Machine Learning Out-of-Equilibrium Phases of Matter

Jordan Venderley,[1] Vedika Khemani,[2] and Eun-Ah Kim[1]
[1]Department of Physics, Cornell University, Ithaca, New York 14853, USA
[2]Department of Physics, Harvard University, Cambridge, Massachusetts 02138, USA

(Received 6 December 2017; revised manuscript received 22 March 2018; published 21 June 2018)

Neural-network-based machine learning is emerging as a powerful tool for obtaining phase diagrams when traditional regression schemes using local equilibrium order parameters are not available, as in many-body localized (MBL) or topological phases. Nevertheless, instances of machine learning offering new insights have been rare up to now. Here we show that a single feed-forward neural network can decode the defining structures of two distinct MBL phases and a thermalizing phase, using entanglement spectra obtained from individual eigenstates. For this, we introduce a simplicial geometry-based method for extracting multipartite phase boundaries. We find that this method outperforms conventional metrics for identifying MBL phase transitions, revealing a sharper phase boundary and shedding new insight on the topology of the phase diagram. Furthermore, the phase diagram we acquire from a single disorder configuration confirms that the machine-learning-based approach we establish here can enable speedy exploration of large phase spaces that can assist with the discovery of new MBL phases. To our knowledge, this Letter represents the first example of a standard machine learning approach revealing new information on phase transitions.

DOI: 10.1103/PhysRevLett.120.257204

The application of machine learning (ML) [1] to central questions in the theory of quantum matter is a rapidly developing research frontier. So far, efforts have been twofold, focusing on (1) representing states compactly [2–5] and (2) identifying and classifying different phases of matter [6–15]. The driving insight here is that the problems of theoretical interest are essentially those of regression in which an exponentially large amount of data must be condensed into a more accessible or meaningful form, e.g., the labeling of wave functions with phases. As neural networks are universal function approximators and facilitate nonlinear regression, neural-network-based ML can effectively distill relevant information from complex data while taking it at face value. This is particularly appealing for phases outside the traditional regression scheme where a local order parameter may not be readily available. Such phases include topological phases and out-of-equilibrium eigenstate phases [16,17] in the context of many-body localization (MBL) [18–24]. Although there has been recent progress in using ML for topological phases [8,13,15] and MBL phases [12,14], extracting phase boundaries in these settings has been a challenging frontier [12,14]. Moreover, the question of whether the same data and architecture can be used to discern multiple phases, especially multiple MBL phases, has been unclear.

MBL generalizes the phenomenon of Anderson localization to the interacting setting, bringing out the interplay of disorder and interactions. Since MBL systems stay out of thermal equilibrium, they can display a host of rich dynamical phenomena [25–30]. Furthermore, it is now known that different varieties of MBL phases (e.g., MBL paramagnets, symmetry breaking MBL phases, topologically ordered MBL phases, etc.)—each showing different patterns of order in individual highly excited many-body eigenstates—can be realized in a given system [16,17,31–33]. With experimental realizations of novel out-of-equilibrium states in MBL settings such as time crystals [34–38], it is all the more important to understand these out-of-equilibrium phases and the associated transitions. Moreover, we need efficient ways to study and discover new MBL phases *without* an *a priori* knowledge of the defining order parameters.

Despite extensive research [20,39–46], a complete theoretical understanding of the MBL transitions is lacking, partially due to the absence of a comprehensive scheme for regression. Although entanglement entropy serves as a useful diagnostic of thermalization (excited eigenstates in the thermal phase are volume-law entangled, while they are only area-law entangled in the MBL phase [20,31]), it appears to be too aggressive a regression since it traces out important entanglement correlations. The structure of these entanglement correlations is expected to be relevant for understanding the nature of the many-body "resonances" that drive the transition out of the MBL phase [40,41,46]. Moreover the entanglement spectra of individual many-body eigenstates must encode the structure of different MBL phases, even when the defining correlation functions are not known *a priori*. While there have been efforts to utilize the full entanglement spectra (ES) [47,48], a complete understanding of how to interpret the ES has

not yet been established. Alternatively, there have been efforts to employ neural networks to extract relevant information from entanglement spectra [12,14], but it has been unclear whether ML has been able to offer any new insights thus far.

In this Letter, we take a first step towards a ML assisted study of MBL phase transitions. Our model is a disordered and interacting transverse-field Ising model (TFIM) which has two distinct many-body localized phases: (1) many-body localized spin glass (MBL-SG) and (2) many-body localized paramagnetic (MBL-PM), in addition to a thermal phase. Using the entanglement spectra of individual eigenstates as our only input to a standard neural network, we are able to locate these phase boundaries with greater precision than standard methods for studying MBL transitions. To do this, we introduce a new geometric approach for interpreting neural network outputs for multipartite classification.

Model.—The TFIM in the presence of disorder and interactions is a "canonical model" for studying novel eigenstate phases [16,17,49]. It has a well-studied non-interacting limit [50] and well-understood descriptions for the paramagnetic and spin glass phases in the different limits. An Ising self-dual variant for an L site chain is [51]

$$H = -\sum_{i=1}^{L}[J_i\sigma_i^z\sigma_{i+1}^z + h_i\sigma_i^x + \lambda(\bar{h}\sigma_i^x\sigma_{i+1}^x + \bar{J}\sigma_i^z\sigma_{i+2}^z)], \quad (1)$$

where σ_i^z are Pauli spin $1/2$ matrices on site i. The couplings $\{J_i\}$ and on site fields $\{h_i\}$ are drawn from log-normal distributions such that the standard deviation of their logarithms is $\delta(\log J) = \delta(\log h) = 1$. Note that this model is equivalent to a disordered interacting fermion model upon a Jordan Wigner transformation, where the interaction strengths are proportional to λ. Finally, $\bar{h}$ and $\bar{J}$ denote the means of $\{J_i\}$ and $\{h_i\}$ [52]. The means $\overline{\log J}$, $\overline{\log h}$, and λ serve as tuning parameters that can be used to drive different phase transitions.

Let us first consider the axis $\lambda = 0$, which is equivalent to a disordered free-fermion model subject to Anderson localization. From the eigenstate order perspective, there are two distinct phases with respect to the global $\mathbb{Z}_2$ Ising symmetry of the model under spin flips $P = \prod_i \sigma_i^x$: the symmetry-broken spin glass (SG) phase for $\bar{J} > \bar{h}$ and the paramagnetic (PM) phase for $\bar{J} < \bar{h}$. Deep in the SG phase, individual many-body eigenstates are macroscopic superpositions (i.e., Schrödinger "cat" states) in the σ^z basis with localized domain walls: $|\alpha\rangle \sim |\uparrow\downarrow\downarrow\uparrow\downarrow\cdots\rangle \pm |\downarrow\uparrow\uparrow\downarrow\uparrow\cdots\rangle$, and the connected correlation function of σ^z shows long-range order with

$$\langle\alpha|\sigma_i^z\sigma_j^z|\alpha\rangle_c \equiv \langle\alpha|\sigma_i^z\sigma_j^z|\alpha\rangle - \langle\alpha|\sigma_i^z|\alpha\rangle\langle\alpha|\sigma_j^z|\alpha\rangle = \pm|c_\alpha|, \quad (2)$$

$|c_\alpha| > 0$ even as $|i-j| \to \infty$. By contrast, for the equilibrium problem in the absence of disorder, a finite density of delocalized domain walls destroys long-range order at any finite temperature in 1D in accordance with Peierls-Mermin-Wagner theorems. Thus, the SG phase in 1D furnishes an example of a model where localization enables a new form of dynamical order that is disallowed in equilibrium [16,17]. On the other hand, the eigenstates deep in the PM phase resemble product states in the σ^x basis, $|\alpha\rangle \sim |\rightarrow\leftarrow\leftarrow\rightarrow\leftarrow\cdots\rangle$ without long-range order, i.e., $\langle\alpha|\sigma_i^z\sigma_j^z|\alpha\rangle_c = 0$. The critical point between these two phases is at the Ising self-dual point, $\overline{\log J} = \overline{\log h}$, and the critical properties for $\lambda = 0$ are described by an infinite randomness fixed point [50].

Once $\lambda \neq 0$, a numerical study over a large number of disorder realizations looking at all the eigenstates is necessary to obtain the phase diagram that now includes the thermal phase. With finite λ, the nature and mechanism of various phase transitions largely remain open questions since the existing theoretical understanding is limited to three extreme regimes in the phase space: (1) $J \gg h$, λ, (2) $h \gg J$, λ, and (3) $\lambda \gg J = h$. In the limits (1) and (2), the Anderson localized SG and PM phases of the non-interacting system generalize to MBL versions of themselves [16,17,49]. On the other hand, in the strongly interacting limit, the system will be in a thermal phase with its excited states exhibiting volume-law entanglement [53]. Finally, since our interactions were chosen to respect the Ising duality, we expect the phase diagram with nonzero λ to still be symmetric about $\overline{\log J} = \overline{\log h}$ (with small corrections for open boundary conditions). Nevertheless, the precise topology of the tripartite phase boundary and the existence or absence of a direct MBL-MBL phase transition [51,54] are hotly debated questions. On the other hand, most existing approaches for detecting phase boundaries rely on the standard deviation of the entanglement entropy (see Fig. 3) with low resolution leaving the physics of the critical regime largely inaccessible.

Neural-network-based approach.—In order to access the information in the entanglement spectra in a holistic manner, we build and employ a feed-forward neural network with a single hidden layer. Our hidden layer contains 200 neurons with sigmoid activation functions. We utilize a cross-entropy cost function with L2 regularization and use a softmax output layer with three neurons, each of which corresponds to one of three possible phases, namely, the SG-MBL, PM-MBL, and thermal phase [see Fig. 1(a)].

We then generate the training and testing data for different disorder configurations of the model (1) on an open chain with 12 sites. Specifically, we use exact diagonalization to obtain all the eigenstates and take the middle quarter of the eigenstates in each Ising symmetry sector to calculate the bipartite entanglement spectra for each eigenstate. The training set consists of three points in the phase space where the phase is known: $\overline{\log J} - \overline{\log h} = \pm 0.8$ with $\lambda = 0.2$ and $\overline{\log J} - \overline{\log h} = 0.0$ with $\lambda = 1.0$. We use 1000 disorder configurations labeled with each of

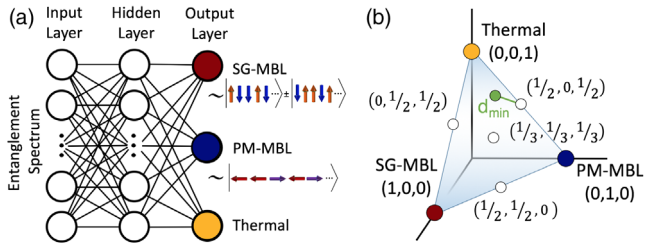

FIG. 1. (a) A depiction of our neural network. (b) The two-simplex codomain of our neural network outputs. Colored circles represent regions with maximal confidence; the white circles represent regions of maximal confusion. The green point represents an example output with its associated $d_{\min}$ marked with a green line.

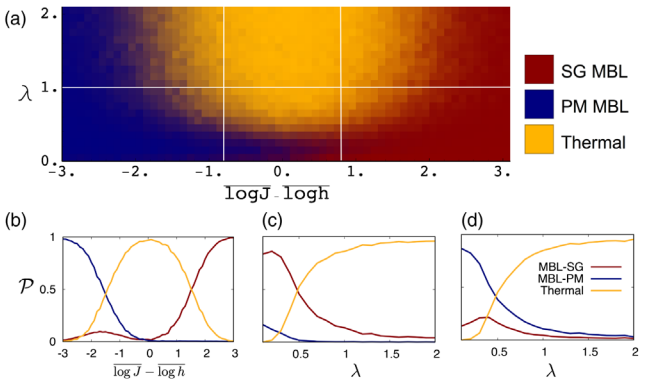

FIG. 2. (a) The phase diagram with average neural network output plotted as an RGB parameter. Cuts marked by white lines in (a) are shown in (b) $\lambda = 1.0$, (c) $\overline{\log J} - \overline{\log h} = 0.8$, and (d) $\overline{\log J} - \overline{\log h} = -0.8$. The sampling width is 0.1 for each parameter.

the three points to train our network using a standard error function to an accuracy of over 90%. The fact that successful training could be reached already points to the fact that our network could extract qualitatively distinct information in the entanglement spectra of eigenstates in the three phases of interest. Additionally, in Appendix A (see Supplemental Material [55]), we look at the ability of a neural network to "discover" new phases by only training on the two MBL phases. In this case, although it knows nothing about the thermal phase, it is nevertheless able to recognize that it does not fit within its learned phase paradigm.

Once the training is complete, we feed the entanglement spectra from each point in the phase space of $(\overline{\log J} - \overline{\log h}, \lambda)$ to the network. The neural network output is a triplet that may be thought of as the network's confidence that the given input is in each phase. Note that all conventional measures require sampling thousands of disorder configurations. On the other hand, we find that averaging the neuron output over just 100 disorder configurations yields a satisfying phase diagram, paving the way for fast scans of large areas of phase space. The purpose of the averaging is to both look into the statistics, as well as to compare with the conventional measure on equal footing. In Fig. 2(a), we plot the average neural network confidence output in the range of $\overline{\log J} - \overline{\log h} \in [-3.0, +3.0]$ and $\lambda \in [0.1, 2.0]$ by representing each component of the triplet with three colors.

The phase diagram in Fig. 2(a) obtained by the neural network displays several satisfying features that are consistent with theoretical insights. First of all, the phase diagram is roughly symmetric about the line $\overline{\log J} - \overline{\log h} = 0.0$ and consistent with the Ising duality of the Hamiltonian [Eq. (1)]. Furthermore, the upward curvature of the phase boundary between the MBL phases and the thermal phase is consistent with the fact that the non-interacting model is most delocalized near the SG-PM transition [50], and hence the transition is most susceptible to thermalization upon adding interactions near the $(\overline{\log J} - \overline{\log h} = 0.0, \lambda = 0)$ point. However, it is evident from the representative line cuts in Figs. 2(b)–2(d) that the variation of the confidence outputs is gradual and broad, masking the precise topology of the phase boundaries. Finite-size scaling is presented in Appendix B (see Supplemental Material [55]) for $L = 8$, 10, and 12, showing the sharpening of the transition with increasing system size.

In order to more precisely study the topology of the phase diagram, we developed a protocol for extracting phase boundaries from multineuron outputs. This approach will extract the phase boundaries in an unbiased way that allows us to quantify the smoothness and width of the transition and compare our results to conventional methods. Our approach is geometrically motivated and uses the fact that the neuron outputs sum to unity in a softmax layer. Specifically, with a softmax N-neuron output, the codomain of the neural network confidence output is an $(N-1)$ simplex embedded in the N-dimensional space of outputs. The points of maximal confusion constitute geometrically notable points on the $(N-1)$ simplex; for $N = 3$, these are the midpoints of the edges and the barycenter. Explicitly, in our present case, the codomain of our neural network is a two simplex and the points of maximum confusion that should naturally belong to the phase boundary [11] are $(1/2, 1/2, 0), (1/2, 0, 1/2), (0, 1/2, 1/2)$, and $(1/3, 1/3, 1/3)$ [see Fig. 1(b)]. Now, for any confidence triplet, one can measure the minimal distance $d_{\min}$ to the set of maximal confusion points. Once we normalize this distance by the maximal possible distance of any point on the simplex to a point of maximal confusion, we obtain a continuous measure of confusion capable of extracting boundaries: $\mathcal{C} \equiv 1 - \bar{d}_{\min}$, where $\bar{d}_{\min}$ denotes the normalized distance. This measure of confusion ranges between $\mathcal{C} = 1$, when the confidence corresponds to one of the maximal confusion points, and $\mathcal{C} = 0$, when the network outputs a particular phase with 100% confidence.

Now, at each point in the phase space, we take the average confidence triplet to evaluate the confusion measure $\mathcal{C}$ as shown in Fig. 3(a). It is notable that our confusion

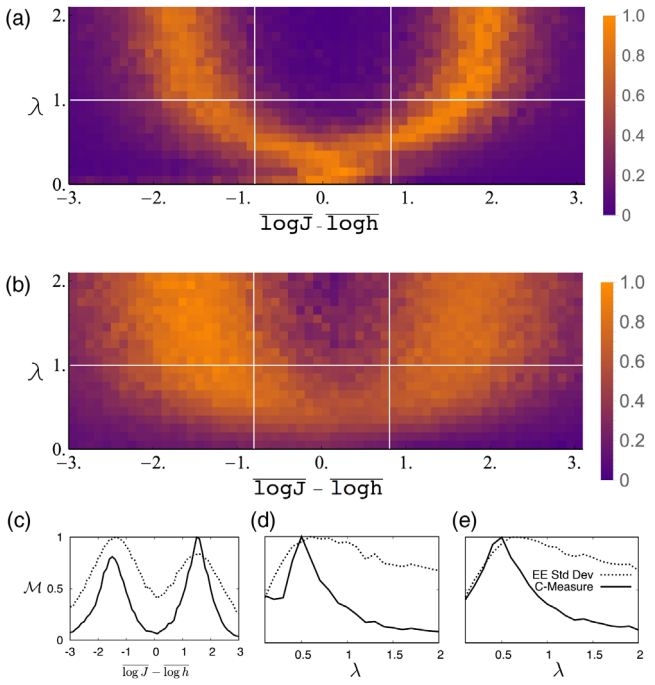

FIG. 3. (a) Our $\mathcal{C}$ measure for extracting phase boundaries (defined in the main text) and (b) the average standard deviation (Std Dev) of the entanglement entropy. The data in each have been normalized by the largest value in the parameter space for meaningful comparison. (c)–(e) The measures plotted in (a) and (b) along the cuts marked in white lines: (c) $\lambda = 1.0$, (d) $\overline{\log J} - \overline{\log h} = +0.8$, and (e) $\overline{\log J} - \overline{\log h} = -0.8$.

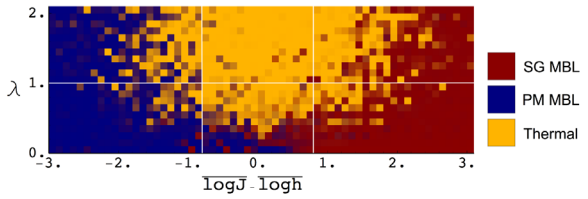

FIG. 4. The 2D phase diagram where the fully trained network has been tested on a single disorder realization.

measure allows us to establish phase boundaries in a manner that is native to the neural network approach. Surprisingly, the phase boundary detected by the neural network has the topology of a "wishbone" with a visible phase boundary between two MBL phases [see Fig. 3(a)] at small λ. This warrants a more exhaustive study of this transition, including finite-size effects in order to probe the existence of a direct SG-MBL to PM-MBL transition.

The $\mathcal{C}$-measure-based extraction of the phase boundary can be contrasted with a more conventional entanglement-entropy-based approach [49]. Since the entanglement entropy (EE) changes from area law to volume law upon transition from a MBL phase to a thermal phase, it is expected that the standard deviation of the EE in eigenstates peaks at the phase boundary [49]. Figure 3(b) shows the standard deviation taken over all disorder samples and the middle quarter of the eigenstates from each sample. As expected, the standard deviation of the EE is peaked at the MBL-thermal boundaries and is aligned with our machine-learning-derived measure. However, two advantages of the neural network $\mathcal{C}$ measure easily stand out. First, the EE-based approach cannot distinguish the boundary between the two area-law MBL phases [see the U-shaped phase boundary in Fig. 3(b)], whereas the neural network is successfully differentiating these [see the wishbone-shaped phase boundary in Fig. 3(a)]. For the MBL-SG problem, one can additionally construct an Edwards-Anderson spin glass order parameter to single out the MBL-SG phase [49]. However, the ability of the neural network to distinguish between different MBL phases using just the ES and no other "prior knowledge" about order parameters can prove useful for future studies of new MBL phases where order parameters might be unknown. Second, the $\mathcal{C}$ measure reveals a markedly sharper phase boundary that enables a better study of its topology [see the line cut comparisons in Figs. 3(c)–3(e)].

Finally, we should remark on the neural networks' ability to see through the noise that is inevitable in studies of disorder effects. Although we have averaged over 100 different disorder configurations to gain statistics in the bulk of this letter, Fig. 4 shows that the neural network can capture the coarse features of the phase diagram even for a single disorder realization. The fact that the neural network has gained a regression scheme alternate to the manual modeling of statistical distributions over disorder realizations implies that one can use it as a tool to quickly explore large areas of phase space to map out new nonequilibrium phase diagrams.

Summary and outlook.—Here we exploit the ability of neural networks to distill characteristic features from noisy data in order to extract information from the entanglement spectra associated with out-of-equilibrium phases. To this end, we built a neural network and employed it to process entanglement spectra from a transverse-field Ising model with disorder, a poster-child model system that can be in one of three distinct out-of-equilibrium phases. Our neural network, being trained with typical data associated with three limiting points in the phase space, was able to output a phase diagram that is consistent with theoretical expectations. Moreover, using a simplicial geometry construction to quantify the network's degree of confusion, we were able to extract the phase boundary with significantly sharper resolution compared to entanglement-entropy-based approaches. Any effort to better understand this transition and/or the possibility of an intervening sliver of thermal phase between the two MBL phases will benefit from a method for obtaining a sharper determination of phase boundaries, which our Letter provides.

The significance of what we have achieved is multifaceted. First, we have demonstrated that a standard neural-network-based approach can give us a sharper look at the

multipartite phase boundary by using the geometric measure of confusion $\mathcal{C}$ that we introduced. This is the first example, to the best of our knowledge, that a neural-network-based approach in a standard setup outperformed the conventional approach in terms of sharper phase boundaries [56]. Our work paves the way for future studies on the nature of MBL phase transitions. Second, by having multiple neuron outputs, we were able to a obtain tripartite phase diagram involving two distinct MBL phases with a *single* measurement. This is valuable even for MBL phases where there are known order parameters [49], as in the model we considered. However, this multineuron output approach will be even more valuable when dealing with new out-of-equilibrium phases without *a priori* knowledge of suitable order parameters.

E.-A. K. and J. V. thank Yi Zhang for discussions. E.-A. K. acknowledges the Simons Fellow in Theoretical Physics Award No. 392182 and DOE support under Award No. DE-SC0010313. E.-A. K. is grateful for the hospitality of the Kavli Institute of Theoretical Physics supported by NSF under Grant No. NSF PHY-1125915, where this work was initiated. J. V. acknowledges NSF support under Grant No. NSF DMR-1308089. V. K. thanks S. Moudgalya and D. Huse for an ongoing collaboration on the model studied in this paper. V. K. is supported by the Harvard Society of Fellows and the William F. Milton Fund.

[1] M. Jordan, Science **349**, 255 (2015).
[2] J. Chen, S. Cheng, H. Xie, L. Wang, and T. Xiang, Phys. Rev. B **97**, 085104 (2018).
[3] D.-L. Deng, X. Li, and S. D. Sarma, Phys. Rev. X **7**, 021021 (2017).
[4] D.-L. Deng, X. Li, and S. D. Sarma, Phys. Rev. B **96**, 195145 (2017).
[5] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, Phys. Rev. B **95**, 041101(R) (2017).
[6] P. Broecker, F. F. Assaad, and S. Trebst, arXiv:1707.00663.
[7] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Sci. Rep. **7**, 8823 (2017).
[8] Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).
[9] L. Wang, Phys. Rev. B **94**, 195105 (2016).
[10] G. Carleo and M. Troyer, Science **355**, 602 (2017).
[11] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
[12] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).
[13] Y. Zhang, R. G. Melko, and E.-A. Kim, Phys. Rev. B **96**, 245119 (2017).
[14] F. Schindler, N. Regnault, and T. Neupert, Phys. Rev. B **95**, 245134 (2017).
[15] T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **85**, 123706 (2016).
[16] D. A. Huse, R. Nandkishore, V. Oganesyan, A. Pal, and S. L. Sondhi, Phys. Rev. B **88**, 014206 (2013).
[17] D. Pekker, G. Refael, E. Altman, E. Demler, and V. Oganesyan, Phys. Rev. X **4**, 011052 (2014).
[18] P. W. Anderson, Phys. Rev. **109**, 1492 (1958).
[19] D. M. Basko, I. L. Aleiner, and B. L. Altshuler, Ann. Phys. (Amsterdam) **321**, 1126 (2006).
[20] A. Pal and D. A. Huse, Phys. Rev. B **82**, 174411 (2010).
[21] M. Žnidarič, T. Prosen, and P. Prelovšek, Phys. Rev. B **77**, 064426 (2008).
[22] V. Oganesyan and D. A. Huse, Phys. Rev. B **75**, 155111 (2007).
[23] R. Nandkishore and D. A. Huse, Annu. Rev. Condens. Matter Phys. **6**, 15 (2015).
[24] E. Altman and R. Vosk, Annu. Rev. Condens. Matter Phys. **6**, 383 (2015).
[25] D. A. Huse, R. Nandkishore, and V. Oganesyan, Phys. Rev. B **90**, 174202 (2014).
[26] M. Serbyn, Z. Papić, and D. A. Abanin, Phys. Rev. Lett. **111**, 127201 (2013).
[27] J. Z. Imbrie, J. Stat. Phys. **163**, 998 (2016).
[28] R. Vasseur, S. A. Parameswaran, and J. E. Moore, Phys. Rev. B **91**, 140202 (2015).
[29] V. Khemani, R. Nandkishore, and S. L. Sondhi, Nat. Phys. **11**, 560 (2015).
[30] S. Gopalakrishnan, M. Müller, V. Khemani, M. Knap, E. Demler, and D. A. Huse, Phys. Rev. B **92**, 104202 (2015).
[31] B. Bauer and C. Nayak, J. Stat. Mech. (2013) P09005.
[32] A. Chandran, V. Khemani, C. R. Laumann, and S. L. Sondhi, Phys. Rev. B **89**, 144201 (2014).
[33] Y. Bahri, R. Vosk, E. Altman, and A. Vishwanath, Nat. Commun. **6**, 7341 (2015).
[34] V. Khemani, A. Lazarides, R. Moessner, and S. L. Sondhi, Phys. Rev. Lett. **116**, 250401 (2016).
[35] D. V. Else, B. Bauer, and C. Nayak, Phys. Rev. Lett. **117**, 090402 (2016).
[36] C. W. von Keyserlingk, V. Khemani, and S. L. Sondhi, Phys. Rev. B **94**, 085112 (2016).
[37] S. Choi, J. Choi, R. Landig, G. Kucsko, H. Zhou, J. Isoya, F. Jelezko, S. Onoda, H. Sumiya, V. Khemani, C. von Keyserlingk, N. Y. Yao, E. Demler, and M. D. Lukin, Nature (London) **543**, 221 (2017).
[38] J. Zhang, P. W. Hess, A. Kyprianidis, P. Becker, A. Lee, J. Smith, G. Pagano, I.-D. Potirniche, A. C. Potter, A. Vishwanath, N. Y. Yao, and C. Monroe, Nature (London) **543**, 217 (2017).
[39] D. J. Luitz, N. Laflorencie, and F. Alet, Phys. Rev. B **91**, 081103 (2015).
[40] R. Vosk, D. A. Huse, and E. Altman, Phys. Rev. X **5**, 031032 (2015).
[41] A. C. Potter, R. Vasseur, and S. A. Parameswaran, Phys. Rev. X **5**, 031033 (2015).
[42] T. Grover, arXiv:1405.1471.
[43] M. Serbyn, Z. Papić, and D. A. Abanin, Phys. Rev. X **5**, 041047 (2015).
[44] X. Yu, D. J. Luitz, and B. K. Clark, Phys. Rev. B **94**, 184202 (2016).
[45] V. Khemani, S. P. Lim, D. N. Sheng, and D. A. Huse, Phys. Rev. X **7**, 021013 (2017).
[46] V. Khemani, D. N. Sheng, and D. A. Huse, Phys. Rev. Lett. **119**, 075702 (2017).
[47] S. D. Geraedts, R. Nandkishore, and N. Regnault, Phys. Rev. B **93**, 174202 (2016).

[48] S. D. Geraedts, N. Regnault, and R. M. Nandkishore, New J. Phys. **19**, 113021 (2017).
[49] J. A. Kjäll, J. H. Bardarson, and F. Pollmann, Phys. Rev. Lett. **113**, 107204 (2014).
[50] D. S. Fisher, Phys. Rev. B **51**, 6411 (1995).
[51] S. Moudgalya, V. Khemani, and D. A. Huse (to be published).
[52] Scaling the interaction terms with $\bar{J}$, $\bar{h}$ ensures that the interactions are not overwhelmed by the noninteracting terms, which can be large because of the log-normal distributions.
[53] D. N. Page, Phys. Rev. Lett. **71**, 1291 (1993).
[54] W. De Roeck and F. Huveneers, Phys. Rev. B **95**, 155129 (2017).
[55] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevLett.120.257204 for an investigation into the neural network's sensitivity to unknown phases and finite-size scaling.
[56] We note that Ref. [14] found that a neural network can outperform conventional methods, after adding a nontrivial input that penalizes lack of confidence to the standard setup in the training process.