

**Few-shot machine learning in the three-dimensional Ising model**Rui Zhang,<sup>1,2,3,\*</sup> Bin Wei,<sup>1,2,\*</sup> Dong Zhang,<sup>1,2</sup> Jia-Ji Zhu,<sup>4,†</sup> and Kai Chang<sup>1,2,‡</sup><sup>1</sup>*SKLSM, Institute of Semiconductors, Chinese Academy of Sciences, 100083 Beijing, China*<sup>2</sup>*Center for Excellence in Topological Quantum Computation, University of Chinese Academy of Sciences, Beijing 100190, China*<sup>3</sup>*COT Design Department, HiSilicon Technologies Co., Ltd., 518129 Shenzhen, China*<sup>4</sup>*School of Science and Laboratory of Quantum Information Technology, Chongqing University of Posts and Telecommunications, 400061 Chongqing, China*

(Received 14 October 2018; revised manuscript received 12 March 2019; published 19 March 2019)

We investigate theoretically the phase transition in a three-dimensional cubic Ising model utilizing state-of-the-art machine learning algorithms. Supervised machine learning models show high accuracies (99%) in phase classification and very small relative errors ( $<10^{-4}$ ) of the energies in different spin configurations. Unsupervised machine learning models are introduced to study the spin configuration reconstructions and reductions, and the phases of reconstructed spin configurations can be accurately classified by a linear logistic algorithm. Based on the comparison between various machine learning models, we develop a few-shot strategy to predict phase transitions in larger lattices from a trained sample in smaller lattices. The few-shot machine learning strategy for a three-dimensional (3D) Ising model enables us to study the 3D Ising model efficiently and provides an integrated and highly accurate approach to other spin models.

DOI: [10.1103/PhysRevB.99.094427](https://doi.org/10.1103/PhysRevB.99.094427)**I. INTRODUCTION**

Machine learning techniques have achieved remarkable progress in many different fields, e.g., image recognition, fraud detection, natural language processing, auto driving, etc., due to their powerful magic to extract features from huge data sets without explicit guidance from a human programmer. In physics, machine learning algorithms have been employed to study many-body physics [1–3], strong correlated systems [4,5], electronic or transport properties [6–8], structure predictions [9–12], phase matter or phase transitions [13–18], and different spin models [14,19–23].

Among the various studied spin models, the Ising model has attracted the most interest, because it is not only rich in physics and mathematics, such as phase transitions [24] and Kolmogorov's zero-one laws [25], but also inspirations in many other fields, such as the social sciences [26] and neuroscience [27]. The possible applications of the Ising model are extremely large. For instance, it can be used for magnetic insulators [28], binary alloys [29–31], a lattice gas model for fluids [32], ferroelectrics [33], biological systems [34], or general demonstrations of statistical mechanics [35].

Of particular interest is the three-dimensional (3D) Ising model, which has been introduced to exactly describe the second generation D-wave quantum computer of 512 qubits. Very recently, quantum supremacy has made the 3D Ising model again a hot topic [36]. Although the Ising model looks very simple, the complexity of the solutions to the model

increases along with the dimensions. Until now, the Ising model has been solved mathematically rigorously in only one- [37] and two-dimensional lattices [38]. In three dimensions or higher-dimensional cases, the analytic solutions remain unavailable. Instead, various approximation methods have been developed to study the Ising model as a last resort, including the conformal bootstrap approach [39–41], and numerical approaches such as Monte Carlo simulation [42] and mean-field theories [43].

It is notable that there exists a reciprocal relationship between the Ising model and neural networks or machine learning algorithms. On the one hand, the Ising model has played a critical role in the development of machine learning techniques. For example, the famous Hopfield network [34] is a typical recurrent artificial network based on the dynamical Ising model. On the other hand, machine learning techniques provide an effective approach to solve the Ising model. Carrasquilla and Melko have investigated the phase transition of a two-dimensional (2D) Ising model recently, utilizing supervised learning models such as neural networks (NNs) and 2D convolutional neural networks (CNNs) [14], and they demonstrated that machine learning techniques can study the phase transition of 2D Ising models with very high accuracy. Besides the supervised learning models, unsupervised learning models such as principal component analysis (PCA) and auto encoder (AE) have also been applied to study the 2D XY model [21,22] and 3D XY model [23]. However, the 3D Ising model still remains unexplored using the machine learning techniques.

In this work, we have studied the phase transition of a ferromagnetic 3D Ising model by Monte Carlo-sampled machine learning techniques. Instead of standard Monte Carlo simulations [42,44,45], the needed parameters are predicted

\*These authors contributed equally to this work.

†Corresponding author: zhujj@cqupt.edu.cn

‡Corresponding author: kchang@semi.ac.cn

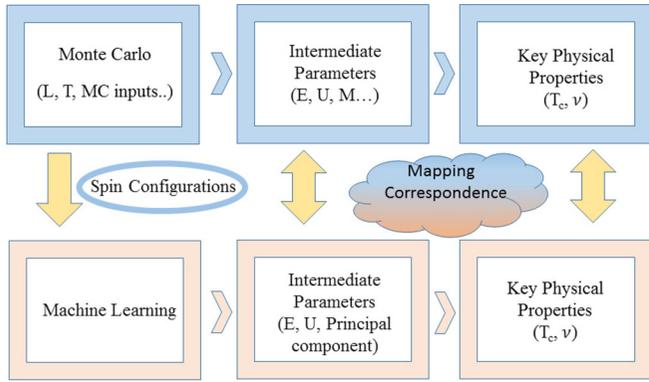


FIG. 1. (Upper panel) Standard Monte Carlo procedure to capture the key physical parameters  $T_c$  and  $\nu$ ; (lower panel) Monte Carlo–based machine learning procedure. The relationship between the Monte Carlo and the machine learning procedures are indicated by the arrows in the middle.

by the state-of-the-art machine learning techniques. In order to compare the accuracy and efficiency of current machine learning algorithms, we perform both supervised and unsupervised machine learning algorithms in the Ising models. Since the computation consumptions increase sharply with increasing lattice sizes, we developed a few-shot machine learning strategy by combining specific supervised and unsupervised machine learning algorithms to predict phase transitions in larger lattices from trained data in smaller lattices. The paper is organized as follows: In Sec. II, we present the 3D Ising model and methods used during the calculations. In Secs. III and IV, we demonstrate the performances of machine learning algorithms and the few-shot machine learning strategy. In Sec. IV, we report the conclusions of our work.

## II. MODELS AND METHODS

The Hamiltonian of the ferromagnetic 3D Ising model in a cubic lattice is given by

$$H = -J \sum_{(i,j)} \sigma_i \sigma_j, \quad (1)$$

with uniform interaction strength  $J = 1$  and binary spin configurations  $\sigma_i \in \{+1 = \uparrow, -1 = \downarrow\}$  on each site  $i$  where  $i$  runs over all the spin sites ( $N$ ). The sum goes over all nearest-neighbor pairs, and the periodic boundary condition is taken in our calculation. The magnetization of a spin sample is defined as  $M = \frac{1}{N} \sum_i \sigma_i$ . In the 3D Ising model studied here, there are  $2^N$  different spin configurations for cubic lattices and two key physical parameters, the transition temperature ( $T_c$ ) and the critical exponent ( $\nu$ ).

To investigate the two key physical parameters, we consider cubic lattices with different sizes ( $L$ ) and adopt a Monte Carlo–sampled machine learning procedure as illustrated in Fig. 1. In Fig. 1, the upper panel indicates a standard Monte Carlo procedure in which all the intermediate parameters, such as energy ( $E$ ), binder cumulant ( $U$ ), magnetization ( $M$ ), etc. for fitting the key physical quantities  $T_c$  and  $\nu$  are

calculated by the standard Monte Carlo algorithm itself. While in the Monte Carlo–sampled machine learning procedure as illustrated in the lower panel, the Monte Carlo method is applied to generate raw data sets to feed into machine learning models. At the first stage, the input data sets will be minimized to spin configurations only. Afterwards, we prove the exact correspondence between the intermediate parameters given by standard Monte Carlo simulation and machine learning models. Finally, one can compare the predicted  $T_c$  and  $\nu$  from both procedures.

The Monte Carlo simulation, implemented by the Swendsen-Wang cluster algorithms (SWCA) [46], is performed to calculate the transition temperatures ( $T_c$ ) by creating 2000 independent spin configurations at each temperature step and to provide initial raw data sets to train the machine learning models. The SWCA algorithm updates a cluster of spins for each step, which reduces the correlation of data remarkably. We take 50,000 update steps between every two outputs of spin configuration in order to reduce the data correlation more efficiently.

Machine learning methods are applied to get the above parameters and are categorized as supervised and unsupervised learning models as follows. Supervised learning is the machine learning task of inferring a function from labeled training data [47], in which each sample is a pair of an input object and its desired label. A supervised learning algorithm learns from the training sets and then produces an inferred function which can be applied to the test sets that have never been trained before. The accuracy of the supervised learning model is determined by comparing the output of the trained model and the original labels on the test sets, which is also an important benchmark to evaluate the performance of the machine learning model. In this case, the input data for the supervised learning models are the spin configurations of the cubic Ising lattice. The corresponding labels are determined by the corresponding  $T$  of the input configurations: the spin configurations are labeled by (1, 0) when  $T < T_c$ , and by (0, 1) when  $T > T_c$ . Three supervised learning models are employed in this work: support vector machine (SVM) [48,49], neural network (NN), and 3D convolution neural network (CNN) [50,51].

Contrary to the supervised learning method, unsupervised learning models are expected to infer a function to describe hidden structure from unlabeled data, which can be a classification or categorization not included in the given physical observability. To solve complicated physical problems, the labels of the target system are usually difficult to get from calculations or experiments, which makes the unsupervised machine learning necessary and useful. Since the input data sets to the unsupervised learning are merely spin configurations, the spin configurations with different labels will fall into different clusters as the unsupervised learning models are well trained. In this work we take two commonly used unsupervised machine learning models, the principal component analysis (PCA) [52,53] and the restricted Boltzmann machine (RBM). All the supervised and unsupervised machine learning models are trained by the spin configurations generated with  $T \in [3, 6]$  in units of  $J/k$  in a total of  $n = 100,000$  configurations for every size.

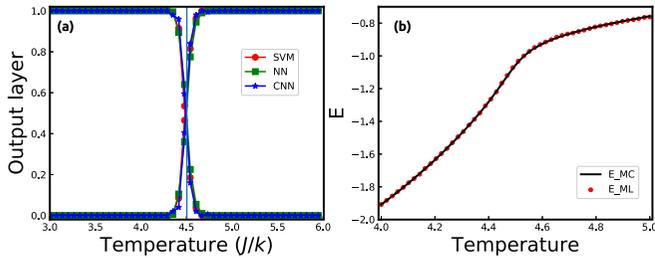


FIG. 2. (a) The average outputs of different machine learning models SVM (red dot), NN (green rectangle), and 3D CNN (blue star) on the test set with  $L = 16$ . (b) The average energies predicted by the 3D CNN model. The black solid line and the red dots are obtained from the Monte Carlo simulations and 3D CNN model, respectively.

### III. CONVENTIONAL MACHINE LEARNING IN A 3D ISING MODEL

First, we perform supervised classifications for spin configurations in the 3D Ising model utilizing SVM, NN, and 3D CNN algorithms. The input data for supervised machine learning models are obtained from the Monte Carlo sampling at different temperatures  $T$  in a finite cubic lattice  $L = 16$ . The spin configurations are binary labeled by their corresponding phases, i.e., high-temperature paramagnetic phase when  $T > T_c$  and the low-temperature ferromagnetic phase when  $T < T_c$  [14]. The data sets are randomly divided into the train and test sets; all the supervised machine learning models are trained in the train sets and predicted in the test sets. In this work, we adopt the SVM model which is implemented in the Scikit-Learn library [54,55] with optimized hyper parameters and build the NN and 3D CNN models in the framework of the Keras library [56].

The numerical results obtained from the SVM, NN, and 3D CNN models are shown in Fig. 2(a). All the supervised machine learning models can classify the high- and low-temperature phases with accuracies about 99%, and the average outputs of the supervised machine learning models cross at the predicted  $T_c$ , which is very close to the Monte Carlo simulations. The result shows that all three supervised machine learning models used in our calculations are effective to distinguish the high- and low-temperature phases, although they are quite different.

Machine learning models can not only classify phases of the 3D Ising model, but they also can predict physical quantities, such as the transition temperature  $T_c$ , the critical exponents  $\nu$ , and the averaged energies of various spin configurations. We train a 3D CNN model to predict energies of different spin configurations [see Fig. 2(b)]. The predicted energies agree very well with the Monte Carlo simulations, and the relative error of the predicted energy is smaller than  $1.0 \times 10^{-4}$ .

Until now, all the input data are directly used to train the supervised machine learning models without any preprocessing. Next, we do some preprocessing to the 3D spin configurations. Two unsupervised machine learning models, PCA and RBM, are considered to reduce dimensions and reconstruct 3D spin configurations, respectively. We take the PCA and Bernoulli RBM implemented in the Scikit-Learn library [54,55].

We consider one principal component when we apply PCA to 3D spin configurations, since we find that the first principal component possesses the greatest explained variance [see the inset of Fig. 3(a)]. In Fig. 3(a), one can find that the outputs  $p1$  are linear with magnetization  $M$ , which means the correlation between  $p1$  and  $M$  is almost 1.0 and the physical observable, magnetization, can be learned from the first PCA of the input data. We also plot the absolute value of  $p1$  as a function of temperature in Fig. 3(b) and find that the lines with different size parameter  $L$  tend to approach each other when the lattice size  $L$  increases. Therefore we take a finite-size scaling analysis to determine the critical exponent  $\nu$  and predict the transition temperature of the 3D Ising model with infinite  $L$  based on the first principal component and the predicted energy [57,58].

The results of finite-size analysis are shown in Figs. 3(c) and 3(d). The critical exponent of the 3D Ising model is a very important physical quantity to describe the critical behavior of the phase transition and still has not been studied utilizing the machine learning models [21–23,59,60]. Here it can be determined based on cubic spin configurations with  $L \in [8, 28]$ .  $K_c = \frac{1}{T_c}$  is obtained by fitting the data with  $L \geq 22$ . The critical exponent  $\nu$  and the critical temperature  $T_c$  are measured to be 0.629 and 4.511 417 ( $K_c = 0.221 659 8$ ), respectively, which agree very well with the results obtained from the renormalization group theory [57,58].

Since  $p1$  is also closely related to the high- and low-temperature phases [see the red and blue scatters shown in Fig. 3(a)], this feature makes it possible to classify different phases based on the PCA outputs. In addition, the RBM is applied to reconstruct the 3D spin configurations. The results obtained from supervised learning models based on the reduced and reconstructed spin configurations are shown in Fig. 4(a). From this figure, one can see clearly that the combined model can be used to distinguish different phases with high enough accuracy. Notice that the RBM-reconstructed 3D Ising model can be simply classified by a logistic linear classifier. The reconstruction process differentiates various spin configurations, and the PCA reduction process makes the classification procedure much more efficient than previous direct classification machine learning models.

A full comparison between preprocessed machine learning models and direct supervised machine learning models is shown in Fig. 4(b). One can find that the overall accuracies increase with the size parameter  $L$ , and the performance of all these machine learning models approach each other. The different behaviors of these models arise from the complexity of these machine learning models, and the preprocessed model could require fewer computation resources without loss of accuracy.

### IV. FEW-SHOT MACHINE LEARNING IN 3D ISING MODEL

Supervised machine learning models trained from the data with lattice size  $L$  cannot be applied to the lattices with  $L + \Delta L$  directly. In order to make it possible to predict the phase transition in a larger size  $L + \Delta L$ , we combine the PCA and the NN together. We first perform PCA to the spin lattice with different sizes (e.g.,  $L = 10, 12, 14, 16, 20, 24$ ), and only the

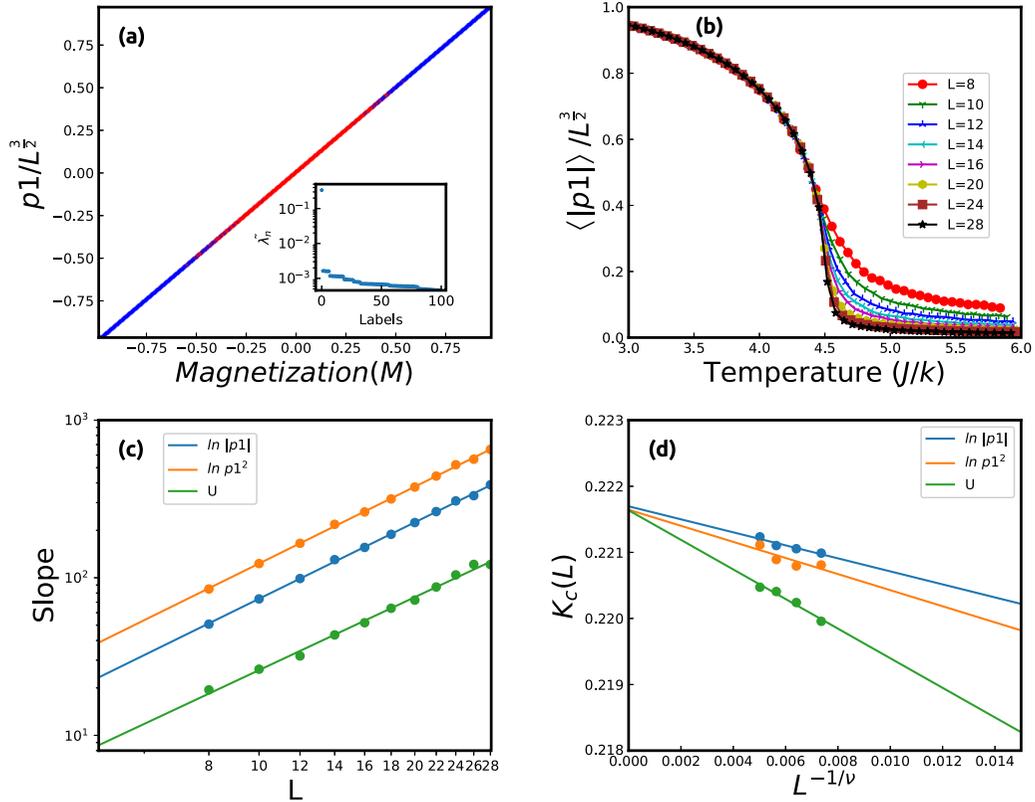


FIG. 3. (a) The output of the first principal component as a function of the magnetization ( $M$ ) in a cubic lattice with  $L = 16$ . The explained variance, i.e., the ratio between the variance of the principal components and the total variance, is shown in the inset of (a). (b) Absolute value of the first PCA output as a function of temperature. (c) Finite-size scaling analysis to determine critical exponent  $\nu$  based on the first principal component and energy; the critical exponent is fitted to be 0.629. (d) Finite-size scaling analysis to determine transition  $K_c$ .

first two leading principal components remain as the input for the next step.

Such a group of preprocessed data makes next-step train sets. Then the train sets are fed to a NN model which consists of an input layer with two units, one hidden layer with 30 units, and at the end the output layer with two units. The performances of the mixed trained model on test sets are shown in Fig. 5(a), and one can find that the mixed trained model performs very well with different size parameters. The overall accuracies of this mixed trained model obtained from the test sets with different  $L$  are shown in Fig. 5(b), and the

performance of the mixed trained model is comparable to the models based on particular  $L$  shown in Fig. 4(b).

After the confirmation of the validity of the mixed trained model for 3D Ising models in known lattice sizes, we extend it to unexplored lattice sizes. By applying the mixed trained model to the spin configurations with  $L = 28$ , we find out that the model can predict the  $L = 28$  lattice with an amazing accuracy of 0.996, therefore verifying the capability of the model to predict critical temperatures in larger lattices which have never been trained. The average outputs at different temperatures are shown in Fig. 6(a), and the predicted

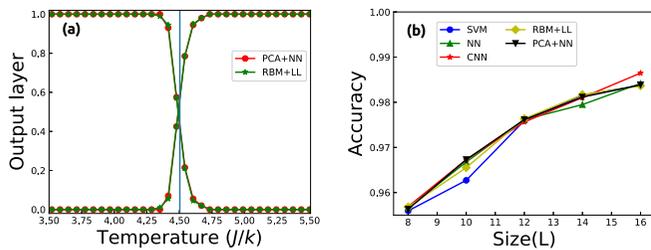


FIG. 4. (a) The average outputs of different RBM combined with a linear logistic classifier and PCA combined with a NN classifier, both with  $L = 16$ . (b) Accuracies as a function of  $L$  for different supervised and combined supervised models.

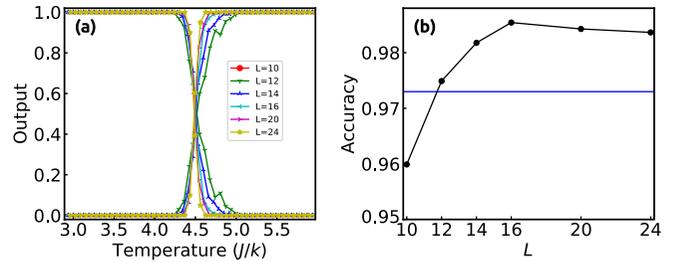


FIG. 5. The performance of the model trained by mixing data with different  $L$ . (a) The average output of the mixed trained model as a function of temperature. (b) The overall accuracies of the mixed trained model tested on different  $L$ .

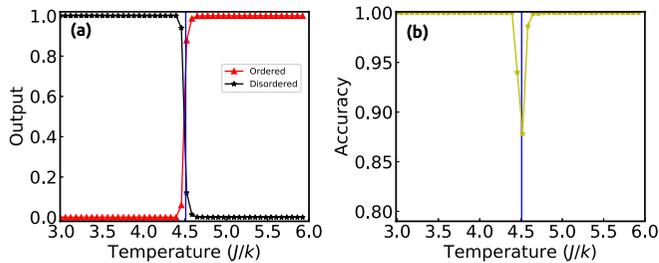


FIG. 6. The performance of the mixed trained model on the unexplored size  $L = 28$ . (a) The average output of the mixed trained model as a function of temperature. (b) The accuracies of the model at different temperatures  $T$ .

phase transition temperature is very close to the standard Monte Carlo simulation [the blue vertical line in Fig. 6(a)]. Figure 6(b) shows that the mixed trained model possesses very high accuracies at different temperatures  $T$ , even when  $T$  is very close to the critical temperature  $T_c$ . Since the prediction of larger lattices can be obtained merely from training smaller samples, the combination of PCA and NN enables us to predict the phase transition in a quite large lattice which is unattainable by standard Monte Carlo procedures. This mixed trained model demonstrates a convincing few-shot machine learning strategy for spin models.

## V. CONCLUSIONS

In conclusion, we performed several machine learning models for 3D Ising models. Based on the state-of-the-art Monte Carlo simulations, we find both supervised and unsupervised machine learning models can extract physical information from raw data. A critical exponent and the transition temperature are accurately fitted by PCA. The combination of RBM and a simple linear logistic classifier is proven to be as powerful as supervised machine learning models for 3D Ising models. We show a type of few-shot machine learning strategy based on PCA and a feed-forward NN classifier. The few-shot machine learning strategy for the 3D Ising model makes it possible to predict the phase transition in a larger lattice based on data from some smaller samples, which also provides an approach to other spin models. Our work presents an integrated and highly accurate machine-learning toolbox and demonstrates a few-shot learning strategy for studying more spin models.

## ACKNOWLEDGMENTS

This work was supported by the NSFC (Grants No. 11434010, No. 61674145 and No. 11404043), the MOST of China (Grants No. 2016YFE0110000, No. 2017YFA0303400, and Grants No. 2015CB921503) and the Chinese Academy of Sciences (Grants No. QYZDJ-SSW-SYS001 and XXH13506-202).

## APPENDIX A: ARCHITECTURES OF NEURAL NETWORKS

In this work we study different deep learning (DL) models where we give the architectures of full connected neural networks (NN) and 3D convolutional neural networks (3D CNN). The full connected NN model consists of four kinds of layers: an input layer, hidden layer, dropout layer, and output layer. The input layer contains  $L^3$  neurons, the hidden layer contains 100 neurons which are activated by the sigmoid function, and the output layer contains two neurons which are activated by softmax functions. We take a dropout rate of 0.1 to prevent overfitting of this model.

The 3D CNN model is composed of a 3D input layer, 3D convolutional layers, a dropout layer, a full connected hidden layer, and the output layer. The 3D convolutional layers apply  $64 \times 2$  filters to the spin configurations, and the full connected hidden layer contains 64 neurons. All the layers are activated by sigmoid functions except for the output layer, which is activated by softmax function. The dropout layer is added between the 3D convolutional layers and the output layer with dropout rate 0.25.

## APPENDIX B: ALGORITHM COMPLEXITY

The mixed trained model consists of two procedures: (1) applying PCA to all the spin configurations of the train set and (2) applying NN to the reduced spin configurations. The determined complexity of this model comes from the PCA step. In this paper we implement PCA utilizing Scikit-Learn with randomized truncated singular value decomposition (SVD) [56,61] where the complexity is  $O(n_{\max}^2 n_{\text{components}})$ . Here  $n_{\max}$  is the max value between the number of spin configurations ( $n_{\text{samples}}$ ) and dimensions of the spin configurations ( $L^3$ ), and  $n_{\text{components}}$  is the number of principal components [56,61]. When the size of the 3D Ising model is big enough ( $L^3 > n_{\text{samples}}$ ), the complexity is proportional to  $L^6$ , which grows rapidly when  $L$  increases.

[1] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).  
 [2] L.-F. Arsenault, A. Lopez-Bezanilla, O. A. von Lilienfeld, and A. J. Millis, *Phys. Rev. B* **90**, 155136 (2014).  
 [3] Z. Cai and J. Liu, *Phys. Rev. B* **97**, 035116 (2018).  
 [4] T. Ohtsuki and T. Ohtsuki, *J. Phys. Soc. Jpn.* **85**, 123706 (2016).  
 [5] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, *Phys. Rev. X* **7**, 031038 (2017).  
 [6] F. Faber, A. Lindmaa, O. A. von Lilienfeld, and R. Armiento, *Int. J. Quantum Chem.* **115**, 1094 (2015).  
 [7] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, *Phys. Rev. Lett.* **108**, 058301 (2012).

[8] A. Lopez-Bezanilla and O. A. von Lilienfeld, *Phys. Rev. B* **89**, 235411 (2014).  
 [9] T. L. Jacobsen, M. S. Jørgensen, and B. Hammer, *Phys. Rev. Lett.* **120**, 026102 (2018).  
 [10] S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, *Nat. Phys.* **12**, 469 (2016).  
 [11] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K.-R. Müller, and O. A. von Lilienfeld, *New J. Phys.* **15**, 095003 (2013).  
 [12] O. Isayev, C. Oses, C. Toher, E. Gossett, S. Curtarolo, and A. Tropsha, *Nat. Commun.* **8**, 15679 (2017).

- [13] G. Torlai and R. G. Melko, *Phys. Rev. B* **94**, 165134 (2016).
- [14] J. Carrasquilla and R. G. Melko, *Nat. Phys.* **13**, 431 (2017).
- [15] Y. Zhang and E.-A. Kim, *Phys. Rev. Lett.* **118**, 216401 (2017).
- [16] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, *Sci. Rep.* **7**, 8823 (2017).
- [17] G. Torlai and R. G. Melko, *Phys. Rev. Lett.* **119**, 030501 (2017).
- [18] P. Ponte and R. G. Melko, *Phys. Rev. B* **96**, 205146 (2017).
- [19] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, *Phys. Rev. B* **95**, 041101 (2017).
- [20] L. Wang, *Phys. Rev. B* **94**, 195105 (2016).
- [21] W. Hu, R. R. P. Singh, and R. T. Scalettar, *Phys. Rev. E* **95**, 062122 (2017).
- [22] C. Wang and H. Zhai, *Phys. Rev. B* **96**, 144432 (2017).
- [23] S. J. Wetzel, *Phys. Rev. E* **96**, 022140 (2017).
- [24] T.-D. Lee and C.-N. Yang, *Phys. Rev.* **87**, 410 (1952).
- [25] A. N. Kolmogorov, *Foundations of the Theory of Probability*, 2nd English ed. (Courier Dover Publications, Mineola, New York, 2018).
- [26] D. Stauffer, *Am. J. Phys.* **76**, 470 (2008).
- [27] E. Schneidman, M. J. Berry II, R. Segev, and W. Bialek, *Nature (London)* **440**, 1007 (2006).
- [28] R. Cowley and W. Buyers, *Rev. Mod. Phys.* **44**, 406 (1972).
- [29] K. Binder, *Z. Phys.* **267**, 313 (1974).
- [30] M. t. Richards and J. Cahn, *Acta Metall.* **19**, 1263 (1971).
- [31] P. C. Clapp and S. C. Moss, *Phys. Rev.* **171**, 754 (1968).
- [32] K. Binder, *Phys. Rev. A* **25**, 1699 (1982).
- [33] R. Blinc and B. Žekš, *Ferroelectrics* **72**, 193 (1987).
- [34] J. J. Hopfield, *Proc. Natl. Acad. Sci. USA* **79**, 2554 (1982).
- [35] R. J. Baxter, *Exactly Solved Models in Statistical Mechanics* (Elsevier, New York, 2016).
- [36] S. Boixo, S. V. Isakov, V. N. Smelyanskiy, R. Babbush, N. Ding, Z. Jiang, M. J. Bremner, J. M. Martinis, and H. Neven, *Nat. Phys.* **14**, 595 (2018).
- [37] E. Ising, *Z. Phys.* **31**, 253 (1925).
- [38] L. Onsager, *Phys. Rev.* **65**, 117 (1944).
- [39] D. Poland and D. Simmons-Duffin, *Nat. Phys.* **12**, 535 (2016).
- [40] S. El-Showk, M. F. Paulos, D. Poland, S. Rychkov, D. Simmons-Duffin, and A. Vichi, *J. Stat. Phys.* **157**, 869 (2014).
- [41] S. El-Showk, M. F. Paulos, D. Poland, S. Rychkov, D. Simmons-Duffin, and A. Vichi, *Phys. Rev. D* **86**, 025022 (2012).
- [42] H. Blote, E. Luijten, and J. R. Heringa, *J. Phys. A: Math. Gen.* **28**, 6289 (1995).
- [43] M. H. Jensen and P. Bak, *Phys. Rev. B* **27**, 6853 (1983).
- [44] U. Wolff, *Phys. Lett. B* **228**, 379 (1989).
- [45] T. Preis, P. Virnau, W. Paul, and J. J. Schneider, *J. Comput. Phys.* **228**, 4468 (2009).
- [46] R. H. Swendsen and J.-S. Wang, *Phys. Rev. Lett.* **58**, 86 (1987).
- [47] M. Mohri, A. Rostamizadeh, and A. Talwalkar, *Foundations of Machine Learning* (MIT Press, Cambridge, MA, 2012).
- [48] C. Cortes and V. Vapnik, *Mach. Learn.* **20**, 273 (1995).
- [49] M. Aizerman, E. Braverman, and L. Rozonoer, *Automat. Remote Control* **25**, 821 (1964).
- [50] W. S. McCulloch and W. Pitts, *Bull. Math. Biophys.* **5**, 115 (1943).
- [51] D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Nature (London)* **323**, 533 (1986).
- [52] K. Pearson, *Philos. Mag.* **2**, 559 (1901).
- [53] I. T. Jolliffe, *Principal Component Analysis* (Springer, New York, 2002).
- [54] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg *et al.*, *J. Mach. Learn. Res.* **12**, 2825 (2011).
- [55] L. Buitinck, G. Louppe, M. Blondel, F. Pedregosa, A. Mueller, O. Grisel, V. Niculae, P. Prettenhofer, A. Gramfort, J. Grobler, *et al.*, [arXiv:1309.0238](https://arxiv.org/abs/1309.0238).
- [56] F. Chollet *et al.*, “Keras”, <https://github.com/keras-team/keras> (2015).
- [57] A. M. Ferrenberg and D. P. Landau, *Phys. Rev. B* **44**, 5081 (1991).
- [58] A. Pelissetto and E. Vicari, *Phys. Rep.* **368**, 549 (2002).
- [59] K. Mills and I. Tamblyn, *Phys. Rev. E* **97**, 032119 (2018).
- [60] C. Wang and H. Zhai, *Front. Phys.* **13**, 130507 (2018).
- [61] N. Halko, P.-G. Martinsson, and J. A. Tropp, *SIAM Rev.* **53**, 217 (2011).