Rapid Communications

Self-organizing maps as a method for detecting phase transitions and phase identification

Albert A. Shirinyan,[1] Valerii K. Kozin,[2,1] Johan Hellsvik,[3,4] Manuel Pereiro,[5] Olle Eriksson,[5,6] and Dmitry Yudin[7]

[1]*ITMO University, Saint Petersburg 197101, Russia*

[2]*Science Institute, University of Iceland, Dunhagi-3, IS-107 Reykjavik, Iceland*

[3]*Nordita, Roslagstullsbacken 23, SE-106 91 Stockholm, Sweden*

[4]*Department of Physics, KTH Royal Institute of Technology, SE-106 91 Stockholm, Sweden*

[5]*Department of Physics and Astronomy, Materials Theory Division, Uppsala University, Box 516, SE-75120 Uppsala, Sweden*

[6]*School of Science and Technology, Örebro University, SE-701 82 Örebro, Sweden*

[7]*Deep Quantum Labs, Skolkovo Institute of Science and Technology, Moscow 121205, Russia*

(Received 27 September 2018; revised manuscript received 2 December 2018; published 10 January 2019)

Originating from image recognition, methods of machine learning allow for effective feature extraction and dimensionality reduction in multidimensional datasets, thereby providing an extraordinary tool to deal with classical and quantum models in many-body physics. In this study, we employ a specific unsupervised machine learning technique—self-organizing maps—to create a low-dimensional representation of microscopic states, relevant for macroscopic phase identification and detecting phase transitions. We explore the properties of spin Hamiltonians of two archetype model systems: a two-dimensional Heisenberg ferromagnet and a three-dimensional crystal, Fe in the body-centered-cubic structure. The method of self-organizing maps, which is known to conserve connectivity of the initial dataset, is compared to the cumulant method theory and is shown to be as accurate while being computationally more efficient in determining a phase transition temperature. We argue that the method proposed here can be applied to explore a broad class of second-order phase-transition systems, not only magnetic systems but also, for example, order-disorder transitions in alloys.

DOI: 10.1103/PhysRevB.99.041108

Introduction. Recently, machine learning has been suggested as a tool to investigate many-body quantum systems [1–31]. In essence, machine learning deals with systems with an extremely huge number of degrees of freedom in data space rather than in phase space of quantum statistics. It is therefore not surprising that even networks of simple architecture can be trained by means of supervised learning to detect very peculiar phases in a variety of systems [5,6], including topological and many-body localization phase transition [14]. The technique relies on sampling a physical system in a weighted way, and projecting the data onto hidden layers which filter out the irrelevant local fluctuations in the system, leaving only the large-scale behavior determining the macroscopic properties [2,24]. Further, it was demonstrated [7] that restricted Boltzmann machines can be used to formulate a very efficient many-body wave-function ansatz depending on a relatively small number of parameters even for a large number of spins ($\sim 10^2$), self-adjusting via gradient descent-based reinforced learning. This allowed for computing both ground states and dynamically evolved states of large many-body systems, with excellent accuracy. Furthermore, this algorithm has been generalized to bosonic and fermionic Hubbard models [10,22]. The application of machine learning to quantum-information problems in condensed-matter physics has also received significant interest recently, opening avenues for the direct experimental observation of the entanglement entropy [25].

In the meantime, there is a growing interest toward a versatile methodology that, on one hand, reduces the dimensionality of the data space, while preserving its topology on the other. In this Rapid Communication we propose a method for determining phase transitions which, in contrast to the previous studies handling Ising-like models, makes it possible to associate the symmetry breaking during a second-order phase transition with a noticeable change in the topology of a certain space. We construct this target space based on a number of microscopic states generated with Monte Carlo simulations for two archetypal examples, namely, a two-dimensional ferromagnet on a square lattice (2DFM) and bcc iron (bcc Fe). We further apply an unsupervised machine learning method in the form of *self-organizing, or Kohonen, maps* (SOMs) and compare the obtained results with those from cumulant method theory. We show that SOMs are able to correctly produce relevant two-dimensional representation of microscopic states, which allows one to visually observe symmetry breaking through a phase transition. The machine learning algorithms proposed here allow for a direct way for determining the critical temperature, while an intuitive interpretation of phase transitions in terms of principal component analysis (PCA) is also possible.

Model systems. For a vast class of magnetic compounds the microscopic description with high level of accuracy can be achieved within the Heisenberg exchange model. Whereas the Mermin-Wagner theorem establishes that an isotropic Heisenberg spin system in two dimensions cannot have a long-range ordering, the addition of anisotropy to the model changes the situation. The case of easy-plane exchange anisotropy is commonly referred to as the XXZ model, a system which akin to the XY model can display a Berezinskii-Kosterlitz-Thouless transition [32]. With easy-axis exchange, or single-ion, anisotropy, the system exhibits a second-order phase

transition [33]. In the following we consider 2DFM on a square lattice with single-site easy-axis anisotropy, as well as the Heisenberg model for Fe in its ground state crystal structure (bcc). For 2DFM we consider a square lattice of classical spins with edge length L, corresponding to a total number $N = L \times L$ of sites, and use periodic boundary conditions. The nearest-neighbor Heisenberg exchange is assumed to have a typical strength for transition-metal systems, $J = 1$ mRyd, and the parameter of easy-axis anisotropy is $K_{\text{anis}} = 0.2$ mRyd (see Supplemental Material, Sec. A [39]). For bcc Fe, one may neglect the tiny magnetocrystalline anisotropy, and the spin Hamiltonian is of pure Heisenberg type. However, the exchange coupling parameters are long ranged, and have here been obtained from first-principles electronic structure calculations. Including up to the fourth coordination shell, we use the same set of exchange couplings $J_1 = 1.3377$ mRyd, $J_2 = 0.7570$ mRyd, $J_3 = -0.0598$ mRyd, and $J_4 = -0.0882$ mRyd as calculated and used in Refs. [34,35] for a $L \times L \times L$ conventional bcc lattice with periodic boundary conditions as a simulation cell, corresponding to a number $N = 2L^3$ of spins (see Supplemental Material, Sec. A [39]).

In general, characterizing a phase transition requires a proper identification of the temperature point where the order parameter $M(T)$ goes to zero. For ferromagnetic Heisenberg-like models the order parameter is defined as the average magnetization per spin and in finite size systems the latter is not sharp enough at high-temperature regime where the role of fluctuations becomes important. To subdue this limitation and correct for finite size scaling one can apply the cumulant crossing method [36,37]. Direct application of this approach to second-order phase transitions suggests that for Ising-like models in the thermodynamic limit the Binder cumulant $U(T) \to 0$ for $T > T_c$, whereas $U(T) \to 2/3$ for $T < T_c$ as the lattice size increases [36]. For large enough systems $U(T)$ for different lattice sizes cross at a fixed point which can be identified with the critical temperature.

Dataset. To generate the appropriate spin configurations, we employed Monte Carlo simulations, with a heat bath algorithm for Heisenberg spin systems [38] as implemented in the UPPASD software [35]. We use 10^5 Monte Carlo steps for equilibration, and 10^6 Monte Carlo steps for the measurement phase. A sampling interval of ten steps was used for averages, susceptibility, total energy, and Binder cumulant measurements. Moreover, a sampling interval of 1000 (or 10 000) steps was used for snapshots of the whole spin configuration, resulting in up to 1000 (or 10 000) snapshots for each system and size. The results of these simulations are presented in the Supplemental Material, Sec. A [39]. We note here, however, that they show a phase transition between ferromagnetic and paramagnetic phases for both 2DFM and bcc Fe. The critical temperature for the 2DFM is estimated from Binder cumulants to be around $T_c \approx 222$ K and for bcc Fe $T_c \approx 915$ K. For the machine learning method, we use as our training set data from the cells with edge length $L = 80$, $L = 120$, and $L = 200$ (with the total number of spins being 6400, 14 400 and 40 000, respectively) for the 2DFM, and $L = 24$, $L = 28$, and $L = 36$ (with 27 648, 43 904, 93 312 spins) for bcc Fe.

Ideology of SOM. A SOM, first introduced by Kohonen [40,41], represents a neural network that performs visualization and clusterization by projecting a multidimensional space onto a lower dimensional one (most often, two-dimensional), and is trained using unsupervised learning. A SOM consists of components called nodes, or neurons, whose number is specified by the analyst. Each node is described by two vectors: the first one is the so-called weight vector, $\boldsymbol{w}$, of the same dimension as the input data, and the second vector, $\boldsymbol{r}$, is the one which gives the coordinates of the node on the map. The Kohonen map is visually displayed using an array of rectangular or hexagonal cells, associated with the respective node. During the training process, depicted schematically in Fig. 1, the weight vectors $\boldsymbol{w}(\boldsymbol{r})$ of the nodes approach the input data: for each observation (sample), the node with the closest weight vector is chosen, and its value moves toward the sample, together with the weight vectors of several neighbor nodes. The update formula for a weight vector $\boldsymbol{w}(\boldsymbol{r})$ is

$$\boldsymbol{w}_{n+1}(\boldsymbol{r}) = \boldsymbol{w}_n(\boldsymbol{r}) + \theta_n(\boldsymbol{r}', \boldsymbol{r}) \cdot \alpha_n \cdot [\boldsymbol{d}_m - \boldsymbol{w}_n(\boldsymbol{r})], \qquad (1)$$

where n is the step index, m stands for an index in the training set, $\boldsymbol{d}_m$ is the sample vector, $\boldsymbol{r}'$ denotes the coordinates of the node with the closest weight vector to the $\boldsymbol{d}_m$, and α_n is a monotonically decreasing learning coefficient. In Eq. (1), $\theta_n(\boldsymbol{r}', \boldsymbol{r})$ is the neighborhood function which depends on the grid distance between the neurons at $\boldsymbol{r}'$ and $\boldsymbol{r}$. In the simplest form it equals 1 for all neurons close enough to $\boldsymbol{r}'$ and 0 otherwise, though the Gaussian function could be an alternative option (regardless of the functional form, the neighborhood function shrinks as n increases). Thus, if two observations are close in the set of input data, they would correspond to nearby nodes on the map. The repeating training process, enumerating the input data, ends when the SOM reaches an acceptable error (predetermined by the analyst), or if a specified number of iterations is done. As a result, the SOM classifies the data into clusters and visually displays the multidimensional input onto a two-dimensional plane, relating the vectors of similar characteristics to neighboring cells (an illustration of training process is shown in Fig. 1).

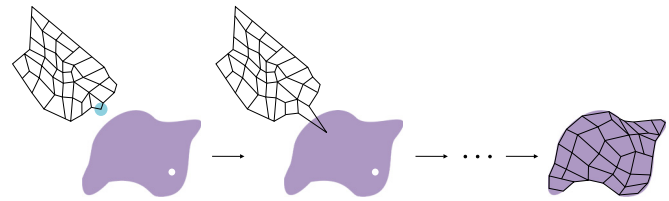

FIG. 1. Training of a self-organizing map (SOM). The distribution of the training data is depicted by a blob, and the small white dot is the current training data chosen from that distribution. At first (leftmost) the SOM nodes are arbitrarily located in the data space. The highlighted node which is nearest to the training data is selected. It is moved toward the training data, as are its neighbors on the grid (but to a lesser degree). After many iterations the grid tends to approximate the data set (rightmost).

Constructing a target space. To apply the methods described above, one has to correctly organize the input data into a target space (in our case, the spin states of the system obtained from Monte Carlo simulations). Importantly, we demand an internal geometry that reflects the regularities that are of interest. The results of Monte Carlo simulations

provide us with an array of K spin states for each selected temperature at which they have been extracted (a higher number of states were generated in the vicinity of the phase transition temperature), and these states are represented by $3N$-dimensional vectors, where N is the total number of lattice sites multiplied by three projections (in x, y, and z directions) of the spins. For convenience, we reshape this array of states into a rectangular $K \times 3N$ matrix,

$$S = \begin{pmatrix} s_x^{1,1} & s_y^{1,1} & s_z^{1,1} & \dots & s_x^{N,1} & s_y^{N,1} & s_z^{N,1} \\ \vdots & & & \ddots & & & \\ s_x^{1,K} & s_y^{1,K} & s_z^{1,K} & \dots & s_x^{N,K} & s_y^{N,K} & s_z^{N,K} \end{pmatrix}, \tag{2}$$

where $s_j^{i,k}$ is the ith-site spin projection on the jth axis in simulation k.

The first and most obvious way to form the target space is to take the rows of the matrix S as its elements, thus obtaining for each temperature a set of K $3N$-dimensional vectors (in other words, to use directly a set of spin states). However, the angular distribution of vectors from any of these sets is practically isotropic, since each spin state obtained using Monte Carlo simulations has a random direction of the average spin, and the only information that we can get by observing the modification of the geometry of such space, is a change of its diameter with temperature. A more practical geometry can be obtained by forming the target space out of the columns of this matrix. In the following, we show that if the system is in the magnetically ordered phase, a clustering of the vectors of the constructed target space takes place, whereas for the disordered phase this does not happen.

Clustering of the target space. Let us fix the temperature T, and consider Monte Carlo step k of the sampling phase. We then consider two columns, $\boldsymbol{y}(i, j)$ and $\boldsymbol{y}'(i', j')$, of the matrix S,

$$\boldsymbol{y}(i, j) = \begin{pmatrix} s_j^{i,1} \\ \vdots \\ s_j^{i,K} \end{pmatrix}, \quad \boldsymbol{y}'(i', j') = \begin{pmatrix} s_{j'}^{i',1} \\ \vdots \\ s_{j'}^{i',K} \end{pmatrix}. \tag{3}$$

To write down $\boldsymbol{y}$ and $\boldsymbol{y}'$ without their arguments, we also fixed the first site number i, its projection j, and do the same with i' and j'. In this way, we have two vectors $\boldsymbol{y}, \boldsymbol{y}' \in \mathbb{R}^K$, which are close, if the Euclidean distance between them is small enough relative to some characteristic value, which, in our case, should be the diameter of the target space. The physical meaning of this proximity is that the given projections of the corresponding lattice site states, described by these vectors, are close in each simulation.

If the vector characterizing the microstate of the system is calculated by averaging over the lattice, and its length increases during some process, then vectors, describing lattice site states become more codirectional, and vice versa, an increase in the proportion of relatively codirectional vectors over the lattice sites leads to an increase in the modulus of the microstate parameter. This obvious reasoning, together with the target space constructed above, forms the basis of the here proposed phase determination method: in the case of a high magnetization, for each simulation k, the projections of the lattice sites spins on the same axis ($j = j'$) are close for the majority of sites i, while such a proximity of different projections ($j \neq j'$) would mean that in a significant part of the observations, the average spin tends to some specific directions, contradicting the isotropic distribution of data, obtained by the Monte Carlo simulation, e.g., represented by the criterion

$$|y_k(i, j) - y'_k(i', j')|_{j=j'} < |y_k(\tilde{i}, \tilde{j}) - y'_k(\tilde{i}', \tilde{j}')|_{\tilde{j} \neq \tilde{j}'}, \tag{4}$$

for all k, j, j', $\tilde{j}$, and $\tilde{j}'$. Thus, we can expect that in the ferromagnetic phase, the vectors in the target space will be grouped into clusters, corresponding to the projections onto x, y, and z axes, while in the paramagnetic phase such clustering should be absent because of the much more isotropic distribution of the lattice site spins. Noteworthy, an intuitively clear PCA-based [42] graphical analysis of the magnetic phase transition unambiguously demonstrates three (for bcc Fe) or two (for 2DFM) well-separated clusters in the target space in the ferromagnetic phase, which are merging together in the disordered phase. This is illustrated in the Supplemental Material, Sec. B [39].

Clustering detection by SOMs. The advantage of using SOMs relies on the fact that they allow one to construct a two-dimensional projection of the multidimensional data distribution, while preserving the topology [43]. For this purpose, the target space vectors are normalized and centered, a SOM of a certain size is selected, a uniform distribution of the weight vectors of the nodes is set, and the SOM is trained according to Eq. (1). We observed best results for square maps of $\sim 15 \times 15$ nodes, providing thus the size for which cluster formation can be clearly distinguished (for more information about choosing the map size, see Supplemental Material, Sec. C [39]). All nodes of the trained map can be divided into two types: those that are not activated even once during training process (so-called *dead neurons*) and those that are activated at least once. In the case of the ferromagnetic phase, when Eq. (4) is fulfilled, groups of activated nodes, separated by a band of dead neurons, are clearly visible on the map [see Figs. 2(a) and 2(d)], which reflects the topology of the target space. The presence of dead neurons is due to the fact that falling close to the middle of the region between clusters, the neuron weight vector, in full agreement with Eq. (1), undergoes a multidirectional displacement during the learning process, caused by alternate attraction from the neighbors that have fallen into different clusters. Such oscillations compensate each other on average, as long as the distance between the clusters is significantly larger than their amplitude, but as the temperature increases, the clusters become closer and this condition breaks down making the position of the weight vector unstable [see Figs. 2(b) and 2(e)]. The probability of attraction to one of the clusters increases, which leads to a sharp decrease in the number of dead neurons and allows us to consider it as a characteristic parameter that specifies the phase of the system [see Figs. 2(c) and 2(f)], i.e., the critical temperature (for a more rigorous mathematical analysis, see Supplemental Material, Sec. D [39]). From Figs. 2(c) and 2(f) we conclude that the Curie temperature $T_c \approx 220$ K for the 2DFM and $T_c \approx 915$ K for bcc Fe, which is in excellent

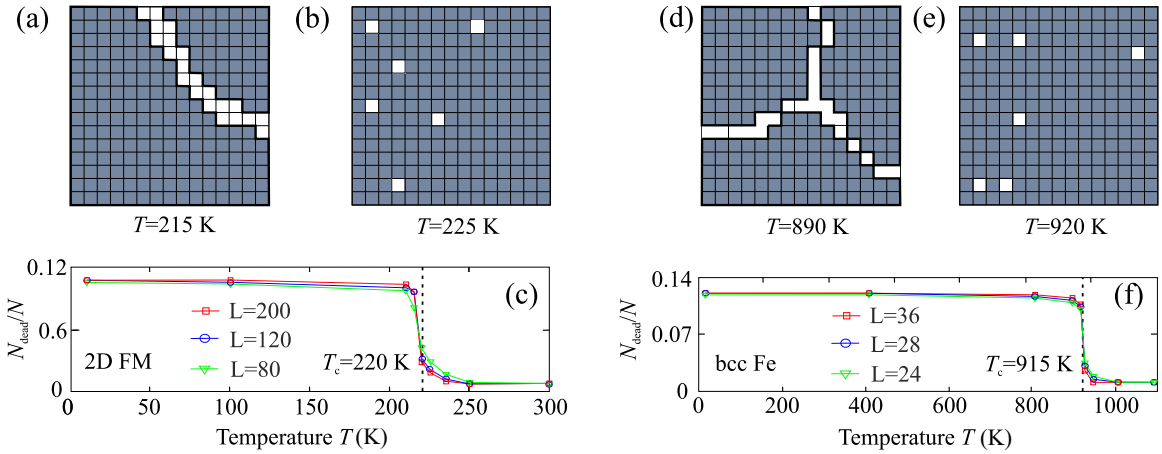

FIG. 2. Neural activity in the SOM close to the phase transition. The gray cells represent the activated neurons, whereas the white ones mark the dead neurons. Below the critical temperature: at $T = 215$ K for 2DFM, $L = 120$ (a) and at $T = 890$ K for bcc Fe, $L = 28$ (d) the well-separated clusters of activated neurons are clearly visible. Above the critical temperature: at $T = 225$ K for 2DFM, $L = 120$ (b) and at $T = 920$ K for bcc Fe, $L = 28$ (e) no clusters are present. The ratio of the dead neurons to the total number of neurons, N_{dead}/N, for different lattice sizes as a function of temperature is shown for 2DFM (c) and bcc Fe (f). The sharp drop of N_{dead}/N indicates the phase transition, revealing minor sensitivity to lattice size.

agreement with the results obtained from the Binder cumulant analysis [39] and also reproduces experimental values of bcc Fe. The sharp change of the SOM neural activity at the critical temperature makes applications of this method more precise as compared to PCA [39] and easier than cumulant method theory, because the simulation of the only one system of representative size is demanded; as one can see from Figs. 2(c) and 2(f), plots of neural activity are almost the same for biggest lattices, representing the fact that obtained results are independent of the model size, since it becomes big enough (a brief explanation is given in the Supplemental Material, Sec. B [39]). Whereas in the Binder cumulant technique, a set of different sized systems are needed.

Conclusions. In this Rapid Communication we proposed an approach for phase detection in systems with second-order transition, where the state is described by a large number of vectors. The method is based on constructing a special multidimensional target space with phase-related topology and an unsupervised learning algorithm of SOMs that is used to determine and visually observe a phase transition. We applied the method to characterize the phase transition and for calculating the critical temperature of a two-dimensional ferromagnet on a square lattice and bcc Fe. Our findings reveal an excellent agreement, being compared with results obtained with the conventional technique of the Binder cumulant theory. As opposed to the cumulant method theory that requires one to scale up the size of a system, the here suggested method allows one to make realistic predictions having Monte Carlo simulations for one copy of the system of a certain size only. The latter makes it possible to further utilize the method for various applications in statistical physics and condensed-matter systems, not only in magnetism but also, for example, for order-disorder transitions in alloy theory. A possible extension of the method proposed here is to provide a deeper understanding of short-range order around phase transitions, where experimental data exists, e.g., based on muon spin spectroscopy. We believe the proposed method can be generalized for the problems of purely quantum-mechanical nature as long as the elements of the corresponding system are described by certain vectors in multidimensional space, while the phase transition is associated with a change in their angular distribution.

Acknowledgments. A.A.S. and D.Y. acknowledge support from the Russian Science Foundation Project No. 17-12-01359. V.K.K. acknowledges support from Horizon2020 RISE project CoExAN. O.E. acknowledges support from the Swedish Research Council (VR), the Knut and Alice Wallenberg Foundation (KAW), the Swedish Foundation for Strategic Research (SSF), the Swedish Energy Agency, eSSENCE, and STandUPP.

[1] L.-F. Arsenault, A. Lopez-Bezanilla, O. A. von Lilienfeld, and A. J. Millis, Machine learning for many-body physics: The case of the Anderson impurity model, Phys. Rev. B **90**, 155136 (2014).

[2] P. Mehta and D. J. Schwab, An exact mapping between the variational renormalization group and deep learning, arXiv:1410.3831.

[3] G. Torlai and R. G. Melko, Learning thermodynamics with Boltzmann machines, Phys. Rev. B **94**, 165134 (2016).

[4] T. Ohtsuki and T. Ohtsuki, Deep learning the quantum phase transitions in random two-dimensional electron systems, J. Phys. Soc. Jpn. **85**, 123706 (2016).

[5] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, Nat. Phys. **13**, 431 (2017).

[6] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Learning phase transitions by confusion, Nat. Phys. **13**, 435 (2017).

[7] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, Science **355**, 602 (2017).

[8] L. Huang and L. Wang, Accelerated Monte Carlo simulations with restricted Boltzmann machines, Phys. Rev. B **95**, 035105 (2017).

[9] Y. Zhang and E.-A. Kim, Quantum Loop Topography for Machine Learning, Phys. Rev. Lett. **118**, 216401 (2017).

[10] H. Saito, Solving the Bose-Hubbard model with machine learning, J. Phys. Soc. Jpn. **86**, 093001 (2017).

[11] 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).

[12] A. Tanaka and A. Tomiya, Detection of phase transition via convolutional neural networks, J. Phys. Soc. Jpn. **86**, 063001 (2017).

[13] W. Hu, R. R. P. Singh, and R. T. Scalettar, Discovering phases, phase transitions, and crossovers through unsupervised machine learning: A critical examination, Phys. Rev. E **95**, 062122 (2017).

[14] F. Schindler, N. Regnault, and T. Neupert, Probing many-body localization with neural networks, Phys. Rev. B **95**, 245134 (2017).

[15] J. Liu, H. Shen, Y. Qi, Z. Y. Meng, and L. Fu, Self-learning Monte Carlo method and cumulative update in fermion systems, Phys. Rev. B **95**, 241104 (2017).

[16] S. J. Wetzel, Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders, Phys. Rev. E **96**, 022140 (2017).

[17] G. Torlai and R. G. Melko, Neural Decoder for Topological Codes, Phys. Rev. Lett. **119**, 030501 (2017).

[18] 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).

[19] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Machine Learning Phases of Strongly Correlated Fermions, Phys. Rev. X **7**, 031038 (2017).

[20] P. Broecker, F. F. Assaad, and S. Trebst, Quantum phase recognition via unsupervised machine learning, arXiv:1707.00663.

[21] N. C. Costa, W. Hu, Z. J. Bai, R. T. Scalettar, and R. R. P. Singh, Principal component analysis for fermionic critical points, Phys. Rev. B **96**, 195138 (2017).

[22] Y. Nomura, A. S. Darmawan, Y. Yamaji, and M. Imada, Restricted Boltzmann machine learning for solving strongly correlated quantum systems, Phys. Rev. B **96**, 205152 (2017).

[23] K. Ch'ng, N. Vazquez, and E. Khatami, Unsupervised machine learning account of magnetic transitions in the Hubbard model, Phys. Rev. E **97**, 013306 (2018).

[24] M. Koch-Janusz and Z. Ringel, Mutual information, neural networks and the renormalization group, Nat. Phys. **14**, 578 (2018).

[25] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Neural-network quantum state tomography, Nat. Phys. **14**, 447 (2017).

[26] P. Zhang, H. Shen, and H. Zhai, Machine Learning Topological Invariants with Neural Networks, Phys. Rev. Lett. **120**, 066401 (2018).

[27] H. Fujita, Y. O. Nakagawa, S. Sugiura, and M. Oshikawa, Construction of Hamiltonians by supervised learning of energy and entanglement spectra, Phys. Rev. B **97**, 075114 (2018).

[28] W.-J. Rao, Z. Li, Q. Zhu, M. Luo, and X. Wan, Identifying product order with restricted Boltzmann machines, Phys. Rev. B **97**, 094207 (2018).

[29] K. Mills and I. Tamblyn, Deep neural networks for direct, featureless learning through observation: The case of two-dimensional spin models, Phys. Rev. E **97**, 032119 (2018).

[30] H. Saito and M. Kato, Machine learning technique to find quantum many-body ground states of bosons on a lattice, J. Phys. Soc. Jpn. **87**, 014001 (2018).

[31] N. Yoshioka, Y. Akagi, and H. Katsura, Learning disordered topological phases by statistical recovery of symmetry, Phys. Rev. B **97**, 205110 (2018).

[32] A. Cuccoli, V. Tognetti, and R. Vaia, Two-dimensional XXZ model on a square lattice: A Monte Carlo simulation, Phys. Rev. B **52**, 10221 (1995).

[33] S. A. Leonel, A. C. Oliveira, B. V. Costa, and P. Z. Coura, Comparative study between a two-dimensional anisotropic Heisenberg antiferromagnet with easy-axis single-ion anisotropy and one with easy-axis exchange anisotropy, J. Magn. Magn. Mater. **305**, 157 (2006).

[34] X. Tao, D. P. Landau, T. C. Schulthess, and G. M. Stocks, Spin Waves in Paramagnetic bcc Iron: Spin Dynamics Simulations, Phys. Rev. Lett. **95**, 087207 (2005).

[35] B. Skubic, J. Hellsvik, L. Nordström, and O. Eriksson, A method for atomistic spin dynamics simulations: Implementation and examples, J. Phys.: Condens. Matter **20**, 315203 (2008).

[36] K. Binder, Finite size scaling analysis of Ising model block distribution functions, Z. Phys. B **43**, 119 (1981).

[37] K. Binder, Critical Properties from Monte Carlo Coarse Graining and Renormalization, Phys. Rev. Lett. **47**, 693 (1981).

[38] J. A. Olive, A. P. Young, and D. Sherrington, Computer simulation of the three-dimensional short-range Heisenberg spin glass, Phys. Rev. B **34**, 6341 (1986).

[39] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevB.99.041108 for technical details of MC simulations, PCA, choosing the maps size, and mathematical grounds of SOM phase detection.

[40] T. Kohonen, Self-organized formation of topologically correct feature maps, Biol. Cybern. **43**, 59 (1982).

[41] T. Kohonen, Analysis of a simple self-organizing process, Biol. Cybern. **44**, 135 (1982).

[42] I. T. Jolliffe, *Principal Component Analysis* (Springer-Verlag, New York, 2002).

[43] T. Graepel, M. Burger, and K. Obermayer, Phase transitions in stochastic self-organizing maps, Phys. Rev. E **56**, 3876 (1997).