Quantum Chaos and Circuit Parameter Optimization

Joonho Kim,[1] Yaron Oz,[2,3] and Dario Rosa[4]

[1] *School of Natural Sciences, Institute for Advanced Study, Princeton, NJ 08540, USA.*
[2] *Simons Center for Geometry and Physics, SUNY, Stony Brook, NY 11794, USA.*
[3] *Raymond and Beverly Sackler School of Physics and Astronomy, Tel-Aviv University, Tel-Aviv 69978, Israel.*
[4] *Center for Theoretical Physics of Complex Systems, Institute for Basic Science (IBS), Daejeon - 34126, Korea*

(Dated: January 6, 2022)

We explore quantum chaos diagnostics of variational circuit states at random parameters and study their correlation with the circuit expressibility and the optimization of control parameters. By measuring the operator spreading coefficient and the eigenvalue spectrum of the modular Hamiltonian of the reduced density matrix, we identify the universal structure of random matrix models in high-depth circuit states. We construct different layer unitaries corresponding to the GOE and GUE distributions and quantify their VQA performance. Our study also highlights a potential tension between the OTOC and BGS-type diagnostics of quantum chaos.

I. INTRODUCTION

The random circuit model provides a framework of hybrid quantum/classical algorithms for solving optimization and learning tasks, formulated as a search for the ground state of k-local Hamiltonians [1–3]. Generic quantum gates create an entanglement between qubits. Entanglement is a valuable resource for achieving quantum advantage, but it also becomes a hurdle for successful optimization of circuit control parameters at the same time, specifically when random circuit states are much more highly entangled than the ground state of the Hamiltonian encoding the task [4–8]. Quantum information in such highly entangled states is scrambled, and a successful adjustment of circuit parameters via local gradient search typically requires over-parametrization [9–12].

Quantum chaos [13] is correlated with information scrambling [14–18] and is in general a feature of interacting dynamical quantum systems [19]. One expects deep random circuit states to be generically chaotic. While the entanglement properties of a state can be quantified by entanglement entropies constructed from the eigenvalues of its reduced density matrix, the quantum chaotic structure of the state is largely diagnosed by measures that depend on the level spacing of the reduced density matrix. The aim of this work is to investigate the chaotic properties of random circuit states, with focus on the relationship between quantum chaos, circuit expressibility, and optimization performance.

A common diagnostic of quantum chaos that characterizes information scrambling is the operator spreading, which can be quantified by the 4-point out-of-time-order correlation function (OTOC) [16, 17, 20]. One considers an operator at time $t = 0$, denoted by $\mathcal{O}(0)$, that acts on a small number of qubits and evolves it to:

$$\mathcal{O}(t) = U(t)^\dagger \mathcal{O}(0) U(t) \ , \tag{1}$$

supported on a larger number of qubits at time step t. This growth is typically ballistic with a characteristic velocity, known as the butterfly velocity, reminiscent of the spread of classical chaotic trajectories. Considering a discrete time evolution driven by the random circuit unitary, this velocity depends on the circuit architecture, i.e., arrangement and type of quantum gate unitaries. However, the operator growth is associated with spectral values of the circuit reduced density matrix themselves, much like the entanglement entropy, but not their level spacings. Let us denote the density matrix of the circuit state by ρ_c and divide n qubits of the quantum register into two subsets A and B of the equal size, $n_A = n_B = \frac{n}{2}$. The modular Hamiltonian $H(\rho_A)$ of the reduced density matrix $\rho_A = \mathrm{Tr}_B \rho_c$ can be written as

$$\rho_A \equiv \frac{e^{-H(\rho_A)}}{Z_A} \ , \tag{2}$$

where $Z_A = \mathrm{Tr}_A \, e^{-H(\rho_A)}$ is the partition function of the modular Hamiltonian. It is indeed the eigenspectrum of $H(\rho_A)$ that encapsulates the entanglement and operator spreading properties of the quantum circuit.

There are by now accumulated pieces of evidence that chaotic properties of Hamiltonian systems reveal themselves in the level spacing distribution of the Hamiltonian energy spectrum [21]. This understanding can be extended to the eigenspectrum of the modular Hamiltonian, diagnosing the chaotic nature of a *quantum state* from its level spacing distribution [22]. We will explore various quantum chaos diagnostics as a function of the circuit depth and show that deep circuit states exhibit the characteristics of random matrix models. Specifically, the level spacing distribution of the modular Hamiltonian, r-statistics, and spectral form factor will manifest the universal structure of the Gaussian Orthogonal Ensemble (GOE) or Gaussian Unitary Ensemble (GUE), depending on types of quantum gates introduced in the random entangling circuit.

The paper is organized as follows. In Section II, we will describe the architecture of layered random circuits used for numerical simulation and briefly review the relationship between the number of circuit layers and VQA performance. Section III will explore the connection between

the operator spreading, a typical diagnostic of quantum chaos, and the optimization efficiency of control variables. We will then study in Section IV the level spacing distributions of the modular Hamiltonians, r-statistics, and their spectral form factors at different circuit depths, showing that all diagnostics match those of random matrix ensembles in the high-depth regime. Section V will conclude with discussion.

II. VQA PERFORMANCE

We begin with specifying the circuit architecture used in this paper and briefly review the relation between the entanglement made by random circuit unitaries and the optimization efficiency of variational quantum algorithms (VQAs) [4–8, 11].

A. Circuit Architecture

Figure 1 illustrates the variational circuit architecture assumed throughout our discussion. The n quantum registers are arranged periodically, $i \simeq i+n$, and acted upon by a chain of two-qubit unitaries. The unitaries are made of single-qubit gates $R(\theta)$ acting on all n distinct qubits, followed by two-qubit entanglers:

$$CZ = \mathrm{diag}\,(+1,+1,+1,-1) \;\text{ or} \tag{3}$$

$$CP = \mathrm{diag}\,(+1,+1,+1,+i) \tag{4}$$

that operate on all adjacent pairs of qubits. Every layer swaps the roles of odd/even qubits, alternating between controlling/controlled and controlled/controlling pairs.

In our numerical simulation, we will consider only two types of single-qubit gates. They are Pauli rotation gates along the y-axis

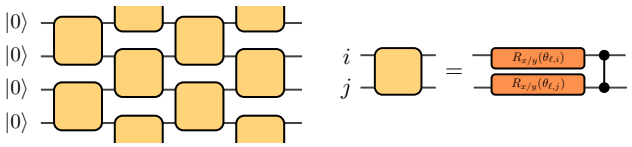
$$R_y(\theta) = \exp(i\sigma_y\theta) = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} , \tag{5}$$

which are real and orthogonal, and those along the x-axis

$$R_x(\theta) = \exp(i\sigma_x\theta) = \begin{pmatrix} \cos\theta & i\sin\theta \\ -i\sin\theta & \cos\theta \end{pmatrix} , \tag{6}$$

which are complex-valued unitary matrices. All the rotation angles are randomly chosen from the uniform distribution $\mathcal{U}(0, 2\pi)$ at circuit initialization. We will use the symbol $\theta_{\ell,i}$ to denote a specific angle that rotates the i'th qubit at the ℓ'th layer, for $1 \le i \le n$ and $1 \le \ell \le L$.

We will consider four different types of layered circuits for numerical experiment. If the single-qubit rotations are all along the x or y axis, followed by the CZ entangler actions, the corresponding circuits will be called $R_x + \mathrm{CZ}$ or $R_y + \mathrm{CZ}$. Two additional variants of circuit structures, dubbed as R_x+CZ+R_y+CZ and R_y+CP, will be examined, where one-qubit gates at odd/even layers alternate between R_x/R_y and where all CZ's are replaced with CP's, respectively. In general, different gate choices will lead to different entanglement and chaos properties.



(a) The circuit architecture (b) The quantum gates

FIG. 1: The variational quantum circuit used in this paper. The n qubits are arranged periodically, initialized in the state $|0\rangle^{\otimes n}$ and acted upon by a chain of the two-qubit unitary gates.

B. Optimization and Expressibility

Random circuit states with a large number L of layers are typically highly entangled. Circuit expressibility, i.e., being able to represent generic states in the Hilbert space, can be achieved for sufficiently deep circuits. However, as quantum typicality makes the energy landscape of VQA Hamiltonians flattened [11, 23], the circuit parameter optimization via local gradient search becomes more difficult with highly entangled circuits [4–8]. A known remedy for the flattened energy landscape is overparametrization of variational ansatz [9, 10, 12, 24], developing multiple steep directions that lead to the robust success of the gradient descent method [11]. This comes with a classical computational cost for storing and manipulating variables.

If a k-local Hamiltonian encodes the task to be solved, so that the corresponding ground state exhibits the area-law entanglement scaling, VQA will perform better with avoiding the region of quantum typicality [4, 5], including the saturation of bipartite entanglement entropy near the maximum value [6, 8]. A canonical example showing this relation is VQA with the one-dimensional Ising Hamiltonian made of the nearest-neighbor spin interaction coupled to a transverse magnetic field [25]. Specifically, we will use two similar Hamiltonians that differ only in the direction of the external magnetic field:

$$H = \sum_{i=1}^{n} \sigma_{z,i}\sigma_{z,i+1} + g\sum_{i=1}^{n} \sigma_{x,i} \quad \text{or} \quad g\sum_{i=1}^{n} \sigma_{y,i} \;, \tag{7}$$

and attempt to reach their ground states at $g = 1$ by optimizing the circuit parameters at different L.

The search for optimal parameters that minimize the energy function will be conducted locally via the Adam optimization algorithm [26]. It is a variant of the plain gradient descent that shows faster convergence in many circumstances, adjusting the step size at each iteration based on the moving average of gradients. We will choose its hyperparameters to be $(\alpha, \beta_1, \beta_2) = (0.05, 0.9, 0.999)$ in all numerical experiments. Per each run, we will allow enough time for convergence towards the ground state by waiting for 5000 steps of the parameter update.

From the collection of 10 independent, repeated runs for each architecture and depth, the overall trend stands

out: A high level of expressibility, measured through the average Renyi entanglement entropy at random parameters saturating around its maximum possible value, has an adverse effect in reaching the ground state of the VQA Hamiltonian. See Figures 2 and 3 for the outputs under four different choices of quantum gates. The orange/blue curves in their left panel display the energy gap from the ground state before/after the circuit parameter optimization as a function of L. Likewise, the orange/blue curves in the right panels represent the Renyi-2 entropy of the reduced density matrix

$$\mathcal{R}_A^2 = -\log \mathrm{tr}_A\, \rho_A^2 \tag{8}$$

obtained through partially tracing out $n/2$ qubits, before/after running the VQA optimization.

It is notable that distinct gate choices lead to different entanglement growth and saturation values. For a particular circuit architecture dubbed as $R_x+\mathrm{CZ}+R_y+\mathrm{CZ}$, where one-qubit gate alternates between R_x and R_y at each layer, the entanglement curve converges to half the saturated value of other circuits. For $R_y+\mathrm{CP}$ model that substitutes CZ entanglers with CP gates, the entanglement growth becomes considerably slower. They result in widening the depth window until approaching the maximum level of entanglement, in which the circuit parameter optimization can likely succeed. Also interestingly, we observe that $R_x+\mathrm{CZ}/R_y+\mathrm{CZ}$ models fail to reach the ground state of the Ising Hamiltonian coupled to the external field along the x/y-axis, respectively.

III. OPERATOR SPREADING

Operator spreading serves as a diagnostic of the chaotic dynamics and information scrambling. It has been extensively studied in the context of random unitary circuits, starting from [27]. In this section, we will examine operator spreading as a function of the circuit depth L that can be regarded as the time t in discrete quantum systems.

Any Hermitian operator $\mathcal{O}(t)$ acting on n qubit systems can be written in the Pauli string basis:

$$\mathcal{O}(t) = \frac{1}{2^{n/2}} \sum_{j_1,\cdots,j_n} h_{j_1,\cdots,j_n}(t)\, \sigma_{j_1}^{(1)} \otimes \cdots \otimes \sigma_{j_n}^{(n)} \ , \tag{9}$$

where

$$h_{j_1,\cdots,j_n}(t) \equiv \frac{1}{2^{n/2}} \mathrm{Tr}(\sigma_{j_1}^{(1)} \otimes \cdots \otimes \sigma_{j_n}^{(n)} \cdot \mathcal{O}(t)) \ . \tag{10}$$

Under the unitary time evolution (1),

$$\begin{aligned} &\mathrm{Tr}\left(\mathcal{O}(t)^\dagger \mathcal{O}(t)\right) = \mathrm{Tr}(\mathcal{O}(0)^\dagger \mathcal{O}(0)) = \\ &= \frac{1}{2^n} \sum_{j_1,\cdots,j_n} |h_{j_1,\cdots,j_n}(t)|^2 = \text{constant} \ . \end{aligned} \tag{11}$$

The size of the operator $\mathcal{O}(t)$ is defined as the size of the region where $\mathcal{O}(t)$ does not commute with an operator $\sigma_a^{(x)}$ located at position $1 \le x \le n$. It can be written as

$$\begin{aligned} \mathcal{C}_a(x,t) &= \frac{1}{2}\mathrm{Tr}(\rho_\infty [\mathcal{O}(t), \sigma_a^{(x)}]^\dagger [\mathcal{O}(t), \sigma_a^{(x)}]) = \\ &= \frac{1}{2}\mathrm{Tr}([\mathcal{O}(t), \sigma_a^{(x)}]^\dagger [\mathcal{O}(t), \sigma_a^{(x)}]) \\ &= \sum_{\substack{j_1,\cdots,j_n \\ j_x \neq 0,a}} 2|h_{j_1,\cdots,j_n}(t)|^2 \ . \end{aligned} \tag{12}$$

We numerically measure it with $a = y$, where the operator $\mathcal{O}(0)$ is the Pauli-x matrix located at $x = n/2$.

Figure 4 visualizes the operator spreading coefficient $\mathcal{C}_y(x,t)$ at different times $t = L$ and positions $1 \le x \le n$, averaged over 50 random circuit instances in a system of $n = 12$ qubits. For comparison of the operator spreading pattern across different quantum circuit architectures, we also draw in Figure 5 the standard deviation of $\mathcal{C}_y(x,t)$ over all $1 \le x \le n$ as a function of L. We observe that the $R_x+\mathrm{CZ}$ and $R_y+\mathrm{CZ}$ unitaries reach saturation around $L \gtrsim 30$, while $R_y+\mathrm{CP}$ takes $L \gtrsim 60$ for complete spread. Furthermore, there is no complete spreading under the $R_x+\mathrm{CZ}+R_y+\mathrm{CZ}$ unitaries even with a large number of circuit layers L. These behaviors are all consistent with the entanglement growth pattern of random circuit states illustrated in Figures 2 and 3, showing a clear correlation between two distinct quantities.

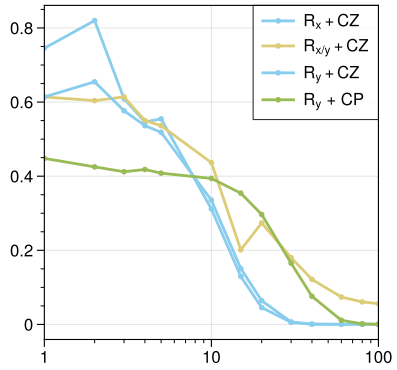
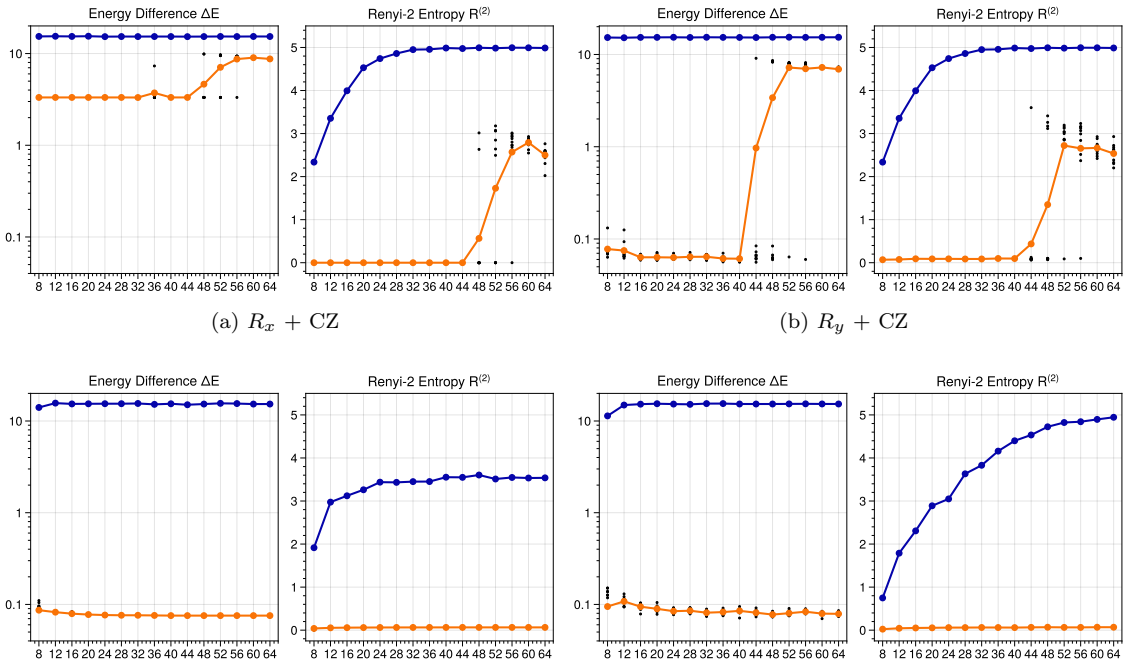


FIG. 5: Standard deviation of $C(x,t)$ over all lattice sites x as a function of the circuit depth t. We see that even at a large number of layers, there is no complete spreading for the R_x + CZ +R_y + CZ circuit. This is consistent with the saturation behaviour of the circuit entanglement in Figure 2 (c) and in Figure 3 (c).

IV. SPECTRAL DIAGNOSTICS OF QUANTUM CHAOS

The Bohigas-Giannoni-Schmit (BGS) conjecture [13, 28] associates the quantum chaotic properties of a system with the correlations between its energy levels. Chaotic Hamiltonians exhibit level correlations in agreement with the predictions of random matrix theory (RMT) [29]. Adjacent eigenvalues show level repulsion and, at larger energy scales, signals of spectral rigidity.



(a) R_x + CZ (b) R_y + CZ

(c) R_x + CZ+R_y + CZ (d) R_y + CP

FIG. 2: Optimization curves for the Ising model coupled to an external field aligned to x-axis (7) at $g = 1$. The plot on the left shows the energy difference from the ground state before (blue curves) and after (orange curves) optimization of the circuit parameters as a function of the number of circuit layers. We see that the entanglement velocity (the slope of the curves) as well as the saturation plateau depend on the type of quantum gates. The plot on the right shows the Renyi-2 entropy before and after optimization of the circuit parameters as a function of the number of circuit layers. (a) R_x one-qubit rotation gate followed by entangling two-qubit CZ gate. (b) R_y one-qubit gate followed by entangling two-qubit CZ gate (c) A sequence of R_x one-qubit rotation gate, entangling two-qubit CZ gate, R_y and CZ. (d) R_y one-qubit gate followed by entangling two-qubit CP gate.

In this section, we will apply three quantum chaos diagnostics, i.e., the level spacing distribution, the r-statistics and the spectral form factor, to the modular Hamiltonian (2) of quantum circuits at varying depth L. The first two diagnoses focus on small energy scales and therefore can determine the presence of level repulsion, one of the most robust indications of quantum chaos [21]. On the other hand, the spectral form factor probes larger energy scales and is mostly a quantifier for spectral rigidity.

A particular care needs to be taken in analysing the eigenvalues of the reduced density matrices ρ_A, since they show unavoidable numerical errors. In order to control the effect of the numerical errors, we have adopted a robust phenomenological procedure, which makes use of the fact that all the eigenvalues of ρ_A must be non-vanishing by definition. Let us denote by $\lambda_{\min}$ the minimum negative eigenvalue among N ensemble realizations for a given value of L. To make sure that we consider only eigenvalues of $\rho_A(L)$ that are not affected by the numerical precision, we impose a cutoff on the spectra by considering only the eigenvalues satisfying the bound:

$$\lambda_i \geq 10|\lambda_{\min}| \ . \tag{13}$$

Such cutoff, when applied at small values of L, removes most of the eigenvalues of $\rho_A(L)$ as most of the eigenvalues are zero at small L. However, this is not the case for larger L, when the RMT structure is clearly visible. The procedure ensures that the eigenvalues kept are robust and not significantly affected by the numerical precision. Out of the significant eigenvalues of $\rho_A(L)$, we compute energy levels E_i of the modular Hamiltonian $H(\rho_A)$.

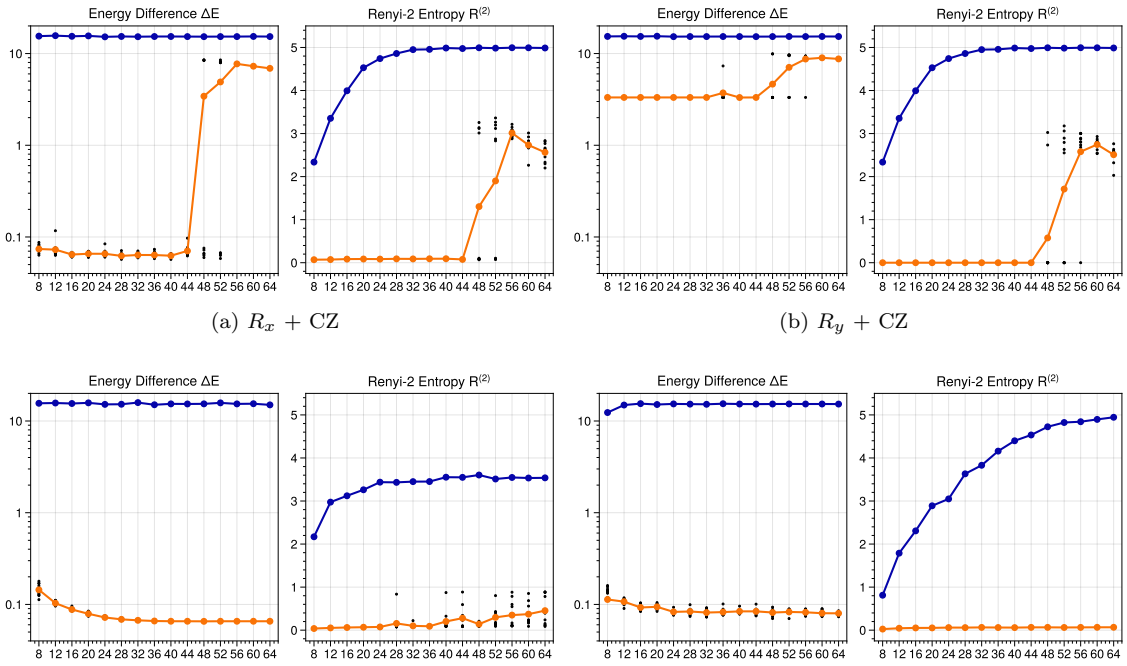
In the next subsections, we will consider only meaningful energy levels, E_i, obtained from the above procedure.

A. Level Spacing Distribution

Roughly speaking, the level spacing distribution measures the probability density for two adjacent eigenvalues to be in the energy distance s, in units of the mean level spacing Δ. The procedure for normalizing all distances in terms of the local mean level spacing is often referred to as unfolding. We unfold the spectrum of the modular Hamiltonian $H(\rho_A)$ by using the following algorithm:

1. Arrange non-degenerate energy levels, E_i, of a modular Hamiltonian $H(\rho_a)$ in ascending order.

2. Compute the staircase function $S(E)$ that enumer-



(a) R_x + CZ (b) R_y + CZ

(c) R_x + CZ+R_y + CZ (d) R_y + CP

FIG. 3: Optimization curves for the Ising model coupled to an external field aligned to y-axis at $g = 1$. In comparison to Figure 2 we see that the circuit R_y + CZ does not optimize well. It shows that the Renyi entropy can be insufficient to diagnose the optimization performance.

ates all eigenstates of $H(\rho_a)$ whose eigenvalues are smaller than or equal to E.

3. Fit a smooth curve that we denote by $\tilde{\rho}(E)$ to the staircase function. To be specific, we used a 12-th order polynomial as the smooth approximation.

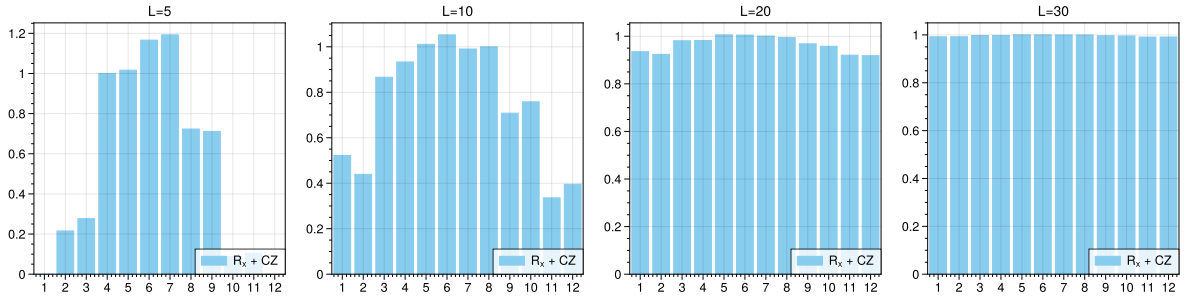
4. Rescale the energy levels E_i as follows:
$$E_i \to e_i = \tilde{\rho}(E_i)\,. \tag{14}$$

5. By construction, the unfolded energy levels e_i must show an approximately uniform distribution with mean level spacing 1. This can be used to ensure if the above procedure has been successful, i.e., by plotting the unfolded levels and checking the flatness of the distribution.

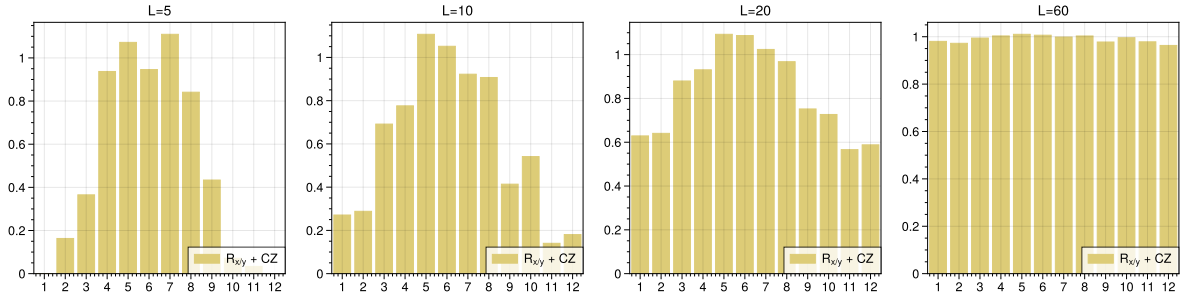
Having obtained the unfolded spectrum, we compute their level spacing, $s_i = e_{i+1} - e_i$, and draw the probability density function $p(s)$ for having two neighbouring eigenvalues separated at a distance s. The level spacing distribution serves as a diagnostic for quantum chaos in Hamiltonian systems. It captures information about the short-range spectral correlations. It thus demonstrates the presence of level repulsion, *i.e.* whether $p(s) \to 0$ as $s \to 0$, which is a common characteristic of random matrix ensembles and particularlly chaotic Hamiltonians.

The level spacing distribution $p(s)$ for integrable systems follows the Poisson distribution
$$p(s) = e^{-s}\,, \tag{15}$$
while for chaotic systems it takes the following form
$$p_\beta(s) = \frac{s^\beta e^{-b_\beta s^2}}{\Gamma(\frac{1+\beta}{2})}\,, \tag{16}$$
where β depends on which universality class of random matrices the chaotic Hamiltonian belongs to [29]: $\beta = 1$ for the Gaussian Orthogonal Ensemble (GOE), $\beta = 2$ for the Gaussian Unitary Ensemble (GUE), and $\beta = 4$ for the Gaussian Symplectic Ensemble (GSE).

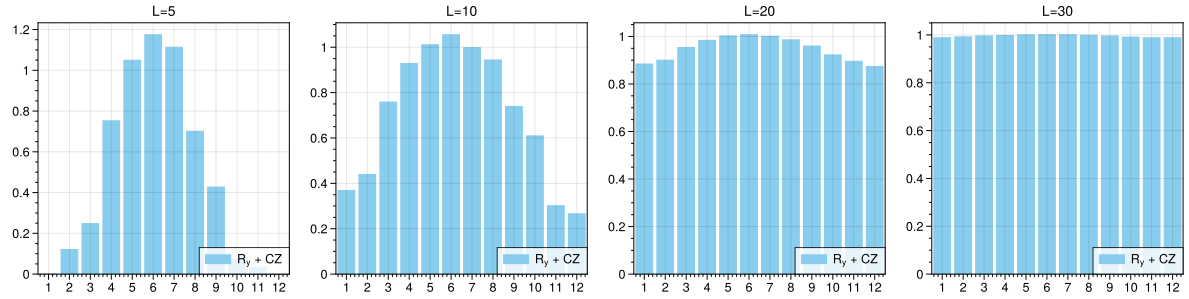
For different types of circuit unitaries defined in Section II with $L = 10$, 30 and 250 layers, we collect 500 random circuit states and draw the corresponding level spacing distributions in Figures 6–8. The modular Hamiltonian of shallow circuit states at $L = 10$ displays a clear departure from the RMT predictions, manifesting a lack of level repulsion. Such distinction is particularly pronounced for the R_y + CP unitary circuit. However, the emergence of random matrix structure becomes evident as stacking more circuit layers. The agreement between empirical level spacing distributions of random circuit states and RMT predictions (16) is already quite obvious at $L = 30$ and further improved at $L = 250$. Note that different choices of unitary gates lead to the emergence of different random matrix ensembles. We observe GUE for the R_x + CZ + R_y + CZ and R_y + CP circuit unitaries and GOE for R_x + CZ and R_y + CZ unitaries. What universality class high-depth random circuit states belong to can be traced from the characteristics of their



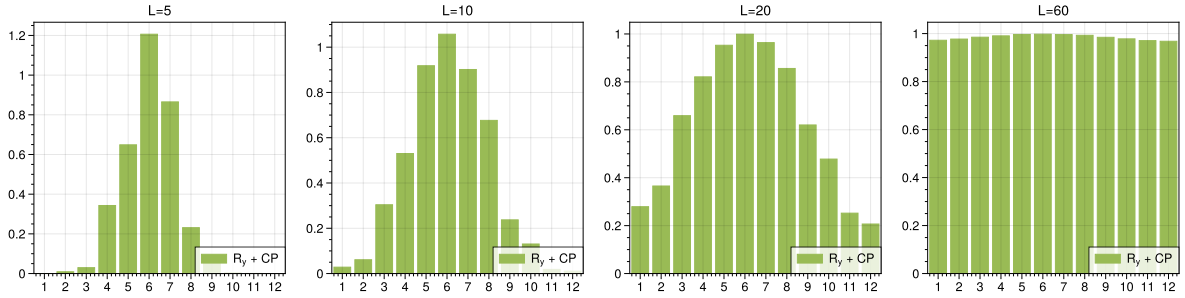
(a) R_x + CZ



(b) R_x + CZ+R_y + CZ



(c) R_y + CZ



(d) R_y + CP

FIG. 4: The operator spreading $\mathcal{C}(x,t)$ for a system of $n = 12$ qubits as a function of $1 \leq x \leq n = 12$ for different number of layers L. The different circuit architectures are: (a) R_x one-qubit rotation gate followed by entangling two-qubit CZ gate. (b) R_y one-qubit gate followed by entangling two-qubit CZ gate (c) A sequence of R_x one-qubit rotation gate, entangling two-qubit CZ gate, R_y and CZ. (d) R_y one-qubit gate followed by entangling two-qubit CP gate. We see a clear correlation between the operator spreading and the entanglement measures of the circuit in Figures 2 and 3.

modular Hamiltonians, or even primarily, from their full density matrices.

We remark that, though the empirical level spacing distribution follows the random matrix theory and exhibits chaotic properties at $L = 30$, the entanglement entropy and operator spreading coefficient have not reached saturation and the VQA optimization works smoothly. This points out a difference between the information scrambling measures based on eigenvalues vs. their spacings of modular Hamiltonians.

B. r-statistics

The previous analysis of the level spacing distribution depends on unfolding the energy spectrum, which is only heuristically defined and has some arbitrariness. Therefore, it would be desirable to have additional diagnostics of quantum chaos that bypass the unfolding procedure. The r-statistics, first introduced in [30], is such a diagnostic tool for short-range correlations, defined without the necessity to unfold the spectrum.

Given the level spacings s_i, defined as the differences between adjacent eigenvalues $\cdots < E_i < E_{i+1} < \cdots$ without unfolding, one defines the following ratios:

$$r_i = \frac{\mathrm{Min}(s_i,\, s_{i+1})}{\mathrm{Max}(s_i,\, s_{i+1})}\,, \tag{17}$$

which are by definition positive numbers between 0 and 1. The ratios r_i take very specific values if the energy levels are the eigenvalues of random matrices: For matrices in GOE, GUE and GSE, the ratios are $r_i \approx 0.53590$, $r_i \approx 0.60266$ and $r_i \approx 0.67617$, respectively. The values become typically smaller for integrable Hamiltonians, approaching $r_i \approx 0.38629$ for a pure Poisson process [31].

From their very definition, we see that the ratios (17) do not require to unfold the spectrum since their dependence on the local density of states is cancelled by taking the ratio between spacings. Moreover, each r_i depends on just three adjacent energy levels, rendering it as a sharp microscopic probe of the chaotic/integrable behavior in a small cluster of spectral values.

Here we use the r-statistics to study the chaotic properties of the entanglement spectra as a function of the number of circuit layers L. Under the equal partitioning of $n = 12$ qubits and with $L = 10, 30, 250$ layers, the numerical values of $\{r_i\}$ are shown in Figure 9 where we observe the transition from Poisson-like to RMT-like values. The low-level eigenstates of the reduced density matrix are more prone to keep their integrable behavior until a sufficient number of entangling layers $L \gtrsim 30$ is reached, where we find the universal GOE/GUE chaotic structure in agreement with the level spacing distribution analysis.

C. Spectral Form Factor

The spectral form factor (SFF) is the Fourier transform of the spectral two-point correlation function [29]. It can be viewed as a long-range observable, since it probes the agreement of a given unfolded spectrum with RMT at energy scales much larger than the mean level spacing. In particular, SFF can detect the presence of spectral rigidity, and is thus a complementary probe of quantum chaos to the distribution of level spacing and the r-statistics that are short-range observables. Formally, one defines the analytically continued partition function

$$Z(\tau) = \mathrm{Tr}\, e^{-i\tau H(\rho_A)}\,, \tag{18}$$

and the spectral form factor is [32]

$$K(\tau) = |Z(\tau)|^2 / Z(0)^2. \tag{19}$$

For a concrete numerical evaluation, we will take the following expression as a robust definition of the spectral form factor [33]:

$$K(\tau) \equiv \frac{1}{Z}\left\langle \left| \sum_i \rho(e_i) e^{-i2\pi e_i \tau} \right|^2 \right\rangle\,. \tag{20}$$

where e_i is the unfolded spectrum of the modular Hamiltonian. The normalization factor $Z = \sum_i |\rho(e_i)|^2$ is chosen to ensure that $K(\tau) \approx 1$ in the limit $\tau \to \infty$. The bracket $\langle \cdots \rangle$ denotes the ensemble average over distinct random circuit realizations. $\rho(e_i)$ is a Gaussian filter [34],

$$\rho(e_i) \equiv \exp\left\{ -\frac{2\,(e_i - \bar{e})^2}{\Gamma^2} \right\}, \tag{21}$$

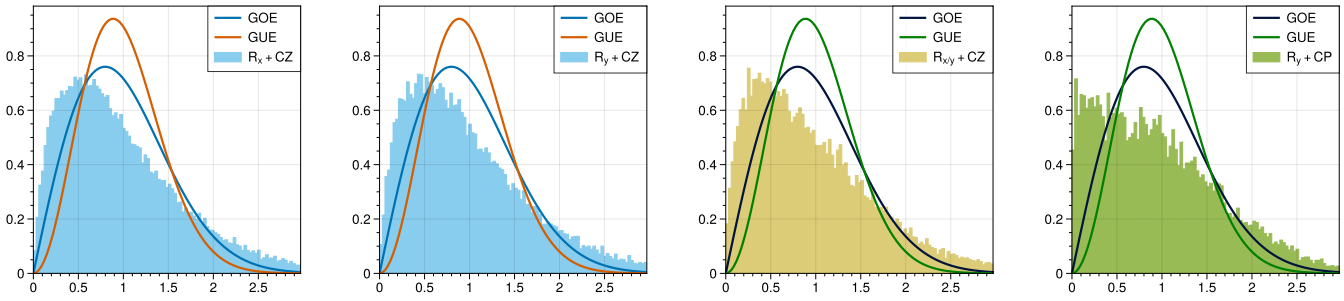
where $\bar{e}$ and Γ^2 denote the mean energy and the variance for each unfolded spectrum. Its purpose is to guarantee that the SFF is mainly affected by eigenvalues located around the mean value of each unfolded spectrum.

The SFF can be computed analytically for the Gaussian ensembles (GOE and GUE). It reads in the thermodynamic limit as [29]

$$\begin{aligned} K_{\mathrm{GOE}}(\tau) &= 2\tau - \tau \ln(1 + 2\tau)\,, \\ K_{\mathrm{GUE}}(\tau) &= \tau, \end{aligned} \tag{22}$$

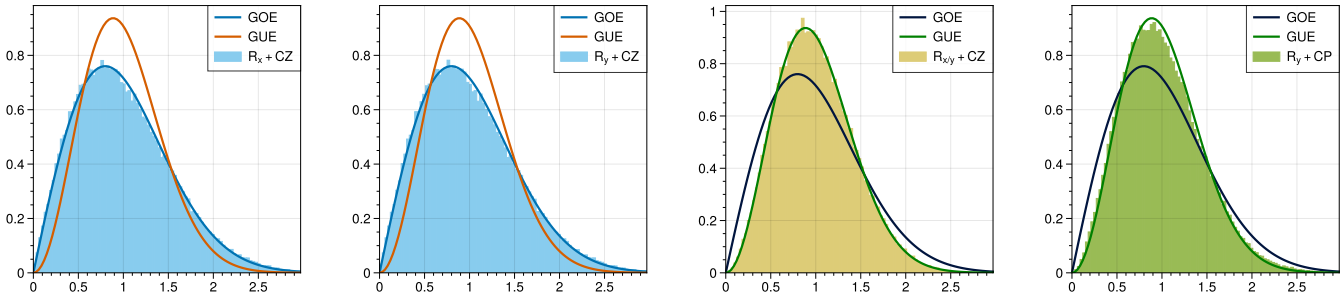
when $0 < \tau < 1$ and $K(\tau) = 1$ when $\tau \geq 1$. Its constancy for $\tau \geq 1$ simply comes from the discreteness of the spectrum and carries no information about spectral correlations. In particular, since the mean level spacing Δ is by construction equal to 1 in the unfolded spectrum, the relevant time scale at which the discreteness of the spectrum becomes relevant is accordingly $\tau \approx 1/\Delta \approx 1$. This scale is usually called the Heisenberg time, τ_{Heis}.

The emergence of the random matrix structure in spectral correlations must be investigated for times shorter than the Heisenberg time, $\tau \leq 1$. The timescale that characterizes the ergodicity of a dynamical system is



(a) R_x + CZ (b) R_y + CZ (c) R_x + CZ+R_y + CZ (d) R_y + CP

FIG. 6: Level spacing distributions at $L = 10$ number of layers. The distributions differ from those of RMTs.



(a) R_x + CZ (b) R_y + CZ (c) R_x + CZ+R_y + CZ (d) R_y + CP

FIG. 7: Level spacing distribution at $L = 30$ number of layers. We see clearly the agreement with the RMT distributions. R_x + CZ and R_y + CZ follow GOE, while R_x + CZ+R_y + CZ and R_y + CP follow GUE.

called the Thouless time, τ_{Thoul}, defined as the time when the SFF of the dynamical system converges to the universal RMT computation. More concretely, it is indicated by the onset of the universal linear ramp as in (22). One expects τ_{Thoul} to decrease by increasing the system size, in ergodic systems, to approach 0 in the thermodynamic limit. In contrast, non-ergodic systems show the absence of linear ramp, *i.e.* $\tau_{\text{Thoul}} \sim \tau_{\text{Heis}} \sim 1$, or more generally, unclear scaling of τ_{Thoul} with respect to the system size.

Following the above discussion, we computed the empirical SFF for different circuit architectures with $L =$ 10, 30 and 250 layers, where the ensemble average is replaced with averaging over 50 random circuit samples. See Figure 10. Inspecting the Thouless time as expanding the system from $n = 12$ to 18 reveals clear indications of ergodicity breaking at $L = 10$, but an expected ergodic behavior for $L = 30$ and 250 layers. They are consistent with the conclusion obtained through the short-range observables in previous subsections.

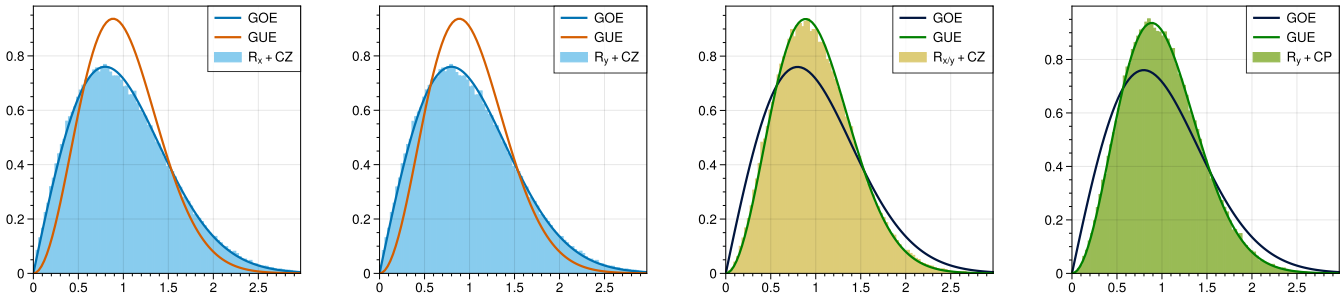
The circuit reduced density matrix is a random matrix by construction. And thus, it may not be surprising that the modular Hamiltonian eigenspectrum exhibits chaotic properties of RMTs. It is interesting, however, to trace the reasons for the GOE and GUE structures to the form of the quantum gates. While random circuit states should generically be in the GUE class, the choice of the gates may generate a modular Hamiltonian whose matrix elements are not complex-valued, but rather real or pure imaginary. In such cases, e.g., for $R_x + \text{CZ}$ and $R_y + \text{CZ}$ unitaries, the eigenspectrum of the corresponding modular Hamiltonian must belong to GOE.

Note that although the level spacing diagnostics show apparent RMT properties at $L = 30$, the local search of optimal circuit parameters still operates well. It indicates that unlike the diagnostic measures based on eigenvalues of the modular Hamiltonian, e.g., entanglement entropies and operator spreading coefficients, the quantum chaos diagnostics constructed from the level spacing of eigenvalues are not precisely correlated with the efficiency of optimizing control variables.

V. DISCUSSION AND OUTLOOK

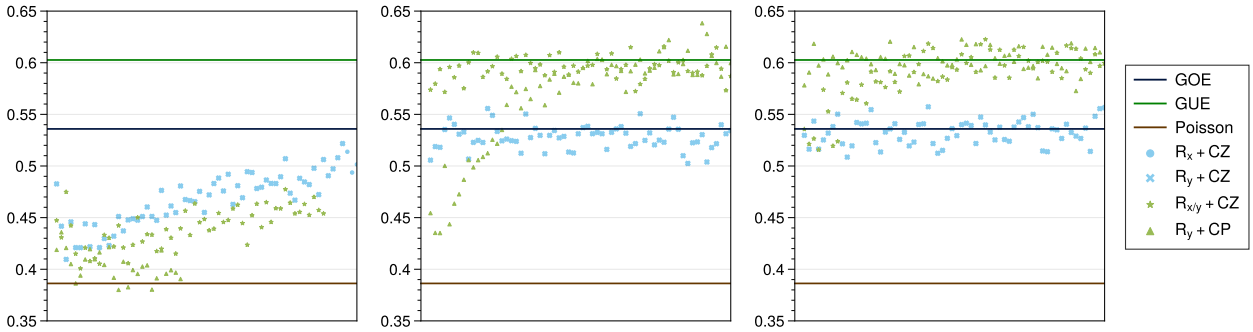
We analyzed the universal chaotic properties of random quantum circuits at different depths and how it correlates to the optimization performance of control variables. Our main focus was on the operator spreading and the level spacing distribution for the eigenspectrum of reduced density matrices. We found that the random circuit wavefunction exhibits the chaotic structure of Gaussian matrix ensembles, which can be either GOE or GUE depending on the type and arrangement of unitary gates.

By changing the direction of the magnetic field coupled



(a) R_x + CZ (b) R_y + CZ (c) R_x + CZ+R_y + CZ (d) R_y + CP

FIG. 8: Level spacing distribution at $L = 250$ number of layers. We see a precise agreement with the RMT distributions. R_x + CZ and R_y + CZ follow GOE, while R_x + CZ+R_y + CZ and R_y + CP follow GUE.



(a) The ratios $\{r_i\}_i$ at $L = 10$ (b) The ratios $\{r_i\}_i$ at $L = 30$ (c) The ratios $\{r_i\}_i$ at $L = 250$

FIG. 9: The r-statistics, for a system with 12 qubits, with equal partitions and with 10 (left), 30 (center) and 250 (right) layers. We observe evident deviations from the expected RMT values for 10 layers, while a substantial agreement with the RMT predictions is obtained at 30 layers. Interesting, the R_y+CP circuit stills show deviations from the RMT values in the low lying modes which disappear at 250 layers.

to the Ising Hamiltonian used in the VQE experiments, we observed the failure of specific GOE-type variational circuits in reaching the ground state. It suggests the expressibility of variational circuits is not determined alone by their capability of creating highly-entangled states.

Both chaos and entanglement follow from the eigenspectrum structure of the reduced density matrix. However, while entanglement and operator spreading are captured by quantities constructed from eigenvalues themselves, there are other measures of quantum chaos beyond the operator spreading which instead relate to their level spacings. We found that the quantum chaos diagnosed by the eigenvalue spacings typically emerges with fewer circuit layers than to come close to the maximum entanglement of random circuit states, which hinders an effective search of optimal circuit variables [4–8].

Such study points to some mismatch between two distinct definitions of quantum chaos, i.e., the BGS conjecture vs. the operator spreading measured by OTOC. Note that the random circuit exhibits the BGS-type chaotic structure before reaching the complete spreading of operators in OTOC. To the best of our knowledge, this is the first example of a genuine many-body setup in which such discrepancy is observed. Previous studies dealt only with single-body examples with classical counterparts [35–41].

As for future studies, it would be interesting to explore the connection between the graph structure of variational circuits, their effectiveness as the eigensolver of distinct Hamiltonians, and the emergence of quantum chaos in random circuit states. A popular measure of information mixing is the k-design state that cannot be distinguished from the Haar random state when considering averages of polynomials of degree not higher than k. It would be useful to investigate the relationship in the framework of random quantum circuits between the k-design structure and the quantum chaos measures that we analyzed. Some results in this direction have been investigated in [42].

Another intriguing line of investigation is to study the non-stabilizerness of variational circuits — often referred to in the literature as *magic* and regarded to be the source of quantum advantage in many computing problems [43, 44]. An explicit measure of magic was recently proposed in [45]. Its relations with quantum chaos was studied in [46]. It would be interesting to better investigate the role of magic in the VQA problems, following [47].

ACKNOWLEDGMENTS

We would like to thank Thi Ha Kyaw, Alexey Milekhin and Jan Olle for valuable discussions. The work of J.K. is supported by the NSF grant PHY-1911298 and the Sivian fund. The work of Y.O. is supported in part by the Israeli Science Foundation Center of Excellence. DR acknowledges the support by the Institute for Basic Science in Korea (IBS-R024-Y2 and IBS-R024-D1).

[1] A. Peruzzo, J. McClean, P. Shadbolt, M.-H. Yung, X.-Q. Zhou, P. J. Love, A. Aspuru-Guzik, and J. L. O'Brien, "A variational eigenvalue solver on a photonic quantum processor," *Nature Communications* **5** no. 1, (Jul, 2014) . http://dx.doi.org/10.1038/ncomms5213.

[2] E. Farhi, J. Goldstone, and S. Gutmann, "A Quantum Approximate Optimization Algorithm," arXiv:1411.4028 [quant-ph].

[3] A. Lucas, "Ising formulations of many np problems," *Frontiers in Physics* **2** (2014) . http://dx.doi.org/10.3389/fphy.2014.00005.

[4] J. R. McClean, S. Boixo, V. N. Smelyanskiy, R. Babbush, and H. Neven, "Barren plateaus in quantum neural network training landscapes," *Nature Communications* **9** no. 1, (Nov, 2018) . http://dx.doi.org/10.1038/s41467-018-07090-4.

[5] M. Cerezo, A. Sone, T. Volkoff, L. Cincio, and P. J. Coles, "Cost-function-dependent barren plateaus in shallow quantum neural networks," arXiv:2001.00550 [quant-ph].

[6] C. O. Marrero, M. Kieferová, and N. Wiebe, "Entanglement induced barren plateaus," arXiv:2010.15968 [quant-ph].

[7] J. Kim, J. Kim, and D. Rosa, "Universal effectiveness of high-depth circuits in variational eigenproblems," *Physical Review Research* **3** no. 2, (Jun, 2021) . http://dx.doi.org/10.1103/PhysRevResearch.3.023203.

[8] J. Kim and Y. Oz, "Entanglement diagnostics for efficient quantum computation," arXiv:2102.12534 [quant-ph].

[9] J. Kim, J. Kim, and D. Rosa, "Universal effectiveness of high-depth circuits in variational eigenproblems," *Physical Review Research* **3** no. 2, (Jun, 2021) . http://dx.doi.org/10.1103/PhysRevResearch.3.023203.

[10] J. Lee, A. B. Magann, H. A. Rabitz, and C. Arenz, "Progress toward favorable landscapes in quantum combinatorial optimization," *Physical Review A* **104** no. 3, (Sep, 2021) . http://dx.doi.org/10.1103/PhysRevA.104.032401.

[11] J. Kim and Y. Oz, "Quantum energy landscape and vqa optimization," arXiv:2107.10166 [quant-ph].

[12] M. Larocca, N. Ju, D. García-Martín, P. J. Coles, and M. Cerezo, "Theory of overparametrization in quantum neural networks," arXiv:2109.11676 [quant-ph].

[13] O. Bohigas, M. J. Giannoni, and C. Schmit, "Characterization of chaotic quantum spectra and universality of level fluctuation laws," *Phys. Rev. Lett.* **52** (Jan, 1984) 1–4. https://link.aps.org/doi/10.1103/PhysRevLett.52.1.

[14] A. I. Larkin and Y. N. Ovchinnikov, "Quasiclassical Method in the Theory of Superconductivity," *Soviet Journal of Experimental and Theoretical Physics* **28** (June, 1969) 1200.

[15] P. Hayden and J. Preskill, "Black holes as mirrors: quantum information in random subsystems," *Journal of High Energy Physics* **2007** no. 09, (Sep, 2007) 120–120. https://doi.org/10.1088/1126-6708/2007/09/120.

[16] Y. Sekino and L. Susskind, "Fast scramblers," *Journal of High Energy Physics* **2008** no. 10, (Oct, 2008) 065–065. https://doi.org/10.1088/1126-6708/2008/10/065.

[17] S. H. Shenker and D. Stanford, "Black holes and the butterfly effect," *Journal of High Energy Physics* **2014** no. 3, (Mar, 2014) . http://dx.doi.org/10.1007/JHEP03(2014)067.

[18] J. Maldacena, S. H. Shenker, and D. Stanford, "A bound on chaos," *Journal of High Energy Physics* **2016** no. 8, (Aug, 2016) . http://dx.doi.org/10.1007/JHEP08(2016)106.

[19] L. D'Alessio, Y. Kafri, A. Polkovnikov, and M. Rigol, "From quantum chaos and eigenstate thermalization to statistical mechanics and thermodynamics," *Advances in Physics* **65** no. 3, (May, 2016) 239–362. http://dx.doi.org/10.1080/00018732.2016.1198134.

[20] B. Kobrin, Z. Yang, G. D. Kahanamoku-Meyer, C. T. Olund, J. E. Moore, D. Stanford, and N. Y. Yao, "Many-body chaos in the sachdev-ye-kitaev model," *Physical Review Letters* **126** no. 3, (Jan, 2021) . http://dx.doi.org/10.1103/PhysRevLett.126.030602.

[21] F. Haake, *Quantum Signatures of Chaos*. Springer-Verlag, Berlin, Heidelberg, 2006.

[22] X. Chen and A. W. W. Ludwig, "Universal Spectral Correlations in the Chaotic Wave Function, and the Development of Quantum Chaos," *Phys. Rev. B* **98** no. 6, (2018) 064309, arXiv:1710.02686 [cond-mat.str-el].

[23] A. Arrasmith, Z. Holmes, M. Cerezo, and P. J. Coles, "Equivalence of quantum barren plateaus to cost concentration and narrow gorges," arXiv:2104.05868 [quant-ph].

[24] B. T. Kiani, S. Lloyd, and R. Maity, "Learning unitaries by gradient descent," arXiv:2001.11897 [quant-ph].

[25] S. Sachdev, *Quantum phase transitions*. Cambridge University Press, Cambridge, second ed. ed., 2011.

[26] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980 [cs.LG].

[27] A. Nahum, S. Vijay, and J. Haah, "Operator spreading in random unitary circuits," *Physical Review X* **8** no. 2, (Apr, 2018) . http://dx.doi.org/10.1103/PhysRevX.8.021014.

[28] T. Guhr, A. Müller–Groeling, and H. A. Weidenmüller, "Random-matrix theories in quantum physics: common concepts," *Physics Reports* **299** no. 4-6, (Jun, 1998) 189–425. http://dx.doi.org/10.1016/S0370-1573(97)00088-4.

[29] M. L. Mehta, *Random Matrices*. 3rd ed., 2004.

[30] V. Oganesyan and D. A. Huse, "Localization of interacting fermions at high temperature," *Phys. Rev. B* **75** (Apr, 2007) 155111. https://link.aps.org/doi/10.1103/PhysRevB.75.155111.

[31] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, "Distribution of the ratio of consecutive level spacings in random matrix ensembles," *Phys. Rev. Lett.* **110** (Feb, 2013) 084101. https://link.aps.org/doi/10.1103/PhysRevLett.110.084101.

[32] J. S. Cotler, G. Gur-Ari, M. Hanada, J. Polchinski, P. Saad, S. H. Shenker, D. Stanford, A. Streicher, and M. Tezuka, "Black holes and random matrices," *Journal of High Energy Physics* **2017** no. 5, (May, 2017) . http://dx.doi.org/10.1007/JHEP05(2017)118.

[33] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, "Quantum chaos challenges many-body localization," *Physical Review E* **102** no. 6, (Dec, 2020) . http://dx.doi.org/10.1103/PhysRevE.102.062144.

[34] H. Gharibyan, M. Hanada, S. H. Shenker, and M. Tezuka, "Onset of random matrix behavior in scrambling systems," *Journal of High Energy Physics* **2018** no. 7, (Jul, 2018) . http://dx.doi.org/10.1007/JHEP07(2018)124.

[35] E. B. Rozenbaum, L. A. Bunimovich, and V. Galitski, "Early-time exponential instabilities in nonchaotic quantum systems," *Phys. Rev. Lett.* **125** (Jul, 2020) 014101. https://link.aps.org/doi/10.1103/PhysRevLett.125.014101.

[36] T. Xu, T. Scaffidi, and X. Cao, "Does scrambling equal chaos?," *Physical Review Letters* **124** no. 14, (Apr, 2020) . http://dx.doi.org/10.1103/PhysRevLett.124.140602.

[37] J. Wang, G. Benenti, G. Casati, and W.-g. Wang, "Complexity of quantum motion and quantum-classical correspondence: A phase-space approach," *Physical Review Research* **2** no. 4, (Nov, 2020) . http://dx.doi.org/10.1103/PhysRevResearch.2.043178.

[38] C. Yin and A. Lucas, "Quantum operator growth bounds for kicked tops and semiclassical spin chains," *Physical Review A* **103** no. 4, (Apr, 2021) . http://dx.doi.org/10.1103/PhysRevA.103.042414.

[39] J. Wang, G. Benenti, G. Casati, and W.-g. Wang, "Quantum chaos and the correspondence principle," *Physical Review E* **103** no. 3, (Mar, 2021) . http://dx.doi.org/10.1103/PhysRevE.103.L030201.

[40] W. Kirkby, D. H. J. O'Dell, and J. Mumford, "False signals of chaos from quantum probes," *Phys. Rev. A* **104** (Oct, 2021) 043308. https://link.aps.org/doi/10.1103/PhysRevA.104.043308.

[41] R. A. Kidd, A. Safavi-Naini, and J. F. Corney, "Saddle-point scrambling without thermalization," *Phys. Rev. A* **103** (Mar, 2021) 033304. https://link.aps.org/doi/10.1103/PhysRevA.103.033304.

[42] J. Haferkamp, F. Montealegre-Mora, M. Heinrich, J. Eisert, D. Gross, and I. Roth, "Quantum homeopathy works: Efficient unitary designs with a system-size independent number of non-clifford gates," 2020.

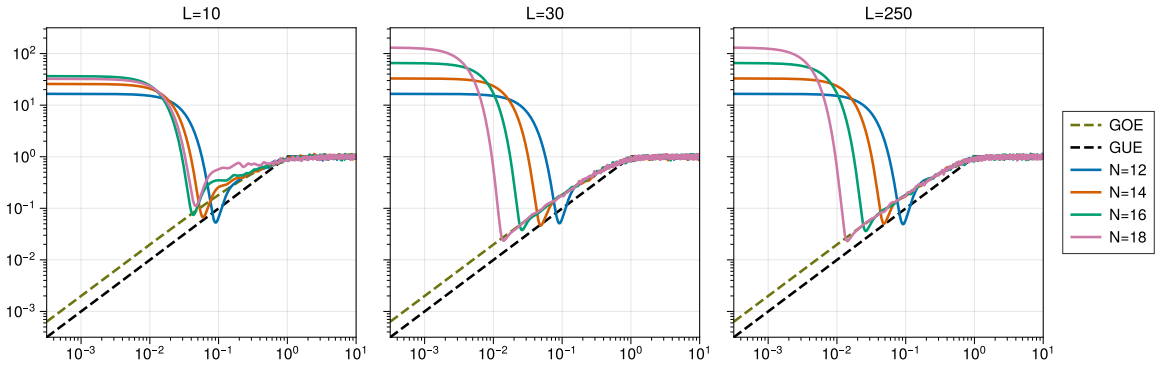
[43] D. Gottesman, "The Heisenberg representation of quantum computers," in *22nd International Colloquium on Group Theoretical Methods in Physics*, pp. 32–43. 7, 1998. arXiv:quant-ph/9807006.

[44] B. Eastin and E. Knill, "Restrictions on transversal encoded quantum gate sets," *Phys. Rev. Lett.* **102** (Mar, 2009) 110502. https://link.aps.org/doi/10.1103/PhysRevLett.102.110502.

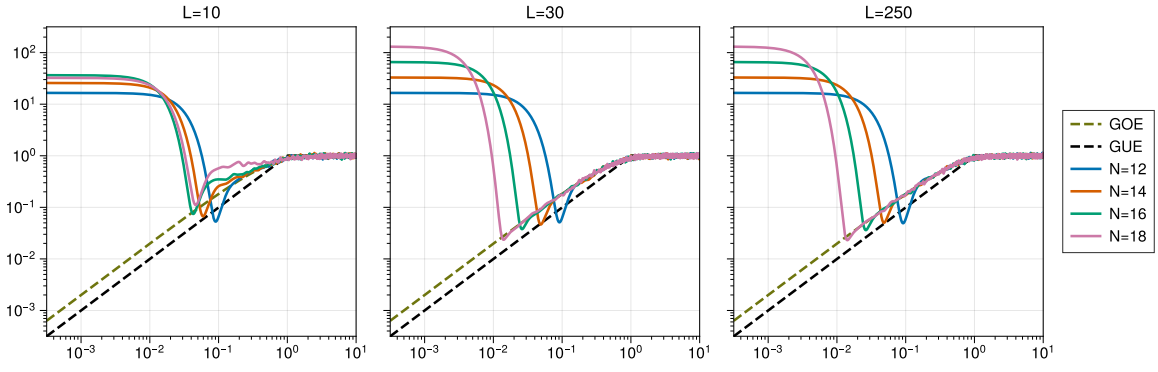
[45] L. Leone, S. F. E. Oliviero, and A. Hamma, "Rényi entropy of magic," 2021.

[46] L. Leone, S. F. E. Oliviero, Y. Zhou, and A. Hamma, "Quantum Chaos is Quantum," *Quantum* **5** (May, 2021) 453. https://doi.org/10.22331/q-2021-05-04-453.

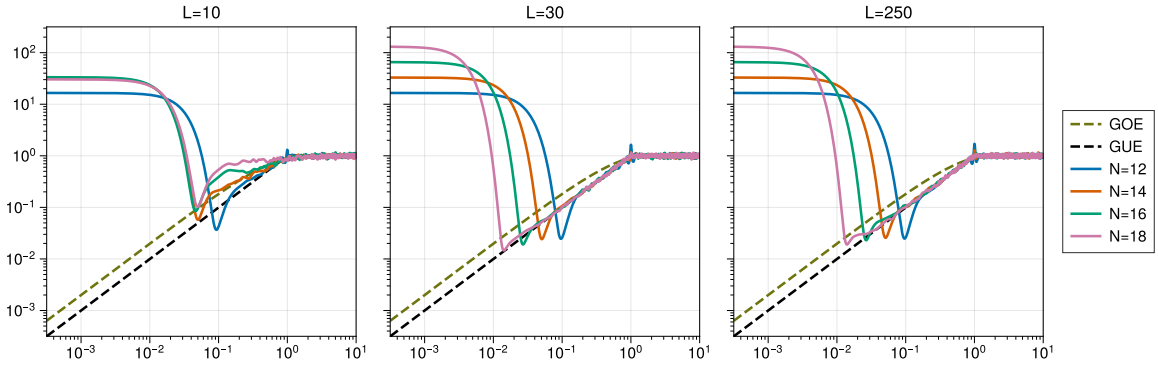
[47] S. Zhou, Z. Yang, A. Hamma, and C. Chamon, "Single t gate in a clifford circuit drives transition to universal entanglement spectrum statistics," *SciPost Physics* **9** no. 6, (Dec, 2020) . http://dx.doi.org/10.21468/SciPostPhys.9.6.087.



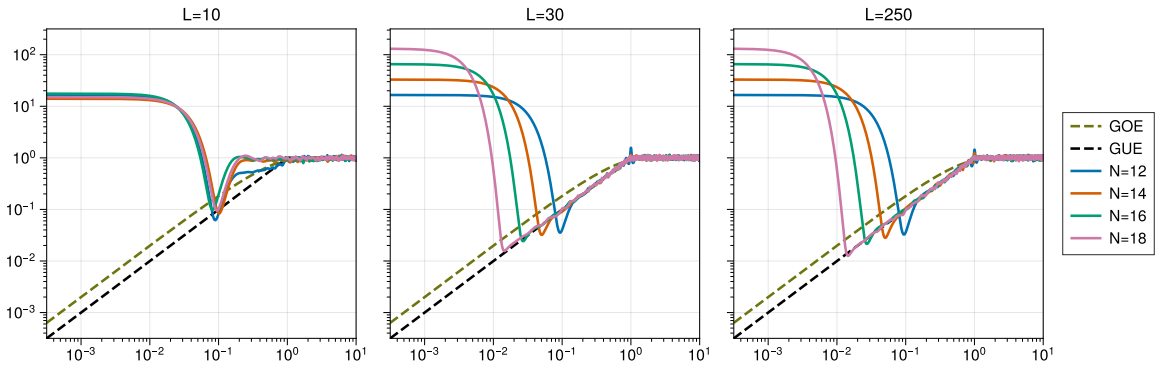
(a) R_x + CZ



(b) R_y + CZ



(c) R_x + CZ+R_y + CZ



(d) R_y + CP

FIG. 10: SFF for 10 (left subfigures), 30 (central subfigures) and 250 (right subfigures) layers. In all cases, the SFFs are very different from the RMT predictions at 10 layers, while starting from 30 layers the agreement is excellent. In particular, the Thouless time, τ_{Thoul}, clearly decreases while increasing the system size. As mentioned in the main text, such a behavior is a signal of the ergodic character of the circuits under investigation.