

**Learning single-particle mobility edges by a neural network based on data compression**Xiao-Dong Bai,<sup>1,2,\*</sup> Jie Zhao<sup>2</sup>, Yu-Yong Han,<sup>2</sup> Jin-Cui Zhao,<sup>3</sup> and Ji-Guo Wang<sup>3</sup><sup>1</sup>*Intelligent Energy Lab, Frontier Research Center, Peng Cheng Laboratory, Shenzhen 518000, China*<sup>2</sup>*Center for Quantum Computing, Peng Cheng Laboratory, Shenzhen 518005, China*<sup>3</sup>*Department of Mathematics and Physics, Shijiazhuang TieDao University, Shijiazhuang 050043, China*

(Received 2 November 2020; revised 19 February 2021; accepted 30 March 2021; published 12 April 2021)

The breaking of the well-known Aubry-André model with a single-particle mobility edge (SPME) is a intriguing subject that has not yet been fully understood. In particular, how to accurately and efficiently recognize a SPME in an optical lattice is currently under active debate. In this work, we develop a data compression-based neural network (DCNN) approach to identify SPMEs in one-dimensional quasiperiodic optical lattice using eigenstates as the sole diagnostic. We find that such method can successfully identify SPMEs of a large system only using a small network trained by the small system data, without onerously and repetitively training a new and large-scale network by massive data of a large system. Furthermore, we show that this method is also applicable to recognize more complex phase transitions, such as many-body localization. Our DCNN approach first paves the way for the development of a generic tool for identifying unexplored phase transitions in large systems.

DOI: [10.1103/PhysRevB.103.134203](https://doi.org/10.1103/PhysRevB.103.134203)**I. INTRODUCTION**

It has been universally believed that the tight-binding version of quasiperiodic optical lattice system is represented by the well-known Aubry-André model capturing all essential localization physics [1–3]. However, recent research has established that this belief is incorrect [4,5]. The system actually can map onto the Aubry-André Hamiltonian only in the deeper lattice, because the nearest-neighbor tight-binding limit is approached [6–9]. But in the shallower lattices, such a model will break down due to the unavoidable appearance of single-particle mobility edge (SPME) which marks a critical energy separating extended and localized states [4,5,10–13]. Hence, investigating the properties of SPME in optical lattice is an essential issue for understanding such unusual phenomena, and it has been reviewed with considerable attention. Its stability, dynamics, and other properties have made pioneering achievements in both Hermitian and non-Hermitian systems [14–19]. But, how to accurately and efficiently identify SPME in optical lattice is still a controversial issue.

Machine learning (ML), as a powerful tool for analyzing data, has recently achieved huge success from industrial applications to fundamental research in physics, cheminformatics, and biology [20–22]. Especially in physics, ML has shown its powerful availability in experimental data analysis [23,24] and classification of phases of matter [25–40], while, among these applications, one of the most interesting subjects is to capture the global properties of localization and thermalization phases of matter directly from local inputs using a neural network (NN) [31,41–44].

There are several crucial results on this subject [31,43,44], but many questions remain. One of the most interesting questions is that a trained NN can only recognize the phases and their transitions of system with the same size. If one wants to identify these properties of other larger systems, such as doubling or tripling the size of that system, it is necessary to retrain a new and large-scale NN with huge amounts of data from this large system. It is not only a serious waste of time but also a huge drain on computing resources. Worse, it may even be impossible to retrain a new NN if the data of that system are too large. Up to now, none of the current NN methods can overcome this deficiency, and there is not yet an effective solution to this problem.

In this paper, we develop a data compression-based neural network (DCNN) approach targeted to solving such a problem. In fact, once the parameters of a system are given, whether SPME exists or not is deterministic and independent on the size of this system, and it can be diagnosed by all eigenstates. This is like the picture ‘1’ no matter it is enlarged or reduced, ‘1’ always means 1 and its characteristics will never change. Based on this inspiration, we here propose a scheme to compress the big data of a large system into small data and then use it as input data of NN trained by the data of a small system to identify the SPMEs of a large system. As a result, we find that this method can successfully identify the SPMEs for large systems. Such approach has avoided onerously and repetitively training for a new and large-scale NN.

**II. MODEL AND METHOD**

The model we explore is the single-particle Hamiltonian of a one-dimensional (1D) quasiperiodic optical lattice, which is

\*baixd@pcl.ac.cn

defined as

$$\mathcal{H}(x) = -\frac{\hbar^2}{2M} \frac{d^2}{dx^2} + \frac{V_p}{2} \cos(2kx) + \frac{V_d}{2} \cos(2\alpha x + \theta), \quad (1)$$

where  $M$  is the mass of atom,  $V_p$  and  $V_d$  are the primary and secondary potentials of lattices, respectively;  $\theta$  is an arbitrary phase shift between the two lattices;  $k = 2\pi/\lambda_p$  denotes the wave-vector of the primary lattice;  $\alpha$  is an irrational number indicating the underlying potential to be quasiperiodic, and for convenience we set  $\alpha = (\sqrt{5} - 1)/2$  in this work. Throughout our discussions we use the lattice constant  $a = \pi/k$  as the length unit and the recoil energy  $E_R = \hbar^2 k^2 / 2M$  as the energy unit.

In order to provide the training data of our ML method, we introduce two key physical quantities. They are the inverse participation ratio (IPR) and the normalized participation ratio (NPR) of all eigenstates in the lowest band, which are defined as following:

$$\text{IPR}^{(n)} = \frac{\sum_m |u_m^n|^4}{(\sum_m |u_m^n|^2)^2}, \quad \text{NPR}^{(n)} = \frac{1}{L \sum_m |u_m^n|^4}, \quad (2)$$

where  $m$  labels the real space (discretized) coordinates,  $u_m^n$  represents the  $n$ -th eigenstate, and  $L$  is the length of the lattice. The average values IPR and NPR of these two quantities have been widely used to identify whether a system is localized or thermalized. In the case of thermalization, IPR vanishes and NPR remains finite; on the contrary, in the case of localization, IPR remains finite but NPR vanishes, whereas in the special intermediate case, both IPR and NPR remain finite. It indicates that SPME exists in such case, where the system is delocalized but nonthermal, and this is the focus of our study.

### III. RESULTS OF A SMALL SYSTEM

In order to develop the DCNN approach, it is necessary to train a convolutional neural network (CNN) by data of a small system to successfully identify the phases of localization, SPME, and thermalization. The structure of the CNN is described in detail in the Supplemental Material [45]. We chose the CNN for the following reasons. First, recent studies have demonstrated that the CNN possesses higher accuracy than other NN methods in dealing with image recognition issues. Our input data, being a matrix of the eigenstates, can actually be viewed as an image, thus we can make full use of this advantage. Second, when an image is zoomed in and out according to certain rules, the characteristics contained therein actually will not change, which makes it possible to compress the data without compromising its main information. It lays a foundation of the DCNN method.

In order to contrast with the results of the NN method, we first present the results of traditional methods by calculating the average values IPR and NPR for all eigenstates in the lowest band of the model defined by Eq. (1), as shown in Fig. 1(b). We here set the size of the small system as  $L = 200$  and the depth of the primary lattice potential as  $V_s = 8$  in this typical case. We can see that depending on the strength of

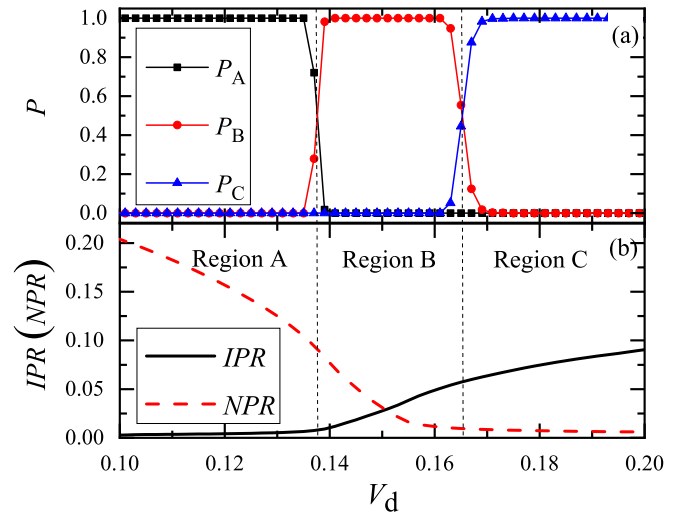


FIG. 1. Comparison of results between NN and the traditional method. (a) Results from NN,  $P_A$ ,  $P_B$ , and  $P_C$  are defined by Eq. (3). Here, the number of instances of the training set is 50000, and  $N_{in}$  is 1000 for every given  $V_d$ . (b) Results from the traditional method, the IPR, and the NPR are defined by Eq. (2). Regions A, B, and C denote the phases of thermalization, SPME, and localization, respectively. These three phases are divided by two vertical dotted lines. Other parameters:  $L = 200$ ,  $V_s = 8$ .

the incommensurate potential  $V_d$ , there exists three distinct phases divided by two vertical dotted lines, which are marked as regions A, B, and C. In region A, all eigenstates in the lowest band remain extended, as indicated by a vanishing IPR, which represents thermalization, while in region C, all eigenstates are localized, as suggested by a vanishing NPR, which predicts the localization. However, in intermediate region B the extended and the localized states are coexistent, and both IPR and NPR become finite. That is, this region corresponds to SPME.

On the other hand, the results of the NN method are shown in Fig. 1(a). The training process is briefly described as follows. We first generate the input data by calculating all eigenstates in the lowest band of the model defined by Eq. (1) using exact diagonalization with a varying global phase  $\theta$  and incommensurate potential  $V_d$ . The training data for each phase are, respectively, generated from the regions labeled by phases A, B, and C, which have been distinguished by both IPR and NPR in Eq. (2). Second, we feed the training data to the input layer, then evolve by the coupling connection mode of CNN between the adjacent layers, where the initial values are randomly chosen according to the log-likelihood cost function. Last, we test the trained network with another independent set of testing data obtained in the same way. If the testing accuracy is over 99%, we consider the training to be a success.

Moreover, in order to clearly present results of the phase transitions calculated by this network, we introduce the parameters  $P_{(A,B,C)}$  defined as the following:

$$P_{(A,B,C)} = \frac{N_{\text{test},A,B,C}}{N_{in}}, \quad (3)$$

where, for every value of  $P_{(A,B,C)}$ , the  $N_{test,A,B,C}$  and  $N_{in}$  are gained by the following processes. We first generate the data by a similar way to calculate all eigenstates in the lowest band of such models using exact diagonalization with a varying global phase  $\theta$ , but  $V_d$  is given. The number of samples in this data set is denoted by  $N_{in}$ , then we feed these data one by one into the input layer to evolve. If the result shows that it is classified as phase A, which indicates the system is thermalized, we add 1 to the counter  $N_{test,A}$ . If the result shows that the system is localized, we add 1 to the counter  $N_{test,C}$ . Otherwise, if the network predicts it is classified as phase B, which means the SPME exists in such system, we add 1 to the counter  $N_{test,B}$ . Last, summarizing the final value of  $N_{test,A,B,C}$  we can calculate the  $P_A$ ,  $P_B$ , and  $P_C$  for a given  $V_d$ , respectively. Obviously,  $P_{(A,B,C)}$  actually respectively represent the probability that NN predicts this system is classified as thermalization, SPME, and localization under a given  $V_d$ . After collating these data sorted by  $V_d$  the resulting phase diagram can be obtained [Fig. 1(a)]. From Fig. 1(a) we can see that once the NN is successfully trained, it not only has the ability to identify each single phase nearly 100% but also has the ability to accurately distinguish their phase boundaries. It is highly consistent with the results of the traditional method described by both IPR and NPR in Fig. 1(b).

#### IV. THE DCNN APPROACH AND RESULTS

In the above section, the small NN has been successfully trained to recognize SPMEs. However, if one wants to identify the phases of other larger systems, a new and large-scale network has to be retrained, because the data size of the large system does not match the size of the input layer of the network trained above. Hence, we here propose our scheme to compress the data of the large system into small data without losing key information so that it is available for the small NN trained above.

We assume that the data of a small system are a matrix by  $n \times n$ , and its elements are marked as  $\{a_{1,1}, a_{1,2}, \dots, a_{n,n-1}, a_{n,n}\}$ , where the first and second subscripts represent labels of rows and columns, respectively. Each column  $\{a_{1,r}, a_{2,r}, \dots, a_{n-1,r}, a_{n,r}\}$  represents an eigenstate, and  $\sum_{i=1}^n |a_{i,r}|^2 = \text{const}$  due to the normalization of eigenstates.

For a large system we assume its data size is enlarged  $q^2$ -times and becomes  $(q \cdot n) \times (q \cdot n)$ , where  $q$  is an integer. The data compression process is as follows. Step 1, we take a matrix  $M_1$  of size  $q \times q$  from this data as the compression object, then remark its elements as  $\{b_{1,1}, b_{1,2}, \dots, b_{q,q-1}, b_{q,q}\}$ , as the magenta small matrices shown in Figs. 2(a) and 2(b), where  $q = 2$  and  $3$ , respectively. The new matrix elements  $b_{1,j}^*$  are then gained by calculating for column elements of  $M_1$  using the following:

$$b_{1,j}^* = \left( \sum_{i=1}^q |b_{i,j}|^2 \right)^{1/2}, \quad j = 1, \dots, q. \quad (4)$$

Thus, the new matrix  $M_2$  can be presented as  $\{b_{1,1}^*, \dots, b_{1,q}^*\}$ . We here choose the way of Eq. (4) to calculate  $b_{1,j}^*$  for the following reasons. First, whether a system is localized or not

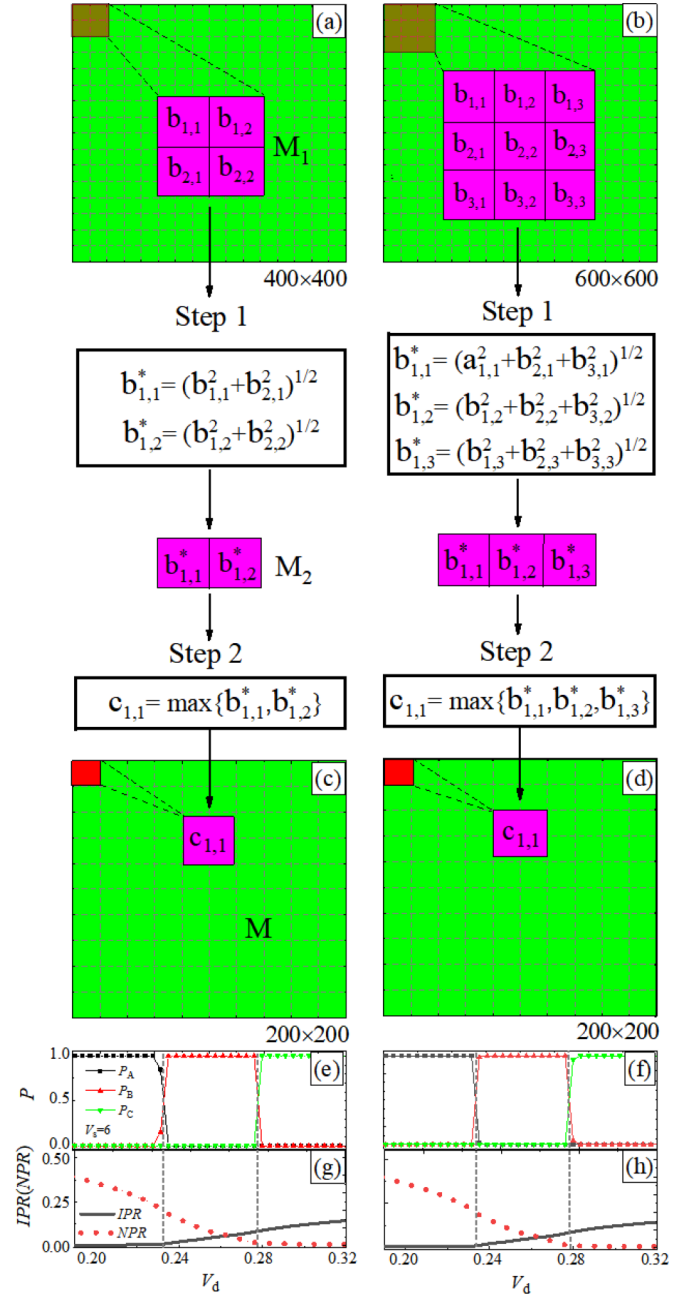


FIG. 2. The schematic of the data compression scheme. (a), (b) Data of larger systems, and  $L = 400$  and  $600$ , respectively.  $M_1$  is taken out in sequence from left to right and top to bottom of the matrix. Steps 1 and 2 illustrate the detailed process of the compression method. (c), (d) Compressed data sets. (e), (f) The results obtained by feeding data of (c) and (d) into small NN trained above. (g), (h) The results from the traditional method, and the three phases are divided by two vertical dotted lines. Other parameters:  $V_s = 6$ ,  $N_{in} = 1000$ .

depends on both the localization and the thermalization characteristics of all eigenstates. That is, it depends on the column elements but not the row elements. Here the column vectors of the matrix represent the eigenstates. So, only the operation for column elements can ensure that the main information will not be lost in these data compression process. Second, the normalization of original eigenstates is  $\sum_{i=1}^n |a_{i,r}|^2 = \text{const}$ ;

thus, such operation of Eq. (4) can ensure that  $b_{1,j}^*$  contains the original eigenstate information to the maximum.

Step 2, we take a maximum of  $M_2$  as a new element  $c_{1,1}$ , i.e.,  $c_{1,1} = \max\{b_{1,1}^*, \dots, b_{1,q}^*\}$ . By the same way, all matrix elements  $c_{i,j}$  ( $i, j = 1, \dots, n$ ) can be calculated by taking every small matrix  $M_1$  with size  $q \times q$  of the large system data as the compression object, then it will generate a new matrix  $M$  with size  $n \times n$ . Now, we have compressed the data of size  $(q \cdot n) \times (q \cdot n)$  into a data of size  $n \times n$ , whose size has been able to match the input layer of NN trained by the data of the small system before. In Figs. 2(a) and 2(d), we take  $q = 2, 3$  as examples to show the entire process of our data compression scheme.

Let's now feed the data compressed above into the input layer of NN trained before to test how well our scheme works [Figs. 2(e) and 2(f)]. As a result, we find that it is highly consistent with the results of traditional method [Figs. 2(g) and 2(h)]. The thermalization, SMPE, and localization of large systems can be successfully identified using our data compression scheme without retraining a new NN. Because our method involves data compression and NN, we call it the DCNN approach.

## V. THE DCNN FOR MANY-BODY LOCALIZATION

Now, we examine whether the DCNN approach is also applicable to other phase classification problems. As a concrete example, we conduct our analysis on a prototype Hamiltonian that has been studied extensively in the Many-Body Localization (MBL) literature: A 1D  $s = 1/2$  spin chain XXZ Hamiltonian with nearest-neighbor interactions,  $H = \sum_{i=1}^{Len} J(s_i^x s_{i+1}^x + s_i^y s_{i+1}^y) + \Delta s_i^z s_{i+1}^z + h_i s_i^x$ , where  $J$  and  $\Delta$  are the coupling constant, and  $h_i$  are random fields uniformly distributed over  $[-h, h]$ ,  $Len$  determines the upper limit of the summation. The total magnetization  $S_z \equiv \sum_i s_i^z$  is a good quantum number, and hence we restrict our calculation for  $S_z = 0$ . We also remark that this Hamiltonian can be mapped onto a Fermi-Hubbard model using a Jordan-Wigner transformation, where  $J$  is equivalent to the hopping coefficient and  $\Delta$  is equivalent to the interaction strength. Thus, with strong enough disorder  $h$  the spin chain is expected to be in the MBL phase for  $\Delta \neq 0$ .

In general, the statistical energy gap distribution is a sensitive indicator of the MBL phase transition, which is determined by the average ratio between the smallest and the largest adjacent energy gaps [46]  $r = \frac{\min\{\delta_{n+1}, \delta_n\}}{\max\{\delta_{n+1}, \delta_n\}}$ , where  $\delta_n = E_n - E_{n-1}$ , and  $E_n$  is the ordered list of many-body energy levels. In the thermalized extended phase,  $r^{\text{th}} = 0.536$ , while in the MBL phase,  $r^{\text{MBL}} = 0.386$  [47]. As shown in Fig. 3(b), the regions A and B are separated by an intersection point of curve lines with different  $Len$ , and denote the thermalized phase and the MBL phase, respectively.

Now we apply the DCNN approach to identify the MBL of the large system using the small NN trained by small system data. First, we choose the data of system with  $Len = 10$  as training data to train a small NN, where the Hilbert space is only 252 and the size of its eigenstates is  $252 \times 252$ , then compress the data of system with  $Len = 12$  and 14 into data of size  $252 \times 252$ , so that their size is able to match the input layer of NN trained before. Last, we feed these compressed

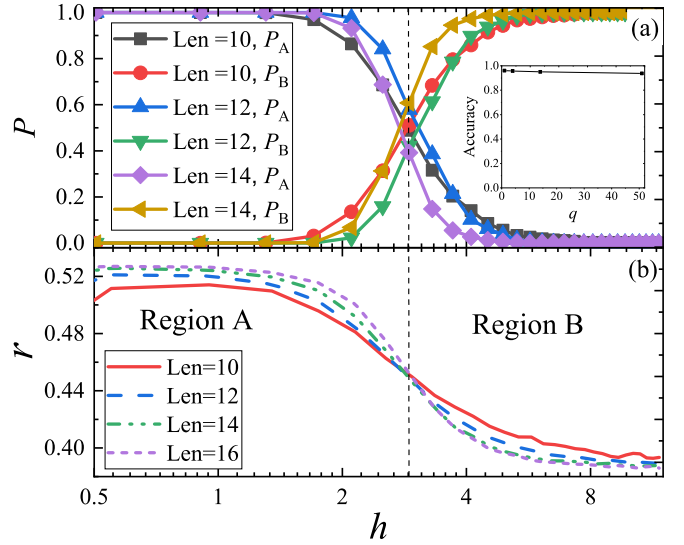


FIG. 3. (a) Results from NN,  $P_A$ , and  $P_B$  represent the probability that the system is classified as thermalized and MBL for a given  $h$ , respectively. Here, the number of instances of the training set is 10000, and  $N_m = 1000$  for every given  $h$ . The inserted figure represents the relationship between the accuracy of the DCNN approach and the compression factor  $q$ . (b) Results from the traditional method;  $r$  is the average ratio between the smallest and the largest adjacent energy gaps. Regions A and B denote the phases of thermalization and MBL, respectively. Other parameters:  $J = 1$  and  $\Delta = 2$ .

data into the input layer of that small NN to evolve. We find that the MBL phases of larger systems can be successfully identified only using this small network by these compressed data [Fig. 3(a)]. In this process the data have been compressed by a factor of  $q = 4$  and 14 for  $Len = 12$  and 14, respectively. Furthermore, we also discuss the relationship between the accuracy of the DCNN approach and the compression factor  $q$ , as shown in the inserted figure of Fig. 3(a). The result shows that, even if the compression factor  $q$  reaches 51 ( $Len = 16$ ), the DCNN method still has high accuracy.

## VI. CONCLUSION

By making full use of the properties of eigenstates, we have developed a DCNN approach for identifying the SPMEs and phase transitions of large systems only using NN trained by small system data. In this method, we can identify the thermalization, SPME, and localization nearly 100% and give clear boundaries among these phases for many times the size of systems, without onerously and repetitively training a new and large-scale network. It not only greatly saves computing resources and time but also makes it possible to solve physical problems by ML method that cannot be calculated because the data set is too large. To the best of our knowledge, compared with the application of other ML methods in physics, our method is the first one that can study the physical characteristics of large systems directly using compressed data without repetitive training, and such method is also applicable for recognizing other more complex phase transitions.

## ACKNOWLEDGMENTS

This work is supported by the National Natural Science Foundation (NSF) of China under Grants No. 11847227,

No. 12005107, and No. 11904242 and the NSF of Hebei province under Grant No. A2019210280.

- 
- [1] G. Roati, C. D'Errico, L. Fallani, M. Fattori, C. Fort, M. Zaccanti, G. Modugno, M. Modugno, and M. Inguscio, *Nature (London)* **453**, 895 (2008).
- [2] M. Schreiber, S. S. Hodgman, P. Bordia, H. P. Lüschen, M. H. Fischer, R. Vosk, E. Altman, U. Schneider, and I. Bloch, *Science* **349**, 842 (2015).
- [3] L. Fallani, J. E. Lye, V. Guarrera, C. Fort, and M. Inguscio, *Phys. Rev. Lett.* **98**, 130404 (2007).
- [4] X. Li, X. Li, and S. Das Sarma, *Phys. Rev. B* **96**, 085119 (2017).
- [5] H. P. Lüschen, S. Scherg, T. Kohlert, M. Schreiber, P. Bordia, X. Li, S. Das Sarma, I. Bloch, *Phys. Rev. Lett.* **120**, 160404 (2018).
- [6] S. Aubry and G. André, *Ann. Israel Phys. Soc.* **3**, 133 (1980).
- [7] P. Bordia, H. Lüschen, S. Scherg, S. Gopalakrishnan, M. Knap, U. Schneider, and I. Bloch, *Phys. Rev. X* **7**, 041047 (2017).
- [8] H. P. Lüschen, P. Bordia, S. Scherg, F. Alet, E. Altman, U. Schneider, and I. Bloch, *Phys. Rev. Lett.* **119**, 260401 (2017).
- [9] Y. Wang, X.-J. Liu, and D. Yu, *Phys. Rev. Lett.* **126**, 080602 (2021).
- [10] J. Biddle, B. Wang, D. J. Priour, Jr., and S. Das Sarma, *Phys. Rev. A* **80**, 021603(R) (2009).
- [11] J. Biddle and S. Das Sarma, *Phys. Rev. Lett.* **104**, 070601 (2010).
- [12] S. Ganeshan, J. H. Pixley, and S. Das Sarma, *Phys. Rev. Lett.* **114**, 146601 (2015).
- [13] H. Yao, H. Khoudli, L. Bresque, and L. Sanchez-Palencia, *Phys. Rev. Lett.* **123**, 070405 (2019).
- [14] G. Semeghini, M. Landini, P. Castilho, S. Roy, G. Spagnolli, A. Trenkwalder, M. Fattori, M. Inguscio, and G. Modugno, *Nat. Phys.* **11**, 554 (2015).
- [15] M. Rossignolo and L. Dell'Anna, *Phys. Rev. B* **99**, 054211 (2019).
- [16] T. Kohlert, S. Scherg, X. Li, H. P. Lüschen, S. Das Sarma, I. Bloch, and M. Aidelsburger, *Phys. Rev. Lett.* **122**, 170403 (2019).
- [17] F. A. An, K. Padavić, E. J. Meier, S. Hegde, S. Ganeshan, J. Pixley, S. Vishveshwara, and B. Gadway, *Phys. Rev. Lett.* **126**, 040603 (2021).
- [18] X. Wei, R. Mondaini, and G. Xianlong, *arXiv:2001.04105* (2020).
- [19] Y. Liu, Y. Wang, X. J. Liu, Q. Zhou, and S. Chen, *Phys. Rev. B* **103**, 014203 (2021).
- [20] M. I. Jordan and T. M. Mitchell, *Science* **349**, 255 (2015).
- [21] Y. LeCun, Y. Bengio, and G. Hinton, *Nature (London)* **521**, 436 (2015).
- [22] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, *Rev. Mod. Phys.* **91**, 045002 (2019).
- [23] R. Biswas, L. Blackburn, J. Cao, R. Essick, K. A. Hodge, E. Katsavounidis, K. Kim, Y.-M. Kim, E.-O. Le Bigot, C.-H. Lee, J. J. Oh, S. H. Oh, E. J. Son, Y. Tao, R. Vaulin, and X. Wang, *Phys. Rev. D* **88**, 062003 (2013).
- [24] B. S. Rem, N. Käming, M. Tarnowski, L. Asteria, N. Fläschner, C. Becker, K. Sengstock, and C. Weitenberg, *Nat. Phys.* **15**, 917 (2019).
- [25] L. Wang, *Phys. Rev. B* **94**, 195105 (2016).
- [26] J. Carrasquilla and R. G. Melko, *Nat. Phys.* **13**, 431 (2017).
- [27] Y. Zhang and E.-A. Kim, *Phys. Rev. Lett.* **118**, 216401 (2017).
- [28] D.-L. Deng, X. Li, and S. Das Sarma, *Phys. Rev. B* **96**, 195145 (2017).
- [29] P. Huembeli, A. Dauphin, P. Wittek, and C. Gogolin, *Phys. Rev. B* **99**, 104106 (2019).
- [30] X.-Y. Dong, F. Pollmann, and X.-F. Zhang, *Phys. Rev. B* **99**, 121104(R) (2019).
- [31] E. P. Van Nieuwenburg, Y.-H. Liu, and S. D. Huber, *Nat. Phys.* **13**, 435 (2017).
- [32] P. Zhang, H. Shen, and H. Zhai, *Phys. Rev. Lett.* **120**, 066401 (2018).
- [33] N. Sun, J. Yi, P. Zhang, H. Shen, and H. Zhai, *Phys. Rev. B* **98**, 085402 (2018).
- [34] P. Huembeli, A. Dauphin, and P. Wittek, *Phys. Rev. B* **97**, 134109 (2018).
- [35] Y.-H. Tsai, M.-Z. Yu, Y.-H. Hsu, and M.-C. Chung, *Phys. Rev. B* **102**, 054512 (2020).
- [36] J. F. Rodriguez-Nieva and M. S. Scheurer, *Nat. Phys.* **15**, 790 (2019).
- [37] N. L. Holanda and M. A. R. Griffith, *Phys. Rev. B* **102**, 054107 (2020).
- [38] Y. Zhang, P. Ginsparg, and E.-A. Kim, *Phys. Rev. Research* **2**, 023283 (2020).
- [39] Y. Long, J. Ren, and H. Chen, *Phys. Rev. Lett.* **124**, 185501 (2020).
- [40] M. S. Scheurer and R.-J. Slager, *Phys. Rev. Lett.* **124**, 226401 (2020).
- [41] F. Schindler, N. Regnault, and T. Neupert, *Phys. Rev. B* **95**, 245134 (2017).
- [42] E. van Nieuwenburg, E. Bairey, and G. Refael, *Phys. Rev. B* **98**, 060301(R) (2018).
- [43] J. Venderley, V. Khemani, and E.-A. Kim, *Phys. Rev. Lett.* **120**, 257204 (2018).
- [44] Y.-T. Hsu, X. Li, D.-L. Deng, and S. Das Sarma, *Phys. Rev. Lett.* **121**, 245701 (2018).
- [45] See Supplemental Material at <http://link.aps.org/supplemental/10.1103/PhysRevB.103.134203> for detailed structure of convolutional neural network and other details.
- [46] V. Oganesyan and D. A. Huse, *Phys. Rev. B* **75**, 155111 (2007).
- [47] Y. Y. Atas, E. Bogomolny, O. Giraud, and G. Roux, *Phys. Rev. Lett.* **110**, 084101 (2013).