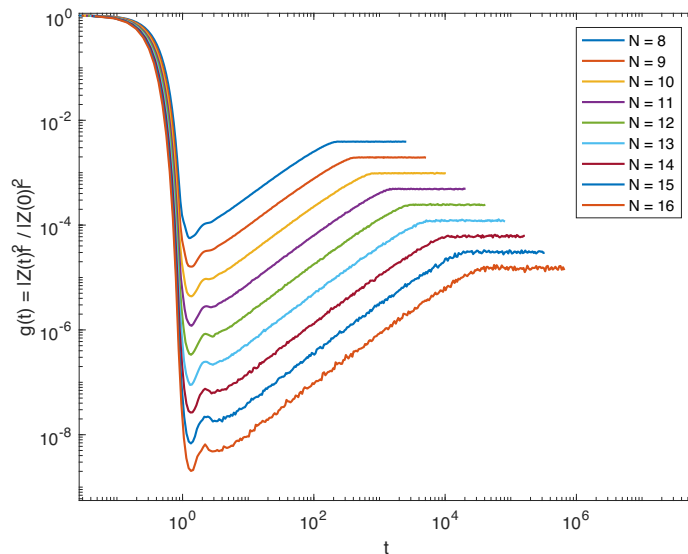
Figure 6. The spectral form factor for the k-local randomly coupled qubits (RCQ) with 8 to 16 qubits. We observe the Gaussian decay at the early times, follow by a 'bump', ramp and plateau.
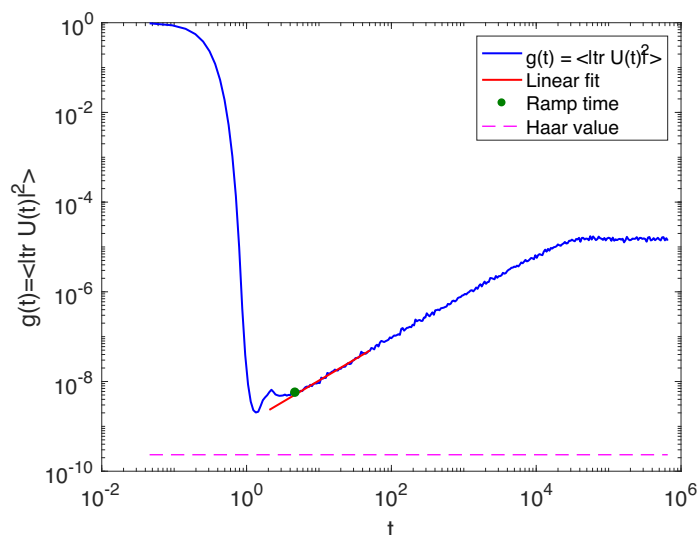
Figure 7. The spectral form factor for the k-local randomly coupled qubits (RCQ) for $N = 16$ qubits is given by the blue curve. We fit a line to the ramp (red line) and determine the time where the fractional error from the linear fit $\epsilon(t)$ equal to $5\,\%, 10\,\%, 20\,\%$ or $30\,\%$. In the plot the green dot represents the ramp time estimate when fractional error is 10 %.

30 %. See figure 7 for an example of an actual fit. We perform this procedure for all the values of N ranging from 8 to 16 and plot the estimates of the ramp time with respect to N. In figure 8, we plot an estimate of the ramp time for four different values of the fractional error. Different values of the error systematically shift the curve.

As mentioned previously our theoretical expectation is that t_{ramp} scales like $\log N$ although for reasons unrelated to the logarithmic behavior of the scrambling time. So
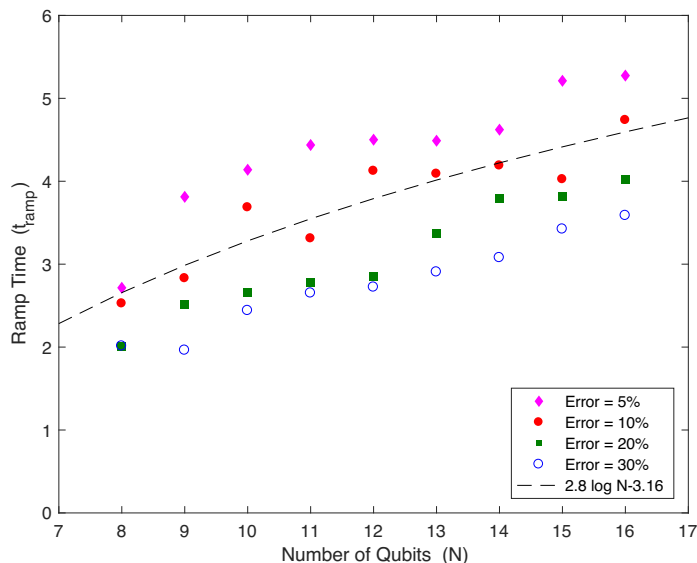
Figure 8. We plot the estimate of t_{ramp} for different N and different choices of fraction errors. The dashed line is the $t(N) = A \log N + B$ fit to the red dots. We find that the best fit for ($\epsilon = 10$ %) is given by $t_{\mathrm{ramp}} = 2.8 \log N - 3.16$.

in figure 8, we fit a function of the form $A \log N + B$ to the data with $\epsilon = 10$ %. (For other values of the fractional error the fit will have the same A, but different values of the offset B.) t_{ramp} is consistent with $A \log N + B$, although it is far from conclusive. A power law like $N^{1/2}$ fits equally well. A more rapid power, faster than N^1 or so, or an exponential dependence, with order one coefficients is disfavored. Larger N would be required to distinguish these functional forms.

To help disentangle the physical mechanisms at work here we now turn to a geometrically local system.

4.2 Local Hamiltonian systems

We will consider two examples of one dimensional local Hamiltonian systems, a nearest neighbor coupled RCQ model and the XXZ chain with random magnetic field.

4.2.1 Local randomly coupled qubits (local-RCQ)

In geometrically local systems (which we will refer to as "local", as opposed to k-local) the various interesting time scales become distinct from one another and so it is easier to diagnose the effect of different physical mechanisms. In a one dimensional local spin system we expect the dissipation time $t_{\mathrm{diss}} \sim 1$, the scrambling time $t_{\mathrm{scr}} \sim N$ and the diffusion time $t_{\mathrm{diff}} \sim N^2$. These time scales vary more quickly with N than in the k-local case making them easier to study numerically. In this system we find that $t_{\mathrm{ramp}} \sim N^2$ and hence conjecture that it is related to diffusion, not scrambling.
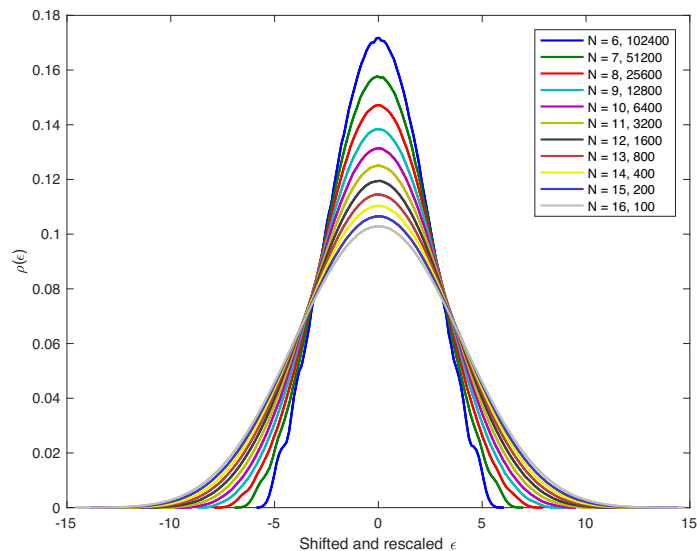
Figure 9. Spectrum of the local spin chain (Local-RCQ) for $N = 6$ to 16, as well as number of samples used for each N. Spectrum appears close to Gaussian as low as $N = 6$, where it starts developing edges characteristic to semi-circle distribution.

Let us start with a one-dimensional local randomly coupled qubit model with nearest-neighbor random interactions,

$$H = \frac{1}{4} \sum_{i=1}^{N-1} \sum_{\alpha,\beta=0}^{3} J_i^{\alpha\beta} \sigma_i^{\alpha} \otimes \sigma_{i+1}^{\beta}, \tag{4.6}$$

where parameters $J_i^{\alpha\beta}$ are normal random variables with mean zero and variance one, so that $\overline{\text{Tr}}H = 0$ and $\overline{\text{Tr}}H^2 = N - 1$. This is analogous to SYK normalization.[16] We will explore the properties of this system numerically.

Energy spectrum

The spectrum of this Hamiltonian was studied numerically [65, 66], and shown to have a Gaussian shape. We repeated the same calculation as a sanity check. The energy spectra for $N = 6, 7, 8, \cdots, 16$ are shown in figure 9. The spectrum becomes closer to Gaussian as N becomes larger. We have also looked at nearest-neighbor level spacing distribution and do not see any hints of a MBL phase anywhere in the spectrum.

Spectral form factor

The spectral form factor is shown in figure 10. We can see that an intermediate region emerges as N grows. Note that this region lasts much longer than in the k-local spin chain. We do not have a theoretical understanding of this region and regard it as an interesting topic for future research. Here we proceed phenomenologically and define the ramp time to be where this region ends and the ramp starts.

---

[16]The Trotterization of the Hamiltonian with this scaling will correspond to N parallel gates acting on the system in an order one time. This is the normalization we will use in our random circuit discussion.
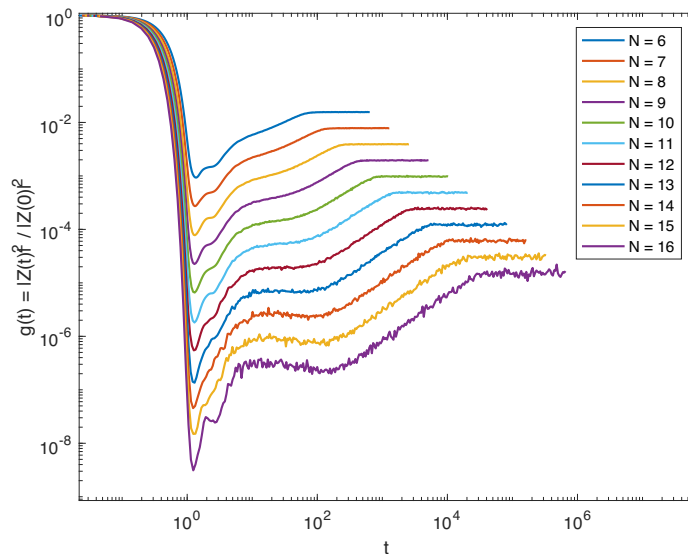
Figure 10. The spectral form factor for the local randomly coupled qubits (Local-RCQ) with 8 to 16 qubits. We observe universal behavior of $g(t)$; Gaussian decay at the early times, follow by an intermediate region, then a ramp and plateau. The time scale where the intermediate region ends and the ramp starts we will call the 'ramp time'.

In order to define the ramp time systematically, we adopted two methods (see figure 11):

- Pick up the points where the fractional error $\epsilon(t)$ defined by (4.5) decreases to certain values, as we have done for the k-local spin chain. We have used $\epsilon(t)$ equal to $5\%, 10\%, 20\%$ and 30%.

- Fit the intermediate region and ramp by powers of t (straight line in a log-log plot) and identify the intersection of two fitting lines with t_{ramp}.

We do this procedure for all the values of N ranging from 8 to 16 and plot the estimates of the ramp time with respect to N. In figure 12, we plot estimates of the ramp time determined in these ways. The ramp time can be fit reasonably well by an ansatz

$$t_{\mathrm{ramp}}(N) = AN^2 + B, \tag{4.7}$$

as shown in figure 12. Here we used t_{ramp} determined from the intersection of the power law fit to the intermediate region and a linear fit to the ramp; other choices lead to more or less the same A, but different values of the offset B. The best fit has $A = 0.85 \pm 0.09$ and $B = -14.57 \pm 6.28$. A linear fit to the graph is significantly worse: a parabolic fit has smaller mean square error and the coefficients are of order one. Hence we conjecture that t_{ramp} should be related to t_{diff} rather than t_{scr}. We will give analytic evidence for this below.

4.2.2 XXZ model

Next let us consider the one-dimensional XXZ Hamiltonian system with random magnetic field. This model has many interesting features, including a very well studied MBL phase [67, 68]. We are interested in the model because it has another conservation law
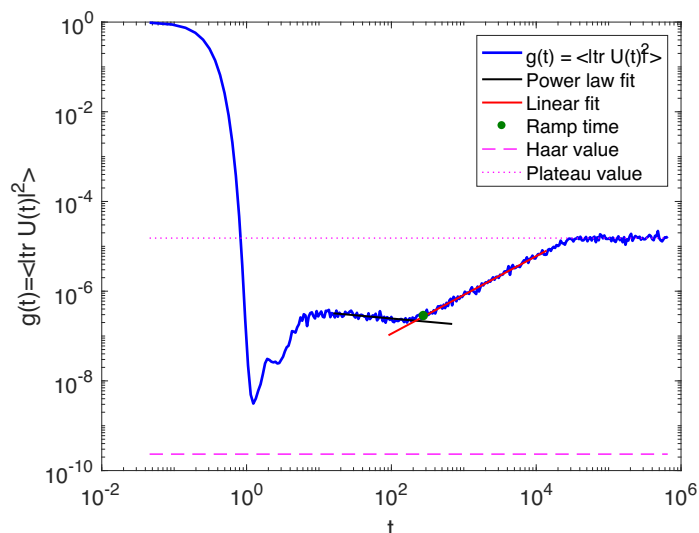
Figure 11. The spectral form factor for the local randomly coupled qubits (Local-RCQ) for $N = 16$ qubits is given by the blue curve. We fit two different lines in the log-log plot to the bump and the ramp (two red lines) and determine, the first, time scale where two red lines intersect and, the second, where fractional error of the blue curve from the best fit of the ramp is ϵ. In the plot the green dot represents the ramp time estimate with the second method, when fractional error is 5 %.
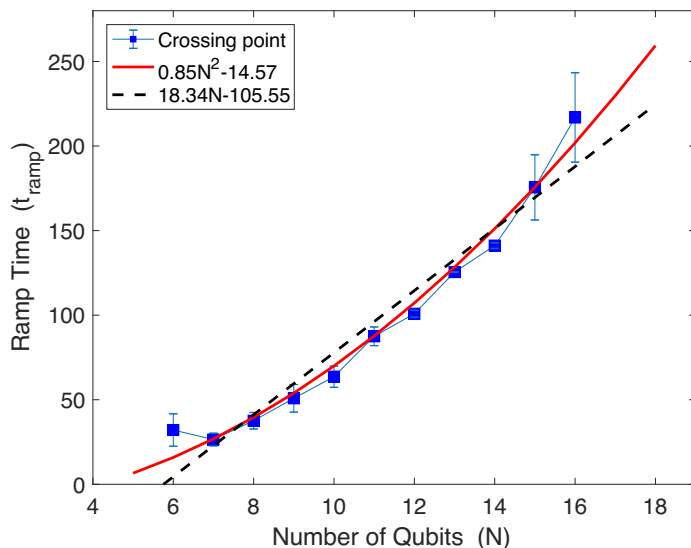
Figure 12. The blue dots represent the estimates of the ramp time from the crossing point of the intermediate region power law fit and the linear fit to the ramp. The error bars illustrate the difference between this estimate and the ramp estimate using an error of $\epsilon = 20\%$. The best fit is given by the red parabola and the alternative, linear fit is show as a black dashed line.
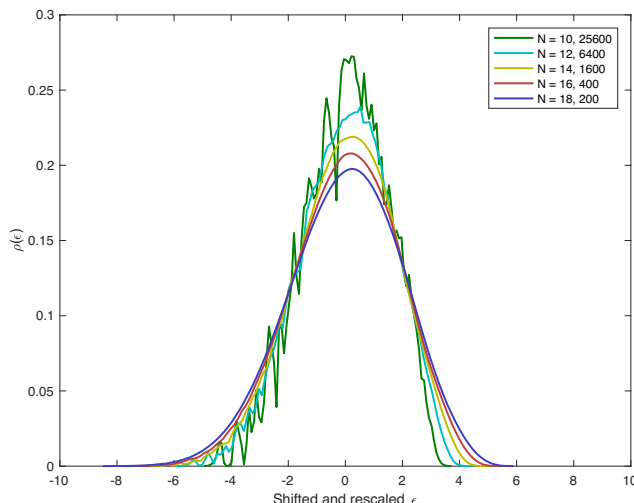
Figure 13. The spectrum of the XXZ Hamiltonian for $N = 10, 12, 14, 16, 18$, $W = 0.5$ and spin zero sector.

besides energy — the z component of spin. This conservation law can be preserved (unlike energy) in a random quantum circuit as we will explore in section 5.2. The random circuit analysis allows us to make the connection between t_{ramp} and t_{diff} manifest. Here we present some preliminary numerical results for the time independent Hamiltonian system. The intermediate region here grows as one approaches the MBL phase. We are unsure of the significance of this result.

The Hamiltonian we consider is

$$H = \sum_{i=1}^{N} \left(\frac{1}{4} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} + \frac{\omega_i}{2} \sigma_i^3 \right) \tag{4.8}$$

where $\vec{\sigma}_i = (\sigma_i^1, \sigma_i^2, \sigma_i^3)$ are Pauli matrices and ω_i, a random magnetic field, is a random number selected from the uniform distribution on the interval $[-W, +W]$. Here we use periodic boundary conditions, so the sums are modulo N. It has been argued that this model is in the MBL phase for $W \geq 2.75$ [68].

This Hamiltonian commutes with the total spin along the z-direction, $S_z^{\mathrm{tot}} = \frac{1}{2} \sum_i \sigma_i^3$. Thus we will need to look into specific spin sectors to detect nearest-neighbor Wigner-Dyson statistics. Unless otherwise specified we will consider the total spin zero sector for an even number of spins.[17]

Energy spectrum

We are interested in the diffusive regime, where the Hamiltonian exhibits chaotic properties and Griffiths effects, contributions of rare many-body localized regions, are not yet significant [69]. So we will investigate this Hamiltonian for $W = 0.5$ which is in the heart of diffusive regime. The energy spectrum is shown in figure 13. We have checked that the nearest-neighbor level correlations are consistent with the GOE ensemble when we look into each conserved spin sector separately.

---

[17]If we do not restrict the spin, we expect a Poisson distribution.
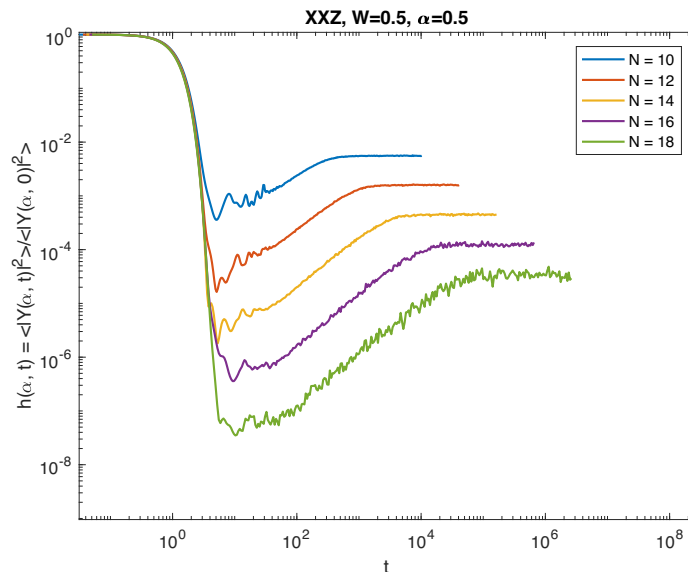
Figure 14. The spectral form factor for the XXZ Hamiltonian with 10 to 18 qubits for the S=0 sector and $W = 0.5$. We observe universal behavior for the $h(\alpha, t)$, Gaussian decay at the early times, flat region, ramp and plateau. The time scale where the 'bump' ends and the ramp starts we will again call ramp time.

Spectral form factor

In this section, we study $|Z|^2$ for the XXZ model (see figure 14). Again we take $W = 0.5$, which is in the heart of the diffusive phase. The qualitative behavior of $|Z|^2$ is similar to the local RCQ, however it is more sensitive to finite N effects in the edge of the spectrum.

As in the local RCQ, the intermediate region lasts longer than in the k-local spin chain. We determine t_{ramp} in exactly the same way as in section 4.2.1. However, for this system the numerical analysis is more subtle. For the small values of W that are in the diffusive regime there are some approximate conservation laws that produce oscillating behavior at early time and make it difficult to determine the ramp time accurately so we do not present results.

If one increases W the system moves toward the MBL phase and a rather long intermediate region appears in the spectral form factor (see figure 15). At this point, we do not have a good theory to explain the shaded region for values of the $W \geq 1$. Perhaps it is related to the subdiffusive regime between ergodic and MBL phases.

These numerical results are inconclusive, but as we will see in section 5.2, the random circuit analogue of this model provides an ideal setup to test our proposal that t_{ramp} and t_{diff} are the same.

5 Random/Brownian quantum circuit models

We now discuss analytic methods. In contrast to few-body systems analytic approaches for time independent Hamiltonian many-body systems have not yet been fully developed. Instead we turn to a related class of systems that display similar thermalization behavior
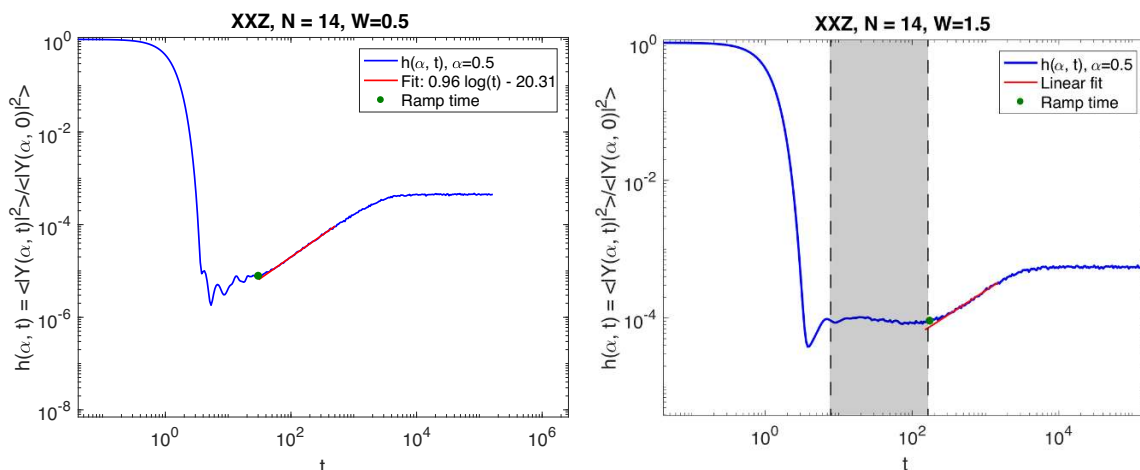
Figure 15. [Left] $h(\alpha = 0.5, t)$ for the XXZ Hamiltonian in the ergodic regime ($W = 0.5$) for $N = 14$ qubits and total spin zero sector. [Right] The $h(\alpha = 0.5, t)$ for the XXZ Hamiltonian in the subdiffusive regime ($W = 1.5$) for $N = 14$ qubits. The shaded region highlights an almost constant region of the spectral form factor whose origin we would like to explore.

and are analytically tractable. These are random quantum circuit and Brownian quantum circuit models.

In Brownian circuit models, the Hamiltonians are chosen randomly from certain ensembles, independently at each small time step dt. After M steps, the unitary evolution will be given by

$$U(t) = e^{-iH_M dt} e^{-iH_{M-1} dt} \cdots e^{-iH_1 dt} \tag{5.1}$$

where $t = Mdt$ and each of the operators H_i is independently drawn from a given ensemble. A random quantum circuit is a discrete version of this where each $e^{-iH_i dt}$ is replaced by a unitary drawn from a suitable ensemble of unitary "gates". If the H_i are chaotic we expect that $U(t)$ performs a kind of random walk on the unitary group $\mathrm{U}(L)$, eventually converging to the Haar random ensemble. The analogous question concerning the onset of random matrix behavior becomes the question about how quickly the spectrum of $U(t)$ starts displaying CUE statistics. A number of recent papers have studied thermalization in various correlations functions using these techniques [18, 22, 41–43].

The conclusion of all this work is that appropriate Brownian/random quantum circuits display known thermalizing behaviors — two point function decay (dissipation); chaotic/scrambling behavior as diagnosed by OTOCs and subsystem density matrices; and diffusion — with mechanisms and time scales that are very much analogous to fixed Hamiltonian systems. We will assume for now that Brownian/random quantum circuits will also display the onset of random matrix behavior via analogous mechanisms and with the same time scales as fixed Hamlitonian systems with the same locality properties and conservation laws. Because of their more efficient randomization we expect that at the least the time scales in these models are a lower bound for those in the coresponding Hamiltonian systems.

All the results we obtain are consistent with the Hamiltonian numerics discussed above. In addition these results are consistent with a heuristic framework for an analytic argument for Hamiltonian systems we discuss in section 6. In any case we find this assumption to be a very plausible one, and now turn to working out its consequences.

Because of the Brownian nature of these systems the naive analog of the spectral form factor $\left\langle |\text{Tr}U(t)|^2 \right\rangle$ decays to the value given by the average using Haar measure, which we call the Haar value, and stays there. We call the time when this value is attained the 'Haar time' t_{Haar}.

But this does not probe the full range of correlations in the spectrum of $U(t)$. To do this imagine running the quantum circuit for a time t, and after that repeating the same time evolution.[18] Then the time evolution from $t = 0$ to $\tilde{t} = kt$ ($k = 1, 2, \cdots$) is described by $U(\tilde{t}) = (U(t))^k$. The analog of the spectral form factor for time $\tilde{t}$ is then

$$u(k,t) \equiv \left\langle |\text{Tr}(U(t))^k|^2 \right\rangle. \tag{5.2}$$

Note that $u(k,t) = k$ for $k \leq L$ when $U(t)$ is drawn from the Haar (CUE) ensemble (see discussions in section 2 and appendix B.1). This is just a discrete version of the ramp. The time t at which the initial moments approach this value, t_{Haar}, is thus the analog of t_{ramp} in the Hamiltonian systems discussed in previous sections. Hopefully the physical meaning of t_{ramp} can be understood if we understand the meaning of t_{Haar}.

An important question is whether t_{Haar} depends on k. We are able to explore $k = 1, 2$ analytically. We then appeal to random matrix theory intuition that suggests the finer grain structure of the eigenvalue distribution is easier to establish in a chaotic system than the coarse grained structure. Since higher k probes finer structure we expect, at least for large enough k, that t_{Haar} will, at the least, not increase as k increases. Our numerics, discussed below, are consistent with this and seem to indicate that $k = 2$ is already in the stable range.[19]

In this section we consider Brownian circuit versions of the RCQ and XXZ models, both local and 2-local.[20] In earlier sections we suggested that for geometrically local systems t_{ramp} is the diffusion time. An essential difference between the random/Brownian quantum circuit and time-independent Hamiltonian system is that the former does not conserve energy and hence in general there is no diffusion. To study the role of diffusion in this context we use a random circuit where spin, not energy, is conserved. In particular we study a random circuit version of the XXZ model introduced by [43, 48] in section 5.2.[21] There we will see it is possible to relate t_{Haar} and hence t_{ramp} to the diffusion time, distinguishing it from the scrambling time. In the 2-local case we show that diffusion happens in order one time, and is not important for setting t_{ramp}.
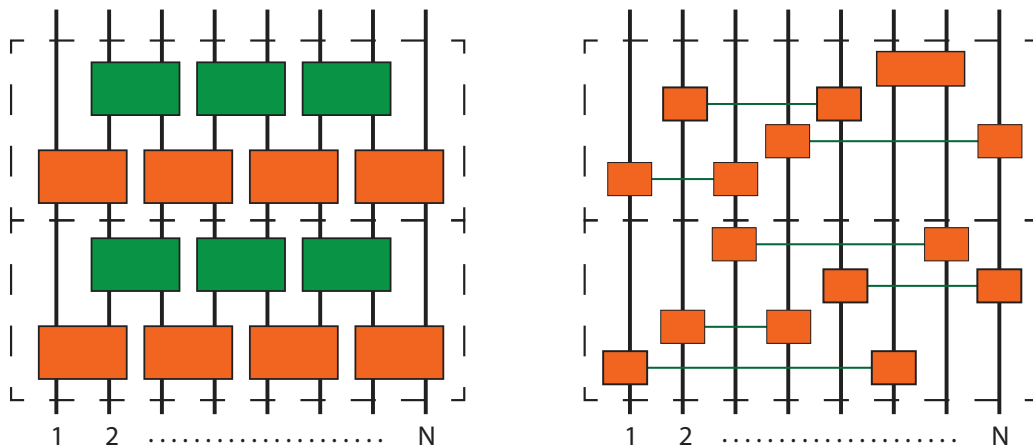
---

[18]In condensed matter physics this is called a "Floquet system".

[19]These results are consistent with analytic results on the SYK brownian circuit [70].

[20]Some results on random banded matrices are presented in appendix C.

[21]We are grateful to David Huse for introducing us to this model and to him and Vedika Khemani for discussions about it.

Figure 16. [Left] Local random circuit evolution on N qubits, each gate is randomly selected from a gateset $\Gamma \in U(4)$. [Right] 2-local random circuit on N qubits, in each step qubits are randomly paired and random gate is applied from a gateset Γ.

5.1 Local and k-local RCQ random/Brownian circuit

5.1.1 Analytic estimate

Here we provide a calculation of the spectral form factor for random circuits. Specifically, we are interested in the behavior of $u(k,t)$ defined by (5.2). We first describe the $k = 1$ case, which is conceptually quite simple.

We use techniques presented in ref. [34]. We will investigate the parallelized, local random circuit model shown in the left part of figure 16, where each step consists of two layers acting on pairs of qubits and each gate is selected at random from some 2 qubit universal gate set $\Gamma \subseteq U(4)$. For general k we can express $\text{Tr}U^k$ as a sum of monomials,

$$\text{Tr}U^k = \sum_{i_1,i_2,\ldots,i_k=1}^{L} U_{i_1 i_2} U_{i_2 i_3} \ldots U_{i_k i_1}, \tag{5.3}$$

where $L = 2^N$ is the dimension of the Hilbert space. Therefore, (5.2) can be written as

$$u(k,t) = \left\langle \left| \text{Tr}U^k \right|^2 \right\rangle = \sum_{i,j} \left\langle U_{i_1 i_2} U_{i_2 i_3} \ldots U_{i_k i_1} U^*_{j_1 j_2} U^*_{j_2 j_3} \ldots U^*_{j_k j_1} \right\rangle. \tag{5.4}$$

Calculating $\left\langle |\text{Tr } U(t)|^2 \right\rangle$

We want to study

$$u(1,t) = \left\langle |\text{Tr}U(t)|^2 \right\rangle = \sum_{i,j=1}^{L} \left\langle U_{ii} U^*_{jj} \right\rangle. \tag{5.5}$$

There are two types of terms we need to study: $i = j$ and $i \neq j$. Without loss of generality we can set $i = 1$.

First let us consider $U_{11} U^*_{11}$. In a natural 0 (up spin) and 1 (down spin) basis, $i = 1$ is $|0,0,\cdots,0\rangle = |0\rangle^{\otimes N}$, and

$$U_{11}(t) = \langle 0,0,\cdots,0| \hat{U}(t) |0,0,\cdots,0\rangle. \tag{5.6}$$

By using a density matrix $\hat{\rho}(t) = \hat{U}(t)\hat{\rho}(0)\hat{U}^\dagger(t)$ with the initial condition

$$\hat{\rho}(0) = |0, 0, \cdots, 0\rangle \langle 0, 0, \cdots, 0| = (|0\rangle \langle 0|)^{\otimes N} \tag{5.7}$$

at $t = 0$, we can write it as

$$U_{11}(t)U_{11}^*(t) = \text{Tr}\Big[\hat{\rho}(t)\hat{\rho}(0)\Big]. \tag{5.8}$$

The density matrix can be expanded in tensor products of strings of Pauli matrices σ_p, where p is a label specifying a Pauli string,[22] as

$$\rho(t) = 2^{-N/2} \sum_p \gamma_t(p)\sigma_p. \tag{5.9}$$

Note that $\text{Tr}(\rho(t)) = 1$ holds for any t due to the unitarity of time evolution. We use $p = 0$ to denote the identity matrix, and so $\gamma_t(0) = 2^{-N/2}$ for any t. By using these expansion coefficients $\gamma_t(p)$, we obtain

$$U_{11}U_{11}^* = 2^{-N}\text{Tr}\left[\sum_p \gamma_t(p)\sigma_p \times \sum_q \gamma_0(q)\sigma_q\right] \tag{5.10}$$

$$= \sum_{p,q} \gamma_t(p)\gamma_0(q)\delta_{pq} \tag{5.11}$$

$$= 2^{-N} + \sum_{p\neq 0} \gamma_t(p)\gamma_0(p). \tag{5.12}$$

We can easily determine $\gamma_0(p)$, by noticing that

$$\hat{\rho}(0) = (|0\rangle \langle 0|)^{\otimes N} = \left(\frac{1}{2}(I + Z)\right)^{\otimes N}, \tag{5.13}$$

where $I = \sigma^0$ and $Z = \sigma^3$. Namely $\gamma_0(p)$ is nonzero only when σ_p is a Pauli string containing only I and Z's (equivalently, when p is a set consisting only from 0 and 3), and all nonzero elements are $2^{-N/2}$. Therefore,

$$\langle U_{11}U_{11}^*\rangle = 2^{-N} + 2^{-N/2}\sum_{p'}\langle\gamma_t(p')\rangle, \tag{5.14}$$

where p' runs through all possible sets of 0 and 3 but $(0, 0, \cdots, 0)$.

For a given choice of circuit the Pauli chain of I's and Z's evolves into other chains with positive and negative weights. On averaging these tend to cancel and $\langle\gamma_t(p')\rangle$ decays exponentially.[23] A pair II is left unchanged by the evolution. So the only invariant chain is all I's. The rate of decay of other chains is proportional to the number of Z's in the

---

[22]More precisely, p is a set of N numbers $(p_1, p_2, \cdots, p_N)$, where $p_i = 0, 1, 2$ or 3, and $\sigma_p = \sigma^{p_1} \otimes \sigma^{p_2} \otimes \cdots \otimes \sigma^{p_N}$. Here σ^0 is the 2×2 identity matrix and $\sigma^{1,2,3}$ are Pauli matrices $\sigma^{x,y,z}$.

[23]In fact for gates chosen randomly from the Haar distribution on U(4) the average $\langle\gamma_t(p')\rangle$ drops to zero in one step.

chain, which we will call d. Lemma 4.1 (section 4.1) of ref. [34] bounds the one and two norms of $\gamma_t(p)$ for the random circuit. Their bound is

$$\sum_{p'} \langle |\gamma_t(p')| \rangle \le \sum_{d=1}^{N} e^{-\Delta t d} \sum_{d(p')=d} |\gamma_0(p')|. \tag{5.15}$$

where $d(p')$ is a number of Pauli-Z's in $\sigma_{p'}$ and Δ is the 'gap parameter' which depends on the choice of the gate set Γ. Thus, we can bound the sum as

$$\sum_{p'} \langle |\gamma_t(p')| \rangle \le \sum_{d=1}^{N} e^{-\Delta t d} \sum_{d(p')=d} |\gamma_0(p')|. \tag{5.16}$$

In our normalization, for both local and k-local RCQ, Δ should be of order 1, independent of the system size N. In our case $\sum_{d(p')=d} |\gamma_0(p')| = 2^{-N/2}\binom{N}{d}$, and hence

$$\langle U_{11} U_{11}^* \rangle \le 2^{-N} + 2^{-N} \sum_{d=1}^{N} e^{-\Delta t d} \binom{N}{d} = 2^{-N} + \left(2^{-N} \left(1 + e^{-\Delta t} \right)^N - 1 \right). \tag{5.17}$$

Next we study the behavior of $U_{11} U_{jj}^*$. The j-th state is represented by a sequence of 0's and 1's. As we will see, the number of 1's, which we denote by q, controls the decay rate. Hence let us denote the j-th state by $\left| 1^q 0^{N-q} \right\rangle$, neglecting the ordering of 0's and 1's. We consider the time evolution of an operator $\hat{O}(t)$, whose initial condition is specified by

$$\hat{O}(0) = |1\rangle \langle j| = \left| 0^N \right\rangle \left\langle 1^q 0^{N-q} \right|. \tag{5.18}$$

By definition,

$$\hat{O}(t) = \hat{U}(t)\hat{O}(0)\hat{U}^\dagger(t), \tag{5.19}$$

and hence

$$U_{11} U_{jj}^* = \left\langle 0^N \right| \hat{U}(t) \left| 0^N \right\rangle \left\langle 1^q 0^{N-q} \right| \hat{U}^\dagger(t) \left| 1^q 0^{N-q} \right\rangle = \mathrm{Tr}\left[\hat{O}(t)\hat{O}^\dagger(0) \right]. \tag{5.20}$$

Noticing that

$$|0\rangle \langle 1| = \frac{1}{2}(\sigma^1 + i\sigma^2) = \frac{1}{2}(X + iY), \tag{5.21}$$

we can write $\hat{O}(0)$ as[24]

$$\hat{O}(0) = 2^{-N/2} \sum_{\alpha_1,...,\alpha_q=1}^{2} \sum_{d=0}^{N} \sum_{i_1 \le ... \le i_d} 2^{-N/2}(+i)^{f(\vec{\alpha})} \sigma_1^{\alpha_1} \otimes ... \otimes \sigma_q^{\alpha_q} \otimes \sigma_{i_1}^3 \otimes \sigma_{i_2}^3 \otimes ... \otimes \sigma_{i_d}^3 \tag{5.22}$$

where $f(\vec{\alpha}) = \sum_{i=1}^{q}(\alpha_i - 1)$ counts the number of σ^2's. By expanding $\hat{O}$ as we have done for $\hat{\rho}$ as

$$\hat{O}(t) = 2^{-N/2} \sum_p \gamma_t(p)\sigma_p, \tag{5.23}$$

---

[24]We have shown the expression for $|j\rangle = |1, 1, \cdots, 1, 0, 0, \cdots, 0\rangle$ for simplicity. Expressions for other orderings can easily be obtained by changing the locations of $\sigma^{1,2}$.

we obtain $\gamma_0(p) = 2^{-N/2}(+i)^{f(\vec{\alpha})}$. Hence

$$|U_{11}U_{jj}^*| = \left|\sum_{p'} \gamma_t(p')\gamma_0^*(p')\right| \leq \sum_{p'} |\gamma_t(p')| \cdot |\gamma_0(p')| \tag{5.24}$$

where we sum over all the Pauli strings of the form $XYY\ldots XIIZZ\ldots Z$, that have X, Y elements for the first q sites and $N-q$ I, Z elements for the remaining sites. The coefficients of these terms will decay as

$$\langle|\gamma_t(p')|\rangle \leq e^{-\Delta t(q+d(p'))}|\gamma_0(p')| \tag{5.25}$$

except for $q + d(p') = 0$, because the 'conversion' can take place anywhere an X, Y or Z is located. (Note that this expression, and hence the following estimate, hold for all choices of the ordering of 0's and 1's in the j-th state.) Thus, for $q > 0$ (i.e. $j \neq 1$),

$$\langle|U_{11}U_{jj}^*|\rangle \leq \sum_{p'} \langle |\gamma_0(p)|^2 \rangle e^{-\Delta t(d(p')+q)}$$

$$= 2^{-N}e^{-\Delta tq} \sum_{d=0}^{N-q} e^{-\Delta td}\binom{N-q}{d}$$

$$= 2^{-N}e^{-\Delta tq}(1 + e^{-\Delta t})^{N-q}. \tag{5.26}$$

Now we can get a good bound for $u(1, t)$,

$$\langle|\mathrm{Tr}U(t)|^2\rangle = \sum_{i=1}^{2^N}\sum_{j=1}^{2^N} \langle U_{ii}U_{jj}^*\rangle = 2^N \sum_{j=1}^{2^N} \langle U_{11}U_{jj}^*\rangle$$

$$\leq 2^N \sum_{q=0}^{N} 2^{-N}\binom{N}{q}e^{-\Delta tq}(1 + e^{-\Delta t})^{N-q}$$

$$= (1 + e^{-\Delta t})^N \sum_{q=0}^{N} \binom{N}{q}\left(\frac{e^{-\Delta t}}{1 + e^{-\Delta t}}\right)^q$$

$$= (1 + 2e^{-\Delta t})^N. \tag{5.27}$$

Hence, for each fixed N, at sufficiently late time,

$$\langle|\mathrm{Tr}U(t)|^2\rangle - 1 \lesssim 2Ne^{-\Delta t}. \tag{5.28}$$

As we have mentioned above, generic gate sets have Δ that is $O(1)$. If this inequality is almost saturated (numerical results explained below suggest this is actually the case), then if we plot $\frac{\langle|\mathrm{Tr}U(t)|^2\rangle - 1}{N} \sim 2e^{-\Delta t}$ as a function of t for various values of N, they should lie on top of each other when N is sufficiently large. The time scale for $\langle|\mathrm{Tr}U(t)|^2\rangle - 1$ to decay as small as ϵ is

$$t_\epsilon \simeq \frac{1}{\Delta} \log\left(\frac{2N}{\epsilon}\right) \sim \log N. \tag{5.29}$$

It is clear that this estimate t_{Haar} does not depend on the connectivity of the circuit.

We should note that we can relate this estimate to the decay of two point functions. As we show in (6.2) of section 6 we can rewrite

$$\left\langle |\mathrm{Tr}\, U(t)|^2 \right\rangle = 1 + \sum_{j=2}^{L^2} \frac{1}{L} \left\langle \mathrm{Tr}\left(U^\dagger(t) P_j U(t) P_j \right) \right\rangle \tag{5.30}$$

$$= 1 + \sum_{j=2}^{L^2} \frac{1}{L} \left\langle \mathrm{Tr}\left(P_j(t) P_j \right) \right\rangle \tag{5.31}$$

where the P_j are a complete basis of Pauli operators acting on the N qubit Hilbert space. The decay of $u(1,t)$ is thus related to the decay of various two point functions. The content of the random circuit analysis is that this sum is dominated by the decay of the P_j containing just one Pauli. This two point function decays like $e^{-\Delta t}$ and there are N such operators. This give the result (5.29).

Sketch of the calculation for $u(2,t) = \left\langle \left| \mathrm{Tr}\, U^2(t) \right|^2 \right\rangle$

In the above section we have estimated the convergence of the first moment $u(1,t)$ and shown the Haar time is $\log N$ for both local and 2-local circuits. To fully understand the approach to random matrix statistics it is important to examine the Haar time for the k-th moment $u(k,t)$. This is related to the study of approximate k designs. The techniques used above become quite challenging for general k and have only been carefully examined for $k = 2$. In the appendix B.2 we carry out an analysis for $u(2,t)$ using the techniques of ref. [34] for a particularly simple gate set, U(4) random two qubit gates. For this gate set $u(1,t)$ converges to Haar in one time step while $u(2,t)$ converges in $\log N$ time. Here we give a summary of the argument.

The analysis for $u(2,t)$ boils down to a Markov process on two copies of a string of N Paulis. When the strings are unequal their weight vanishes in one time step. For a more general gate set the $k = 1$ rate would control their decay. When the strings are equal, a different Markov process involving transitions between pairs of Paulis controls the convergence. Again the strings with very few Z's have the slowest convergence, and cause a deviation from the Haar value of 2 of the form

$$u(2,t) \sim 2 + N e^{-\tilde{\Delta} t} \,. \tag{5.32}$$

So $u(2,t)$ becomes ϵ close to Haar in a time $\frac{1}{\tilde{\Delta}} \log \frac{N}{\epsilon}$. Again we see a $\log N$ dependence but here $\tilde{\Delta}$ is different from the Δ discussed above. It is determined by the Markov chain on pairs of Paulis. We note that the gap of the full Markov chain does not control the convergence of this quantity, as explained in appendix B.2.

We can again interpret these results in terms of two point function decay. The quantity described in (5.30) vanishes in one step for U(4) gates. But we can define "two step" correlators using equation (6.2) of section 6 to rewrite

$$\left\langle |\mathrm{Tr}\, U^2(t)|^2 \right\rangle = 1 + \sum_{j=2}^{L^2} \frac{1}{L} \left\langle \mathrm{Tr}\left((U^\dagger(t))^2 P_j U^2(t) P_j \right) \right\rangle \tag{5.33}$$
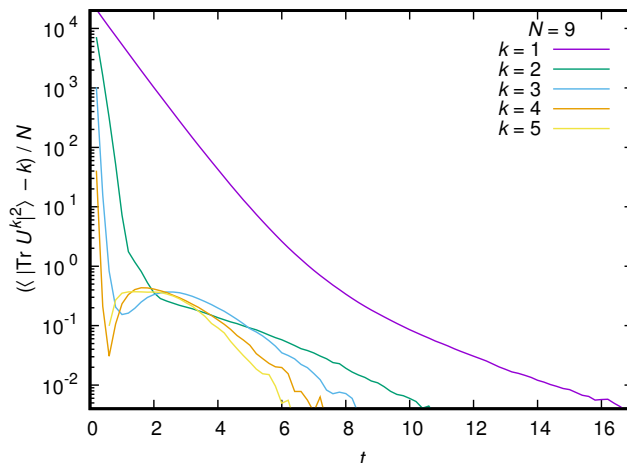
Figure 17. Plot of $(\langle|\mathrm{Tr}U^k|^2\rangle - k)/N$ for $k = 1, 2, 3, 4, 5$ against t for the 2-local RCQ chain with $dt = 0.2$. $N = 9$, 10^5 samples. These exhibit exponential decays with rates that do not decrease as k increases.

These "two step" two point functions will decay exponentially and the slowest ones will be the operators with a single Pauli. The analysis in appendix B.2 shows that these decay like $e^{-\tilde{\Delta}t} = (1/15)^t$.

5.1.2 Numerical results for Brownian circuit

2-local RCQ Brownian circuit

We start with the Brownian circuit made from the 2-local RCQ Hamiltonian introduced in section 4.1.2. In the following, 3000, 20000 and 10^5 samples have been used for $N = 11, 10$ and $N \leq 9$ cases, respectively. In figure 17, $u(k,t)$ minus the Haar value, $\langle|\mathrm{Tr}U^k(t)|^2\rangle - k$, is plotted for $N = 9$ and $dt = 0.2$.[25] We can see clear exponential decays at early time and late time.

As shown in the left panel of figure 18, the early time decay for $k = 1$ can be fitted by an exponential function of the time $\exp(-CNt)$, in which $C \simeq 0.189$ is a constant independent of N. Therefore, the early-time decay to $\langle|\mathrm{Tr}U(t)|^2\rangle - 1 \sim O(1)$ takes place within order one time.

For late time, $\langle|\mathrm{Tr}U(t)|^2\rangle - 1 \sim Ne^{-\Delta t}$, with $\Delta = O(N^0)$, is expected. As shown in the right panel of figure 18, the late time decays are consistent with $\Delta \simeq 0.5$, when N is sufficiently large.

In figure 19, we used the same fitting ansatz for $k > 1$, as $(\langle|\mathrm{Tr}U^k(t)|^2\rangle - k)/N \sim e^{-\Delta_k t}$. For each k, we can see convergence to exponential decay at large N. The exponent Δ_k increases with k (see also figure 17),[26] which means the convergence to the Haar values at larger k happens earlier. Therefore, after t_{Haar} is determined from $k = 1$, higher moments have already converged.

---

[25]As long as dt is not too large, the specific value of dt does not affect the result. For details regarding this point, especially in what sense dt-dependence becomes trivial, see appendix C.2.

[26]Δ_1 actually is approximately the same as Δ_2.
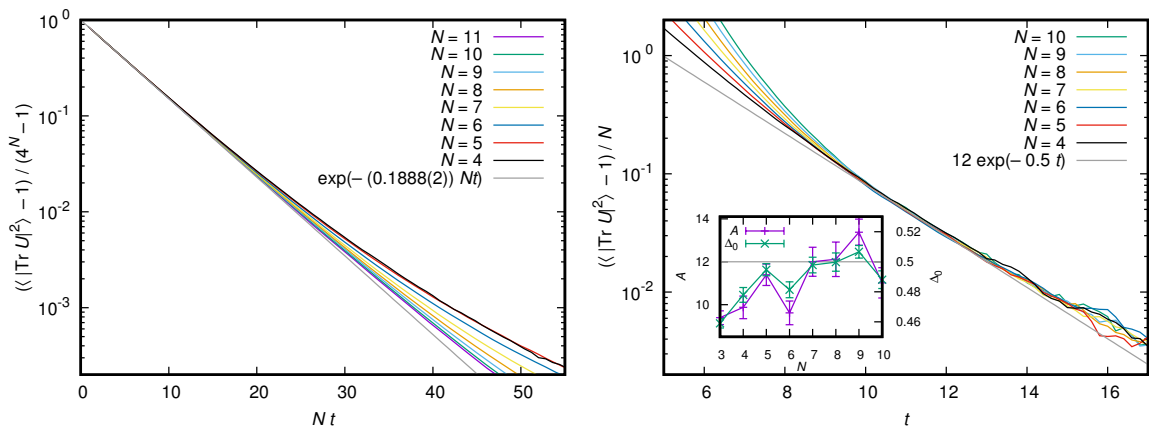
Figure 18. [Left] Plot of $(\langle|\mathrm{Tr}U|^2\rangle - 1)/(4^N - 1)$ against Nt for the RCQ chain with $dt = 0.2$ for $N = 11, 10, 9, 8, 7, 6, 5, 4$ qubits. The result of a single parameter fit by an exponential function of Nt is also shown. [Right] Plot of $(\langle|\mathrm{Tr}U|^2\rangle - 1)/N$ against t for the RCQ chain with $dt = 0.2$ for $N = 10, 9, 8, 7, 6, 5, 4$ qubits. The inset shows the result of fit for $t \in [10 : 14]$ with $Ae^{-\Delta_0 t}$ by taking A and Δ_0 as fitting parameters. The line corresponding to $(A, \Delta_0) = (12, 0.5)$ is also shown in the main plot.
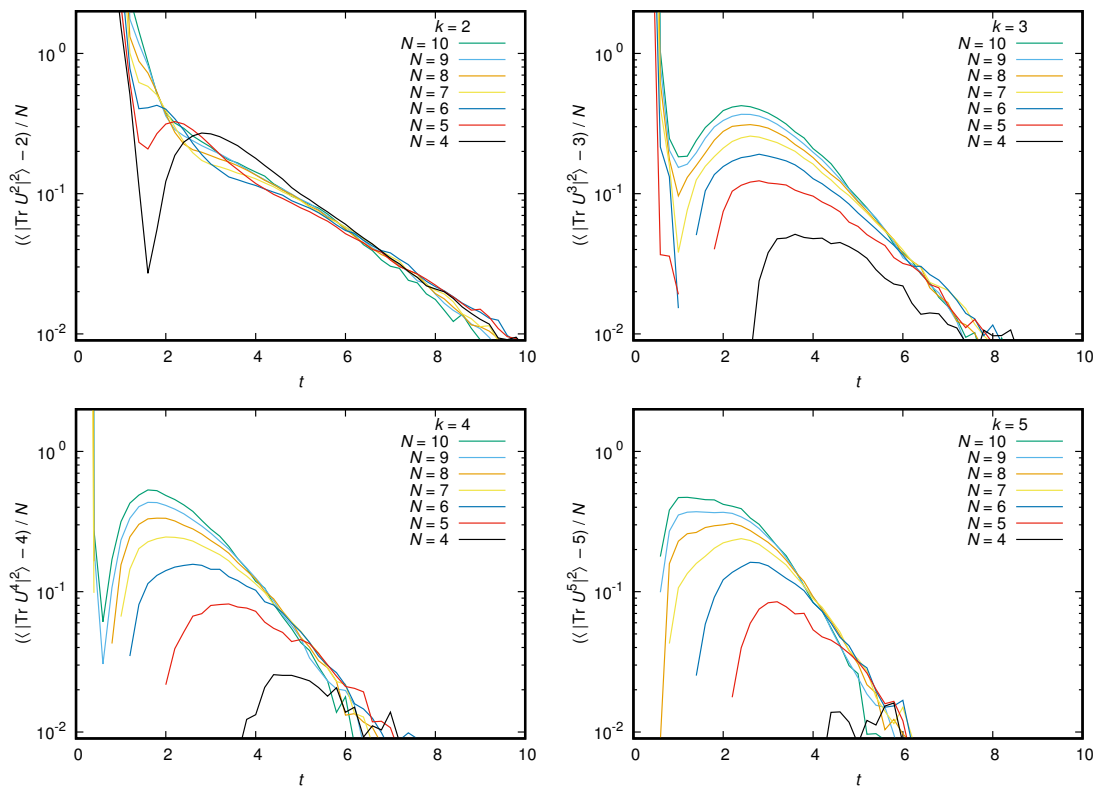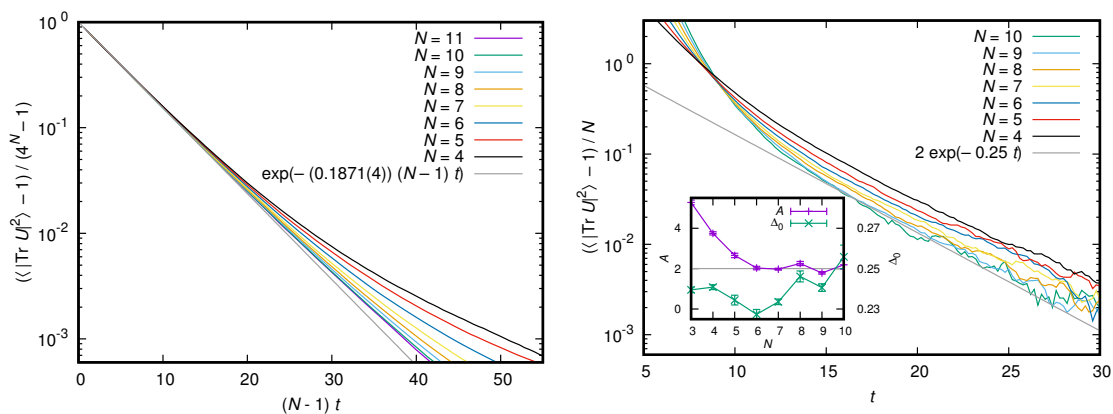

Figure 19. Plot of $(\langle|\mathrm{Tr}U^k(\tilde{t})|^2\rangle - k)/N$ against t for various values of N, $k = 2, 3, 4, 5$.
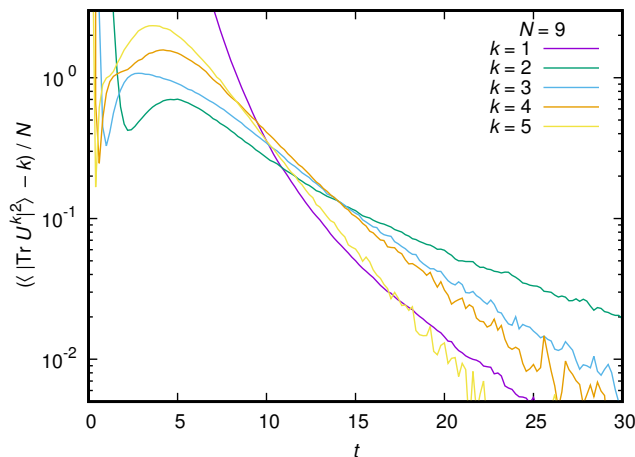
Figure 20. Plot of $\langle |\mathrm{Tr}U^k|^2\rangle - k$ for $k = 1, 2, 3, 4, 5$ for the local RCQ chain with $dt = 0.2$ for $N = 9$ qubits. The decay time scale appears to grow with k for $k \geq 2$.


Figure 21. [Left] Plot of $(\langle |\mathrm{Tr}U|^2\rangle - 1)/(4^N - 1)$ against $(N-1)t$ for the local RCQ chain with $dt = 0.2$ for $N = 11, 10, 9, 8, 7, 6, 5, 4$ qubits. [Right] Plot of $(\langle |\mathrm{Tr}U|^2\rangle - 1)/N$ for the local RCQ chain with $dt = 0.2$ for $N = 10, 9, 8, 7, 6, 5, 4$ qubits.

Local RCQ Brownian circuit

Next we consider the Brownian circuit made from the local RCQ Hamiltonian introduced in section 4.2.1. We expect the same decay pattern as in the 2-local RCQ. We have used the same number of samples as in the 2-local RCQ Brownian circuit case.

As shown in the left panel of figure 21, the early-time decay can be fit by $(\langle |\mathrm{Tr}U(t)|^2\rangle - 1) \simeq (4^N - 1) \cdot e^{-c(N-1)t}$ where $c \simeq 0.187$. Therefore, the early-time decay to $\langle |\mathrm{Tr}U(t)|^2\rangle - 1 \sim O(1)$ takes place within order one time. As shown in figure 20, right panel of figure 21 and in figure 22, the late time decay appears to be consistent with $\tilde{c}Ne^{-\Delta_k t}$ for $k = 1, 2, 3, 4$ and 5, where $\tilde{c}$ is an order one constant. These observations are the same as the case of the 2-local Brownian circuit. Note that here Δ_1 is larger than Δ_2. This is not surprising given the analysis above which shows they result from somewhat different mechanisms. But for $k > 2$, $\Delta_k > \Delta_2$ indicating that $k = 2$ is already in the stable range for determining the approach to random matrix behavior.
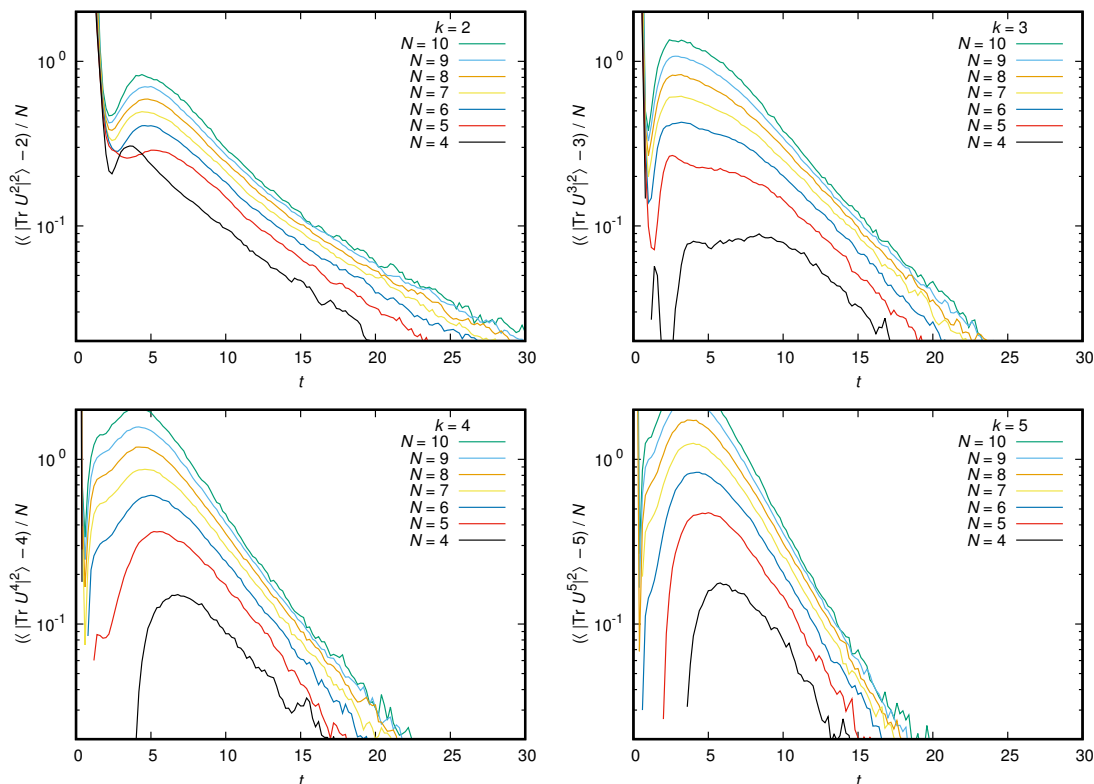
Figure 22. Plot of $(\langle |\mathrm{Tr} U^k|^2 \rangle - k)/N^2$ against t for the local RCQ chain with $dt = 0.2$ for $N = 10, 9, 8, 7, 6, 5, 4$ qubits, $k = 2, 3, 4, 5$. An exponential fit of the $N = 9$ curve is also shown.

5.2 XXZ random circuit and diffusion

In this section, we will calculate $u(k, t)$ defined in (5.2) for an XXZ type random circuit introduced in [43, 48] which has a conserved quantity and so helps clarify the relationship between t_{Haar} and t_{diff}. The quantum circuit is composed of a brickwork array of two-qubit gates as in previous examples (see figure 16). In order to motivate this circuit, let us consider a Brownian circuit made from the Hamiltonian (4.8) (which we repeat here) with time varying couplings:

$$H = \sum_{i=1}^{N} \left(\frac{1}{4} \vec{\sigma}_i \cdot \vec{\sigma}_{i+1} + \frac{\omega_i}{2} \sigma_i^3 \right).$$

The total spin along the z-direction, $S_z^{\mathrm{tot}} = \frac{1}{2} \sum_i \sigma_i^3$, commutes with the Hamiltonian regardless of the value of the random magnetic field, so Brownian circuit time evolution commutes with S_z^{tot}, and hence the total spin is conserved. The random circuit under consideration is the discrete analog of this Brownian circuit.

Because the part of the Hamiltonian relevant for the interaction between i-th and j-th sites, $\left(\frac{1}{4} \vec{\sigma}_i \cdot \vec{\sigma}_j + \frac{\omega_i}{2} \sigma_i^3 + \frac{\omega_j}{2} \sigma_j^3 \right)$, commutes with $\sigma_i^3 + \sigma_j^3$, the unitary gate which describe

the infinitesimal time evolution takes the following form,

$$
G_{i,j} = \begin{pmatrix} U_{11}^{(1)} & 0 & 0 & 0 \\ 0 & U_{11}^{(2)} & U_{12}^{(2)} & 0 \\ 0 & U_{21}^{(2)} & U_{22}^{(2)} & 0 \\ 0 & 0 & 0 & U_{11}^{(3)} \end{pmatrix},
\tag{5.34}
$$

where $U^{(1)}$ and $U^{(3)}$ are independent random elements of U(1) and $U^{(2)}$ is a Haar random unitary element in U(2). The indices i and j indicates the locations of the qubits. Here $U^{(1)}$ and $U^{(3)}$ act on the spin $+1$ sector ($|0,0\rangle$) and the spin -1 sector ($|1,1\rangle$), respectively, and $U^{(2)}$ acts on the spin 0 sector ($|0,1\rangle$ and $|1,0\rangle$).

Because of the symmetry we use the Pauli basis $\{I, \sigma^+, \sigma^-, Z\}$, where

$$
I = |0\rangle\langle 0| + |1\rangle\langle 1|, \qquad\qquad Z = |0\rangle\langle 0| - |1\rangle\langle 1|,
\tag{5.35}
$$

$$
\sigma^+ = \frac{1}{2}(X + iY) = |0\rangle\langle 1|, \qquad \sigma^- = \frac{1}{2}(X - iY) = |1\rangle\langle 0|.
\tag{5.36}
$$

In this notation, the sum of the spins at the i-th and j-th sites is

$$
\frac{\sigma_i^3 + \sigma_j^3}{2} = \frac{Z_i I_j + I_i Z_j}{2}.
\tag{5.37}
$$

The unitary gate can be expressed as

$$
\begin{aligned}
G_{i,j} = {}& \frac{U_{11}^{(1)}}{4}(I_i + Z_i)(I_j + Z_j) \\
& + \frac{U_{11}^{(2)}}{4}(I_i + Z_i)(I_j - Z_j) + U_{12}^{(2)}\sigma_i^+\sigma_j^- + U_{21}^{(2)}\sigma_i^-\sigma_j^+ + \frac{U_{22}^{(2)}}{4}(I_i - Z_i)(I_j + Z_j) \\
& + \frac{U_{11}^{(3)}}{4}(I_i - Z_i)(I_j - Z_j).
\end{aligned}
\tag{5.38}
$$

We now consider the adjoint action of $G_{i,j}$ on pairs of Paulis in the string. This has the following properties [43]: (for simplicity, we suppress i, j indices):

1. II, ZZ and $(IZ + ZI)/2$ are invariant.

2. $(IZ - ZI)/2$, $\sigma^+\sigma^-$ and $\sigma^-\sigma^+$ will mix among each other with equal probability. After Haar averaging, these terms vanish.

3. $\sigma^+\sigma^+$ and $\sigma^-\sigma^-$ get random phases $U_{11}^{(1)}U_{11}^{(3)*}$ and $U_{11}^{(1)*}U_{11}^{(3)}$, respectively. Because of these phases, the averages become zero.

4. The raise-by-one operators σ^+I, σ^+Z, $I\sigma^+$ and $Z\sigma^+$ will transition between each other with equal probability. The averages of these linear combinations are zero. The same will happen to the lower-by-one operators.

The important point here is that any chain other than II, ZZ and $(IZ + ZI)/2$ are completely randomized after one application of the unitary gate, so that they average

to zero immediately. The transitions between the remaining I, Z chains form a Markov process. Note that unlike the U(4) gate the evolution is a Markov chain already at $k = 1$. This Markov chain can be understood more easily if we use another basis

$$P = |0\rangle\langle 0| = \frac{I + Z}{2}, \qquad M = |1\rangle\langle 1| = \frac{I - Z}{2}.$$ (5.39)

In this basis, PP and MM are invariant, while PM and MP are mapped to $\frac{1}{2}PM + \frac{1}{2}MP$. The Markov property is now evident.

5.2.1 $\left\langle |\mathrm{Tr}\, U(t)|^2 \right\rangle$ in the XXZ random circuit

Let us use the setup above to estimate the $k = 1$ moment,

$$\left\langle |\mathrm{Tr}U(t)|^2 \right\rangle = \sum_{i,j=1}^{2^N} \left\langle U_{ii} U_{jj}^* \right\rangle.$$ (5.40)

We first consider the 'local' model defined above, which has nearest-neighbor interactions via the gates $G_{i,i+1}$.

As in the case of RCQ, we can write the spectral form factor as

$$\left\langle U_{ii} U_{jj}^* \right\rangle = \left\langle \mathrm{Tr}\left[U_t |i\rangle\langle j| U_t^\dagger |j\rangle\langle i| \right] \right\rangle$$ (5.41)

where as before $|i\rangle\langle j|$ can be expressed as a sum over strings of σ-matrices. Let us use the $M, P, \sigma^\pm$ basis. When $i \neq j$, at least one $\sigma^\pm$ appears in the chain. Then, as we have seen, as soon as $\sigma^\pm$ is hit by a random gate, $U_t (\sigma\sigma\cdots) U_t^\dagger$ averages to zero. Hence contributions from $i \neq j$ terms disappear in one time step. Therefore we only need to consider $i = j$.

First let us consider the case with a single M,

$$\left\langle U_{ii} U_{ii}^* \right\rangle = \left\langle \mathrm{Tr}\left[U_t (PP\cdots PMP\cdots P) U_t^\dagger (PP\cdots PMP\cdots P) \right] \right\rangle.$$ (5.42)

We need to see how $U_t (PP\cdots PMP\cdots P) U_t^\dagger$ evolves. After averaging MP and PM evolve to $\frac{PM+MP}{2}$, while PP is invariant. Hence, when the random gate hits MP, the chain remains $MPP\cdots P$ with probability $1/2$, and changes to $PMP\cdots P$ with probability $1/2$. So M performs a random walk in the string. Because there is no preferred direction there is no net drift of the walker. We analyze this system in detail below but for now make a few qualitative remarks.

The probability distribution of M locations spreads diffusively so $\langle U_{ii} U_{ii}^* \rangle$, given by the overlap with the initial state, decays like $\langle U_{ii} U_{ii}^* \rangle \sim 1/\sqrt{t}$ at early time. At late time it is uniformly spread on the string and eventually $\langle U_{ii} U_{ii}^* \rangle$ converges to $1/N$. For general i there are q M's and $N - q$ P's. Then the diffusive decay at early time is $\left(1/\sqrt{t}\right)^q$ (when $q \ll N$). The late-time value is $1/\binom{N}{q}$.

In order to give a precise analysis we identify M and P with two dimensional vectors,

$$|M\rangle = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{and} \qquad |P\rangle = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$ (5.43)

In other words we identify $|M\rangle$ and $|P\rangle$ to be two states of a qubit. Let us denote by $P_{i,j}$ a unitary operator that acts on the vector space of N-qubits and swaps the states in i,j positions:

$$\begin{aligned}
P_{i,j}\left(|P\rangle_i \otimes |M\rangle_j\right) &= |M\rangle_i \otimes |P\rangle_j\,, \\
P_{i,j}\left(|M\rangle_i \otimes |P\rangle_j\right) &= |P\rangle_i \otimes |M\rangle_j\,, \\
P_{i,j}\left(|M\rangle_i \otimes |M\rangle_j\right) &= |M\rangle_i \otimes |M\rangle_j\,, \\
P_{i,j}\left(|P\rangle_i \otimes |P\rangle_j\right) &= |P\rangle_i \otimes |P\rangle_j\,.
\end{aligned} \tag{5.44}$$

A single step of our process can be identified with the action of

$$\widetilde{Q}_{\text{loc}} = \left[\bigotimes_{i\in\text{even}} \frac{1}{2}\Big(P_{i,i+1}+I\Big)\right] \cdot \left[\bigotimes_{i\in\text{odd}} \frac{1}{2}\Big(P_{i,i+1}+I\Big)\right]. \tag{5.45}$$

Instead of analyzing this Markov process, we will approximate it by N repetitions of a process where in each step a random site i is selected and the gate is applied between i and $i+1$. This should approximate the original parallelized circuit at long times and in particular we expect it to give the same N scaling of the gap.[27] The Markov operator for this process is

$$Q_{\text{loc}} = \left[\frac{1}{N}\sum_{i=1}^{N} \frac{1}{2}\Big(P_{i,i+1}+I\Big)\right]^{\otimes N}. \tag{5.46}$$

The number r of M's in a string is conserved by this process. We will use the subscript r to denote the restricted action of $Q_{\text{loc,r}}$ in this subspace. This process has already been studied extensively. It is related to an interchange model on cards studied by Diaconis and Ng [71].[28] They point out that this model can be mapped to a Heisenberg spin chain

$$Q_{\text{loc}} = \left[\frac{3}{4}I - \frac{1}{2N}H_{\text{XXX}}\right]^{\otimes N}, \text{where} \quad H_{\text{XXX}} = -\frac{1}{2}\sum_{i=1}^{N}\vec{\sigma}_i \cdot \vec{\sigma}_{i+1}\,. \tag{5.47}$$

This connection becomes clear by noting that

$$P_{i,j} = \frac{1}{2}\left(I + \vec{\sigma}_i \cdot \vec{\sigma}_j\right). \tag{5.48}$$

It is known that the interchange model for local and 2-local cases defines a reversible and irreducible Markov chain. This process has a uniform stationary distribution on the $\binom{N}{r}$ different strings. The largest eigenvalue of $Q_{\text{loc,r}}$ corresponding to the stationary distribution will be $\lambda_{0,r} = 1$. and all the eigenvalues, as we will discuss, will be between 0 and 1.

---

[27]We confirm numerically in the next section that the analytic results for the Markov dynamics of Q_{loc} are fully consistent with the numerical results on the original parallelized circuit $\widetilde{Q}_{\text{loc}}$.

[28]We thank Persi Diaconis for bringing this work to our attention.

Now we go through a standard Markov chain analysis. Suppose we have an initial string with r M's, $|v_r\rangle = |MPPM \cdots MP\rangle$. Then, we can express

$$\left\langle \text{Tr} \left[U_t \left(MPPM \cdots MP \right) U_t^\dagger \left(MPPM \cdots MP \right) \right] \right\rangle = \langle v_r | Q_{\text{loc}}^t | v_r \rangle \qquad (5.49)$$

Denote by $\mathcal{H}_r$ the subspace of the Hilbert space spanned by vectors that have exactly r elements of M. Let us denote by $\lambda_{a,r}$ the eigenvalues of $Q_{\text{loc},r}$ for $a = 0, 1, \ldots, \binom{N}{r} - 1$ in decreasing order; $1 = \lambda_{0,r} > \lambda_{1,r} \geq \lambda_{2,r} \geq \cdots \geq \lambda_{\binom{N}{r}-1,r} > 0$.

The stationary state $|\lambda_{0,r}\rangle$ is the eigenvector with the eigenvalue $\lambda_{0,r} = 1$ and has uniform weight on each string.

Now, we can calculate $u(k = 1, t)$ for each sector. Recalling the notation (5.39) $M = |1\rangle\langle 1|$ and $P = |0\rangle\langle 0|$, the number r of M's is the number of 1's in the binary representation of the state $|i\rangle$ appearing in (5.42). In terms of the XXZ model, 0 and 1 are spin up $(+1/2)$ and down $(-1/2)$, so this state $|i\rangle$ has total spin $S_z = \frac{1}{2}(N-r) - \frac{1}{2}r = (N/2) - r$. Therefore,

$$\left\langle |\text{Tr} U(t)|^2 \right\rangle_{S_z = N/2 - r} = \sum_{i \in \{S_z = N/2 - r\}} \langle U_{ii} U_{ii}^* \rangle = \sum_{v_r} \langle v_r | Q_{\text{loc},r}^t | v_r \rangle \qquad (5.50)$$

$$= \text{Tr} \, Q_{\text{loc},r}^t = \sum_{a=0}^{\binom{N}{r}-1} \lambda_{a,r}^t \qquad (5.51)$$

where the $|v_r\rangle$ that are summed over are all the permutations of the vector $|MPPM \cdots MP\rangle$ with r M's.

We now define a gap parameter Δ_r by $e^{-\Delta_r} = \lambda_{1,r}$. At large N this is just the gap in the H_{XXX} spin chain.[29] We can give a simple lower bound on $u(1,t)$ by just keeping the first two terms of (5.51)

$$u(1,t) > 1 + e^{-\Delta_r t} \,. \qquad (5.52)$$

So the time to get ϵ close to the Haar value[30] is lower bounded by

$$t_\epsilon > \frac{1}{\Delta_r} \log \frac{1}{\epsilon} \,. \qquad (5.53)$$

The gap is determined by the properties of the XXX spin chain which is one of the canonical examples of an integrable system. Reference [71] quotes the following results.[31] For $r = 1, 2$ the gap is the same and is given by $\Delta_1 = \Delta_2 = \frac{c}{N^2}$ where c is an order one constant. Further it is conjectured that the gap is the same for all r. Our numerics (see the following section) support this. Using this result we have

$$t_\epsilon > \frac{N^2}{c} \log \frac{1}{\epsilon} \,. \qquad (5.54)$$

As we discussed earlier the N^2 dependence is a signal of diffusion.

---

[29]Up to an order one constant.

[30]t_ϵ is a precise version of t_{Haar}.

[31]We have not exhaustively studied the literature on this model. It may well be the case that more analytic results are known.

A sharp upper bound is more subtle to obtain. The terms ignored in (5.52) contribute an N dependent offset to the time required to reach the Haar value. We define t_{offset} by

$$t_\epsilon = t_{\text{offset}} + \frac{N^2}{c} \log \frac{1}{\epsilon} \,. \tag{5.55}$$

We can begin to estimate t_{offset} by approximating the XXX chain spectrum. In the large N finite r limit the spectrum is dominated by r free magnons, each with nonrelativistic spectrum

$$E(p) \sim p^2/2 \,. \tag{5.56}$$

Here p is a continuum momentum, where we assume $\frac{1}{N} \ll p \ll 1$.

Assuming $r \ll N$ and $t \ll N^2$ we can approximate

$$u(1,t) \approx \frac{1}{r!} \left(N \int dp\, e^{-p^2 t/2} \right)^r \sim \frac{1}{r!} \left(\frac{N}{\sqrt{t}} \right)^r \,. \tag{5.57}$$

This becomes order one at a time $t_{\text{offset}} \sim \frac{N^2}{r^2}$. Here we see that the diffusional dynamics is related to the nonrelativistic dispersion relation (5.56) of the XXX chain.

But this is not the slowest dynamics in this system. It is known that there are bound states of b magnons for all $b \leq r$. These have "mass" of order b and hence a spectrum (see [72])

$$E(p) \sim p^2/(2b) \,. \tag{5.58}$$

These bound states diffuse more slowly because of their greater mass. Consider the case when all r magnons are bound up into bound states of size b, producing r/b "particles". This makes the following contribution to $u(1,t)$:

$$u(1,t) \supset \frac{1}{(r/b)!} \left(N \int dp\, e^{-p^2 t/(2b)} \right)^{r/b} \sim \frac{1}{(r/b)!} \left(\frac{N}{\sqrt{t/b}} \right)^{r/b} \tag{5.59}$$

This interesting function of time becomes order one at a time $t_{\text{offset}} \sim N^2 b^3/r^2$. This suggests we can make $t_{\text{offset}} \sim N^2$ by choosing $b \sim r^{2/3}$. In fact it seems we can make t_{offset} parametrically longer, $\sim N^2 r$, by choosing the largest allowed value $b \sim r$. For $r \sim N$ this would give $t_{\text{offset}} \sim N^3$. However, when $t \sim N^2$ the discreteness of the spectrum becomes important and these approximations are no longer valid. Instead we turn to numerics (discussed below) which indicate that t_{offset} is independent of r for large enough r and hence $t_{\text{offset}} \sim N^2$.[32] Our best estimate becomes

$$t_\epsilon = \tilde{c} N^2 + \frac{N^2}{c} \log \frac{1}{\epsilon} \,. \tag{5.60}$$

Again we point out that this N^2 dependence is parametrically slower than the scrambling time for this system which is N [43, 48], corresponding to ballistic propagation.

---

[32] We cannot rule out a weak r dependence in t_{offset} from these numerics so for example an $N^2 \log(N/C)$ behavior with a large C coefficient is possible.

We now turn to the 2-local XXZ random circuit. The analysis is analogous to the local system and we end up with a Markov chain. We first consider a process where at each step a gate acts on one randomly selected pair of qubits. One parallel step of the full process corresponds to N repetitions of this. The Markov operator for this process is

$$Q_{\text{2-loc}} = \left[\frac{1}{\binom{N}{2}} \sum_{i<j}^{N} \frac{1}{2} \left(P_{i,j} + I \right) \right]^{\otimes N}. \tag{5.61}$$

The spectrum of this operator differs in important ways from the geometrically local case giving a different N scaling for t_ϵ.

The spectrum of $Q_{\text{2-loc,r}}$ is related to the gap of the Heisenberg model on a complete graph. More specifically, we express

$$Q_{\text{2-loc}} = \left[\frac{3}{4} I - \frac{1}{2N} \widetilde{H}_{\text{XXX}} \right]^{\otimes N}, \quad \text{where} \quad \widetilde{H}_{\text{XXX}} = -\frac{N}{\binom{N}{2}} \frac{1}{2} \sum_{i<j}^{N} \vec{\sigma}_i \cdot \vec{\sigma}_j. \tag{5.62}$$

Then, we use a simple identity

$$\frac{1}{2} \sum_{i<j}^{N} \vec{\sigma}_i \cdot \vec{\sigma}_j = \left(\frac{1}{2} \sum_{i=1}^{N} \vec{\sigma}_i \right)^2 - \frac{1}{4} \sum_{i=1}^{N} \vec{\sigma}_i^2 = \left(\frac{1}{2} \sum_{i=1}^{N} \vec{\sigma}_i \right)^2 - \frac{3N}{4} = \vec{S}^2 - \frac{3N}{4} \tag{5.63}$$

where we use $\vec{S}^2$ to denote the total spin squared operator. So we find

$$\widetilde{H}_{\text{XXX}} = \frac{N}{\binom{N}{2}} \left(\frac{3N}{4} I - \vec{S}^2 \right) \tag{5.64}$$

Recall that $S_z = \frac{1}{2} \sum_i^N \sigma_i^3$ and it commutes with $\vec{S}^2$. In the basis where both are diagonal $\vec{S}^2$ has eigenvalues $s(s+1)$, where $s = 0, 1, \cdots, N/2$ for even N and S_z has eigenvalues $s_z = -s, -s+1, \cdots, +s$. Note that $r = \frac{N}{2} - s_z$. The eigenvalues of $\widetilde{H}_{\text{XXX}}$ are given by

$$E_{(s,s_z)} = \frac{3N^2}{4} \frac{1}{\binom{N}{2}} - \frac{N}{\binom{N}{2}} s(s+1) \tag{5.65}$$

Note that the eigenvalue is independent of s_z and only depends on s. In a given r sector, that is for a given value of s_z, all values of s such that $\frac{N}{2} \geq s \geq s_z$ contribute.

At large N the important eigenvalues $Q_{\text{2-loc}}$ are given by $\lambda_a = e^{-E_a}$. So we can write

$$u(1,t) = \sum_a e^{-E_a t} \tag{5.66}$$

and the gap is given by $E_1 - E_0$. For all r sectors E_0 corresponds to $s = \frac{N}{2}$ and E_1 to $s = \frac{N-2}{2}$. So at large N the gap is N independent.

$$\Delta_r = E_1 - E_0 \to 2. \tag{5.67}$$

Diffusion on this completely connected graph takes place in order one time. The degeneracy of this first excited state is N [73] and so a lower bound for $u(1, t)$ is

$$u(1, t) > 1 + N \cdot e^{-2t} \tag{5.68}$$

giving a Haar time of

$$t_\epsilon > \frac{1}{2} \log N + \frac{1}{2} \log \epsilon \tag{5.69}$$

In fact this is an accurate estimate. The small gaps grow linearly in level, $E_n - E_0 = 2n$ and the degeneracies grow like $\binom{N}{n} - \binom{N}{n-1}$.[33] So the full formula is well approximated by

$$u(1, t) \approx (1 + e^{-2t})^N \cdot (1 - e^{-2t}) + e^{-2t(N+1)} . \tag{5.70}$$

The Haar time is again given by (5.69). Note again that this logarithmic dependence is due to a different mechanism than the logarithmic scrambling time.

One can study $u(2, t)$ for this model using the techniques discussed in appendix B.2. We do not perform a full analysis of this case, but note that the "unequal chain" part of equation (B.15) (the last term) will converge slowly because of the same diffusive effects described above. This indicates that the time scale to converge to the Haar value for $u(2, t)$ will be parametrically the same as in the $u(1, t)$ analysis.

5.2.2 Numerical results for XXZ Markov chain

We conclude this section by showing some numerical results for local XXZ random circuits. For the local XXZ Markov chain the upper left panel of figure 23 illustrates that for a fixed value of N the low lying eigenvalues are independent of the value of r, indicating that bound state effects of the kind described in (5.59) are not important for the final approach to Haar. We can see that the offset is roughly $\sim N^2$ by looking at the behavior of $u(1, t)$ in the lower panel. We plot the $u(1, t)$ for $r = N/2$ (i.e., $S_z = 0$) sector for different values of N. Note that when the time is rescaled by N^2 the results align completely. This suggests that the best numerical fit to the time to get ϵ close to Haar is given by

$$t_\epsilon \sim c_1 N^2 \left[c_2 + \log \frac{1}{\epsilon} \right] . \tag{5.71}$$

where c_1 and c_2 are constants independent of N. However, we would like to emphasize that we cannot rule out a possibility of a weak N dependence in the coefficient c_2.

6 Towards a Hamiltonian estimate

In this section we make some preliminary remarks about an analytic estimate of the ramp time in Hamiltonian systems. We use the approach of [49, 50] and rewrite the spectral

---

[33]Note that this is the degeneracy of the energy levels for each sector of fixed spin in Z direction (which we label by s_z or r).
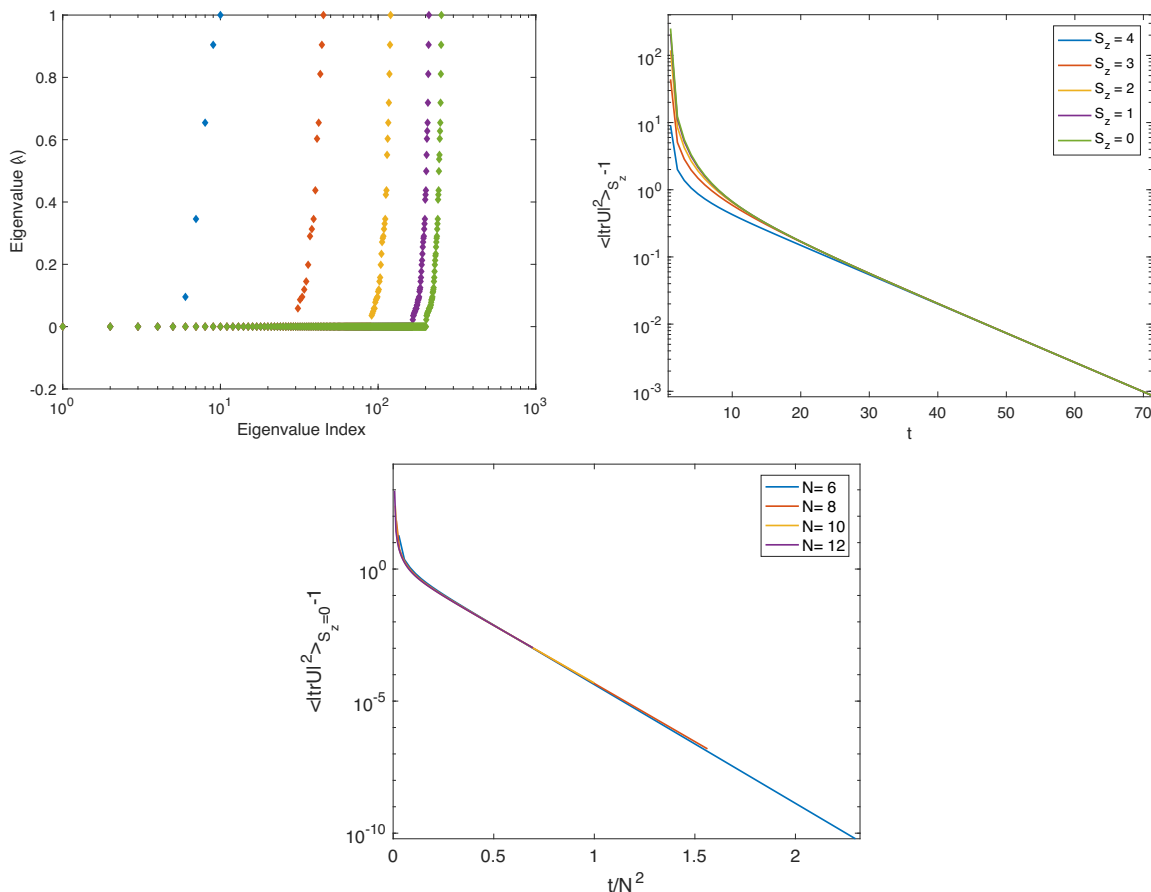
Figure 23. Numerical results for local XXZ random circuit. [Top-Left] The eigenvalues of $Q_{\text{loc},r}$ for $r = 0, 1, 2, 3, 4$ for $N = 8$. The second largest eigenvalues take the same value. [Top-Right] The $u(1,t)$ for $r = 0, 1, 2, 3, 4$ at $N = 8$. They all converge to the same exponential decay in late times. [Bottom] The spectral form factors for several values of N in the largest sector, where $r = N/2$ (or $S_z = 0$).

form factor as a sum over correlation functions. We first discuss qubit systems. Rewriting the operator trace in terms of all Pauli operators P_j [49] one has for any operator $\widehat{O}$

$$I \cdot \text{Tr}(\widehat{O}) = \frac{1}{L} \sum_{j=1}^{L^2} P_j \widehat{O} P_j \tag{6.1}$$

Introducing another operator $\widehat{Q}$ and taking the trace on both sides we get

$$\text{Tr}(\widehat{Q}) \cdot \text{Tr}(\widehat{O}) = \frac{1}{L} \sum_{j=1}^{L^2} \text{Tr}\left(\widehat{Q} P_j \widehat{O} P_j\right) \tag{6.2}$$

If we select $\widehat{Q} = \widehat{O}^\dagger = e^{+iHt}$ and average over the ensemble for H, we have a rewrite of the infinite temperature spectral form factor

$$\left\langle \text{Tr}(e^{+iHt}) \cdot \text{Tr}(e^{-iHt}) \right\rangle = 1 + \sum_{j=2}^{L^2} \frac{1}{L} \left\langle \text{Tr}\left(e^{+iHt} P_j e^{-iHt} P_j\right) \right\rangle \tag{6.3}$$

To understand the approach to the ramp we must understand the decay of the two point correlators of all $L^2 - 1$ non-identity Pauli strings ($\langle P_j(t)P_j\rangle$). We intuitively expect that operators made out of more Paulis will decay faster.[34] This is just what the random circuit analysis shows as well (see discussion in section 5.1). So we expect the leading contribution to come from the Pauli strings with a single non-identity element — there are $3N$ such strings. Each of these simple two point functions will decay exponentially $\sim e^{-mt}$ for some gap parameter m. Therefore, the spectral form factor can be approximated by

$$\langle \text{Tr}(e^{-iHt}) \cdot \text{Tr}(e^{+iHt}) \rangle \sim 1 + 3N \cdot e^{-mt} \tag{6.4}$$

so we expect $\log N$ scaling for the ramp time.[35] In a local system with diffusion we expect power law decay of correlations which would result in power law ramp times. The connection of these estimate with the random circuit calculations is clear. However, there are several issues that are significantly more subtle here.

Factorization

We assumed here that Pauli strings with d non-identity elements decay fast enough to ignore. In large N systems expect factorization

$$\langle (\sigma_1\sigma_2\ldots\sigma_d)(t)(\sigma_1\sigma_2\ldots\sigma_d)(0)\rangle \tag{6.5}$$
$$\approx \langle\sigma_1(t)\sigma_1(0)\rangle\langle\sigma_2(t)\sigma_2(0)\rangle\ldots\langle\sigma_d(t)\sigma_d(0)\rangle + \mathcal{O}\left(\frac{1}{N}\right)$$

So we expect the early time behavior of (6.5) to be $\sim e^{-dmt}$. As we have seen in the random circuit analyses this exponential behavior in d is sufficiently fast to let the $d = 1$ terms control the estimate.

But in these systems there is operator mixing at higher order in $\frac{1}{N}$ between d Pauli operators and operators with fewer Paulis that will cause (6.5) to decay more slowly at long time. The characteristic crossover time is of order $\log N$,[36] which is the time scale of importance here. It is crucial that the operator product coefficients be sufficiently small at large N to make these slowly decaying terms subleading. This should be the case in large N systems. For instance in SYK the amplitude for 3 fermions to mix with 1 is order $\frac{1}{N^{3/2}}$ (assuming $q = 4$). In large N gauge theories the amplitude for 2 gluons to mix with 1 is order $\frac{1}{N}$. If one has a diagrammatic expansion then heuristically at least the same scalings that give a smooth large N limit ensure that no large N dependence occurs in the long time decay. Making these arguments precise and general will be crucial in turning this discussion into a sharp statement.[37]

---

[34]We have verified this numerically for the 2-local RCQ. But because the bump which we do not understand appears in the correlators we must regard these results as provisional.

[35]In fact the early time decay will hit the ramp rather then reaching all the way to an order 1 value. Naively this suggests that we equate $Ne^{-mt} = t$. Solving this will give a dominant $\log N$ term as well as a subleading $\log \log N$ piece.

[36]Operator growth in 2-local systems provides a nice way of visualizing this.

[37]Reproducing the power law slope of the SYK spectral form factor with its exponentially large coefficient seems especially subtle.

Edge

As we have seen a basic feature of the spectral form factor in Hamiltonian systems is the ability of the "edge" of low energy states to give a contribution that swamps the small ramp signal until exponentally late time.[38] We gave a method of removing this contribution by using Gaussian envelope in the modified quantity $|Y(t)|^2$ defined in equation (3.12). We need to do something similar here. It turns out the most useful quantity is a hybrid:

$$\langle Y(\alpha,t)Z(\beta,t)^* \rangle = \text{Tr}\left(e^{-\alpha H^2}e^{-iHt}\right) \cdot \text{Tr}\left(e^{-\beta H}e^{-iHt}\right) \tag{6.6}$$

This one sided Gaussian window is sufficient to remove the sharp edge.

To write this as a sum of correlation functions choose operators of the form $\widehat{Q} = e^{-\alpha H^2}e^{-iHt}$ and $\widehat{O} = e^{-\beta H}e^{-iHt}$. We then have

$$\langle Y(\alpha,t)Z(\beta,t)^* \rangle = \frac{1}{L}\sum_{j=1}^{L^2}\left\langle \text{Tr}\left(e^{-\alpha H^2}P_j(t)e^{-\beta H}P_j\right)\right\rangle . \tag{6.7}$$

In the infinite temperature limit $\beta = 0$, we will get

$$\langle Y(\alpha,t)Z(\beta=0,t)^* \rangle = \frac{1}{L}\sum_{j=1}^{L^2}\left\langle \text{Tr}\left(e^{-\alpha H^2}P_j(t)P_j\right)\right\rangle \tag{6.8}$$

Essentially the same analysis goes through for Majorana fermions. Denote by ψ_i Majorana fermions $i = 1, 2, \cdots, 2N$. The operators live in $L = 2^N$ dimensional Hilbert space. Let us denote by

$$Q_I = \psi_{i_1}\psi_{i_2}\cdots\psi_{i_\omega} \tag{6.9}$$

where $I = \{i_1, i_2, \ldots, i_\omega\}$, ω is the weight of the operator and $i_1 < i_2 < \cdots < i_\omega$. There are total of L^2 such strings (including the identity) and we will use index $I = \{i_1, i_2, \ldots, i_\omega\} \to j$ to denote them and ω_j will be the weight of the j-th element. One can check that for any operator $\widehat{O}$ we can establish following identity

$$I \cdot \text{Tr}(\widehat{O}) = \frac{1}{L}\sum_{j=1}^{L^2} 2^{\omega_j} \cdot Q_j^\dagger \widehat{O} Q_j \tag{6.10}$$

The reason we have a dagger is that Q_j can be both Hermitian and anti-Hermition. The factor 2^{ω_j} is added because $\psi_i^2 = I/2$ for each Majorana in standard normalization. Using the same insertion as in the Pauli case we establish

$$\langle Y(\alpha,t)Z(\beta=0,t)^* \rangle = \frac{1}{L}\sum_{j=1}^{L^2} 2^{\omega_j}\left\langle \text{Tr}\left(e^{-\alpha H^2}Q_j^\dagger(t)Q_j\right)\right\rangle \tag{6.11}$$

where the sum is over all the ordered Majorana fermion strings.

---

[38]This feature is absent in random circuits since energy is not conserved and a system cannot stay near the edge.

Thus, the error from the Haar value will depend on the decay of the two-point function of the simplest one fermion operators in the "α ensemble"

$$\langle \psi_i(t)\psi_i \rangle_\alpha . \tag{6.12}$$

In SYK this correlator can be determined by Laplace transforming the known large N finite temperature two point function. Using an $\alpha \sim \frac{1}{N}$ which effectively removes the edge, we find results are dominated by the infinite temperature correlator which has $m \sim 1$ (in units of J). There are N such operators so we find a decay of the form Ne^{-Jt} giving a ramp time of order $\frac{1}{J} \log N$, consistent with our expectations.

There is an interesting consequence of this link for the q dependence of t_{ramp} in SYK.[39] Using the standard scaling the two point function decays like $e^{-(\mathcal{J}t/q)}$ at large q [12, 17]. So $t_{\text{ramp}} \sim \frac{\log N}{q}$. The chaos exponent is q independent so the scrambling time is $\sim \log N$, parametrically longer than the ramp time.

7   Closing remarks

We have argued above, both in the random quantum circuit and Hamiltonian context, that the scale for the ramp time t_{ramp} is set by the decay of N simple two point functions. The scrambling time is determined by the decay of the OTOC. Both of these quantities are driven by the underlying mechanism of operator growth under time evolution [18, 21, 74], but they diagnose different aspects of the growing operator.

This is made very clear in recent work on SYK [75].[40]   These authors define the probability $P_s(t)$ of having an operator of size s (i.e., containing s majorana fermions) at time t. The two point function of a single fermion is the amount of the operator remaining in that fermion $P_1(t)$. This decays exponentially with the dissipation time scale t_{diss}. The contribution to the spectral form factor is then $Ne^{-t/t_{\text{diss}}}$. This becomes order one at $t_{\text{ramp}} \sim t_{\text{diss}} \log N$. The (anti) commutator squared grows like the average size of the operator $\frac{1}{N}\sum_s sP_s(t)$. This behaves like $\frac{e^{\lambda_L t}}{N}$ where λ_L is the chaos exponent, the quantum analog of the Lyapunov exponent. This becomes order one at the scrambling time $t_{\text{scr}} \sim \frac{1}{\lambda_L} \log N$. We see that these two time scales, while both displaying $\log N$ behavior, describe quite different aspects of the growing operator. The prefactors of $\log N$ can be parametrically different, as mentioned above. At large q, $t_{\text{diss}} \sim \beta/q$ while $\frac{1}{\lambda_L} \sim \beta$. There are many holographic examples that display this parametric separation as well.

In this paper we have studied the onset of random matrix behavior in strongly chaotic many-body systems. But it remains to give a convincing argument for random matrix behavior over the full range of energy and time scales present. For example it is an important task to give a general argument for presence of the full ramp/plateau structure in the spectral form factor. One approach that has been suggested in [2, 76] is to extend the ideas of Dyson Brownian motion used with great effect in [76] to many-body systems. Further exploring this approach is an interesting subject for future research. One important issue

---

[39]Recall that q is the number of fermions coupled together in the interaction term.
[40]We thank Alex Streicher for explaining these results to us.

will be to understand the role of the comparatively large $\mathcal{O}(1/N)$ (rather than $\mathcal{O}(1/L)$) fluctuations in eigenvalue density that are characteristic of many-body systems.

Another interesting approach has been developed recently in [3]. They apply ideas of single-body random hopping problems to the exponentially large Hilbert space of the many-body system. Their approach does not get the value of the ramp time we argue for but we think this technique may well be useful to describe the later time ramp and plateau.

This approach is closely connected to the observation in appendix F of [8] that the connected contributions to $\langle \mathrm{Tr} H^k \mathrm{Tr} H^k \rangle$ where H is the SYK Hamiltonian have the perturbative SYK value for k fixed as N gets large, but at $k^* \approx N \log N$ this quantity assumes the value it would have if H were a random matrix in the second quantized Hilbert space. Naively converting this to a time scale (remembering that the width of the spectrum is $\sim \sqrt{N}$) one finds $t \sim \sqrt{N} \log N$, the scale found in [3].

The result of the calculation in [3] is a formula rather similar to (6.3), a sum over an exponentially large number of more and more rapidly decaying terms. The leading correction is of the form $Ne^{-\mu t}$ much like (6.4). But here $\mu \sim 1/\sqrt{N}$, corresponding to a parametrically slower decay rate than in the SYK version of (6.4). It is this slow decay that causes the large estimate $t_{\mathrm{ramp}} \sim \sqrt{N} \log N$. There are no perturbative modes in SYK with such a slow decay rate[41] so their significance is a puzzle to which we hope to return.[42]

One of the motivations for thinking about these strongly chaotic, scrambling, systems is their connection to the physics of quantum black holes. One important thing to understand is which phenomena are important in both small and large AdS black holes and which only occur in large ones. The time scale for evaporation of a small black hole is order S (where S is the black hole entropy). The ramp time we have found for the k-local systems appropriate to the black hole horizon is order $\log S$, parametrically smaller than the evaporation time. These effects will be exponentially subdominant in simple observables at these short times so studying them will be difficult, but they are present and are a signature of the nonperturbative quantum dynamics of these systems.

Perhaps the most important question from the gravitational point of view is the bulk dual in the sense of AdS/CFT of these random matrix phenomena [8]. In SYK a related question is the explanation of these phenomena in terms of the G, Σ collective fields. We have made progress on this problem which we hope to report on in the near future [70].

Acknowledgments

---

[41]There are power-law decays in correlators due to the sharp edge in the density of states [61]. But for quantities focussed in the bulk of the band like (6.11) these decays should be exponentially suppressed because of the small matrix elements between bulk states and edge states.

[42]The methods being used in [70] should be helpful.

In accordance with institutional policy the data used in the preparation of this paper is available to other scientists on request.

## A Additional numerical results

A.1 Number variance in SYK

García-García and Verbaarschot studied the number variance in the SYK model; see figure 9 and figure 10 in [7]. They observed a deviation from the RMT result at $K \sim N^2$ [26]. In section 3.1 we have explained that this very small value of K is due to the large fluctuations in the variance of the energy distribution. As a preliminary check we have implemented a slightly improved unfolding procedure:

1. Shift the origin of the energy spectrum so that $\langle E \rangle = 0$ for each sample.

2. Then we rescale the eigenvalues for each sample so that the variance for each sample is the same as $\langle \text{Tr} H^2 \rangle$, the average variance.

3. Then we unfold each sample using the density profile determined by $\langle \rho(E) \rangle$.

4. Then we compute $\Sigma^2(K)$.

Steps (1) and (2) are not implemented in [7]. The number variance calculated in this manner is shown in figure 24, for $N = 32$ and $N = 34$. The agreement with RMT persists to larger K, roughly a factor of two larger, than in [7]. Thus we see that the number variance is sensitive to the unfolding procedure.

As discussed in the text we expect that RMT behavior extends to an exponentially large value of K, much larger than shown even by these improved results. Our unfolding procedure is still rather crude. As a first step one could try to fully unfold the spectrum sample by sample. But it is unclear to us whether this is even in principle sufficient. At a minimum an exponentially large number of samples would be required to get a sufficiently accurate value of the variance.

For this reason the number variance seems not to be the best quantity to probe long range spectral rigidity in many-body systems. As discussed in the text the spectral form factor is less sensitive to such errors.
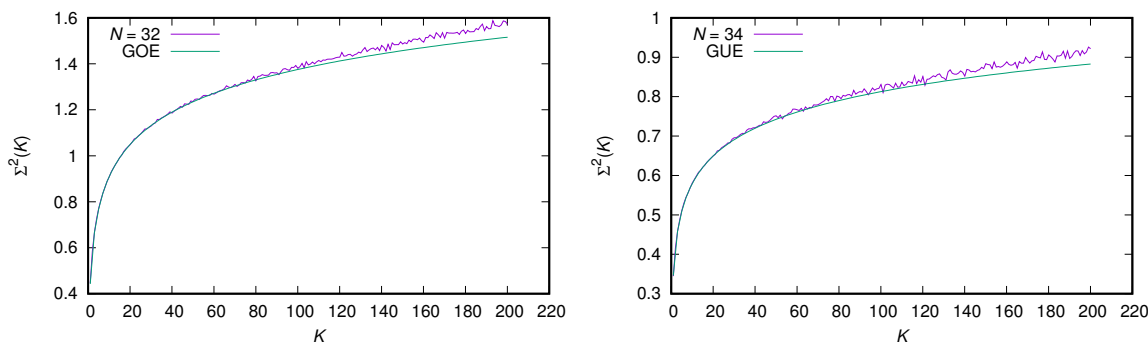
Figure 24. $\Sigma^2(K)$ for the SYK model, $N = 32$ (250 samples) and $N = 34$ (110 samples), obtained using the unfolded spectra after shifting and rescaling. The curves for corresponding random matrix ensembles, GOE and GUE, are also plotted.

A.2 Estimating $t_{\min}$ in the SYK model

In this appendix we give our detailed results for determining $t_{\min}$ in the SYK model numerically. The algorithm used to select an optimal α is as follows: in the second panel of figure 3 we have varied α with step size 0.5, and observed that some α between $2.5 < \alpha < 3.0$ gives the smallest value of $t_{\min}$; at $\alpha \le 2.5$ the oscillation due to the sharp edges survives; the oscillation disappears at $\alpha = 3.0$, and after then the min moves toward later time. By closer inspection we conclude that α for the smallest $t_{\min}$ is 2.9 ± 0.1. In figure 4, we plotted the near-dip region of $g(t)$ and $h(\alpha = 2.5, t)$. $h(\alpha = 2.5, t)$ is rather smooth near the dip and $h(\alpha \ge 3.0, t)$ has a single minimum. Hence we can estimate the min time by choosing the value of t in the data which gives a single minimum of $h(\alpha, t)$. The neighboring data points can give the edges of the error bar for α. We should emphasize that this procedure is somewhat arbitrary.

It is hard to determine the N-dependence of t_{ramp} from table 1. The fact that $t_{\min}$ and hence an upper bound for t_{ramp} is nonmonotonic, a physically unlikely result, indicates the uncertainty in the procedure. The best we can say is that t_{ramp} is certainly far smaller than the dip time t_{dip} and increases at most slowly with N. Assuming order one coefficients a power law faster than linear or an exponential behavior (as suggested in [7]) is disfavored.

A.3 Eigenvalue behavior for 2-local RCQ

In order to test the spectral rigidity of the energy spectrum, we performed the unfolding with steps 1–4 explained in section A.1. The shifted-and-rescaled energy spectrum (obtained by performing steps 1, 2 and 3), before the unfolding, is shown in the right panel of figure 5. Unlike the SYK model, the edge of the spectrum is not sharp. Due to this, the slope of SFF decays much faster, as we will see shortly. The nearest-level separation obtained from the unfolded spectrum is plotted in the left panel of figure 25. A good agreement with GUE ensemble at large N can be seen. The number variance $\Sigma^2(K)$ is plotted in the right panel of figure 25. At $N = 16$, an agreement with RMT can be seen only at $K \lesssim 10$.

N	smallest $t_{\min}$	α for smallest $t_{\min}$
10	9.1 ± 0.1	4.1 ± 0.1
12	8.7 ± 0.1	4.2 ± 0.1
14	10.1 ± 0.1	4.9 ± 0.1
16	12.9 ± 0.1	4.7 ± 0.1
18	12.2 ± 0.1	3.9 ± 0.1
20	11.5 ± 0.1	3.7 ± 0.1
22	11.2 ± 0.1	3.5 ± 0.1
24	11.7 ± 0.1	3.5 ± 0.1
26	14.5 ± 0.1	3.3 ± 0.1
28	13.4 ± 0.1	3.0 ± 0.1
30	13.0 ± 0.1	2.9 ± 0.1
32	12.5 ± 0.1	2.9 ± 0.1
34	12.1 ± 0.1	2.8 ± 0.1

Table 1. The value of smallest $t_{\min}$, defined as the value of t when $|Y(\alpha, \beta = 0, t)|^2 / |Y(\alpha, \beta = 0, t = 0)|^2$ has a single minimum as a function of t for the smallest α. The value of such α is also given. The number of samples is $240000 \times 2^{(14-N)/2}$ for $N = 10, 12, 14$, $600000 \times 2^{(18-N)/2}$ for $N = 16, 18$, $120000 \times 2^{(22-N)/2}$ for $N = 20, 22$, 48000 for $N = 24$, 10000 for $N = 26$, 3000 for $N = 28$, 1000 for $N = 30$, 1500 for $N = 32$ and 110 for $N = 34$.
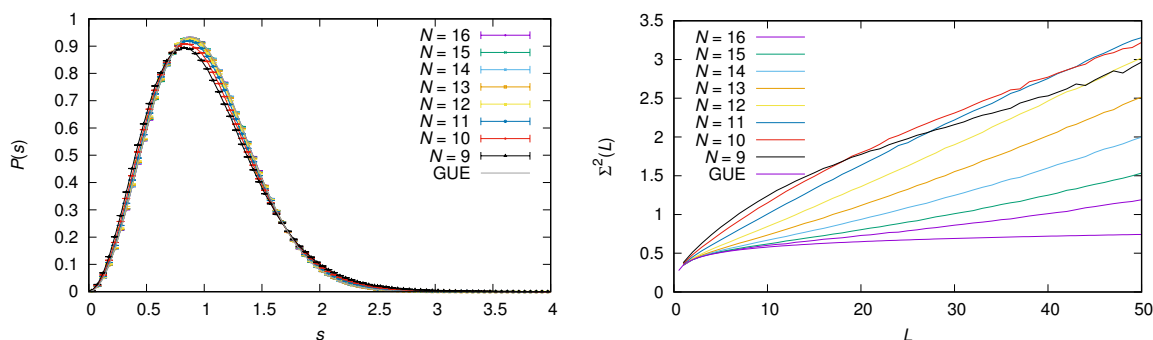
Figure 25. Unfolded level separation distribution $P(s)$ for the RCQ model for $N = 9, 10, \ldots, 16$ compared against the Padé approximant for GUE, and the number variance $\Sigma^2(K)$. L on the right panel corresponds to K in the text.

B   Analytic results for random quantum circuit

In this appendix we collect several analytic results which complement the discussion in section 5.

B.1 k-th moment of the Haar measure

If we take the average with respect to the Haar measure on $\mathrm{U}(L)$, $u(k, t) = \left\langle |\mathrm{Tr}U^k(t)|^2 \right\rangle$ is equal to k for $k \leq L$ and is equal to L for $k > L$. For $k = 1$ and $k = 2$, it can be seen by

contracting the indices in (see e. g. ref. [41])

$$\langle U_{i_1 i_2} U^*_{j_1 j_2} \rangle_{\mathrm{Haar}} = \frac{1}{L} \delta_{i_1 j_1} \delta_{i_2 j_2} \tag{B.1}$$

and

$$\langle U_{i_1 i_2} U^*_{j_1 j_2} U_{i_3 i_4} U^*_{j_3 j_4} \rangle_{\mathrm{Haar}} = \frac{1}{L^2 - 1} \left[\delta_{i_1 j_1} \delta_{i_2 j_2} \delta_{i_3 j_3} \delta_{i_4 j_4} + \delta_{i_1 j_3} \delta_{i_2 j_4} \delta_{i_3 j_1} \delta_{i_4 j_2} \right. \tag{B.2}$$
$$\left. - \frac{1}{L} (\delta_{i_2 j_2} \delta_{i_4 j_4} \delta_{i_1 j_3} \delta_{i_3 j_1} + \delta_{i_1 j_1} \delta_{i_3 j_3} \delta_{i_2 j_4} \delta_{i_4 j_2}) \right],$$

as

$$\langle |\mathrm{Tr} U(t)|^2 \rangle_{\mathrm{Haar}} = \sum_{i_1, i_2, j_1, j_2} \delta_{i_1 i_2} \delta_{j_1 j_2} \cdot \frac{1}{L} \delta_{i_1 j_1} \delta_{i_2 j_2} = 1 \tag{B.3}$$

and

$$\langle |\mathrm{Tr} U^2(t)|^2 \rangle_{\mathrm{Haar}} = \sum_{i_1, i_2, j_1, j_2 = 1}^{L} \frac{1}{L^2 - 1} \left[\delta_{i_1 j_1} \delta_{i_2 j_2} + \delta_{i_1 j_2} \delta_{i_2 j_1} - \frac{1}{L} 2 \delta_{i_1 i_2 j_1 j_2} \right] \tag{B.4}$$

$$= \sum_{i,j=1}^{L} \frac{2}{L^2 - 1} - \frac{1}{L} \sum_{i=1}^{L} \frac{2}{L^2 - 1} \tag{B.5}$$

$$= 2 \frac{L^2}{L^2 - 1} - \frac{2}{L^2 - 1} = 2. \tag{B.6}$$

A proof for generic k is discussed in section 3.2 (also see [63]).

B.2 Calculating $\langle |\mathrm{Tr}\, U^2(t)|^2 \rangle$ in RCQ random circuit

In this section we study

$$u(2, t) = \langle |\mathrm{Tr} U^2(t)|^2 \rangle = \sum_{i_1, i_2, j_1, j_2 = 1}^{L} \langle U_{i_1 i_2} U_{i_2 i_1} U^*_{j_1 j_2} U^*_{j_2 j_1} \rangle. \tag{B.7}$$

for the RCQ random circuit using the techniques of [34]. The generalization to the Brownian circuit is straightforward. Here we specialize to the U(4) gateset. We need to estimate two kinds of contributions:

- $\langle U_{ii} U_{ii} U^*_{jj} U^*_{jj} \rangle$. These terms start out at 1 and decay to their Haar value which from (B.2) has a large L value of $2/L^2$ for $i = j$. It is zero for $i \neq j$.

- $\langle U_{ij} U_{ji} U^*_{ij} U^*_{ji} \rangle$ with $i \neq j$. These terms start at zero and reach their Haar value $\sim 1/L^2$.

Estimating $\left\langle U_{ij}U_{ji}U_{ij}^*U_{ji}^*\right\rangle$ $(i \neq j)$

Without loss of generality we can set $i = 1$ and consider $\left\langle U_{1j}U_{j1}U_{1j}^*U_{j1}^*\right\rangle$ $(j \neq 1)$. These terms start out at zero and reach their Haar value $\sim 1/L^2$. In the $0, 1$-basis, the $|1\rangle$ state is expressed by $|0^N\rangle$. In the same way as the discussion below equation (5.17), we take the state j to be of the form $|1^q 0^{N-q}\rangle$, again without loss of generality. The argument is a simple generalization of the one for $\left\langle U_{ii}U_{jj}^*\right\rangle$, presented below (5.17). We start from the initial state on two copies of the system,

$$\hat{\rho}^{(2)}(0) = \left(\left|1^q 0^{N-q}\right\rangle\left\langle 1^q 0^{N-q}\right|\right) \otimes \left(\left|0^N\right\rangle\left\langle 0^N\right|\right). \tag{B.8}$$

Then, we can express

$$U_{1j}U_{j1}U_{1j}^*U_{j1}^* = \mathrm{Tr}\left[\hat{\rho}^{(2)}(t)\tilde{\hat{\rho}}^{(2)}(0)\right], \tag{B.9}$$

where

$$\tilde{\hat{\rho}}^{(2)}(0) = \left(\left|0^N\right\rangle\left\langle 0^N\right|\right) \otimes \left(\left|1^q 0^{N-q}\right\rangle\left\langle 1^q 0^{N-q}\right|\right). \tag{B.10}$$

We will use $\tilde{\gamma}_0(p_1, p_2)$ to denote the Pauli coefficients for $\tilde{\hat{\rho}}^{(2)}(0)$. Recalling that

$$|0\rangle\langle 0| = \frac{1}{2}(I + Z) \quad \text{and} \quad |1\rangle\langle 1| = \frac{1}{2}(I - Z), \tag{B.11}$$

we can see that $\tilde{\hat{\rho}}^{(2)}(0)$ consists of the Pauli strings made only of I and Z. Therefore,

$$\left\langle U_{1j}U_{j1}U_{1j}^*U_{j1}^*\right\rangle = \sum_{p_1', p_2'} \gamma_t(p_1', p_2')\tilde{\gamma}_0(p_1', p_2'), \tag{B.12}$$

where p_1' and p_2' run through the Pauli strings made only of I and Z. That $\left\langle U_{1j}U_{j1}U_{1j}^*U_{j1}^*\right\rangle$ is zero at $t = 0$ means $\gamma_t(p_1', p_2')\tilde{\gamma}_0(p_1', p_2')$ cancel with each other at early time (due to sign differences). In order to see the time evolution of the second and third terms, we need to understand how the unitary gates selected from $\Gamma = \mathrm{U}(4)$ act on two copies of Pauli strings $\sigma_{p_1'}$ and $\sigma_{p_2'}$. We write them as $\sigma_{p_1'} = \sigma_{p_{1,1}'} \otimes \sigma_{p_{1,2}'} \otimes \cdots \otimes \sigma_{p_{1,N}'}$ and $\sigma_{p_1'} = \sigma_{p_{2,1}'} \otimes \sigma_{p_{2,2}'} \otimes \cdots \otimes \sigma_{p_{2,N}'}$, respectively. Because each gate acts on Pauli matrices at two sites i and j, let us focus on these two sites: $\sigma_{p_{1,i}'} \otimes \sigma_{p_{1,j}'}$ and $\sigma_{p_{2,i}'} \otimes \sigma_{p_{2,j}'}$. The gate acts on $(\sigma_{p_{1,i}'} \otimes \sigma_{p_{1,j}'}) \otimes (\sigma_{p_{2,i}'} \otimes \sigma_{p_{2,j}'})$ as

$$\left\langle \left\{ G_{i,j}(\sigma_{p_{1,i}'} \otimes \sigma_{p_{1,j}'})(G_{i,j})^\dagger \right\} \otimes \left\{ G_{i,j}(\sigma_{p_{2,i}'} \otimes \sigma_{p_{2,j}'})(G_{i,j})^\dagger \right\} \right\rangle_{G \in \Gamma}$$

$$= \sum_{p''} c_{p',p''} \left(\sigma_{p_{1,i}''} \otimes \sigma_{p_{1,j}''} \right) \otimes \left(\sigma_{p_{2,i}''} \otimes \sigma_{p_{2,j}''} \right). \tag{B.13}$$

Here $G_{i,j} \in \Gamma = \mathrm{U}(4)$ is a unitary gate and the Haar average over the gate set Γ has been taken. The coefficients $c_{p',q'}$ are [34]

$$c_{p',p''} = \begin{cases} 0, & \left(p_{1,i}', p_{1,j}'\right) \neq \left(p_{2,i}', p_{2,j}'\right) \text{ or } \left(p_{1,i}'', p_{1,j}''\right) \neq \left(p_{2,i}'', p_{2,j}''\right), \\ 1, & \left(p_{1,i}', p_{1,j}'\right) = \left(p_{2,i}', p_{2,j}'\right) = \left(p_{1,i}'', p_{1,j}''\right) = \left(p_{2,i}'', p_{2,j}''\right) = (0,0) \\ \frac{1}{15}, & \left(p_{1,i}', p_{1,j}'\right) = \left(p_{2,i}', p_{2,j}'\right) \neq (0,0) \text{ and } \left(p_{1,i}'', p_{1,j}''\right) = \left(p_{2,i}'', p_{2,j}''\right) \neq (0,0). \end{cases} \tag{B.14}$$

In words, unequal strings are mapped to zero in one step. When strings are equal and II they remain unchanged. When strings are equal and not II they are uniformly sprayed among the 15 non II configurations. (When the gate set is a subset of $U(4)$ the first, unequal case, is given by $1 - \Delta$ where Δ is the gap of the gate set. These unequal configurations die away exponentially, like the $k = 1$ discussion in the text. The last coefficient is also modified with no qualitative effect.)

By using this expression, we can estimate the time dependence both for local and 2-local circuits. We first focus on the local circuit.

Note that we can write the sum in eq. (B.12) as

$$\langle U_{1j} U_{j1} U_{1j}^* U_{j1}^* \rangle = 4^{-N} + 2^{-N} \sum_{p' \neq 0} \gamma_t(p', p') + 2^{-N} \sum_{p'_1 \neq p'_2} \gamma_t(p'_1, p'_2). \tag{B.15}$$

We can bound the second and third terms separately. For the U(4) gateset the third term in the sum is set to zero after a single step of the random circuit. One can check that the terms in the second term have the same sign, so we can rewrite the second term as

$$(*2) = 4^{-N} \sum_{p' \neq 0} 2^N \cdot |\gamma_t(p', p')| \tag{B.16}$$

where p' is a sum of all $2^N - 1$ Pauli strings of I's and Z's. For each fixed p', we have a Markov chain that starts at position p' (note that $2^N \cdot |\gamma_0(p', p')| = 1$)and converges to a uniform distribution on $4^N - 1$ strings.

$$\lim_{t \to \infty} 2^N \cdot |\gamma_t(p', p')| = \frac{1}{4^N - 1} \tag{B.17}$$

This implies that $\lim_{t \to \infty} (*2) \approx 4^{-N} \cdot 2^{-N}$, which is exponentially suppressed relative to the Haar value $(= 4^{-N})$. So we need only focus on the early time decay of this quantity. In particular to establish ϵ closeness to Haar for fixed ϵ we do not need to consider the final decay governed by the gap of the full Markov chain.

Because of the large number of configurations, and in particular the 15 possible two Pauli strings not equal to II, the return probability in this chain is small and can be ignored (up to fractional errors of $1/15$). So we can think of the configuration space as a large tree. The slowest decaying N qubit strings are those with lots of II substrings and hence very few Z's. We can ignore the rare cases where the Z's are adjacent. Then the overlap (B.16) decays with a rate proportional to the number of Z's. The estimate becomes analogous to that presented in section 5.1.

As in section 5.1 we will find a time of order $\log \frac{N}{\epsilon}$ for the sum of the slowest terms to be small enough to get ϵ close to the Haar value. At this time each of the slowest chains, those with a few Z's, have only spread to polynomially many other chains. So the gap of the full Markov process should not come into play.

The initial decay rate of the Pauli string with d Z's will be given by $e^{-\tilde{\Delta} d}$ per step of the process where we write the probability for a two qubit string to remain in a given non

II configuration as $e^{-\tilde{\Delta}} = \frac{1}{15}$. Therefore we have

$$(*2) = 4^{-N} \sum_{p' \neq 0} 2^N \cdot |\gamma_t(p', p')| \approx 4^{-N} \sum_{d=1}^{N} \binom{N}{d} e^{-\tilde{\Delta} dt} \tag{B.18}$$

$$= 4^{-N} \left((1 + e^{-\tilde{\Delta} t})^N - 1 \right) \tag{B.19}$$

Putting everything together and assuming N is sufficiently large, we conclude

$$\langle U_{1j} U_{j1} U_{1j}^* U_{j1}^* \rangle \approx 4^{-N} (1 + N e^{-\tilde{\Delta} t}) \qquad (j \neq 1). \tag{B.20}$$

That is the strings with just one Z dominate this decay.

Estimating $\left\langle U_{ii} U_{ii} U_{jj}^* U_{jj}^* \right\rangle$

We consider $i = 1$, without loss of generality. Again in the $0, 1$-basis, $|1\rangle$ is expressed by $|0^N\rangle$. In the same way as the discussion below (5.17), we take the state $|j\rangle$ to be of the form $|1^q 0^{N-q}\rangle$.

The argument is a simple generalization of the one for $\left\langle U_{ii} U_{jj}^* \right\rangle$, presented below (5.17). We start from the initial state on two copies of the system,

$$\hat{O}^{(2)}(0) = \left(|0^N\rangle \langle 1^q 0^{N-q}| \right) \otimes \left(|0^N\rangle \langle 1^q 0^{N-q}| \right). \tag{B.21}$$

The time evolution of $\hat{\rho}$ is defined by

$$\hat{O}^{(2)}(t) = \left(U(t) |0^N\rangle \langle 1^q 0^{N-q}| U^\dagger(t) \right) \otimes \left(U(t) |0^N\rangle \langle 1^q 0^{N-q}| U^\dagger(t) \right). \tag{B.22}$$

Then, by construction,

$$U_{11} U_{jj}^* U_{11} U_{jj}^* = \text{Tr} \left[\hat{O}^{(2)}(t) \hat{O}^{(2)\dagger}(0) \right]. \tag{B.23}$$

We expand $\hat{O}^{(2)}$ by Pauli strings as

$$\hat{O}^{(2)}(t) = 2^{-N} \sum_{p_1, p_2} \gamma_t(p_1, p_2) \sigma_{p_1} \otimes \sigma_{p_2}. \tag{B.24}$$

The coefficients at $t = 0$ are given by

$$\gamma_0(p_1', p_2') = 2^{-N} (+i)^{f(\vec{\alpha}(p_1')) + f(\vec{\alpha}(p_2'))}, \tag{B.25}$$

when p_1' and p_2' are Pauli strings whose first q elements are X or Y and the remaining $N - q$ elements are I or Z. For other p_1 and p_2, $\gamma_0(p_1, p_2) = 0$. We can now write

$$U_{11} U_{jj}^* U_{11} U_{jj}^* = \sum_{p_1', p_2'} \gamma_t(p_1', p_2') \gamma_0^*(p_1', p_2') \tag{B.26}$$

We can split the sum into three parts

$$\langle U_{11} U_{jj}^* U_{11} U_{jj}^* \rangle = 4^{-N} \delta_{1j} + \sum_{p' \neq 0} \gamma_t(p', p') \gamma_0^*(p', p') + \sum_{p_1' \neq p_2'} \gamma_t(p_1', p_2') \gamma_0^*(p_1', p_2'). \tag{B.27}$$

The strings we have chosen have q X or Ys. We denote the number of Z's by d, where d can runs from zero to $n - q$. When $j \neq 1$ or equivalently $(q \geq 1)$ we note that for every string with an X at a given position there is an identical string with a Y at that position. Within the second term, contributions of string with X and Y in the same position will differ by a sign, due to the fact that there are two copies of the system and each one will contribute an i for each copy of Y. So the contributions of all such strings cancel exactly. This implies that when $j \neq 1$ the second term in (B.27) will be set to zero in one step. Again for the U(4) gateset, the third term in the sum is also set to zero after a single step of the random circuit.

In the remaining case $(j = 1)$, we will only have strings of I and Z and the evolution will be analogous to the previous section. In fact we not need to repeat this estimate because there is only one such term, $i = j = 1$, which makes an exponentially small contribution to the final value of $u(2, t)$.

Summing up all the contributions

We can rewrite the sum as (recall $L = 2^N$ and assume $t > 1$)

$$
\begin{aligned}
u(2, t) = \left\langle |\mathrm{Tr} U^2(t)|^2 \right\rangle &= L \sum_{j=1}^{L} \left\langle U_{11} U_{11} U_{jj}^* U_{jj}^* \right\rangle + 2L \sum_{j=2}^{L} \left\langle U_{1j} U_{j1} U_{1j}^* U_{j1}^* \right\rangle. \\
&\lesssim \frac{1}{L} \left(1 + N \cdot e^{-\tilde{\Delta} t} \right) + 2 \left(1 + N \cdot e^{-\tilde{\Delta} t} \right) \\
&\simeq 2 + 2N \cdot e^{-\tilde{\Delta} t}.
\end{aligned}
\tag{B.28}
$$

The factor of 2 comes from the i, j symmetry in the sum. Note that this result becomes ϵ close to the Haar value in times $\frac{1}{\tilde{\Delta}} \log(N/\epsilon)$.

For the 2-local system the above bound goes through without modification. There is one subtlety however. The results of Harrow-Low [34] show that full Markov chain convergence occurs in order $\log N$ time, indicating that the entire Pauli string configuration space is covered in that time. So the tree approximation may no longer be valid. But by this time our quantity is exponentially close to the Haar value (B.17) and so there should be no increase in the estimate.

Alternate strategy

Another method for calculation the $u(2, t)$ is to use the rewrite presented in equation (6.2),

$$
u(2, t) = \left\langle \mathrm{Tr} U_t^2 \cdot \mathrm{Tr}(U_t^\dagger)^2 \right\rangle = 1 + \sum_{j=2}^{L^2} \frac{1}{L} \left\langle \mathrm{Tr} \left(U_t^2 P_j (U_t^\dagger)^2 P_j \right) \right\rangle
\tag{B.29}
$$

where the sum is over all the non-identity Pauli strings. In the second term of the above formula, we translate the average over product of circuits to an average on two copies of the system followed by a swap operation $\mathcal{F}$.

$$
\left\langle \mathrm{Tr} \left(U_t^2 P_j (U_t^\dagger)^2 P_j \right) \right\rangle = \left\langle \mathrm{Tr} \left[\left((U_t P_j U_t^\dagger) \otimes (U_t^\dagger P_j U_t) \right) \mathcal{F} \right] \right\rangle
\tag{B.30}
$$

We have not analyzed this systematically, but observe that averaging can be done independently for each gate and so a new, perhaps more complicated, chain will result. The question then is to understand how coefficients of the Pauli strings decay in this chain for each application of a gate. If we ignore backtracking and approximate the chain evolution by a tree, as above, the question reduces to computing the decay rate for single gate from U(4).

The average of the U(4) gateset on a non-identity Pauli string AB $(A, B = I, X, Y, Z)$ dotted with the same string can be computed and is,

$$\int_{U(4)} dU \ \text{Tr}\Big(U^2 AB (U^\dagger)^2 AB\Big) = \frac{1}{15}. \tag{B.31}$$

This means that in the local and 2-local circuit the strings with d non-identity elements will decay like $e^{-d\tilde{\Delta}t}$, where again $e^{-\tilde{\Delta}} = 1/15$. This indicates that the decay is dominated by the slowest decaying terms which have $d = 1$. The total number of such terms is $3N$, giving the $\log N$ result described above.

There is a small subtlety here. There are $d = 2$ terms with the two Paulis adjacent to each other. These decay like a $d = 1$ term. However there are only $O(N)$ terms of this kind so this will just produce an addition to the numerical factor in front of N. This is a random circuit version of the factorization issue discussed in section 6. Therefore, the Haar time scaling, $\sim \log \frac{N}{\epsilon}$, is unaffected.

C Random band matrix

C.1 Hamiltonian evolution

We consider an $L \times L$ 'band matrix' M with width w. Rather than introducing a sharp cutoff, we introduce a Gaussian suppression factor,

$$M_{ij} = M_{ij}^{(0)} \cdot \frac{e^{-\frac{(i-j)^2}{2w^2}}}{\sqrt{w}}, \tag{C.1}$$

where $M_{ij}^{(0)}$ is the standard GUE ensemble.

By definition, banded random matrix becomes usual RMT at $w \gtrsim L$. It is not chaotic when the band is too narrow, namely at $w \lesssim \sqrt{L}$. In the following we consider large $w \geq \sqrt{L}$.

By using Y with appropriate value of α, we can see the emergence of the ramp. As shown in figures 26 and 27, if we plot $(w/L)^2 |Y^2|$ against $(w/L)^2 \cdot t$, we can see a convergence to a smooth curve in a wide range, across the onset of the ramp. Therefore, regardless of the detail of the procedure to define t_{ramp}, it must scale as $t_{\text{ramp}} \sim (L/w)^2$. This scaling was derived in [25], where the authors showed that the energy scale at which sine kernel description of the two-point function breaks down is the same as the Thouless energy, which is the inverse of the time-scale for a local perturbation to diffuse through the entire system, namely t_{diff}. Note that the earlier growth of $(w/L)^2 |Y^2|$ is slower, $\sim t^{1/2}$. Our numerical plots are consistent with this scaling, as can be seen in figure 27.

It would be instructive to mention the difference between the banded matrices and local or k-local Hamiltonians. The band matrix mimics a Hamiltonian acting on L-state
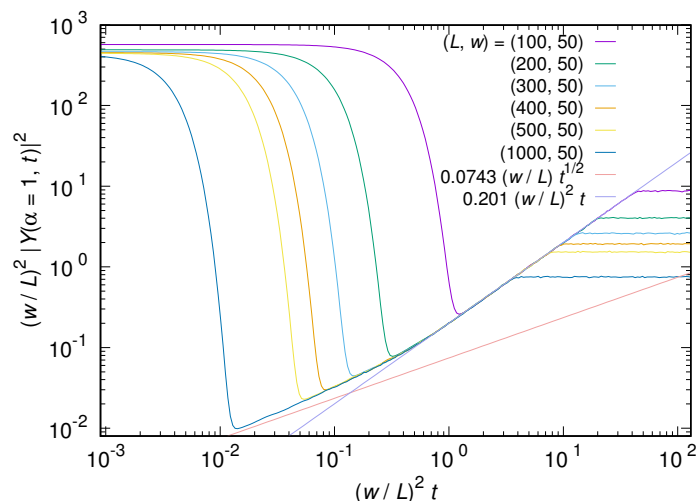
Figure 26. $(w/L)^2|Y(\alpha = 1, t)|^2$ vs $(w/L)^2 \cdot t$, $L = 100, 200, 300, 400, 500, 1000$, $w = 50$, $\alpha = 1$. 10000 samples have been taken.
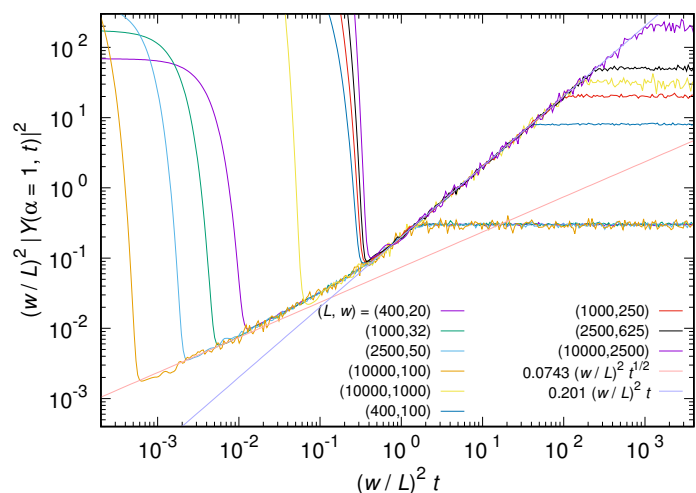

Figure 27. The value of $(w/L)^2|Y(\alpha = 1, t)|^2$ plotted against $(w/L)^2 \cdot t$ for various (L, w) satisfying $\sqrt{L} \leq w \leq L/4$, $\alpha = 1$. The number of samples with each value of w is $10^6/L = 2500, 1000, 400, 100$ for $L = 400, 1000, 2500, 10000$ cases. For $w \ll L/4$ the initial growth indicates that $|Y(\alpha = 1, \tau)|^2 \sim (L/w)\sqrt{t}$, while for all (L, w) plotted, the growth for larger t before the plateau agrees with $|Y(\alpha = 1, t)|^2 \sim t$.

Hilbert space, which connects w nearby states. This looks 'local' if we identify the states with a single particle hopping on L lattice sites, but it is useful to note that this is rather different from a many-body local Hamiltonian, say a local spin system. In the latter, in a natural local basis, the Hamiltonian is sparse but the nonzero entries are not aligned near the diagonal as in the banded matrix.

In local Hamiltonians with N spins, each row and column has $L = 2^N$ entries, among which there are $O(N)$ nonzero elements. This is much sparser than a banded matrix with $w \sim \sqrt{L}$, but it can already be chaotic. The time scales are also very different; the ramp
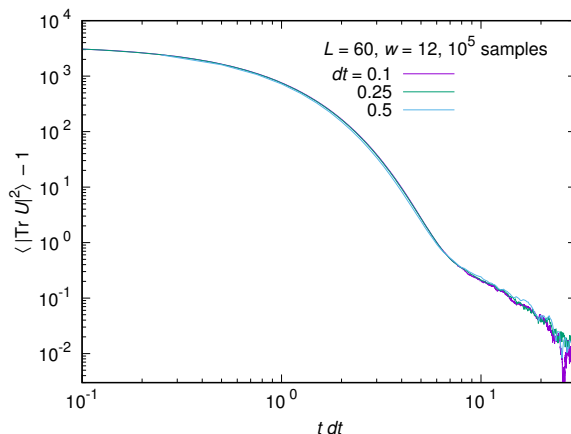
Figure 28. Plot of $\langle |\mathrm{Tr}U|^2 \rangle - 1$ against $t\,dt$ for $dt = 0.1, 0.25, 0.5$ for Brownian circuits of random band matrices with $L = 60$ and $w = 12$. 10^5 samples have been taken for each choice of dt.

time for the local Hamiltonian, $t_{\mathrm{ramp}} \sim N^2 \sim (\log L)^2$ (see section 4.2), is much shorter than $(L/w)^2$ when $w \sim \sqrt{L}$.

C.2 Brownian circuit

The Brownian circuit of random band matrices can describe diffusion, because the number of particles — which is one — is conserved. For this reason, as we will see, we can confirm $t_{\mathrm{Haar}} \sim (L/w)^2 \sim t_{\mathrm{diff}}$.

The numerical procedure is the same: we calculate $U(t) = \prod_{k=1}^n e^{-iH_k dt}$, where $t = n \cdot dt$ and H_k are $L \times L$ random band matrices with a width w. From this we calculate the $\langle |TrU|^2 \rangle$.

First let us see that the specific choice of dt is not important. When dt is sufficiently small, if we take $tdt = n(dt)^2$ to be the horizontal axis, the dt-dependence is gone; see figure 28. This scaling can be understood if the time evolution is described as a random walk with step-size wdt. After n steps, the typical distance from the starting point should be $\sqrt{n}wdt$, if the random walk picture is true. Then for fixed L and w the dt-dependence should disappear when the horizontal axis is $n(dt)^2$. Below we fix dt to be 0.5.

In figure 29 we have plotted $\langle |\mathrm{Tr}U|^2 \rangle - 1$ against $(t(w/L)^2)/(L^2 - 1)$ for fixed $w = 25$ and various L, and for a fixed L/w. In both cases we can see an exponential decay. The exponent is a function of L/w, as we can see from the right panel of figure 29.

As we can see from figure 30, we can show the exponential decay at late time as well, $\langle |\mathrm{Tr}U|^2 \rangle - 1 \sim e^{-2t(w/L)^2}$. Note that the exponent is different from the early time. If we define t_{Haar} as the time $\langle |\mathrm{Tr}U|^2 \rangle - 1$ reaches to a certain value, say 0.1, then this scaling leads to $t_{\mathrm{Haar}} \sim (L/w)^2$.

It is possible to understand this scaling from the random walk picture. When the average distance from the starting point, which is $\sqrt{t}w$, is of order L, it is natural to expect that the transfer matrix U is almost a random unitary. Therefore $\sqrt{t_{\mathrm{Haar}}}w \sim L$, or equivalently, $t_{\mathrm{Haar}} \sim (L/w)^2$.
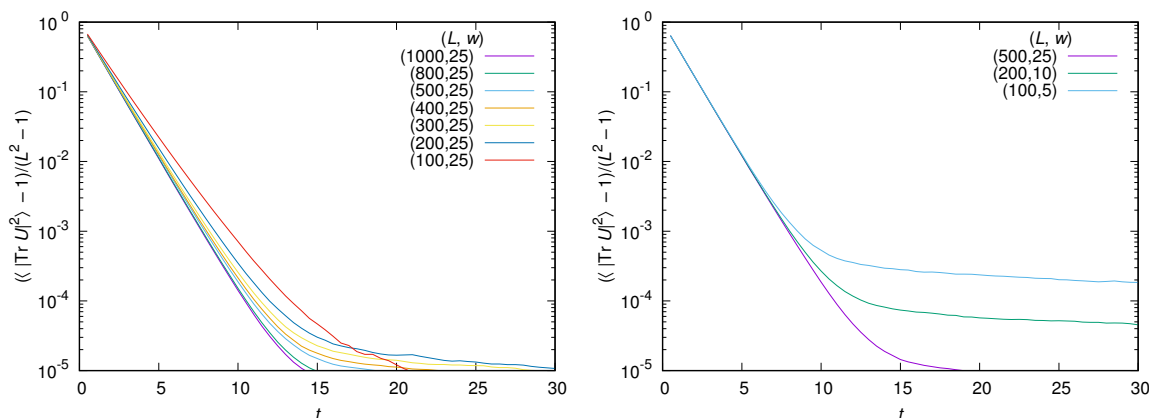
Figure 29. Plot of $(\langle |\mathrm{Tr}U|^2 \rangle - 1)/(L^2 - 1)$ against t for $w = 25$ and various values of L (left) and for $L/w = 20$ with $L = 500, 200, 100$ (right). $dt = 0.5$ is fixed.
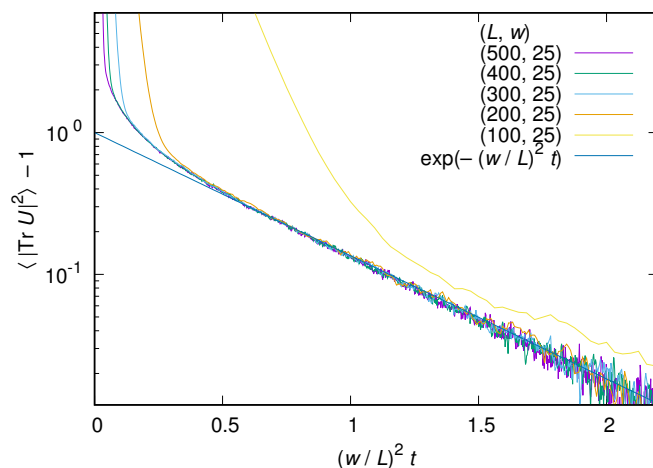
Figure 30. Log plot of $\langle |\mathrm{Tr}U|^2 \rangle - 1$ against $t(w/L)^2$ for $w = 25$ and various values of L. $dt = 0.5$ is fixed. 10^5 samples have been used for $L = 300, 200, 100$.

References

[1] F. Haake, *Quantum Signatures of Chaos*, Springer-Verlag New York, Inc. (2006).

[2] M. Serbyn and J.E. Moore, *Spectral statistics across the many-body localization transition*, *Phys. Rev.* **B 93** (2016) 041424 [`arXiv:1508.07293`].

[3] A. Altland and D. Bagrets, *Quantum ergodicity in the SYK model*, *Nucl. Phys.* **B 930** (2018) 45 [`arXiv:1712.05073`] [INSPIRE].

[4] A. Chan, A. De Luca and J.T. Chalker, *Spectral statistics in spatially extended chaotic quantum many-body systems*, `arXiv:1803.03841` [INSPIRE].

[5] P. Kos, M. Ljubotina and T. Prosen, *Many-body quantum chaos: Analytic connection to random matrix theory*, *Phys. Rev.* **X 8** (2018) 021062 [arXiv:1712.02665] [INSPIRE].

[6] 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.* **B 95** (2017) 115150 [arXiv:1602.06964] [INSPIRE].

[7] A.M. García-García and J.J.M. Verbaarschot, *Spectral and thermodynamic properties of the Sachdev-Ye-Kitaev model*, *Phys. Rev.* **D 94** (2016) 126010 [arXiv:1610.03816] [INSPIRE].

[8] J.S. Cotler et al., *Black Holes and Random Matrices*, *JHEP* **05** (2017) 118 [arXiv:1611.04650] [INSPIRE].

[9] S. Sachdev and J. Ye, *Gapless spin fluid ground state in a random, quantum Heisenberg magnet*, *Phys. Rev. Lett.* **70** (1993) 3339 [cond-mat/9212030] [INSPIRE].

[10] A. Kitaev, *A Simple Model of Quantum Holography*, talks at KITP (2015) [http://online.kitp.ucsb.edu/online/entangled15/kitaev/] [http://online.kitp.ucsb.edu/online/entangled15/kitaev2/].

[11] A. Kitaev and S.J. Suh, *The soft mode in the Sachdev-Ye-Kitaev model and its gravity dual*, *JHEP* **05** (2018) 183 [arXiv:1711.08467] [INSPIRE].

[12] J. Maldacena and D. Stanford, *Remarks on the Sachdev-Ye-Kitaev model*, *Phys. Rev.* **D 94** (2016) 106002 [arXiv:1604.07818] [INSPIRE].

[13] A.I. Larkin and Y.N. Ovchinnikov, *Quasiclassical method in the theory of superconductivity*, *Sov. Phys. JETP* **28** (1969) 1200 [http://www.jetp.ac.ru/cgi-bin/dn/e_028_06_1200.pdf].

[14] P. Hayden and J. Preskill, *Black holes as mirrors: Quantum information in random subsystems*, *JHEP* **09** (2007) 120 [arXiv:0708.4025] [INSPIRE].

[15] Y. Sekino and L. Susskind, *Fast Scramblers*, *JHEP* **10** (2008) 065 [arXiv:0808.2096] [INSPIRE].

[16] A. Almheiri, D. Marolf, J. Polchinski, D. Stanford and J. Sully, *An Apologia for Firewalls*, *JHEP* **09** (2013) 018 [arXiv:1304.6483] [INSPIRE].

[17] A. Kitaev, *Hidden Correlations in the Hawking Radiation and Thermal Noise*, talk given at *Fundamental Physics Prize Symposium*, Nov. 10, 2014, and Stanford SITP seminars Nov. 11, 2014 and Dec. 18, 2014 [https://www.youtube.com/watch?v=OQ9qN8j7EZI].

[18] N. Lashkari, D. Stanford, M. Hastings, T. Osborne and P. Hayden, *Towards the Fast Scrambling Conjecture*, *JHEP* **04** (2013) 022 [arXiv:1111.6580] [INSPIRE].

[19] S.H. Shenker and D. Stanford, *Black holes and the butterfly effect*, *JHEP* **03** (2014) 067 [arXiv:1306.0622] [INSPIRE].

[20] S.H. Shenker and D. Stanford, *Multiple Shocks*, *JHEP* **12** (2014) 046 [arXiv:1312.3296] [INSPIRE].

[21] D.A. Roberts, D. Stanford and L. Susskind, *Localized shocks*, *JHEP* **03** (2015) 051 [arXiv:1409.8180] [INSPIRE].

[22] S.H. Shenker and D. Stanford, *Stringy effects in scrambling*, *JHEP* **05** (2015) 132 [arXiv:1412.6087] [INSPIRE].

[23] J. Maldacena, S.H. Shenker and D. Stanford, *A bound on chaos*, *JHEP* **08** (2016) 106 [arXiv:1503.01409] [INSPIRE].

[24] B. Altshuler and B. Shklovskii, *Repulsion of energy levels and conductivity of small metal samples*, *Sov. Phys. JETP* **64** (1986) 127 [http://www-thphys.physics.ox.ac.uk/talks/CMTjournalclub/sources/AltshulerShklovskii.pdf].

[25] L. Erdős and A. Knowles, *The Altshuler-Shklovskii Formulas for Random Band Matrices I: the Unimodular Case*, *Commun. Math. Phys.* **333** (2015) 1365.

[26] A.M. García-García, Y. Jia and J.J.M. Verbaarschot, *Universality and Thouless energy in the supersymmetric Sachdev-Ye-Kitaev Model*, *Phys. Rev.* **D 97** (2018) 106003 [arXiv:1801.01071] [INSPIRE].

[27] T. Banks, L. Susskind and M.E. Peskin, *Difficulties for the Evolution of Pure States Into Mixed States*, *Nucl. Phys. B* **244** (1984) 125.

[28] J. Emerson, E. Livine and S. Lloyd, *Convergence conditions for random quantum circuits*, *Phys. Rev.* **A 72** (2005) 060302 [quant-ph/0503210].

[29] R. Oliveira, O.C.O. Dahlsten and M.B. Plenio, *Efficient Generation of Generic Entanglement*, *Phys. Rev. Lett.* **98** (2007) 130502 [quant-ph/0605126].

[30] J. Emerson, Y.S. Weinstein, M. Saraceno, S. Lloyd and D.G. Cory, *Pseudo-random unitary operators for quantum information processing*, *Science* **302** (2003) 2098.

[31] D. Gross, K. Audenaert and J. Eisert, *Evenly distributed unitaries: On the structure of unitary designs*, *J. Math. Phys.* **48** (2007) 052104 [quant-ph/0611002].

[32] L. Arnaud and D. Braun, *Efficiency of producing random unitary matrices with quantum circuits*, *Phys. Rev.* **A 78** (2008) 062329 [arXiv:0807.0775].

[33] A. Harrow and R. Low. *Efficient quantum tensor product expanders and k-designs*, *Lect. Notes Comput. Sci.* **5687** (2009) 548 [arXiv:0811.2597].

[34] A.W. Harrow and R.A. Low, *Random quantum circuits are approximate 2-designs*, *Commun. Math. Phys.* **291** (2009) 257 [arXiv:0802.1919].

[35] W.G. Brown and L. Viola, *Convergence rates for arbitrary statistical moments of random quantum circuits*, *Phys. Rev. Lett.* **104** (2010) 250501 [arXiv:0910.0913].

[36] W. Brown and O. Fawzi, *Scrambling speed of random quantum circuits*, arXiv:1210.6644 [INSPIRE].

[37] F. Brandao, A. Harrow and M. Horodecki, *Local random quantum circuits are approximate polynomial-designs*, *Commun. Math. Phys.* **346** (2016) 397 [arXiv:1208.0692].

[38] E. Onorati, O. Buerschaper, M. Kliesch, W. Brown, A.H. Werner and J. Eisert, *Mixing properties of stochastic quantum Hamiltonians*, *Commun. Math. Phys.* **355** (2017) 905 [arXiv:1606.01914] [INSPIRE].

[39] Y. Nakata, C. Hirche, M. Koashi and A. Winter, *Efficient Quantum Pseudorandomness with Nearly Time-Independent Hamiltonian Dynamics*, *Phys. Rev.* **X 7** (2017) 021006 [arXiv:1609.07021] [INSPIRE].

[40] L. Banchi, D. Burgarth and M.J. Kastoryano, *Driven Quantum Dynamics: Will It Blend?*, *Phys. Rev.* **X 7** (2017) 041015 [arXiv:1704.03041] [INSPIRE].

[41] A. Nahum, S. Vijay and J. Haah, *Operator Spreading in Random Unitary Circuits*, *Phys. Rev.* **X 8** (2018) 021014 [arXiv:1705.08975] [INSPIRE].

[42] C. von Keyserlingk, T. Rakovszky, F. Pollmann and S. Sondhi, *Operator hydrodynamics, OTOCs and entanglement growth in systems without conservation laws*, *Phys. Rev.* **X 8** (2018) 021013 [arXiv:1705.08910] [InSPIRE].

[43] V. Khemani, A. Vishwanath and D.A. Huse, *Operator spreading and the emergence of dissipation in unitary dynamics with conservation laws*, arXiv:1710.09835 [InSPIRE].

[44] J. Preskill, *Quantum computing and the entanglement frontier*, arXiv:1203.5813 [InSPIRE].

[45] C. Neill et al., *A blueprint for demonstrating quantum supremacy with superconducting qubits*, *Science* **360** (2018) 195 [arXiv:1709.06678].

[46] A.W. Harrow and A. Montanaro, *Quantum computational supremacy*, *Nature* **549** (2017) 203.

[47] A. Bouland, B. Fefferman, C. Nirkhe, U. Vazirani *Quantum Supremacy and the Complexity of Random Circuit Sampling*, arXiv:1803.04402.

[48] T. Rakovszky, F. Pollmann and C.W. von Keyserlingk, *Diffusive hydrodynamics of out-of-time-ordered correlators with charge conservation*, arXiv:1710.09827 [InSPIRE].

[49] D.A. Roberts and B. Yoshida, *Chaos and complexity by design*, *JHEP* **04** (2017) 121 [arXiv:1610.04903] [InSPIRE].

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

[51] F.J. Dyson, *Statistical theory of the energy levels of complex systems. I*, *J. Math. Phys.* **3** (1962) 140 [InSPIRE].

[52] M.L. Mehta, *Random matrices*, vol. 142, Academic Press (2004).

[53] T. Guhr, A. Müller-Groeling and H.A. Weidenmuller, *Random matrix theories in quantum physics: Common concepts*, *Phys. Rept.* **299** (1998) 189 [cond-mat/9707301] [InSPIRE].

[54] J. Flores, M. Horoi, M. Müller and T.H. Seligman, *Spectral statistics of the two-body random ensemble revisited*, *Phys. Rev.* **E 63** (2001) 026204 [cond-mat/0006144] [InSPIRE].

[55] A. Altland and D. Bagrets, *Quantum ergodicity in the SYK model*, *Nucl. Phys.* **B 930** (2018) 45 [arXiv:1712.05073] [InSPIRE].

[56] L. Erdős and D. Schröder, *Phase transition in the density of states of quantum spin glasses*, *Math. Phys. Anal. Geom.* **17** (2014) 441 [arXiv:1407.1552].

[57] E. Brézin and S. Hikami, *Spectral form factor in a random matrix theory*, article *Phys. Rev.* **E 55** (1997) 4067 [cond-mat/9608116].

[58] E. Brézin and S. Hikami, *Extension of level-spacing universality*, *Phys. Rev.* **E 56** (1997) 264 [InSPIRE].

[59] K. Papadodimas and S. Raju, *Local Operators in the Eternal Black Hole*, *Phys. Rev. Lett.* **115** (2015) 211601 [arXiv:1502.06692] [InSPIRE].

[60] A. Gaikwad and R. Sinha, *Spectral Form Factor in Non-Gaussian Random Matrix Theories*, arXiv:1706.07439 [InSPIRE].

[61] D. Bagrets, A. Altland and A. Kamenev, *Sachdev-Ye-Kitaev model as Liouville quantum mechanics*, *Nucl. Phys.* **B 911** (2016) 191 [arXiv:1607.00694] [InSPIRE].

[62] D. Stanford and E. Witten, *Fermionic Localization of the Schwarzian Theory*, *JHEP* **10** (2017) 008 [arXiv:1703.04612] [InSPIRE].

[63] F. Haake, H.-J. Sommers and J. Weber, *Fluctuations and ergodicity of the form factor of quantum propagators and random unitary matrices*, *J. Phys.* **A 32** (1999) 6903.

[64] P. Diaconis and S.N. Evans, *Linear Functionals of Eigenvalues of Random Matrices*, *Trans. Am. Math. Soc.* **353** (2001) 2615.

[65] J.P. Keating, N. Linden and H.J. Wells, *Random matrices and quantum spin chains*, *Markov Process. Related Fields* **21** (2015) 537 [arXiv:1403.1114].

[66] J.P. Keating, N. Linden and H.J. Wells, *Spectra and Eigenstates of Spin Chain Hamiltonians*, *Commun. Math. Phys.* **338** (2015) 81 [arXiv:1403.1121].

[67] A. Pal and D.A. Huse, *The many-body localization transition*, arXiv:1003.2613.

[68] D.J. Luitz, N. Laflorencie and F. Alet, *Many-body localization edge in the random-field Heisenberg chain*, *Phys. Rev.* **B 91** (2015) 081103 [arXiv:1411.0660].

[69] K. Agarwal, E. Altman, E. Demler, S. Gopalakrishnan, D.A. Huse and M. Knap, *Rare-region effects and dynamics near the many-body localization transition*, *Annalen Phys.* **529** (2017) 1600326 [arXiv:1611.00770].

[70] P. Saad, S.H. Shenker and D. Stanford, *A semiclassical ramp in SYK and in gravity*, in progress.

[71] L.L. Ng, *Heisenberg Model, Bethe Ansatz and Random Walks*, Senior Honors Thesis, Harvard University (1996) [https://services.math.duke.edu/ ng/math/papers/senior-thesis.pdf].

[72] M. Karabach, G. Müller, H. Gould and J. Tobochnik, *Introduction to the Bethe Ansatz I*, *Comput. Phys.* **11** (1997) 36 [cond-mat/9809162].

[73] J.R.G. Mendonca, *Exact eigenspectrum of the symmetric simple exclusion process on the complete, complete bipartite and related graphs*, *J. Phys.* **A 46** (2013) 295001 [arXiv:1207.4106].

[74] E.H. Lieb and D.W. Robinson, *The finite group velocity of quantum spin systems*, *Commun. Math. Phys.* **28** (1972) 251 [INSPIRE].

[75] D.A. Roberts, D. Stanford and A. Streicher, *Operator growth in the SYK model*, *JHEP* **06** (2018) 122 [arXiv:1802.02633] [INSPIRE].

[76] J.T. Chalker, I.V. Lerner and R.A. Smith, *Random Walks through the Ensemble: Linking Spectral Statistics with Wave-Function Correlations in Disordered Metals*, *Phys. Rev. Lett.* **77** (1996) 554 [INSPIRE].