


Machine learning wave functions to identify fractal phasesTilen Čadež ^{*}*Center for Theoretical Physics of Complex Systems, Institute for Basic Science (IBS), Daejeon 34126, Korea*Barbara Dietz [†], Dario Rosa [‡], and Alexei Andreev [§]*Center for Theoretical Physics of Complex Systems, Institute for Basic Science (IBS), Daejeon 34126, Korea and Basic Science Program, Korea University of Science and Technology (UST), Daejeon 34113, Republic of Korea*Keith Slevin ^{||}*Department of Physics, Graduate School of Science, Osaka University, Toyonaka, Osaka 560-0043, Japan*Tomi Ohtsuki [¶]*Physics Division, Sophia University, Chiyoda, Tokyo 102-8554, Japan*

(Received 12 June 2023; accepted 18 October 2023; published 17 November 2023)

We demonstrate that an image recognition algorithm based on a convolutional neural network provides a powerful procedure to differentiate between ergodic, nonergodic extended (fractal), and localized phases in various systems: Single-particle models, including random-matrix and random-graph models, and many-body quantum systems. We propose an efficient procedure in which the network is successfully trained on a small data set of only 500 wave functions (images) per class for a single model which exhibits these phases. The trained network is then used to classify phases in the other models. We discuss the strengths and limitations of the approach.

DOI: [10.1103/PhysRevB.108.184202](https://doi.org/10.1103/PhysRevB.108.184202)**I. INTRODUCTION**

Computational physics has effectively addressed numerous challenges in solid-state physics. The recent advances in machine learning techniques [1] including deep learning [2] make it a natural choice for tackling complex problems in physics. Indeed, since circa 2016 [3,4], there has been a surge of interest in applying machine learning methods to problems in condensed-matter physics [5,6].

One crucial application of machine learning is the extraction of features from data. However, electron states in random systems often exhibit intricate features. Neural networks obtained through supervised training, which have demonstrated immense potential in image recognition [2], are anticipated to be effective for analyzing electron wave behavior in random systems as images [7].

Random free-electron systems exhibit the Anderson-type metal-insulator transition, also known as the Anderson transition [8–10]. Analyzing the wave functions of random quantum systems can be challenging due to the significant fluctuations in their distribution. Nevertheless, a trained convolutional

neural network (CNN) has been proven to successfully detect such quantum phase transitions [7,11–14]. As is well known in the condensed-matter literature, the Anderson problem, although theoretically very interesting, deals with an idealized situation in which interactions are absent. Therefore, it is of crucial importance to determine whether Anderson localization is robust when interactions are present and, ultimately, in a genuine quantum many-body setup, the latter being usually dubbed as *many-body localization* (MBL) [15–17]. In recent years considerable theoretical [18–39] and experimental [40–51] progress in the study of MBL has been achieved. Therefore, it is natural to apply the same CNN approach to detect the *Anderson-type* MBL transition. Such a task has been addressed from various perspectives and by training the neural networks with different kinds of data in the past few years [52–56].

Some models exhibit localized and extended phases. Anderson's model of localization for the three-dimensional (3D) case and above is a well-known example. In contrast with such models, it is generally believed, although still under intense debate, that certain models, including numerous many-body interacting systems, can exhibit an intermediate phase depending on the associated disorder parameter in which eigenstates are neither fully delocalized nor localized. In the intermediate phase, the states are characterized by nontrivial *fractal* or *multifractal* dimensions. The role of (multi)fractality in the Anderson localization transition has been studied in single-particle models [10,57–66] as well as many-body systems [67–73] based on an analysis,

^{*}tilencadez@ibs.re.kr[†]barbara@ibs.re.kr[‡]dario.rosa85@gmail.com[§]aalexei@ibs.re.kr^{||}slevin.keith.sci@osaka-u.ac.jp[¶]ohtsuki@sophia.ac.jp

which requires a large amount of data. A procedure, like the one proposed in this work, which is capable of identifying (multi)fractal properties of wave functions based on small data sets is thus advantageous.

We have explored the capacity of a CNN to recognize the presence or absence of a fractal phase, both in single-particle and in interacting many-body systems, in addition to the localized and extended phases. To this end, we train the CNN by means of the eigenstates of the generalized Rosenzweig-Porter (gRP) model [74,75], a random-matrix model, which can be seen as a disordered single-particle system and for which the presence of a fractal phase has been determined analytically [75]. After the CNN has been successfully trained to identify the ergodic, fractal, and localized phases in the gRP model, we apply the same CNN, without additional training, to single-particle, random, and many-body systems. In this way, we assess the capability of the CNN in generalizing the acquired knowledge to new models and situations, a *generalization capability*, that, in the context of MBL, has been studied in Ref. [76]. As we will discuss in detail, our results show that the trained CNN can indeed be used on new and unknown models to successfully verify the presence of the three phases learned on the training model. We will also demonstrate that this result can be reached by using an amount of data being *significantly lower* than that required when employing standard numerical techniques, though with less accurate estimates of the critical disorder. All in all, our results will show that a CNN, trained on a well-known physical model, can be considered as an efficient *preliminary tool* to charter the phase diagram of a given model.

The paper is organized as follows. The methods, namely the CNN and the exact diagonalization, are introduced in Sec. II. The various models of interest are discussed in Sec. III and their phase diagrams are schematically shown in Fig. 1. The generalized Rosenzweig-Porter model, introduced in Sec. III A, is used to train and test the CNN, as given in Sec. III B. That network is then applied to other models in Sec. III C. Finally, the results are discussed and the conclusions are given in Sec. IV.

II. METHODS

Convolutional neural networks are networks which take a certain input in the form of a single or multiple arrays, process it, and produce an output, based on the task to be fulfilled. In the case of image recognition the CNN takes an image as an input and as an output classifies the content of the image. A famous example is the handwritten digit recognition [77], where the task is to correctly recognize handwritten digits. Here we use the CNN to recognize different phases of matter [4,11,13], specifically the ergodic extended phase, the nonergodic extended (fractal) phase, and a localized phase in various systems. Typical CNNs are composed of several main parts: Convolutional layers, pooling layers, and standard dense layers. The first two of these offer an improvement compared to the simpler artificial neural networks, consisting of only dense layers, while also reducing the number of parameters to be optimized. We refer the interested reader to Ref. [7] for further details on supervised learning of disordered quantum systems.

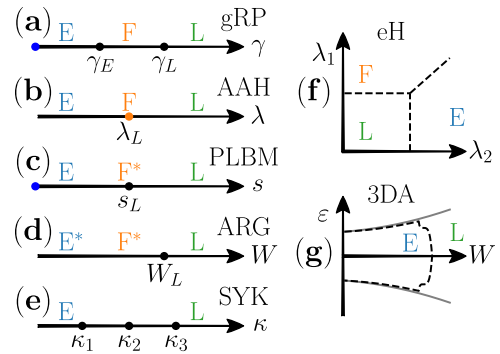


FIG. 1. Schematic phase diagrams of the considered models. The CNN is trained and tested on the gRP model. The corresponding phase diagram is shown in (a). The same network is then applied to (b) the Aubry-André-Harper model (AAH), (c) the power-law banded matrices (PLBM), (d) the Anderson model on random graphs (ARG), (e) the mass-deformed Sachdev-Ye-Kitaev model (SYK), (f) the extended Harper's model (eH), and (g) the 3D Anderson model (3DA). In the left and right panels the dots and dashed lines, respectively, mark known phase transitions (see main text). Blue dots mark the parameter value where the model Hamiltonian is a member of the GOE. The orange dot in (b) signifies that (multi)fractal states are present in the AAH only at the transition point. The letters E , F , L denote the ergodic, fractal, localized phases, respectively, which are the output classes of the CNN. The asterisk in the phase diagrams of the PLBM model and the Anderson model on random graphs signifies that the corresponding phases are observed in the finite-size systems that are considered.

To obtain the eigenfunctions of various models we use the exact diagonalization by solving the equation $H|\psi_\mu\rangle = \varepsilon_\mu|\psi_\mu\rangle$, which is written as $|\psi_\mu\rangle = \sum_i \psi_\mu(i)|i\rangle$, where $\psi_\mu(i)$ are the coefficients of the μ th eigenfunction in the basis spanned by the states $|i\rangle$. Unless stated otherwise, we choose the computational basis. The input data given to the CNN are the squares of the absolute value of the eigenstate coefficients $|\psi_\mu(i)|^2$ (the probabilities of site occupations). For comparison we also calculate the *inverse participation ratio* (IPR) \mathcal{I} , given by $\mathcal{I} = \langle \sum_i |\psi_\mu(i)|^4 \rangle$, where the average can be either over the eigenstates μ in a chosen energy window or over different disorder realizations or both. The output classes of the CNN are the probabilities that an eigenfunction belongs to the ergodic (E), fractal (F), and localized (L) states, respectively.

III. MODELS AND RESULTS

We consider a set of models that exhibit a transition from extended to localized phases. In some of these models the transition takes place via an intermediate fractal phase. In other models, such as Anderson's model of localization, a fractal regime (as opposed to a phase) can be observed as a consequence of a finite-size effect.

A. Generalized Rosenzweig-Porter model (gRP)

To train and test the CNN we use the eigenstates of the gRP model [74,75], which comprises Hermitian random matrices whose elements are Gaussian distributed with zero mean. The

variances of the diagonal and off-diagonal elements, denoted by σ_d^2 and σ_{off}^2 , respectively, are defined as

$$\sigma_d^2 = \langle H_{nn}^2 \rangle = \frac{1}{2N}, \quad \sigma_{off}^2 = \langle H_{nm}^2 \rangle = \frac{1}{4N\gamma+1}. \quad (1)$$

Here, the parameter γ determines the strength of the off-diagonal matrix elements compared to that of the diagonal ones. In this work we consider real matrices implying that for $\gamma = 0$ they are members of the Gaussian orthogonal ensemble (GOE). In Ref. [75] the phase diagram, which is schematically shown in Fig. 1(a), was established. It was shown that the states around the band center exhibit three distinct phases, namely, an ergodic phase for $\gamma < 1$, an extended nonergodic phase for $1 < \gamma < 2$, and a localized phase for $\gamma > 2$. At $\gamma_E = 1$ and $\gamma_L = 2$ they undergo continuous ergodic and Anderson transitions, respectively. The characteristics of the recently discovered extended nonergodic phase is the occurrence of fractal eigenstates whose fractal dimension equals $2 - \gamma$. The phase diagram was confirmed and the properties of the model were further studied recently [78–87]. The ergodic and localized phases can also be determined using simple criteria [88,89], as shown in Appendix E. Here we use the three distinct phases of the gRP model as the output classes of the CNN.

B. CNN training and testing on the gRP model

For training the CNN we use the eigenstates obtained from diagonalizing the gRP model. We use $N \times N$ matrices with $N = 2048$ and provide the absolute-value square of the eigenstate coefficients, that is, the occupation probabilities of the sites in the computational basis, to the input layer. For each of the three phases we extract a single eigenstate corresponding to the eigenenergy closest to the band center, which is at energy 0, for in total 3×500 (ergodic, fractal, localized) random-matrix realizations and use them as input training data set. During the training 90% of the input data are used as a training set and the remaining 10% as the validation set. We observed that the performance improved when applying for all the eigenfunctions a cyclic permutation to the component indices such that the maximum occupation is at the center of the computational basis. Accordingly, assuming periodic boundaries, we applied this procedure to all the data considered. The output layer classifies the ergodic, fractal, and localized phases in terms of probabilities for each phase. The objective of this work is to first use the CNN to classify these three phases in the trained model itself and then to apply the same network to various systems: Single-particle models, including random-matrix models, random-graphs models, and many-body quantum systems. In this way, we will test the ability of CNNs to serve as *diagnostic* tool, i.e., as a tool to uncover the presence of interesting phase diagrams in new and unknown models.

The network architecture consists of two convolutional layers, each followed by a pooling layer where we utilize a max pooling strategy. We flatten the data and apply a dense layer after the second pooling layer, followed by a rectified linear unit (ReLU) activation function. Finally, a second dense layer is applied followed by a softmax activation [7]. Further details and the network hyperparameters are provided in

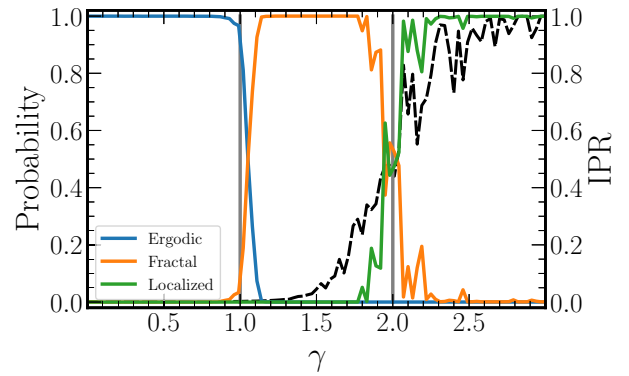


FIG. 2. Testing the trained CNN on the gRP model. The probabilities of each of the phases as well as the average IPR (black dashed line) are shown. The gray vertical lines indicate the analytical values γ_E and γ_L for the ergodic and Anderson transitions.

Appendix A. We use the categorical cross entropy between the output probabilities as a loss function during the training for each of the three phases and the corresponding labels.

After the training, we test the CNN on a new set of data generated as follows. We consider a sequence of values of $\gamma \in [0, 3]$ in which γ is increased in steps of 0.03. For each γ we generate five different random-matrix realizations. For each realization we input to the CNN the eigenstate with energy closest to the band center and average the probabilities output by the CNN over the five realizations. The resulting probabilities for the three phases are shown in Fig. 2. The CNN successfully recognizes each of the phases with probability close to 1. For both the phase transitions, the precision of the determination of the critical value of γ is about 10%. For comparison, we also plot the average IPR. We see that, for a given system size and pool of eigenstates, the different phases are more clearly discernible with the CNN.

C. Generalization capability: Application to other models

Having demonstrated that a CNN can successfully classify the phases of the gRP model, we next demonstrate a *generalization capability*: A CNN trained on one model, the gRP model, can classify phases of other models, *without any retraining of the model*.

1. Aubry-André-Harper model (AAH)

The Aubry-André-Harper (AAH) model [90] is a one-dimensional (1D) single-particle model of a particle hopping along a tight-binding chain with homogeneous hopping (set to 1) and a quasiperiodic onsite potential, with the Hamiltonian given by

$$H\psi_n = V_n\psi_n + \psi_{n+1} + \psi_{n-1}. \quad (2)$$

The potential at site n is $V_n = \lambda \cos(2\pi\alpha n + \varphi)$, with a modulation parameter α chosen as an irrational number. We chose for its value the inverse golden mean $\alpha = (\sqrt{5} - 1)/2$. The angle $\varphi \in [0, 2\pi)$ is a simple phase shift of the modulation. We use it to generate multiple disorder realizations in finite-size systems. The model has two phases, metallic with extended eigenstates for $\lambda < 2$ and insulating with localized

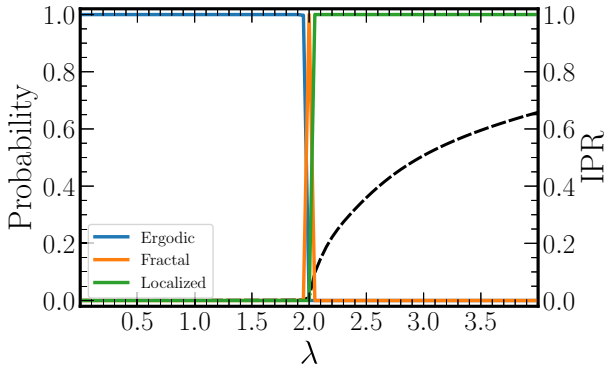


FIG. 3. Testing the trained CNN on the AAH model. The probabilities of each of the phases are shown as well as the average IPR (dashed line). The Anderson transition takes place at $\lambda_L = 2$.

eigenstates for $\lambda > 2$. At the Anderson transition ($\lambda_L = 2$) the eigenstates are multifractal. Furthermore, the two phases are connected via a duality transformation as can be seen by looking at the Fourier transform of the Hamiltonian in Eq. (2). All the eigenstates have the same characteristic length, given by the localization length $\xi = 1/\log(\lambda/2)$ where $\lambda > 2$ [90]. The phase diagram is shown in Fig. 1(b).

We test the CNN, previously trained on the gRP model, as follows. We consider a sequence of values of $\lambda \in [0, 4]$ in which λ is increased in steps of 0.05. For each λ we generate five quasidisorder realizations. For each realization we input to the CNN the eigenstate with energy closest to the energy 0 and average the probabilities output by the CNN over the five realizations. We use the same system size $N = 2048$ as for the gRP model and impose open boundary conditions. As can be seen in Fig. 3, the CNN recognizes the two distinct phases and also correctly identifies the transition point itself as being fractal, all with probabilities close to 1. The transition region is sharper than in the gRP model. To get a clearer picture of its structure, we reduced the λ step to 0.005 and increased the number of realizations to 50. The results are plotted in Fig. 4. The curves for the localized and fractal phases cross at approximately $\lambda = 2.012$. For this value of λ , the localization length is $\xi \approx 167$. This is approximately one order of magnitude smaller than the system size N .

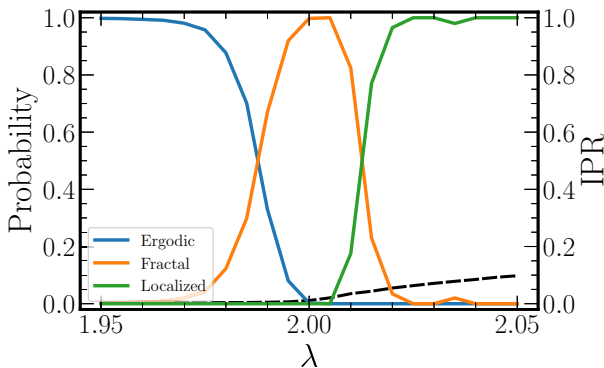


FIG. 4. Zoom into the transition region of the AAH model. The probabilities of each of the phases are shown as well as the average IPR (dashed line).

2. Extended Harper's model (eH)

For the extended Harper's model (eH) the phase diagram was recently established analytically [91] and is schematically shown in Fig. 1(f). It is a one-dimensional nearest-neighbor hopping model. The Hamiltonian is given by

$$\begin{aligned} H\psi_n &= V_n\psi_n + t_n\psi_{n+1} + t_{n-1}^*\psi_{n-1}, \\ V_n &= 2 \cos(2\pi\alpha n + \varphi), \\ t_n &= \lambda_1 e^{-2i\pi\alpha(n+1/2)-i\varphi} + \lambda_2 + \lambda_3 e^{2i\pi\alpha(n+1/2)+i\varphi} \end{aligned} \quad (3)$$

with the modulated onsite potential V_n and modulated nearest-neighbor hoppings t_n . As in the AAH model we choose $\alpha = (\sqrt{5} - 1)/2$. In studies of the eH model, disorder realizations are often generated by taking φ to be randomly distributed on $[0, 2\pi)$. However, below we consider only a single realization and accordingly set $\varphi = 0$. The three parameters $\lambda_1, \lambda_2, \lambda_3$ determine the phase diagram of the model. Without loss of generality we restrict to the cases $\lambda_2 \geq 0$, $\lambda_1 + \lambda_3 \geq 0$, and $\lambda_i > 0$ for at least one of $i = 1, 2, 3$. We consider the symmetric case $\lambda_1 \equiv \lambda_3$ where $t_n = 2\lambda_1 \cos[2\pi\alpha(n + 1/2)] + \lambda_2$ and the Hamiltonian in Eq. (3) is purely real.

It was shown in Ref. [91] that in the symmetric case of a purely real Hamiltonian the spectrum of the eH model belongs to one of the three distinct cases:

- (i) Region I: Localized eigenfunctions and pure point spectrum: $0 \leq \lambda_1 + \lambda_3 \leq 1$, $0 < \lambda_2 \leq 1$.
- (ii) Region II: Extended eigenfunctions and purely absolutely continuous spectrum: $0 \leq \lambda_1 + \lambda_3 \leq \lambda_2$, $\lambda_2 \geq 1$.
- (iii) Region III: Fractal eigenfunctions and purely singular continuous spectrum: $\max\{1, \lambda_2\} \leq \lambda_1 + \lambda_3$, $\lambda_2 > 0$.

The transition lines between the three regions exhibit fractal eigenstates. In the nonsymmetric case, that is, for $\lambda_1 \neq \lambda_3$, regions I and II persist, whereas the eigenstates of region III become extended.

To reduce the number of parameters, we consider a closed loop in the two-parameter space λ_1, λ_2 , which transverses all the three distinct phases of the model. We parametrized the loop by an angle ζ so that

$$\begin{aligned} \lambda_1 &= 0.5 + r_0 \sin(\zeta), \\ \lambda_2 &= 1.0 + 2r_0 \cos(\zeta). \end{aligned} \quad (4)$$

We set $r_0 = \frac{1}{4}$ and increased ζ from 0 to 2π in steps of 0.02π . For each point on the loop, we test the CNN on *the full spectrum* of eigenfunctions. We use open boundary conditions, for which localized edge states appear in the band gaps. The results are presented in Fig. 5. The CNN recognizes the ergodic and localized phases with probabilities close to 1 for the majority of the states and correctly identifies about 85% of the fractal states. The transition regions are sharp. The localized edge states are also successfully identified by the CNN.

Recently, a related model, called the generalized AAH model, was studied [92] using a topological machine learning technique of persistent homology. The authors successfully distinguished the localized, extended, and critical phases within the model.

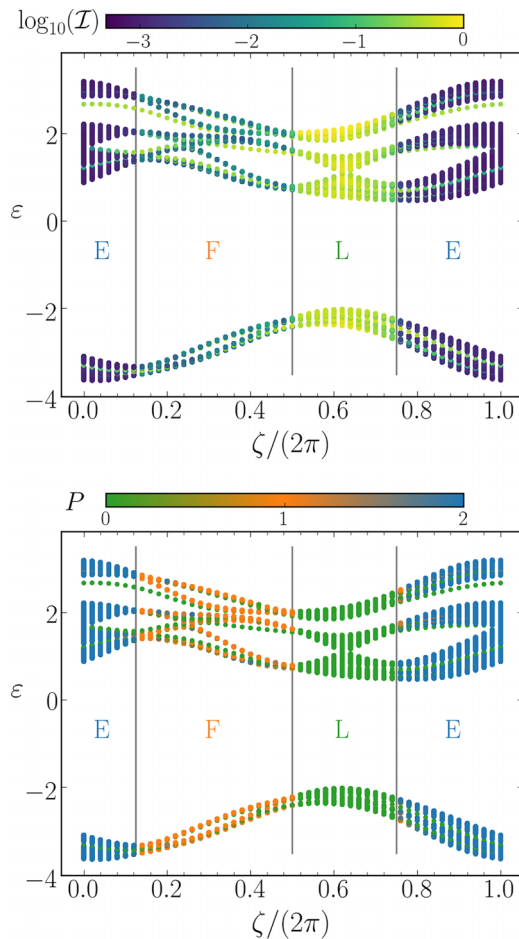


FIG. 5. Testing the trained CNN on the eH model. Upper panel: Logarithm of the IPR while traversing a loop in the parameter space for the full energy spectrum ε . Lower panel: A simple function of the probabilities associated with the different phases obtained from the CNN, $P = 0 \cdot P_L + 1 \cdot P_F + 2 \cdot P_E$, where P_L, P_F, P_E are the probabilities for localized (L), fractal (F), and ergodic (E) phases, respectively [12]. Gray vertical lines mark the transition values. Note the presence of localized states in the ergodic and fractal phases. They originate from the open boundary conditions.

3. Three-dimensional Anderson model (3DA)

A paradigmatic model of metal-insulator transitions is Anderson's model of localization [8–10]. It describes a single-particle hopping on a tight-binding lattice. We consider a three-dimensional (3D) cubic lattice, with homogeneous hopping (set to unity) and a random onsite potential disorder. The onsite potentials are independently and uniformly distributed on $[-W/2, W/2]$. The parameter W defines the disorder strength. The model has been extensively studied numerically [63,64,93–99]. Its phase diagram is shown schematically in Fig. 1(g). At the phase boundary, an Anderson transition, where eigenstates are multifractal, separates localized and extended eigenstates. In Refs. [12,14] the 3DA was recently studied using the CNN, and the authors demonstrated that the CNN can efficiently recognize the metallic and Anderson localized phases.

We simulate a $16 \times 16 \times 8$ lattice and impose periodic boundary conditions. We increase W through the range $[0,35]$

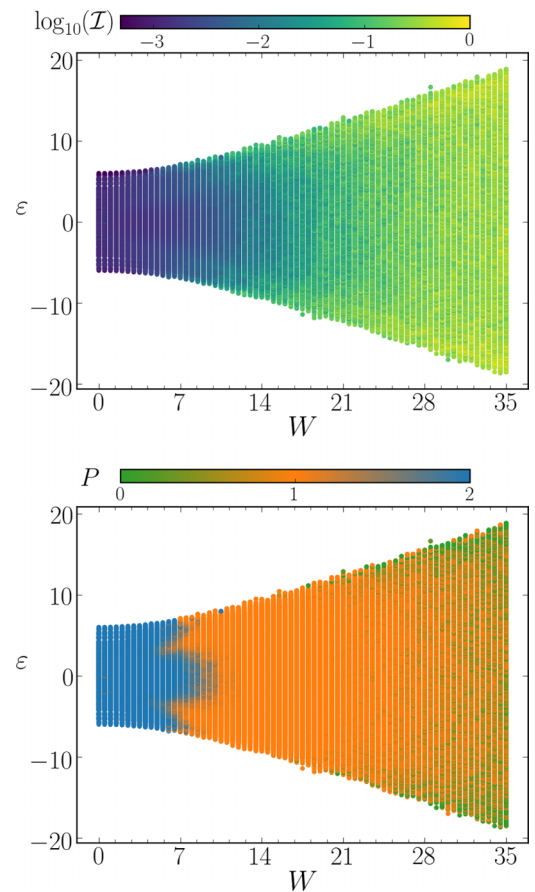


FIG. 6. Testing the trained CNN on the 3DA model. Upper panel: The logarithm of the IPR for different disorder strengths W and for the full energy spectrum ε . Lower panel: Probability P for the different phases obtained using CNN. Here, P is defined in the caption of Fig. 5. Note the very strong finite-size effects (orange colored points in the lower panel) attributed to the small system size of $16 \times 16 \times 8$ yielding $N = 2048$.

in steps of 0.5. For each W , we generate a single realization and compute the entire spectrum by exact diagonalization. While the lattice has $N = 2048$ states, which matches the input layer of the CNN, we need to flatten the data. In doing so, some spatial information is lost.

In Fig. 6 we show the results. The CNN correctly recognizes the ergodic phase for $W < 7$, however, it incorrectly classifies most of the states for $7 < W < 30$ as being fractal and only at the stronger disorder $W > 30$ the eigenstates are gradually identified as localized.

One possible explanation for this failure is that it is a finite-size effect. At the band center $\varepsilon = 0$, the Anderson transition occurs at a critical disorder $W_c = 16.54 \pm 0.01$ [97]. Estimates for the correlation and localization lengths for the band center have been tabulated in Ref. [95]. The correlation length is approximately 1 lattice spacing for $W = 10$ and the localization length is approximately 2 lattice spacings for $W = 30$. Given the results above for the gRP model and the dimensions of the systems simulated here, an explanation in terms of a finite-size effect is plausible. Another possibility is that the failure is due to the loss of spatial information that results from the flattening of the data that is dictated by the

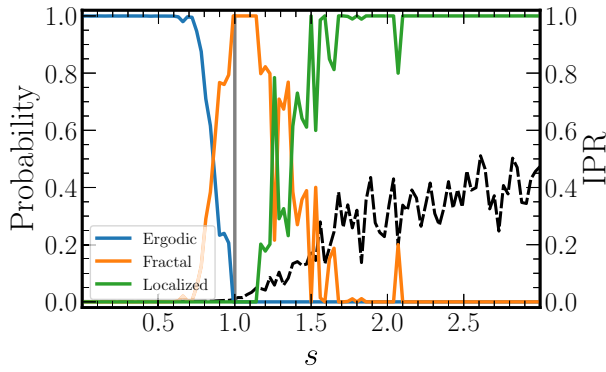


FIG. 7. Testing the trained CNN on the PLBM model. The probabilities of each of the phases are shown as well as the average IPR (black dashed line). The gray vertical line marks the analytical value s_L for the transition from extended to localized states.

gRP model, which was used to train the CNN, and for which there is no concept of a spatial lattice.

4. Power-law random banded matrices (PLBM)

The power-law random banded matrices [100] that we consider are real symmetric matrices with Gaussian distributed elements with zero mean and variances

$$\sigma_d^2 = \langle H_{nn}^2 \rangle = \frac{1}{2N}, \quad \sigma_{off}^2 = \langle H_{nm}^2 \rangle = \frac{1}{4N} a^2(|n-m|). \quad (5)$$

Here, N is the matrix dimension and $a(|n-m|)$ is a function of the distance $r = |n-m|$ from the diagonal, which at large distances decreases according to a power law $a(r) \sim r^{-s}$ for $r \gg 1$ with $s \geq 0$. We adopt the function $a(r)$ introduced in the original work on power-law random banded matrices [100],

$$a(r) = \begin{cases} 1, & r \leq b \\ (\bar{r}/b)^{-s}, & r > b \end{cases} \quad (6)$$

where we define $\bar{r} = \min(r, N-r)$ to reduce boundary effects and b is an additional parameter. It was shown analytically [100] that for $b \gg 1$ the model exhibits a phase transition from extended to localized eigenstates as a function of s at $s = 1$. However, the case $b = 1$ shows an anomalously large critical region around the transition point $s_L = 1$ [101] where multifractal eigenstates persist up to extremely large system sizes. We focus on that case.

We test the CNN as follows. We increase s through the range $[0, 3]$ in steps 0.03. For each s we generate five realizations, extract a single eigenstate closest to the band center, and input them in turn to the CNN. We then average the probabilities output by the CNN over the five realizations. The probabilities for each of the phases and the average IPR are shown in Fig. 7. In good agreement with other numerical results [101] the CNN recognizes the ergodic and localized phases, with the intermediate fractal regime.

5. Anderson model on random graphs (ARG)

The Anderson model on random graphs (ARG) has been extensively studied recently [102–112]. Here we consider two types of graphs: The first is a variant of the celebrated

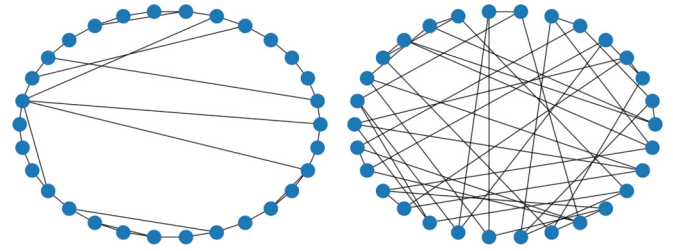


FIG. 8. Examples of the small-world network with $p = 0.25$ (left) and an RRG (right), both have 30 nodes.

small-world networks [113–116] and the second is the random regular graph (RRG). The corresponding Hamiltonian describing a single particle on a tight-binding lattice with onsite potential disorder, where the lattice is defined by the choice of the graph, can be written as

$$H\psi_n = \varepsilon_n\psi_n + \sum_m A_{nm}\psi_m. \quad (7)$$

Here ε_n is the onsite potential and A is the adjacency matrix of the random graph with hopping set to 1. For the small-world network [116] we use the nearest-neighbor hopping and add additional long-range connections among two random sites with probability p , so that the average node degree is $2 + 2p$. Note that the case $p = 0$ corresponds to the 1D Anderson model. For the RRG the node degree is a fixed number v , which we choose to be $v = 3$. In order to compare our results to the literature, we use Gaussian onsite potentials with zero mean and variance W for the small-world network and a uniform distribution $\varepsilon_n \in [-W/2, W/2]$ for the RRG. The random graphs were generated by the NETWORKX library [117]. Examples of random graphs that we considered are shown in Fig 8.

The exact phase diagrams of these models continue to be a subject of intense debate. For the case of the Anderson model on the RRG it has been unequivocally established that for sufficiently strong disorder $W_L \approx 18.17$, Anderson localization occurs. Several works have reported the existence of a nonergodic extended (multifractal) phase for intermediate disorder strengths [102–104] as sketched in Fig. 1(d), while others [105–107, 118] argue that any nonergodic behavior is due to strong finite-size effects, whose scale diverges exponentially at both sides of the transition. It is worth mentioning also the supposition stated in Ref. [108] that there is no ergodic but only a fractal phase. For the small-world networks the critical properties depend on two length scales that diverge with different critical exponents, which leads to strong finite-size effects [109–111]. Here we are interested in relatively small system sizes where all studies agree that there is a range of disorder strength exhibiting a regime of multifractal states.

In Fig. 9 we show the results of testing the CNN with the wave functions of the Anderson model on random graphs. We increase W through $[0, 3]$, $[0, 7]$, $[0, 30]$, in steps of 0.025, 0.05, 0.3 for Figs. 9(a), 9(b), 9(c), respectively. For each W , we generate 50 realizations of the disorder. For each realization we input the wave function with energy closest to $\varepsilon = 0$ to the CNN. We then average the probabilities

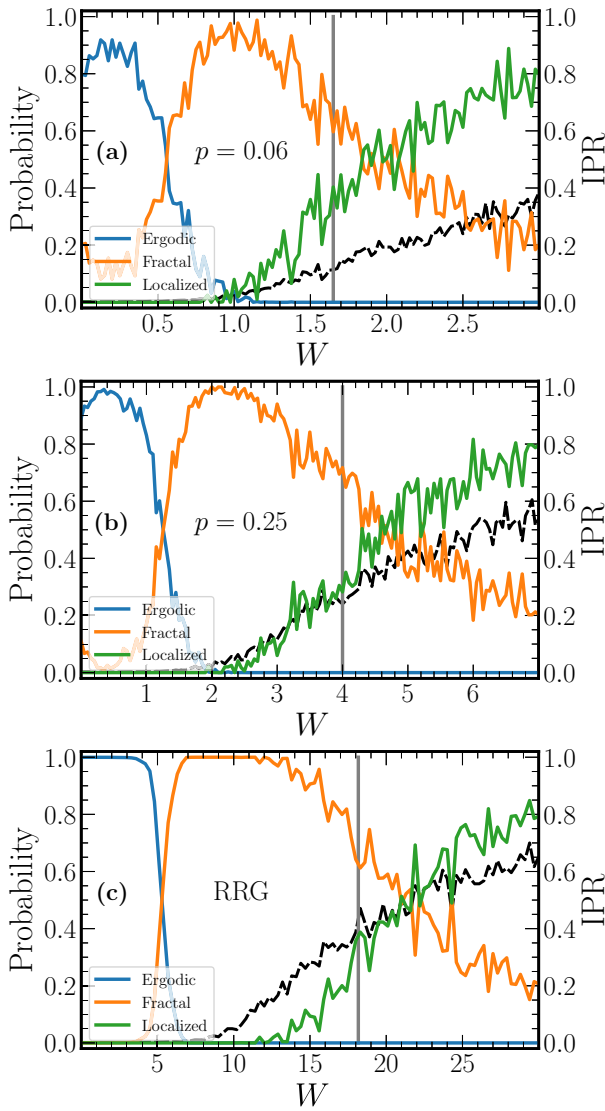


FIG. 9. Testing the trained CNN for the Anderson model on the small-world network with $p = 0.06$, 0.25 and the RRG are shown in (a), (b), and (c), respectively. The probabilities of each of the phases are shown as well as the average IPR (black dashed line) as a function of the onsite disorder strength W . The gray vertical line indicates the numerical value for the Anderson transition W_L .

output by the CNN over the 50 realizations. The values of the disorder where the Anderson transition occurs are given by $W_L \approx 1.65, 4.0$ for the small-world networks [111] and $W_L \approx 18.17$ for the RRG. These values are indicated by gray vertical lines. The CNN identifies ergodic, fractal, and localized phases. Note the similarity of the results obtained for different random graphs. The ergodic transition is found at small disorder strengths of about $W \sim 0.3W_L$, whereas the Anderson transition is blurred. Identifying the crossing point of the probabilities P_F and P_L with the Anderson localization transition yields that the CNN overestimates the value of W where it occurs. There are several possible explanations for this discrepancy: (i) the criterion $P_F = P_L$ overestimates the transition, which has already been seen in the case of the AAH model, or (ii) due to the finite-size effects the multifractal

regime extends further into the localized phase (see Fig. 14 of Ref. [111]).

6. Mass deformed Sachdev-Ye-Kitaev model (SYK)

A modification of the Sachdev-Ye-Kitaev (SYK) model [119–121] was introduced recently, called *mass-deformed SYK model* [122–127]. It is a many-body model of an even number N_M of interacting Majorana fermions $\hat{\chi}_i$, $i = 1, \dots, N_M$, which obey the Clifford algebra $\{\hat{\chi}_i, \hat{\chi}_j\} = \delta_{ij}$. The Hamiltonian comprises one-body and two-body parts, both being all-to-all connected,

$$\begin{aligned} \hat{\mathcal{H}} &\equiv \frac{2}{\sqrt{N_M}} \hat{\mathcal{H}}_4 + \kappa \hat{\mathcal{H}}_2, \\ \hat{\mathcal{H}}_4 &= - \sum_{i < j < k < l} J_{ijkl} \hat{\chi}_i \hat{\chi}_j \hat{\chi}_k \hat{\chi}_l, \\ \hat{\mathcal{H}}_2 &= i \sum_{i < j} J_{ij} \hat{\chi}_i \hat{\chi}_j, \end{aligned} \quad (8)$$

where the coupling constants J_{ijkl} and J_{ij} are Gaussian distributed with zero mean and variances $6/N_M^3$ and $1/N_M$, respectively. The Hamiltonian does not preserve the number of particles but their parity, yielding $N = 2^{N_M/2-1}$ for the dimension of the relevant Hilbert space.

The model is analytically solvable in the thermodynamic limit [126] and the analytical predictions have been verified numerically in finite systems [126–128]. In Ref. [126] the authors identified four regimes in the phase diagram of the model¹ that can be distinguished by the localization properties of the eigenstates in the Fock space, when studied in the eigenbasis of the one-body term. In regime I ($\kappa < \kappa_1$) the eigenstates are ergodically extended over the full Fock space. In regimes II ($\kappa_1 < \kappa < \kappa_2$) and III ($\kappa_2 < \kappa < \kappa_3$) the eigenstates are still ergodically extended, implying, in particular, that they do not have any fractal properties as extensively discussed in Ref. [127], but their extension is over energy shells whose dimension still scales exponentially in N . In regime IV ($\kappa > \kappa_3$) the eigenstates are localized in the Fock space. Interestingly, as a consequence of the one-body term being all to all, the eigenstates are fully delocalized in all regimes in the computational basis implying that localization properties *must* be studied in the eigenbasis defined by the eigenstates of the one-body term (see Appendix C for further details). This is confirmed by the procedure used in this work. The values of the regime boundaries are $\kappa_1 = \sqrt{(N_M - 2)(N_M - 3)/(2N_M^3)}$, $\kappa_2 = \sqrt{N_M} \kappa_1$, and $\kappa_3 = Z/\sqrt{8\rho} \mathcal{W}(2\sqrt{\pi}Z)$, with $\rho = \binom{N_M}{4}$, $Z = \binom{N_M/2}{4}$, and $\mathcal{W}(x)$ denoting the Lambert function. The schematic phase diagram is shown in Fig. 1(e).

We test the CNN on 5 eigenfunctions expanded in the eigenbasis of $\hat{\mathcal{H}}_2$ for each value of $\kappa \in [0.1, 1000]$ on a logarithmic scale for 200 values of κ . We use $N_M = 24$, so that the

¹Note that our notation differs from that used in Ref. [126]. It is the same as in Ref. [128]. The mapping between the two notations is given in Appendix A of Ref. [128].

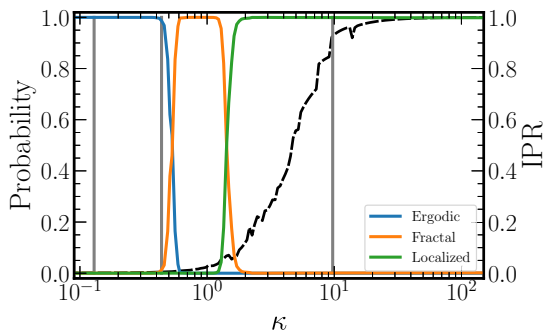


FIG. 10. Testing the trained CNN on the mass-deformed SYK model. The probabilities of each of the phases are shown as well as the average IPR (dashed black line) as a function of the parameter κ , which is plotted on a logarithmic scale. The gray vertical lines indicate the analytical values κ_1 , κ_2 , κ_3 for the transitions.

Hilbert space has dimension 2048. As can be seen in Fig. 10, the CNN identifies three distinct phases, ergodic, fractal, and localized. It should be stressed once again that the model *does not* have a genuine fractal phase [127]. Nevertheless, it is reasonable to expect that the CNN, once required to classify the intermediate regime and knowing only about the existence of three phases, finds highest similarity of it with the previously learned fractal phase of the gRP model. More quantitatively, the ergodic transition is found at $\kappa_E \approx 0.5$, which is slightly larger than the value of κ_2 , while the Anderson transition is observed at $\kappa_L \sim 2$ which underestimates the analytical value. However, it is interesting to note that the values κ at the transitions κ_E and κ_L identified by the CNN coincide well with those where another quantity of interest in the study of many-body systems, namely, the *adiabatic gauge potential*, which is related to the *fidelity susceptibility* [129–132], shows maximally chaotic behavior [128], the onset being close to κ_E , whereas it exhibits a peak at κ_L . Similar behavior is observed in the gRP model [87,133].

IV. DISCUSSION AND CONCLUSIONS

In this work we have trained a CNN to identify ergodic, fractal, and localized states in the gRP model. The model was chosen for its simplicity, one of the simplest models displaying localization, and because its phase diagram is known exactly [75]. Moreover, in the fractal phase of the gRP model the states have a fractal dimension given as $2 - \gamma$, implying that we use in the training set for the fractal phase states with fractal dimensions between 0 and 1. In contrast, generally, models may exhibit a (multi)fractal phase (or regime) consisting of multifractal states, whose dimensions are peaked around a certain value. We expect that the network trained on such models has difficulties in identifying fractal states with dimensions close to either 0 or 1.

The main result of this work is a demonstration of the generalization capability. The same network was applied to diverse systems: Single-particle models, including random-matrix models, random graphs, and a many-body quantum system. The training set of only 500 states per class and testing sets of several states per test point suffice to discern the three

phases for a fixed system size, making the network efficient. Thus, it can be effectively used also for full spectra.

On the other hand, we have found that improving the precision is not straightforward. For example, using larger training sets (5000 states for each class in the training) or larger system sizes ($N = 4096, 8192$) produces results similar to the ones presented. Thus, more detailed analyses such as finite-size scaling are hard to perform. Steps in this direction are outlined in Appendixes C and D. From those results we can see that an estimate of the critical point is feasible, however, the universal features such as critical exponents require much larger amount of data as compared to what we have used in this work. A similar conclusion was reported recently in an attempt to use a neural network for the detection of the many-body localization transition [53].

Note that in our approach we use relatively small system sizes, implying that it may exhibit strong finite-size effects. This is particularly evident in the cases of the 3DA model, the PLBM models, and the Anderson model on random graphs. Yet, comparable sizes are frequently used in numerical simulations, so especially for these the results deduced from the CNN provide relevant information on wave-function localization properties. Furthermore, in all these models the characteristic length scales corresponding to the three phases become large in a non-negligible range around the parameter values where the transition takes place, implying that there the interpretation of phases or locating the transition can be complicated like, e.g., for the Anderson model on an RRG. Nevertheless, our approach provides a good qualitative agreement with other procedures as outlined in this work.

Note added. Recently, we became aware of a work [134], where the authors use an artificial neural network to study delocalized, multifractal, and localized phases in a variant of the AAH model, the long-range AAH model. They construct a multilayer perceptron, which is a dense neural network that is trained on the long-range AAH and tested on the AAH and the training model itself, implying that our study of generalization capabilities is obviously more extensive. Another disadvantage as compared to the gRP model is that in the long-range AAH the multifractal or localized phases coexist with the delocalized one. This is a drawback for preparing high-quality training sets. Above all, in distinction to the AAH, the gRP model is an all-to-all model, which thus can be applied to various types of systems, including single-particle systems like the AAH, as demonstrated in this work.

ACKNOWLEDGMENTS

A.A., B.D., D.R., and T.Č. acknowledge financial support from the Institute for Basic Science (IBS) in the Republic of Korea through the Project No. IBS-R024-D1. T.O. and K.S. were supported by JSPS KAKENHI Grant No. 19H00658. T.O. was supported by JSPS KAKENHI Grant No. 22H05114.

APPENDIX A: CNN HYPERPARAMETERS

As described in Sec. III B we use a simple CNN which we constructed using KERAS [135] as the front end and TENSORFLOW [136] as the back end. The CNN parameters related to

TABLE I. The structure and hyperparameters of the CNN used. The total number of trainable parameters is 245 808.

Layer class	Filters	Kernel Size	Stride	Output size
Input				(1, 2048)
Convolutional 1	64	256	1	(64, 1793)
Pooling 1		2	2	(64, 896)
Convolutional 2	16	128	1	(16, 769)
Pooling 2		2	2	(16, 384)
Dense 1				(16)
Dense 2				(3)

its structure, such as the number of different layers, their sizes, etc., are known as *hyperparameters*. We have tried several CNN architectures and found that the one given in Table I performed best. The convolutional layers apply filtering on the data using a number of *kernels* of given sizes. The number of kernels that have been employed is provided in the table under filters. The stride is a translation for which the kernel moves along the data. The output size corresponds to that of the data in each of the layers. In the pooling layers we have used max pooling. No padding was used. To minimize the categorical cross entropy we have used the Adam optimizer.

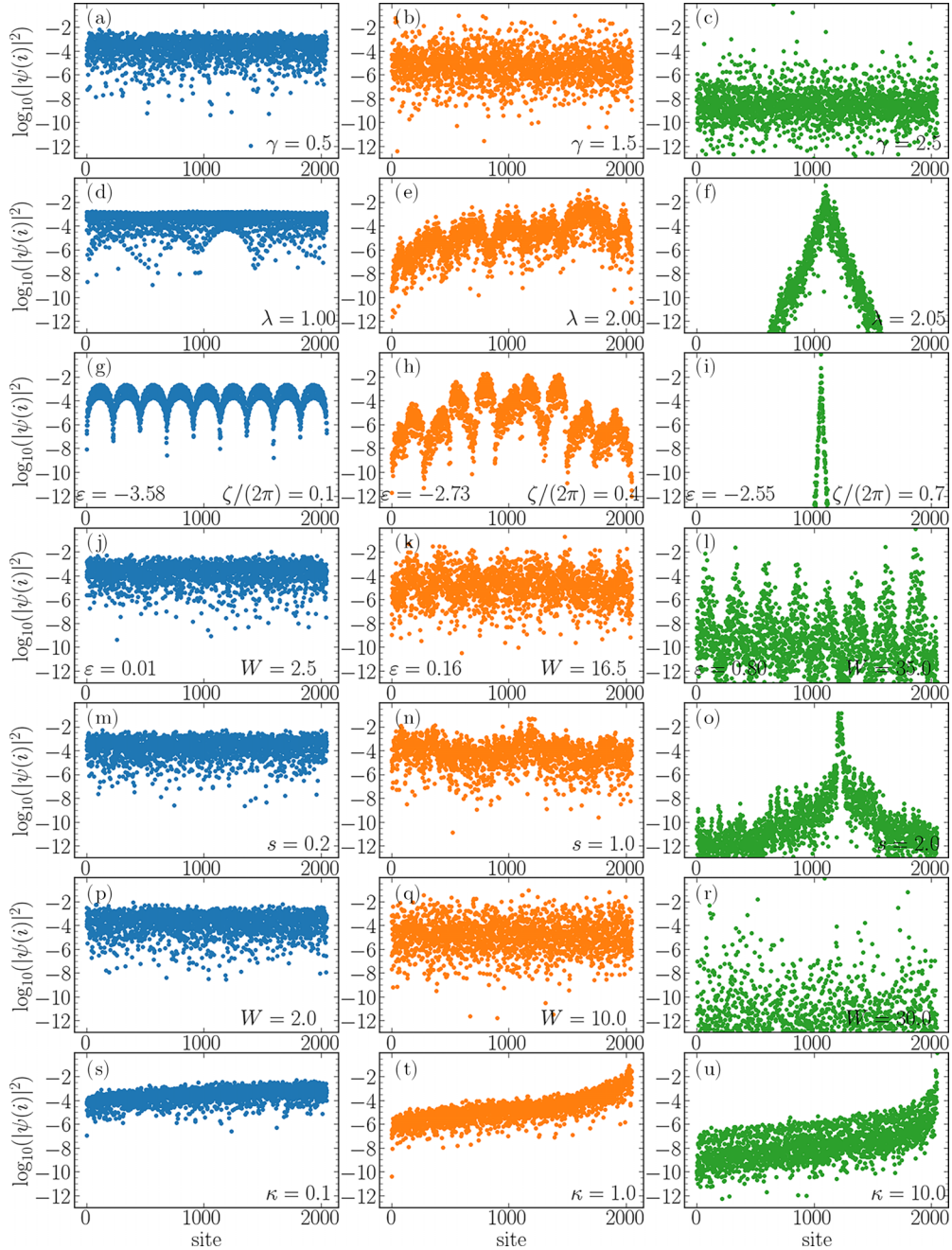


FIG. 11. Examples for the eigenvector occupations for (a)–(c) the gRP model, (d)–(f) the AAH model, (g)–(i) the eH model, (j)–(l) the 3DA model, (m)–(o) the PLBM model, (p)–(r) the Anderson model on RRG, and (s)–(u) the mass-deformed SYK model. Note the logarithm scale for the y axis. The values of the parameters used are given in panels. See main text for details.

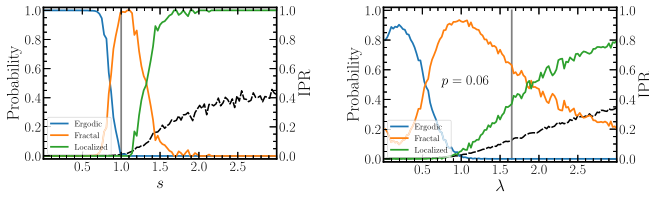


FIG. 12. Generalization capability on the PLBM (left panel) and the Anderson model on a small-world network with $p = 0.06$ (right panel) on a larger data set reduces fluctuations. We used 50 and 500 realizations per each point for left and right panel, respectively.

APPENDIX B: EXAMPLES OF WAVE FUNCTIONS

Examples of the eigenvector occupations, the raw data that we use as input for CNN, are shown in Fig. 11. Note that these are the original states that are not centered. The left, middle, and right panels display states that were classified by the CNN as extended, fractal, and localized, respectively. Results are shown for (a)–(c) the gRP model, (d)–(f) the AAH model, (g)–(i) the eH model, (j)–(l) the 3DA model, (m)–(o) the PLBM model, (p)–(r) the Anderson model on RRGs, and (s)–(u) the mass-deformed SYK model. For the gRP, PLBM, and Anderson models on RRGs we used the computational basis. In the single-particle models (AAH, eH, and 3DA) the computational basis is the real-space basis. For the mass-deformed SYK model we use the eigenbasis of $\hat{\mathcal{H}}_2$. The probability distributions of the gRP and PLBM were studied recently in Refs. [83,88]. In the ergodic regime they are given by the Porter-Thomas distribution [137–139]. For the eH and 3DA models we show states from the lowest band (see Fig. 5) and from the vicinity of the band center, respectively.

APPENDIX C: ADDITIONAL RESULTS

Here we show some additional results that address four points: (i) larger testing data set, (ii) larger training data set, (iii) possible further training improvements, and (iv) the importance of the choice of basis.

Especially in the cases of the PLBM and the Anderson models on random graphs the testing data averaged over 5 and 50 states, as shown in Figs. 7 and 9, respectively, fluctuate strongly. In Fig. 12 we demonstrate that increasing the testing data amount by a factor of 10 reduces the fluctuations considerably. However, the transition values of the associated parameter do not change up to the currently achievable precision.

Regarding the supervised machine learning the amount of training data can be important since it can improve the performance of the trained network. In the main text we have

trained the network on only 500 images per phase. Here we used the same architecture as given in Appendix A and trained the network on 10 times more data, again based on the gRP model. The results of testing are shown in Fig. 13. We can see that the performance of the CNN can improve slightly when tested on the same gRP model. However, the generalization capability does not seem to improve.

Aside from increasing the training data set we considered other activation functions and/or increasing the validation set to achieve improvement of training. Exponential rectified linear unit (eLU) as an activation function has been shown to produce better accuracy and to converge faster [140] compared to other activation functions including ReLU, which is used in this work. During multiple training attempts with eLU we observe on average (i) reduced validation loss with similar training loss, thus reducing overfitting as compared to the training using the ReLU activation and (ii) faster convergence, while the accuracy and validation accuracy is similar to that achieved with ReLU. The loss (validation loss) is the average cross-entropy loss S_c between the predictors and labels L_v in the training (validation) set during the training, $S_c = -\sum_v L_v \ln(P_v)$, with the predictors being the probabilities P_v predicted by the CNN, where $v = E, F, L$ denote the output classes. In Fig. 14 we compare three training cases for different activation functions and ratios of the validation data set. As a test case, denoted ReLU, we consider the same validation ratio of 10% and ReLU activation function as in all the other cases in this work. The case called eLU shows results using the eLU activation function and the validation ratio of 10%. The last case denoted val shows results obtained using ReLU activation and an increased validation ratio of 20%. For each of the three cases we performed 70 different training attempts. Compared to the hyperparameters given in Table I, we have changed only the number of filters in the convolutional 1 layer, from 64 down to 16, thus reducing the total number of trainable parameters to 135 216 and speeding up the training. We have used 500 wave functions per phase for the total training set. In Fig. 14 we show one testing result for each of the three cases considered. Similar performance is observed for all three cases.

Finally, we note that the localization properties depend crucially on the choice of basis. We show exemplary for the mass-deformed SYK model in Fig. 15 the CNN testing results when the chosen basis is the computational basis of the full Hamiltonian (8) rather than the eigenbasis of $\hat{\mathcal{H}}_2$.

APPENDIX D: COMBINING MODELS

For a more systematic study, improved accuracy, and estimates of the errors of the CNN predictions we employ

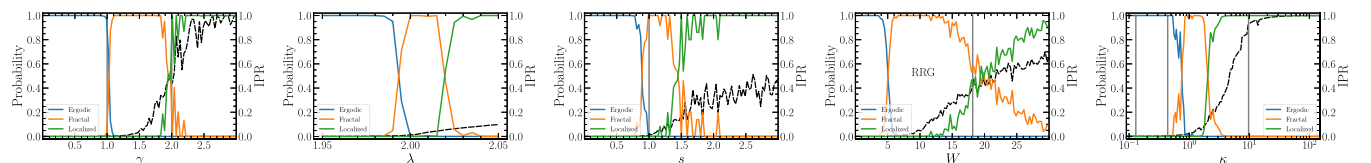


FIG. 13. Training the CNN on a larger data set (5000 eigenfunctions per phase) slightly improves testing results on the same model, as seen in the leftmost panel. However, the generalization capability does not necessarily improve as can be seen in the cases of (from left to right) AAH model, PLBM, Anderson model on RRG, and the mass-deformed SYK model. The same testing data as in the main text were used.

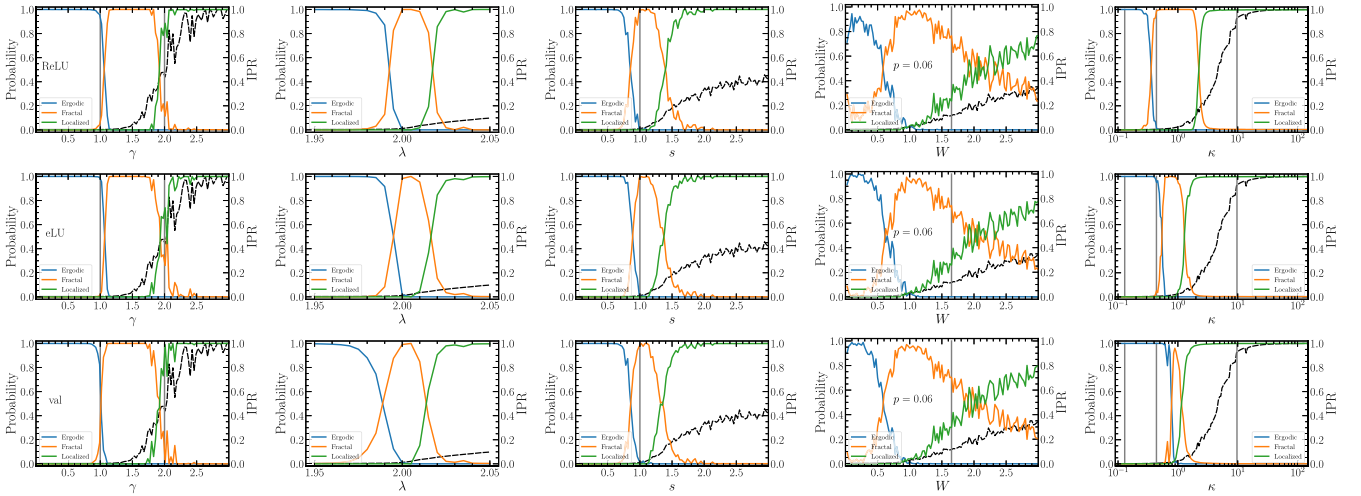


FIG. 14. Testing results after training the CNN on 500 eigenfunctions per phase for three different cases, ReLU, eLU, and increased validation val, from top to bottom, respectively. The testing is performed on gRP model, AAH model, PLBM, Anderson model on small-world network with $p = 0.06$, and the mass-deformed SYK model, from left to right, respectively. Except for the small-world network case (where we have rather used 50 out of 500 states per disorder strength as used in Fig. 12) the same testing data as in the main text was used.

a simple scheme of combining multiple trained models g_i , known as bootstrap aggregation [6,141], or in short bagging. To obtain the trained models we prepare a training set consisting of a total of 26 100 wave functions from the gRP model corresponding to the energy closest to 0, each from a different disorder realization. All states are labeled by the class E , F , or L they belong to. For the training we randomly select 5000 wave functions from each of the three classes and perform 70 training procedures. We choose 10 successful trainings for each of the system sizes $N = 2048, 4096, 8192$, thus giving $m = 10$ training models g_1, \dots, g_{10} per system size. The structure and hyperparameters of the CNN used are the same as given in Table I, except the number of filters in the convolutional 1 layer, which was reduced to 16 (thus reducing also the output sizes in the convolutional and pooling layers). Note that due to the overlap in the training data sets, these trained models are partially correlated. Each training model g_i , when applied on the states for one of the systems considered, gives probabilities $P_{E,i}, P_{F,i}, P_{L,i}$ that the state is ergodic, fractal, or localized, respectively. Bagging treats the

training models as independent and averages over the trained models, namely, the prediction \bar{P}_v for the state to be in the v th class ($v = E, F, L$) is $\bar{P}_v = 1/m \sum_{i=1}^m P_{v,i}$.

In Fig. 16 we apply the bagging to the AAH model. We consider three different system sizes and show separate predictors $P_{F,i}$ and the averages \bar{P}_v . Note the different λ ranges used. In Fig. 17 we show the averaged probabilities in one plot. From the crossings of $\bar{P}_E(\lambda_E) = \bar{P}_F(\lambda_L) = \bar{P}_L(\lambda_L)$ we estimate the ergodic and localization transitions to be $\lambda_E = 1.9897, 1.9911, 1.9943$ and $\lambda_L = 2.023, 2.0124, 2.0064$ for the system sizes $N = 2048, 4096, 8192$, respectively.

Alternatively we can estimate the transition points from the condition $P_{F,i} = 0.5$ for each model g_i and obtain the estimates $\langle \lambda_E \rangle = 1.9888 \pm 0.0015, 1.99122 \pm 0.00094, 1.99451 \pm 0.00064$ and $\langle \lambda_L \rangle = 2.0219 \pm 0.0018, 2.0118 \pm 0.0012, 2.00654 \pm 0.00052$ for the system sizes $N = 2048, 4096, 8192$, respectively. We show the results in Fig. 18. From the linear fit and extrapolation we obtain for the transition points the estimates $\lambda_E = 1.9957 \pm 0.0011$ and $\lambda_L = 2.0014 \pm 0.0012$, where the errors are the standard deviations of 5000 transition estimates obtained via Monte Carlo sampling of synthetic data from their averages $\langle \lambda_E \rangle$ and $\langle \lambda_L \rangle$. However, we believe that these errors are underestimated, due to (i) the correlations in the trained models g_i and (ii) using the same testing data on all trained models. Nevertheless, the results are consistent for both transitions, whose values actually coincide, with the known transition [90] $\lambda_L = 2$. Furthermore, we expect that the difference between the crossing and the predicted critical point should be proportional to $N^{-1/\nu}$ and $\nu = 1$ for AAH model. We can see that up to the errors the results are consistent with the expected scaling.

As another example we consider the Anderson model on the small-world network with $p = 0.06$. In Fig. 19 we show the bagging results with the focus on the localization transition. Again we consider three different system sizes and show separate predictors $P_{F,i}$ and the averages \bar{P}_v .

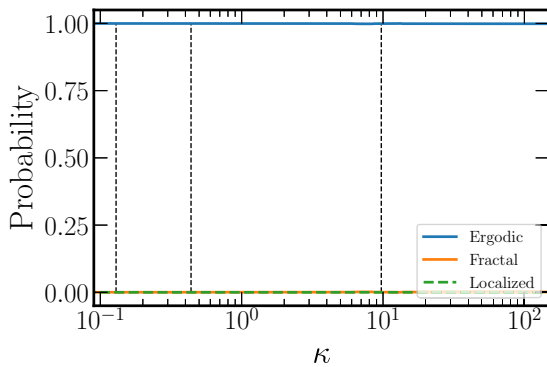


FIG. 15. Testing results of CNN when applied to data from the computational basis of the mass-deformed SYK model, showing complete delocalization.

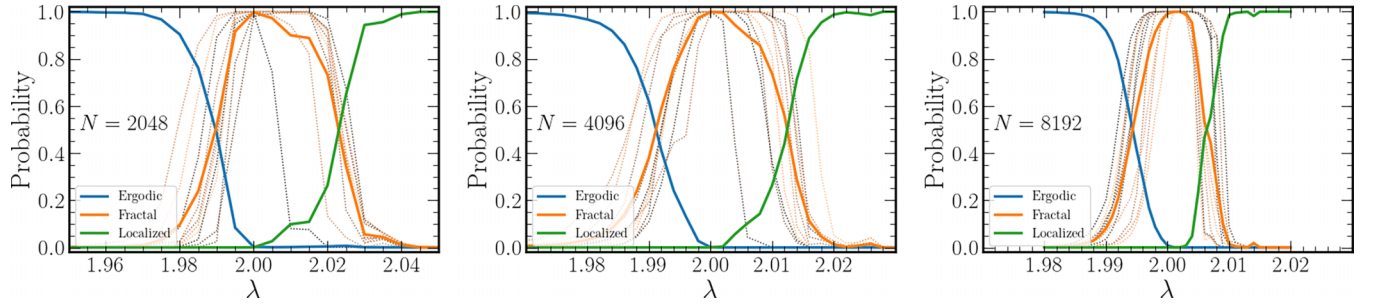


FIG. 16. The averaged probabilities \bar{P}_v for ($v = E, F, L$) for the bagging predictions based on 10 training models for each system size $N = 2048, 4096, 8192$ for the AAH model, while the training was done on the gRP model. With dotted lines we show the 10 probabilities $P_{F,i}$.

From the crossings of $\bar{P}_F(W_L) = \bar{P}_L(W_L)$ we can estimate the localization transition to be $W_L = 2.12, 2.12, 1.98$ for the system sizes $N = 2048, 4096, 8192$, respectively. If alternatively we estimate $W_{L,i}$ for each of the trained models g_i from the condition $P_{F,i} = 0.5$ we obtain $\langle W_L \rangle = 2.13 \pm 0.06, 2.13 \pm 0.04, 1.99 \pm 0.05$, where we give mean and standard error of $W_{L,i}$ for the system sizes $N = 2048, 4096, 8192$, respectively.

APPENDIX E: CRITERIA FOR ERGODICITY AND LOCALIZATION FOR FULL RANDOM MATRICES

The rule of thumb criteria for ergodicity and localization in dense matrices are based on the following sums of the averages of matrix elements [88,89]:

$$S_q(N) = \frac{1}{NA^q} \sum_{n,m=1}^N \langle |H_{nm}|^q \rangle, \quad (\text{E1})$$

where $q = 1, 2$, $A = \sqrt{\langle |H_{nm}|^2 \rangle} = \sigma_d$ and N is the dimension of the matrix. The criteria are as follows:

- (i) The Anderson localization criterion states that when $\lim_{N \rightarrow \infty} S_1(N) < \infty$ the eigenstates are localized.
- (ii) The ergodicity criterion states that when $\lim_{N \rightarrow \infty} S_2(N) \rightarrow \infty$ the eigenstates are ergodic.
- (iii) Additionally, a sufficient condition for full ergodicity is met [89] if $\lim_{N \rightarrow \infty} S_1(N) \rightarrow \infty$, $\lim_{N \rightarrow \infty} S_2(N) \rightarrow \infty$,

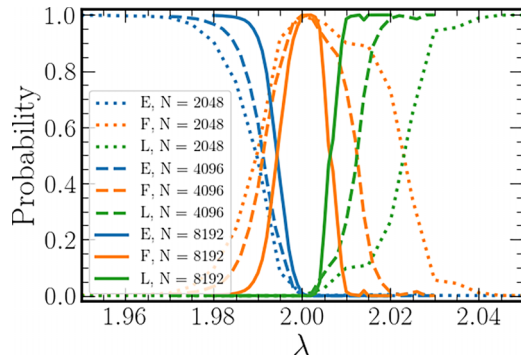


FIG. 17. The averaged probabilities \bar{P}_v for ($v = E, F, L$) for the bagging predictions based on 10 training models for each system size $N = 2048, 4096, 8192$ for the AAH model, while the training was done on the gRP model.

and $\lim_{N \rightarrow \infty} \bar{S}(N) \rightarrow \infty$, where

$$\bar{S}(N) = \frac{(\sum_m \langle |H_{nm}|^2 \rangle_t)^2}{A^2 S_2(N)}, \quad (\text{E2})$$

and the typical value is given as $\langle |H_{nm}|^2 \rangle_t = \exp(\ln \langle |H_{nm}|^2 \rangle)$. For a Gaussian distribution given by

$$\mathcal{P}(x) = \exp[-x^2/(2\sigma^2)]/\sqrt{2\pi\sigma^2}, \quad (\text{E3})$$

the moments can be calculated exactly $\langle |x|^q \rangle = 2^{q/2} \sigma^q \Gamma(q/2 + 1/2)/\sqrt{\pi}$ for $q > -1$, with the gamma

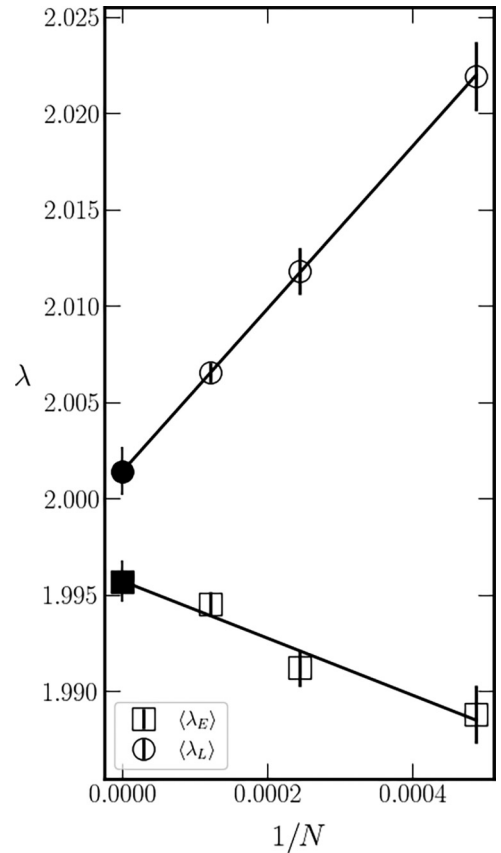


FIG. 18. Estimates of the ergodic (squares) and localization (circles) transitions for the AAH model as function of the inverse of the system size. The extrapolation to the thermodynamic limit gives the estimates for the transition points, marked with full symbols. The errors are estimated via Monte Carlo sampling of synthetic data.

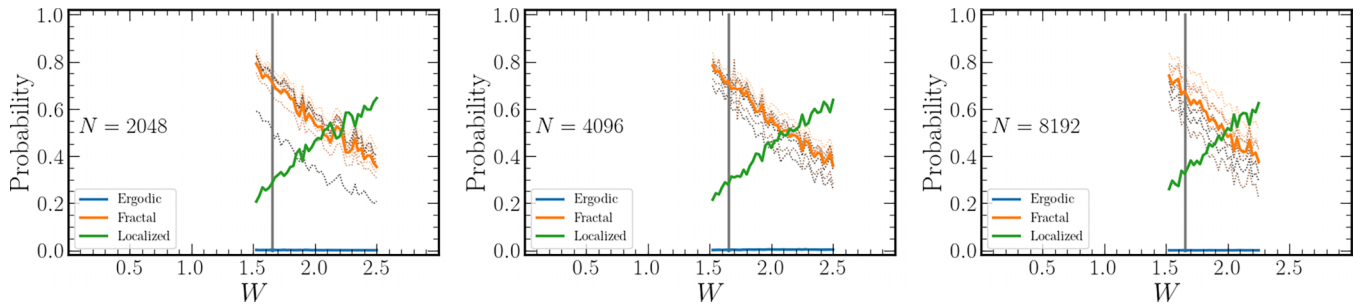


FIG. 19. The averaged probabilities \bar{P}_v for ($\nu = E, F, L$) for the bagging predictions based on 10 training models for each system size $N = 2048, 4096, 8192$ for the Anderson model on the small-world network with $p = 0.06$, while the training was done on the gRP model. With dotted lines we show the 10 probabilities $P_{F,i}$.

function $\Gamma(a) = \int_0^\infty x^{a-1} \exp(-x) dx$, while the typical value of the second moment is $\langle |x|^2 \rangle_t = \sigma^2 / [2 \exp(\gamma_E)]$, with γ_E being the Euler-Mascheroni constant. Note that since $\bar{S} = S_2 / [2 \exp(\gamma_E)]$ for Gaussian distributions, the criterion for full ergodicity coincides with the criterion for ergodicity.

For the gRP model, the variances are given in Eq. (1) and we get $S_1(N) = \sqrt{2/\pi} [1 + 1/\sqrt{2}(N-1)N^{-\gamma/2}]$ and $S_2(N) = 1 + 1/2(N-1)N^{-\gamma}$. Taking the limit $N \rightarrow \infty$ and using the above criteria follows the phase diagram of gRP.

For the PLBM, the variances are given in Eq. (5) and we get $S_1(N) = \sqrt{2/\pi} [1 + 1/\sqrt{2} H_{N,s}]$ and $S_2(N) = 1 + 1/2 H_{N,2s}$, where $H_{N,s}$ are the generalized harmonic numbers. For $s > 1$

we have $\lim_{N \rightarrow \infty} H_{N,s} = \zeta(s)$, where $\zeta(s)$ is the Riemann zeta function, whereas for $0 < s \leq 1$ the limit can be bounded $0 < I_s < \lim_{N \rightarrow \infty} H_{N,s} < 1 + I_s$, with a simple integral $I_s = \int_1^\infty x^{-s} dx$. The integral diverges logarithmically for $s = 1$ whereas $\int_1^N x^{-s} dx = N^{1-s}/(1-s) - 1/(1-s)$ for $0 < s < 1$. The Anderson localization criterion then gives $\lim_{N \rightarrow \infty} S_1(N) = \sqrt{2/\pi} [1 + 1/\sqrt{2} \zeta(s)]$ for $s > 1$, whereas the ergodic criterion is satisfied for $s \leq \frac{1}{2}$. The eigenvector distribution in the region $\frac{1}{2} < s < 1$ was shown [88] to have no anomalous scaling, implying that the fractal dimensions are the same as for the nonbanded random matrices and thus that this phase is ergodic.

- [1] T. M. Mitchell, *Machine Learning* (McGraw-Hill Education, 1997).
- [2] Y. LeCun, Y. Bengio, and G. Hinton, Deep learning, *Nature (London)* **521**, 436 (2015).
- [3] G. Torlai and R. G. Melko, Learning thermodynamics with boltzmann machines, *Phys. Rev. B* **94**, 165134 (2016).
- [4] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, *Nat. Phys.* **13**, 431 (2017).
- [5] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, *Rev. Mod. Phys.* **91**, 045002 (2019).
- [6] P. Mehta, M. Bukov, C.-H. Wang, A. G. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, A high-bias, low-variance introduction to machine learning for physicists, *Phys. Rep.* **810**, 1 (2019).
- [7] T. Ohtsuki and T. Mano, Drawing phase diagrams of random quantum systems by deep learning the wave functions, *J. Phys. Soc. Jpn.* **89**, 022001 (2020).
- [8] P. Anderson, Absence of diffusion in certain random lattices, *Phys. Rev.* **109**, 1492 (1958).
- [9] B. Kramer and A. MacKinnon, Localization: Theory and experiment, *Rep. Prog. Phys.* **56**, 1469 (1993).
- [10] F. Evers and A. D. Mirlin, Anderson transitions, *Rev. Mod. Phys.* **80**, 1355 (2008).
- [11] T. Ohtsuki and T. Ohtsuki, Deep learning the quantum phase transitions in random two-dimensional electron systems, *J. Phys. Soc. Jpn.* **85**, 123706 (2016).
- [12] T. Ohtsuki and T. Ohtsuki, Deep learning the quantum phase transitions in random electron systems: Applications to three dimensions, *J. Phys. Soc. Jpn.* **86**, 044708 (2017).
- [13] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Machine learning quantum phases of matter beyond the fermion sign problem, *Sci. Rep.* **7**, 8823 (2017).
- [14] T. Mano and T. Ohtsuki, Phase diagrams of three-dimensional anderson and quantum percolation models using deep three-dimensional convolutional neural network, *J. Phys. Soc. Jpn.* **86**, 113704 (2017).
- [15] B. L. Altshuler, Y. Gefen, A. Kamenev, and L. S. Levitov, Quasiparticle lifetime in a finite system: A nonperturbative approach, *Phys. Rev. Lett.* **78**, 2803 (1997).
- [16] D. Basko, I. Aleiner, and B. Altshuler, Metal-insulator transition in a weakly interacting many-electron system with localized single-particle states, *Ann. Phys.* **321**, 1126 (2006).
- [17] D. A. Abanin, E. Altman, I. Bloch, and M. Serbyn, Colloquium: Many-body localization, thermalization, and entanglement, *Rev. Mod. Phys.* **91**, 021001 (2019).
- [18] I. V. Gornyi, A. D. Mirlin, and D. G. Polyakov, Interacting electrons in disordered wires: Anderson localization and low- t transport, *Phys. Rev. Lett.* **95**, 206603 (2005).
- [19] V. Oganesyan and D. A. Huse, Localization of interacting fermions at high temperature, *Phys. Rev. B* **75**, 155111 (2007).
- [20] M. Žnidarič, T. Prosen, and P. Prelovšek, Many-body localization in the heisenberg xxz magnet in a random field, *Phys. Rev. B* **77**, 064426 (2008).
- [21] J. H. Bardarson, F. Pollmann, and J. E. Moore, Unbounded growth of entanglement in models of many-body localization, *Phys. Rev. Lett.* **109**, 017202 (2012).
- [22] A. Pal and D. A. Huse, Many-body localization phase transition, *Phys. Rev. B* **82**, 174411 (2010).

- [23] M. Serbyn, Z. Papić, and D. A. Abanin, Local conservation laws and the structure of the many-body localized states, *Phys. Rev. Lett.* **111**, 127201 (2013).
- [24] B. Bauer and C. Nayak, Area laws in a many-body localized state and its implications for topological order, *J. Stat. Mech.: Theory Exp.* (2013) P09005.
- [25] D. A. Huse, R. Nandkishore, and V. Oganesyan, Phenomenology of fully many-body-localized systems, *Phys. Rev. B* **90**, 174202 (2014).
- [26] V. Ros, M. Müller, and A. Scardicchio, Integrals of motion in the many-body localized phase, *Nucl. Phys. B* **891**, 420 (2015).
- [27] D. J. Luitz, N. Laflorencie, and F. Alet, Many-body localization edge in the random-field heisenberg chain, *Phys. Rev. B* **91**, 081103(R) (2015).
- [28] R. Mondaini and M. Rigol, Many-body localization and thermalization in disordered hubbard chains, *Phys. Rev. A* **92**, 041601(R) (2015).
- [29] M. Yan, H.-Y. Hui, M. Rigol, and V. W. Scarola, Equilibration dynamics of strongly interacting bosons in 2d lattices with disorder, *Phys. Rev. Lett.* **119**, 073002 (2017).
- [30] M. Yan, H.-Y. Hui, and V. W. Scarola, Dynamics of disordered states in the bose-hubbard model with confinement, *Phys. Rev. A* **95**, 053624 (2017).
- [31] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, Quantum chaos challenges many-body localization, *Phys. Rev. E* **102**, 062144 (2020).
- [32] J. Šuntajs, J. Bonča, T. Prosen, and L. Vidmar, Ergodicity breaking transition in finite disordered spin chains, *Phys. Rev. B* **102**, 064207 (2020).
- [33] P. Sierant, D. Delande, and J. Zakrzewski, Thouless time analysis of anderson and many-body localization transitions, *Phys. Rev. Lett.* **124**, 186601 (2020).
- [34] P. Sierant, M. Lewenstein, and J. Zakrzewski, Polynomially filtered exact diagonalization approach to many-body localization, *Phys. Rev. Lett.* **125**, 156601 (2020).
- [35] D. Abanin, J. Bardarson, G. De Tomasi, S. Gopalakrishnan, V. Khemani, S. Parameswaran, F. Pollmann, A. Potter, M. Serbyn, and R. Vasseur, Distinguishing localization from chaos: Challenges in finite-size systems, *Ann. Phys.* **427**, 168415 (2021).
- [36] A. Morningstar, L. Colmenarez, V. Khemani, D. J. Luitz, and D. A. Huse, Avalanches and many-body resonances in many-body localized systems, *Phys. Rev. B* **105**, 174205 (2022).
- [37] F. Evers, I. Modak, and S. Bera, Internal clock of many-body delocalization, *Phys. Rev. B*, **108**, 134204 (2023).
- [38] J. Dieplinger and S. Bera, Finite-size prethermalization at the chaos-to-integrable crossover, *Phys. Rev. B* **107**, 224207 (2023).
- [39] D. M. Long, P. J. D. Crowley, V. Khemani, and A. Chandran, Phenomenology of the prethermal many-body localized regime, *Phys. Rev. Lett.* **131**, 106301 (2023).
- [40] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, Observation of many-body localization of interacting fermions in a quasirandom optical lattice, *Science* **349**, 842 (2015).
- [41] J.-y. Choi, S. Hild, J. Zeiher, P. Schauß, A. Rubio-Abadal, T. Yefsah, V. Khemani, D. A. Huse, I. Bloch, and C. Gross, Exploring the many-body localization transition in two dimensions, *Science* **352**, 1547 (2016).
- [42] J. Smith, A. Lee, P. Richerme, B. Neyenhuis, P. W. Hess, P. Hauke, M. Heyl, D. A. Huse, and C. Monroe, Many-body localization in a quantum simulator with programmable random disorder, *Nat. Phys.* **12**, 907 (2016).
- [43] P. Bordia, H. P. Lüschen, S. S. Hodgman, M. Schreiber, I. Bloch, and U. Schneider, Coupling identical one-dimensional many-body localized systems, *Phys. Rev. Lett.* **116**, 140401 (2016).
- [44] P. Bordia, H. Lüschen, S. Scherg, S. Gopalakrishnan, M. Knap, U. Schneider, and I. Bloch, Probing slow relaxation and many-body localization in two-dimensional quasiperiodic systems, *Phys. Rev. X* **7**, 041047 (2017).
- [45] H. P. Lüschen, P. Bordia, S. Scherg, F. Alet, E. Altman, U. Schneider, and I. Bloch, Observation of slow dynamics near the many-body localization transition in one-dimensional quasiperiodic systems, *Phys. Rev. Lett.* **119**, 260401 (2017).
- [46] P. Roushan, C. Neill, J. Tangpanitanon, V. M. Bastidas, A. Megrant, R. Barends, Y. Chen, Z. Chen, B. Chiaro, A. Dunsworth, A. Fowler, B. Foxen, M. Giustina, E. Jeffrey, J. Kelly, E. Lucero, J. Mutus, M. Neeley, C. Quintana, D. Sank *et al.*, Spectroscopic signatures of localization with interacting photons in superconducting qubits, *Science* **358**, 1175 (2017).
- [47] C. Neill, P. Roushan, K. Kechedzhi, S. Boixo, S. V. Isakov, V. Smelyanskiy, A. Megrant, B. Chiaro, A. Dunsworth, K. Arya, R. Barends, B. Burkett, Y. Chen, Z. Chen, A. Fowler, B. Foxen, M. Giustina, R. Graff, E. Jeffrey, T. Huang *et al.*, A blueprint for demonstrating quantum supremacy with superconducting qubits, *Science* **360**, 195 (2018).
- [48] K. Xu, J.-J. Chen, Y. Zeng, Y.-R. Zhang, C. Song, W. Liu, Q. Guo, P. Zhang, D. Xu, H. Deng, K. Huang, H. Wang, X. Zhu, D. Zheng, and H. Fan, Emulating many-body localization with a superconducting quantum processor, *Phys. Rev. Lett.* **120**, 050507 (2018).
- [49] M. Rispoli, A. Lukin, R. Schittko, S. Kim, M. E. Tai, J. Leonard, and M. Greiner, Quantum critical behaviour at the many-body localization transition, *Nature (London)* **573**, 385 (2019).
- [50] T. Kohlert, S. Scherg, X. Li, H. P. Lüschen, S. Das Sarma, I. Bloch, and M. Aidelsburger, Observation of many-body localization in a one-dimensional system with a single-particle mobility edge, *Phys. Rev. Lett.* **122**, 170403 (2019).
- [51] Q. Guo, C. Cheng, Z.-H. Sun, Z. Song, H. Li, Z. Wang, W. Ren, H. Dong, D. Zheng, Y.-R. Zhang, R. Mondaini, H. Fan, and H. Wang, Observation of energy-resolved many-body localization, *Nat. Phys.* **17**, 234 (2021).
- [52] F. Schindler, N. Regnault, and T. Neupert, Probing many-body localization with neural networks, *Phys. Rev. B* **95**, 245134 (2017).
- [53] H. Théveniaut and F. Alet, Neural network setups for a precise detection of the many-body localization transition: Finite-size scaling and limitations, *Phys. Rev. B* **100**, 224202 (2019).
- [54] W.-J. Rao, Machine learning for many-body localization transition*, *Chin. Phys. Lett.* **37**, 080501 (2020).
- [55] R. Kausar, W.-J. Rao, and X. Wan, Learning what a machine learns in a many-body localization transition, *J. Phys.: Condens. Matter* **32**, 415605 (2020).
- [56] F. Kotthoff, F. Pollmann, and G. De Tomasi, Distinguishing an anderson insulator from a many-body localized phase through space-time snapshots with neural networks, *Phys. Rev. B* **104**, 224307 (2021).

- [57] H. Obuse, A. R. Subramaniam, A. Furusaki, I. A. Gruzberg, and A. W. W. Ludwig, Multifractality and conformal invariance at 2d metal-insulator transition in the spin-orbit symmetry class, *Phys. Rev. Lett.* **98**, 156802 (2007).
- [58] F. Evers, A. Mildenerger, and A. D. Mirlin, Multifractality at the quantum hall transition: Beyond the parabolic paradigm, *Phys. Rev. Lett.* **101**, 116803 (2008).
- [59] H. Obuse, A. R. Subramaniam, A. Furusaki, I. A. Gruzberg, and A. W. W. Ludwig, Boundary multifractality at the integer quantum hall plateau transition: Implications for the critical theory, *Phys. Rev. Lett.* **101**, 116802 (2008).
- [60] H. Obuse, A. R. Subramaniam, A. Furusaki, I. A. Gruzberg, and A. W. W. Ludwig, Conformal invariance, multifractality, and finite-size scaling at anderson localization transitions in two dimensions, *Phys. Rev. B* **82**, 035309 (2010).
- [61] L. J. Vasquez, A. Rodriguez, and R. A. Romer, Multifractal analysis of the metal-insulator transition in the three-dimensional anderson model. i. symmetry relation under typical averaging, *Phys. Rev. B* **78**, 195106 (2008).
- [62] A. Rodriguez, L. J. Vasquez, and R. A. Romer, Multifractal analysis with the probability density function at the three-dimensional anderson transition, *Phys. Rev. Lett.* **102**, 106406 (2009).
- [63] A. Rodriguez, L. J. Vasquez, K. Slevin, and R. A. Römer, Critical parameters from a generalized multifractal analysis at the anderson transition, *Phys. Rev. Lett.* **105**, 046403 (2010).
- [64] A. Rodriguez, L. J. Vasquez, K. Slevin, and R. A. Römer, Multifractal finite-size scaling and universality at the anderson transition, *Phys. Rev. B* **84**, 134209 (2011).
- [65] L. Ujfalusi and I. Varga, Quantum percolation transition in three dimensions: Density of states, finite-size scaling, and multifractality, *Phys. Rev. B* **90**, 174203 (2014).
- [66] L. Ujfalusi and I. Varga, Finite-size scaling and multifractality at the anderson transition for the three wigner-dyson symmetry classes in three dimensions, *Phys. Rev. B* **91**, 184206 (2015).
- [67] A. Bäcker, M. Haque, and I. M. Khaymovich, Multifractal dimensions for random matrices, chaotic quantum maps, and many-body systems, *Phys. Rev. E* **100**, 032117 (2019).
- [68] N. Macé, F. Alet, and N. Laflorencie, Multifractal scalings across the many-body localization transition, *Phys. Rev. Lett.* **123**, 180601 (2019).
- [69] D. J. Luitz, I. M. Khaymovich, and Y. B. Lev, Multifractality and its role in anomalous transport in the disordered XXZ spin-chain, *SciPost Phys. Core* **2**, 006 (2020).
- [70] F. Pietracaprina and N. Laflorencie, Hilbert-space fragmentation, multifractality, and many-body localization, *Ann. Phys.* **435**, 168502 (2021), special Issue on Localization 2020.
- [71] P. Sierant and X. Turkeshi, Universal behavior beyond multifractality of wave functions at measurement-induced phase transitions, *Phys. Rev. Lett.* **128**, 130605 (2022).
- [72] J. Sutradhar, S. Ghosh, S. Roy, D. E. Logan, S. Mukerjee, and S. Banerjee, Scaling of the fock-space propagator and multifractality across the many-body localization transition, *Phys. Rev. B* **106**, 054203 (2022).
- [73] C. Cheng, Many-body localization in clean chains with long-range interactions, *Phys. Rev. B* **108**, 155113 (2023).
- [74] N. Rosenzweig and C. E. Porter, "repulsion of energy levels" in complex atomic spectra, *Phys. Rev.* **120**, 1698 (1960).
- [75] V. E. Kravtsov, I. M. Khaymovich, E. Cuevas, and M. Amini, A random matrix model with localization and ergodic transitions, *New J. Phys.* **17**, 122002 (2015).
- [76] C. Beetar, J. Murugan, and D. Rosa, Neural networks as universal probes of many-body localization in quantum graphs, [arXiv:2108.05737](https://arxiv.org/abs/2108.05737).
- [77] Y. LeCun, B. Boser, J. Denker, D. Henderson, R. Howard, W. Hubbard, and L. Jackel, Handwritten digit recognition with a back-propagation network, in *Advances in Neural Information Processing Systems*, edited by D. Touretzky (Morgan-Kaufmann, Burlington, MA, 1989), Vol. 2.
- [78] B. Landon, P. Sosee, and H.-T. Yau, Fixed energy universality of dyson brownian motion, *Adv. Math.* **346**, 1137 (2019).
- [79] D. Facchetti, P. Vivo, and G. Biroli, From non-ergodic eigenvectors to local resolvent statistics and back: A random matrix perspective, *Europhys. Lett.* **115**, 47003 (2016).
- [80] K. Truong and A. Ossipov, Eigenvectors under a generic perturbation: Non-perturbative results from the random matrix approach, *Europhys. Lett.* **116**, 37002 (2016).
- [81] C. Monthus, Multifractality of eigenstates in the delocalized non-ergodic phase of some random matrix models: Wigner-weisskopf approach, *J. Phys. A: Math. Theor.* **50**, 295101 (2017).
- [82] P. von Soosten and S. Warzel, Non-ergodic delocalization in the rosenzweig-porter model, *Lett. Math. Phys.* **109**, 905 (2019).
- [83] E. Bogomolny and M. Sieber, Eigenfunction distribution for the rosenzweig-porter model, *Phys. Rev. E* **98**, 032139 (2018).
- [84] M. Pino, J. Tabanera, and P. Serna, From ergodic to non-ergodic chaos in rosenzweig-porter model, *J. Phys. A: Math. Theor.* **52**, 475101 (2019).
- [85] G. D. Tomasi, M. Amini, S. Bera, I. M. Khaymovich, and V. E. Kravtsov, Survival probability in Generalized Rosenzweig-Porter random matrix ensemble, *SciPost Phys.* **6**, 014 (2019).
- [86] R. Berkovits, Super-poissonian behavior of the rosenzweig-porter model in the nonergodic extended regime, *Phys. Rev. B* **102**, 165140 (2020).
- [87] M. A. Skvortsov, M. Amini, and V. E. Kravtsov, Sensitivity of (multi)fractal eigenstates to a perturbation of the hamiltonian, *Phys. Rev. B* **106**, 054208 (2022).
- [88] E. Bogomolny and M. Sieber, Power-law random banded matrices and ultrametric matrices: Eigenvector distribution in the intermediate regime, *Phys. Rev. E* **98**, 042116 (2018).
- [89] I. M. Khaymovich, V. E. Kravtsov, B. L. Altshuler, and L. B. Ioffe, Fragile extended phases in the log-normal rosenzweig-porter model, *Phys. Rev. Res.* **2**, 043346 (2020).
- [90] S. Aubry and G. André, Analyticity breaking and Anderson localization in incommensurate lattices, *Ann. Israel Phys. Soc* **3**, 18 (1980).
- [91] A. Avila, S. Jitomirskaya, and C. A. Marx, Spectral theory of extended harper's model and a question by Erdős and Szekeres, *Invent. Math.* **210**, 283 (2017).
- [92] Y. He, S. Xia, D. G. Angelakis, D. Song, Z. Chen, and D. Leykam, Persistent homology analysis of a generalized Aubry-André-Harper model, *Phys. Rev. B* **106**, 054210 (2022).
- [93] J. L. Pichard and G. Sarma, Finite size scaling approach to anderson localisation, *J. Phys. C: Solid State Phys.* **14**, L127 (1981).

- [94] A. MacKinnon and B. Kramer, One-parameter scaling of localization length and conductance in disordered systems, *Phys. Rev. Lett.* **47**, 1546 (1981).
- [95] A. MacKinnon and B. Kramer, The scaling theory of electrons in disordered solids: Additional numerical results, *Z. Phys. B* **53**, 1 (1983).
- [96] K. Slevin and T. Ohtsuki, Corrections to scaling at the anderson transition, *Phys. Rev. Lett.* **82**, 382 (1999).
- [97] K. Slevin and T. Ohtsuki, Critical exponent for the anderson transition in the three-dimensional orthogonal universality class, *New J. Phys.* **16**, 015012 (2014).
- [98] K. Slevin and T. Ohtsuki, Critical exponent of the anderson transition using massively parallel supercomputing, *J. Phys. Soc. Jpn.* **87**, 094703 (2018).
- [99] J. Šuntajs, T. Prosen, and L. Vidmar, Spectral properties of three-dimensional anderson model, *Ann. Phys.* **435**, 168469 (2021), special issue on Philip W. Anderson.
- [100] A. D. Mirlin, Y. V. Fyodorov, F.-M. Dittes, J. Quezada, and T. H. Seligman, Transition from localized to extended eigenstates in the ensemble of power-law random banded matrices, *Phys. Rev. E* **54**, 3221 (1996).
- [101] E. Cuevas, V. Gasparian, and M. Ortuño, Anomalous large critical regions in power-law random matrix ensembles, *Phys. Rev. Lett.* **87**, 056601 (2001).
- [102] A. De Luca, B. L. Altshuler, V. E. Kravtsov, and A. Scardicchio, Anderson localization on the bethe lattice: Non-ergodicity of extended states, *Phys. Rev. Lett.* **113**, 046806 (2014).
- [103] B. L. Altshuler, E. Cuevas, L. B. Ioffe, and V. E. Kravtsov, Nonergodic phases in strongly disordered random regular graphs, *Phys. Rev. Lett.* **117**, 156601 (2016).
- [104] V. Kravtsov, B. Altshuler, and L. Ioffe, Non-ergodic delocalized phase in anderson model on bethe lattice and regular graph, *Ann. Phys.* **389**, 148 (2018).
- [105] K. S. Tikhonov, A. D. Mirlin, and M. A. Skvortsov, Anderson localization and ergodicity on random regular graphs, *Phys. Rev. B* **94**, 220203(R) (2016).
- [106] K. S. Tikhonov and A. D. Mirlin, Critical behavior at the localization transition on random regular graphs, *Phys. Rev. B* **99**, 214202 (2019).
- [107] G. Biroli, A. K. Hartmann, and M. Tarzia, Critical behavior of the anderson model on the bethe lattice via a large-deviation approach, *Phys. Rev. B* **105**, 094202 (2022).
- [108] M. Pino, Scaling up the anderson transition in random-regular graphs, *Phys. Rev. Res.* **2**, 042031(R) (2020).
- [109] I. García-Mata, O. Giraud, B. Georgeot, J. Martin, R. Dubertrand, and G. Lemarié, Scaling theory of the anderson transition in random graphs: Ergodicity and universality, *Phys. Rev. Lett.* **118**, 166801 (2017).
- [110] I. García-Mata, J. Martin, R. Dubertrand, O. Giraud, B. Georgeot, and G. Lemarié, Two critical localization lengths in the anderson transition on random graphs, *Phys. Rev. Res.* **2**, 012020(R) (2020).
- [111] I. García-Mata, J. Martin, O. Giraud, B. Georgeot, R. Dubertrand, and G. Lemarié, Critical properties of the anderson transition on random graphs: Two-parameter scaling theory, kosterlitz-thouless type flow, and many-body localization, *Phys. Rev. B* **106**, 214202 (2022).
- [112] P. Sierant, M. Lewenstein, and A. Scardicchio, Universality in Anderson localization on random graphs with varying connectivity, *SciPost Phys.* **15**, 045 (2023).
- [113] S. Milgram, The small world problem, *Psychol. Today* **1**, 61 (1967).
- [114] D. J. Watts and S. H. Strogatz, Collective dynamics of small-world networks, *Nature (London)* **393**, 440 (1998).
- [115] M. E. J. Newman and D. J. Watts, Scaling and percolation in the small-world network model, *Phys. Rev. E* **60**, 7332 (1999).
- [116] M. E. J. Newman, C. Moore, and D. J. Watts, Mean-field solution of the small-world network model, *Phys. Rev. Lett.* **84**, 3201 (2000).
- [117] A. A. Hagberg, D. A. Schult, and P. J. Swart, Exploring network structure, dynamics, and function using networkx, in *Proceedings of the 7th Python in Science Conference*, edited by G. Varoquaux, T. Vaught, and J. Millman (2008), pp. 11–15.
- [118] C. Vanoni, B. L. Altshuler, V. E. Kravtsov, and A. Scardicchio, Renormalization group analysis of the anderson model on random regular graphs, [arXiv:2306.14965](https://arxiv.org/abs/2306.14965).
- [119] S. Sachdev and J. Ye, Gapless spin-fluid ground state in a random quantum heisenberg magnet, *Phys. Rev. Lett.* **70**, 3339 (1993).
- [120] J. Maldacena and D. Stanford, Remarks on the sachdev-ye-kitaev model, *Phys. Rev. D* **94**, 106002 (2016).
- [121] J. Polchinski and V. Rosenhaus, The spectrum in the sachdev-ye-kitaev model, *J. High Energy Phys.* **04** (2016) 001.
- [122] A. M. García-García, B. Loureiro, A. Romero-Bermúdez, and M. Tezuka, Chaotic-integrable transition in the sachdev-ye-kitaev model, *Phys. Rev. Lett.* **120**, 241603 (2018).
- [123] J. Kim and X. Cao, Comment on chaotic-integrable transition in the sachdev-ye-kitaev model, *Phys. Rev. Lett.* **126**, 109101 (2021).
- [124] A. M. García-García, B. Loureiro, A. Romero-Bermúdez, and M. Tezuka, García-garcía *et al.* Reply., *Phys. Rev. Lett.* **126**, 109102 (2021).
- [125] T. Nosaka, D. Rosa, and J. Yoon, The Thouless time for mass-deformed SYK, *J. High Energy Phys.* **09** (2018) 041.
- [126] F. Monteiro, T. Micklitz, M. Tezuka, and A. Altland, Minimal model of many-body localization, *Phys. Rev. Res.* **3**, 013023 (2021).
- [127] F. Monteiro, M. Tezuka, A. Altland, D. A. Huse, and T. Micklitz, Quantum ergodicity in the many-body localization problem, *Phys. Rev. Lett.* **127**, 030601 (2021).
- [128] D. K. Nandy, T. Čadež, B. Dietz, A. Andreanov, and D. Rosa, Delayed thermalization in the mass-deformed sachdev-ye-kitaev model, *Phys. Rev. B* **106**, 245147 (2022).
- [129] P. Sierant, A. Maksymov, M. Kuś, and J. Zakrzewski, Fidelity susceptibility in gaussian random ensembles, *Phys. Rev. E* **99**, 050102(R) (2019).
- [130] A. Maksymov, P. Sierant, and J. Zakrzewski, Energy level dynamics across the many-body localization transition, *Phys. Rev. B* **99**, 224202 (2019).
- [131] M. Pandey, P. W. Claeys, D. K. Campbell, A. Polkovnikov, and D. Sels, Adiabatic eigenstate deformations as a sensitive probe for quantum chaos, *Phys. Rev. X* **10**, 041017 (2020).
- [132] D. Sels and A. Polkovnikov, Dynamical obstruction to localization in a disordered spin chain, *Phys. Rev. E* **104**, 054105 (2021).
- [133] T. Čadež, D. K. Nandy, D. Rosa, A. Andreanov, and B. Dietz, The rosenzweig-porter model revisited for the three Wigner-Dyson symmetry classes (unpublished).

- [134] A. Ahmed, A. Nelson, A. Raina, and A. Sharma, Phase classification in the long-range Harper model using machine learning, *Phys. Rev. B*, **108**, 155128 (2023).
- [135] F. Chollet *et al.*, KERAS, <https://github.com/keras-team/keras>.
- [136] M. Abadi, A. Agarwal, P. Barham, E. Brevdo, Z. Chen, C. Citro, G. S. Corrado, A. Davis, J. Dean, M. Devin, S. Ghemawat, I. Goodfellow, A. Harp, G. Irving, M. Isard, Y. Jia, R. Jozefowicz, L. Kaiser, M. Kudlur, J. Levenberg *et al.*, TENSORFLOW: Large-scale machine learning on heterogeneous systems, software available from tensorflow.org.
- [137] C. E. Porter, *Statistical Theories of Spectra: Fluctuations* (Academic, New York, 1965).
- [138] M. L. Mehta, *Random Matrices* (Elsevier, Amsterdam, 2004).
- [139] F. Haake, S. Gnutzmann, and M. Kuś, *Quantum Signatures of Chaos* (Springer, Heidelberg, 2018).
- [140] D.-A. Clevert, T. Unterthiner, and S. Hochreiter, Fast and accurate deep network learning by exponential linear units (elus), [arXiv:1511.07289](https://arxiv.org/abs/1511.07289).
- [141] L. Breiman, Bagging predictors, *Mach. Learn.* **24**, 123 (1996).