

# Identifying Product Order with Restricted Boltzmann Machines

Wen-Jia Rao,<sup>1,2</sup> Zhenyu Li,<sup>1</sup> Qiong Zhu,<sup>3</sup> Mingxing Luo,<sup>1</sup> and Xin Wan<sup>1,2</sup>

<sup>1</sup>*Zhejiang Institute of Modern Physics, Zhejiang University, Hangzhou 310027, China*

<sup>2</sup>*Collaborative Innovation Center of Advanced Microstructures, Nanjing 210093, China and*

<sup>3</sup>*International Center for Quantum Materials, Peking University, Beijing 100871, China*

(Dated: October 17, 2019)

Unsupervised machine learning via a restricted Boltzmann machine is an useful tool in distinguishing an ordered phase from a disordered phase. Here we study its application on the two-dimensional Ashkin-Teller model, which features a partially ordered product phase. We train the neural network with spin configuration data generated by Monte Carlo simulations and show that distinct features of the product phase can be learned from non-ergodic samples resulting from symmetry breaking. Careful analysis of the weight matrices inspires us to define a nontrivial machine-learning motivated quantity of the product form, which resembles the conventional product order parameter.

PACS numbers:

## I. INTRODUCTION

One of the central tasks of condensed matter physics is to identified phases and phase transitions. The conventional approach introduces the concept of order parameter, which is a quantity that vanishes in a disordered phase but nonzero in the adjacent ordered phase. The value of the order parameter can be used to identify the transition between the two phases, and plays a central role in the Landau theory of phase transitions. The fluctuations of the order parameter, which is closely related to its dimension and symmetry, are crucial in understanding the corresponding phase transition. In practice, the choice of the order parameter is not unique, but may not be obvious sometimes, such as in metal-insulator transitions.

Recent developments in machine learning (ML)<sup>1</sup> have found growing applications in the study of phases and phase transitions.<sup>2–13</sup> In these studies computer algorithms identify patterns in the configurations of physical systems just as they recognize images in the field of artificial intelligence. Utilizing the knowledge learnt from data analysis, one can also use machine learning schemes to improve existing numerical algorithms.<sup>14–20</sup> Despite all these successful applications, the power and limitation of ML remains to be understood. It is tempting to seek connections to fundamental concepts in physics, such as symmetry, locality, and renormalization group.<sup>21–23</sup>

Among various machine learning schemes, the restricted Boltzmann machine (RBM) bears the closest analogy to physical systems. The joint probability distribution of the model is a Boltzmann distribution whose energy functional describes the couplings between a visible layer and a hidden layer of spins. RBMs can be used as generative models in machine vision or language processing to extract high-level features. Obvious applications in physics include representing probability distributions, such as the Boltzmann distribution in calculating partition functions, the probability density of wave functions, or complex wave functions themselves. For example, Torlai and Melko<sup>3</sup> applied RBMs to study the thermodynamics of classical Ising models. Amin *et al.*<sup>33</sup> further generalized the RBM approach to study quantum models. Carleo

and Troyer<sup>26</sup> demonstrated that using RBM representations as variational wave functions one can approach even lower ground state energies than with methods based on tensor networks. In models of stabilizer codes, RBMs can be shown to represent exact ground states.<sup>28,29</sup> The connection between the representative power of RBM and tensor network states has been explored.<sup>30–32</sup> Morningstar and Melko<sup>34</sup> found that the shallow RBM is more efficient than its deep generalizations in representing physical probability distributions, at least for Ising systems near criticality.

In this paper we apply an unsupervised learning with RBMs to the  $N_c$ -color Ashkin-Teller (AT) model on a square lattice. One motivation is that the AT model features a fully disordered paramagnetic phase and a partially ordered product phase. The two bear strong similarities in spin configurations of any single color and both possess a large entropy. It is, therefore, an interesting question whether the RBM can distinguish the two phases. In addition, the conventional order parameter of the product phase is constructed by the product of two spins of different color on the same lattice site. The operator, however, is not directly present in the energy functional of the RBM. One wonders how ML can capture the product order with the nontrivial order parameter. To answer these questions, we organize our paper as follows. In Sec. II we describe the physical AT model and the RBM neural network model. We explain how we train the RBM with spin configurations from Monte Carlo simulations. The results of the training are presented in Sec. III. We conclude with discussions of the results and on possible further directions in Sec. IV.

## II. MODEL AND METHOD

We consider the homogeneous  $N_c$ -color AT model on a two-dimensional (2D) square lattice. In the AT model each lattice site hosts  $N_c$  colors (or species) of Ising spins, which are coupled through the Hamiltonian

$$H_{\text{AT}} = -K_2 \sum_{\langle i,j \rangle} \sum_{\alpha=1}^{N_c} \sigma_i^\alpha \sigma_j^\alpha - K_4 \sum_{\langle i,j \rangle} \sum_{\alpha < \beta} \sigma_i^\alpha \sigma_i^\beta \sigma_j^\alpha \sigma_j^\beta. \quad (1)$$

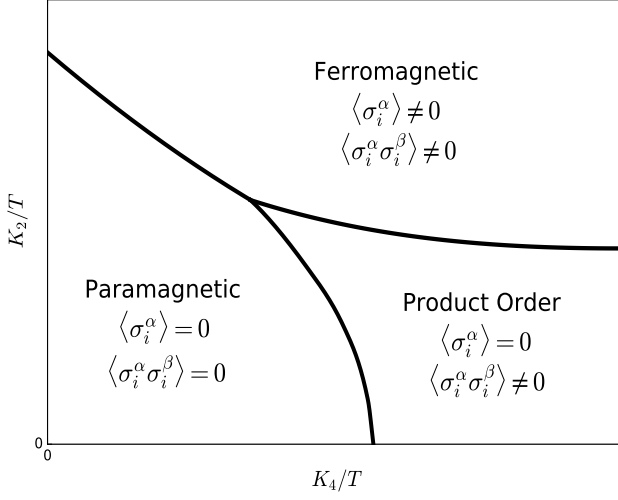


FIG. 1: Illustration of the phase diagram of the  $N_c$ -color AT model. The intercept along the  $K_2$  axis is  $K_2/T = \ln(1+\sqrt{2})/2 \approx 0.4407$ , which is the 2D square-lattice Ising critical point and independent of the number of colors  $N_c$ . The intercept along the  $K_4$  axis is  $N_c$ -dependent. For the two-color AT model the tricritical point is self-dual at  $(K_t/T, K_t/T)$ , where  $K_t/T = (\ln 3)/4 \approx 0.2746$ .

where  $i$  and  $j$  are lattice site indices while  $\alpha$  and  $\beta$  are color indices. We restrict ourselves to the parameter space with  $K_2 > 0$  and  $K_4 > 0$ . The first term describes  $N_c$  independent copies of the 2D nearest-neighbor Ising models, while the second term couples different species with nearest-neighbor four-spin interaction. The 2D square-lattice Ising model has a continuous phase transition from ordered (ferromagnetic) to disordered (paramagnetic) phases at  $K_2/T = \ln(1+\sqrt{2})/2 \approx 0.4407$ , which can be characterized by a local order parameter  $\langle \sigma_i \rangle$ . For large  $K_2$  energy dominates and the Ising system is in the ferromagnetic phase, while for small  $K_2$  entropy dominates and the system is disordered. In the presence of the four-spin interaction, a new phase emerges when entropy competes favorably to the two-spin interaction energy, but not to the four-spin interaction energy. Accordingly, the phase can be characterized by a non-zero product order parameter  $\langle O_i^{\alpha\beta} \rangle \equiv \langle \sigma_i^\alpha \sigma_i^\beta \rangle \neq 0$ , even though there is no order among individual species, i.e.  $\langle \sigma_i^\alpha \rangle = 0$ . This partially ordered phase is known as the product phase. The partially ordered product phase still possesses a large spin-orientation entropy  $[O(N \ln 2)]$ , which is of the same order as that of the disordered phase. Based on the spin configuration of any single color, one cannot distinguish the product phase from the paramagnetic phase. The generic phase diagram for the homogeneous  $N_c$ -color AT model is illustrated in Fig. 1. In particular, the AT model with  $N_c = 2$  can be solved exactly at the self-dual tricritical point  $(K_2, K_4) = (K_t, K_t)$ , where  $K_t/T = (\ln 3)/4 \approx 0.2746$ .<sup>35</sup> In the neural network context, the AT spin glass has been considered as a generalization of the Hopfield model for the storage and retrieval of embedded patterns.<sup>36</sup>

The RBM for the AT model is illustrated in Fig. 2.  $N$  lattice

sites, each with  $N_c$  physical Ising spins (or nodes)  $\sigma_i^\alpha$ , form a visible layer, while an additional  $M$  Ising spins (or nodes)  $h_j$  form a hidden layer. Local fields  $a_i^\alpha$  and  $b_j$  are applied to the visible and hidden nodes, respectively. In a RBM couplings  $w_{ij}^\alpha$  exist as edges only between two nodes in different layers, hence the modifier restricted. We emphasize that we color-code the  $N_c$  visible spin species, as well as the corresponding edges, in Fig. 2 by different colors (blue for  $\alpha = 1$ , red for  $\alpha = 2$ , green for  $\alpha = 3$ , etc.) and will present our results with such a color scheme. Mathematically, this graph describes a joint probability distribution

$$p_\lambda(\sigma, \mathbf{h}) = \frac{e^{-E_\lambda(\sigma, \mathbf{h})}}{Z_\lambda}, \quad (2)$$

where

$$E_\lambda(\sigma, \mathbf{h}) = - \sum_{i, \alpha} a_i^\alpha \sigma_i^\alpha - \sum_j b_j h_j - \sum_{i, \alpha, j} \sigma_i^\alpha w_{ij}^\alpha h_j. \quad (3)$$

Here, the subscript  $\lambda$  stands for the collection of RBM parameters  $\{a, b, w\}$ , and  $Z_\lambda$  is the normalizing partition function. The probability distribution for the visible nodes is then

$$p_\lambda(\sigma) = \sum_{\mathbf{h}} p_\lambda(\sigma, \mathbf{h}) \equiv \frac{1}{Z_\lambda} \exp(-E_\lambda(\sigma)) \quad (4)$$

where the model energy functional

$$E_\lambda(\sigma) = - \sum_{i, \alpha} a_i^\alpha \sigma_i^\alpha - \sum_j \log(1 + e^{b'_j}) \quad (5)$$

and the effective field for the  $j$ th hidden spin

$$b'_j = b_j + \sum_{i, \alpha} \sigma_i^\alpha w_{ij}^\alpha. \quad (6)$$

It has been proved<sup>40,41</sup> that such a distribution can describe an arbitrary probability distribution provided that the number of hidden nodes is sufficiently large. In the present study, we feed spin configurations of the AT model to the RBM with the objective to train the RBM distribution  $p_\lambda(\sigma)$  to represent the Boltzmann distribution of the classical Hamiltonian as closely as possible; in other words, we approximate the original Hamiltonian with the effective RBM energy functional  $E_\lambda(\sigma)$  in Eq. 5. Previously, the RBM approach has been applied to the classical Ising model in both one and two dimensions.<sup>3</sup> For the AT model, a visible spins contain an additional color index and, therefore, the coupling between the visible spin and any hidden node also contains the corresponding color index. As we will show later, the correlation between the couplings of a hidden node to a pair of visible spins with the same site index but different color indices is crucial for the understanding of the partially ordered product phase.

To obtain an optimized set of parameters  $\lambda \equiv (a, b, w)$ , we define the standard KL divergence

$$KL(\lambda) = \sum_{\sigma} p_B(\sigma) \log \frac{p_B(\sigma)}{p_\lambda(\sigma)} \quad (7)$$

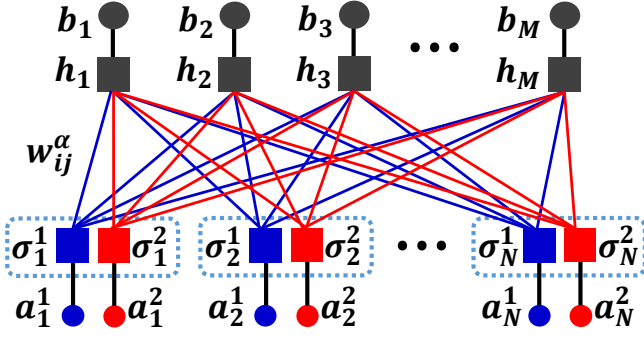


FIG. 2: Illustration of the RBM for the two-color AT model. The dotted boxes represent  $N$  lattice sites, on each of which there are two Ising spins  $\sigma_i^\alpha$  with different color  $\alpha$  (blue and red squares for  $\alpha = 1, 2$ ). They comprise the visible layer. The hidden layer consists of  $M$  hidden Ising spins  $h_j$  (dark grey boxes). There are no intralayer couplings, but spins in the two layers are coupled via edges  $w_{ij}^\alpha$ . External field  $a_i^\alpha$  (blue and red dots) and  $b_j$  (dark grey dots) are introduced for the visible and hidden layers, respectively. This graph structure can be easily generalized to the  $N_c$ -color AT model with  $N_c > 2$ .

as a cost function to judge how close  $p_\lambda(\sigma)$  is to the Boltzmann distribution function

$$p_B(\sigma, T) = \frac{1}{Z_{\text{AT}}} e^{-H_{\text{AT}}(\sigma)/T}, \quad (8)$$

where the partition function

$$Z_{\text{AT}} = \text{Tr} \left[ e^{-H_{\text{AT}}(\sigma)/T} \right]. \quad (9)$$

It can be shown that  $KL(\lambda) \geq 0$ ; the equality holds only when  $p_B$  coincides with  $p_\lambda$ . Hence, *training* is an optimization procedure that updates parameters  $\lambda$  by reducing the KL divergence.

We follow the standard machine learning procedure to train the RBM machine with the spin configurations of the AT model generated by Monte Carlo simulation. If we denote the data set by  $D \equiv \{\sigma^{(1)}, \sigma^{(2)}, \dots, \sigma^{(D)}\}$ , we thus replace the Boltzmann distribution by the probability distribution  $p_{\text{data}}(\sigma) = \frac{1}{D} \sum_{\sigma' \in D} \delta(\sigma, \sigma')$ . The KL divergence is then simply

$$KL(\lambda) = -\frac{1}{D} \sum_{\sigma' \in D} \log p_\lambda(\sigma') - H(p_{\text{data}}) \quad (10)$$

where

$$H(p_{\text{data}}) = -\frac{1}{D} \sum_{\sigma' \in D} \log \left( \frac{1}{D} \sum_{\sigma'' \in D} \delta(\sigma', \sigma'') \right) \quad (11)$$

The separation is convenient, because only the first term in the KL divergence depends on  $\lambda$  and needs to be updated during training, while the second term  $H(p_{\text{data}})$ , or the entropy of the data set, needs to be computed only once.

For the optimization of the KL divergence in the presence of a large data set, we adopt the stochastic gradient descent

(SGD) approach. The approach involves repeated calculations of the gradient of the log-likelihood  $\log p_\lambda(\sigma)$  with respect to any of the parameters  $\lambda_k \equiv a_i^\alpha, b_j$ , or  $w_{ij}^\alpha$ ,

$$\nabla_{\lambda_k} \log p_\lambda(\sigma) = -\nabla_{\lambda_k} E_\lambda(\sigma) + \sum_{\sigma'} \frac{e^{-E_\lambda(\sigma')}}{Z_\lambda} \nabla_{\lambda_k} E_\lambda(\sigma'). \quad (12)$$

Note that the evaluation of the normalization  $Z_\lambda$  in the second term involves a summation over exponentially large number of configurations, hence impossible for practical calculations of the log-likelihood gradient. Instead, we adopt the  $k$ -step contrastive divergence ( $\text{CD}_k$ ),<sup>38</sup> which approximates the gradient locally around the training data.

With these methods and approximations, we start from a randomly chosen initial parameters  $\lambda^{(0)}$  and update through

$$\lambda^{(n+1)} = \lambda^{(n)} - \zeta \nabla_{\lambda} KL(\lambda^{(n)}) \quad (13)$$

where the coefficient  $\zeta$  is the learning rate, whose value needs to be carefully chosen to balance the speed to explore the parameter space and the stability. After sufficiently long steps, we terminate the training and explore the resulting parameter set  $\lambda$  and try to associate certain patterns to various phases, in particular, to the product phase.

### III. MACHINE-LEARNING RESULTS

In the following RBM study we use 200 hidden nodes and initiate  $\lambda^{(0)}$  with uniform distribution within  $[-0.03, 0.03]$ . For the training, we choose the  $\text{CD}_{20}$  approximation and a fixed learning rate  $\zeta = 0.03$ . We have checked that the qualitative results we present below are robust against the variation of these super parameters within reasonable range.

#### A. Similarities between the product and paramagnetic phases

We begin our discussion on the identification of the product phase of the two-color AT model in an  $8 \times 8$  lattice. We first perform Monte Carlo sampling to generate a data set of  $10^5$  configurations for the following two parameter sets: (i)  $K_2/T = 0.1$  and  $K_4/T = 0.1$  in the paramagnetic phase, and (ii)  $K_2/T = 0.1$  and  $K_4/T = 1$  in the product phase. We choose to compare the two cases both with  $\langle \sigma_i^\alpha \rangle = 0$ . The parameters are chosen that the states are sufficiently far away from the phase boundary, but we have ensured that the ergodicity in (ii) is not broken in our simulation. We then feed the data into the RBM to learn the Boltzmann distribution of the AT model.

In Figure 3(a) and (b) we compare the histograms of the interlayer coupling coefficients  $w_{ij}^\alpha$  for the two parameter sets in the paramagnetic and product phases. We break the histograms of the couplings into two colors according to their connections to the visible spins with the corresponding colors (blue for  $\alpha = 1$  and red for  $\alpha = 2$ ). The histograms show a similar bell-shape in both paramagnetic and product

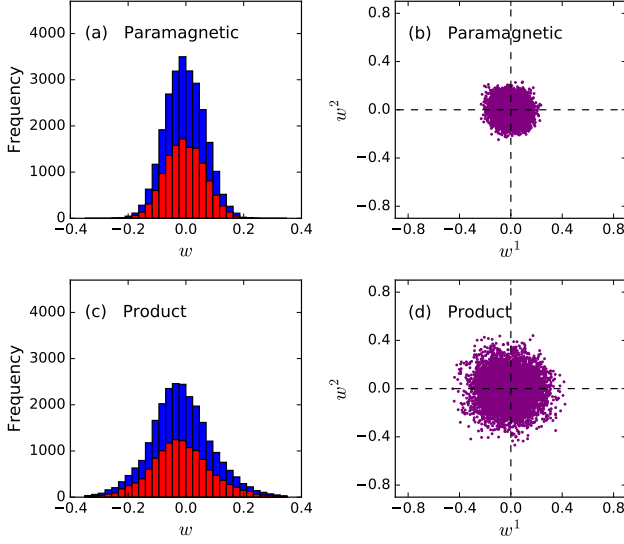


FIG. 3: Histogram of the coupling coefficients  $w_{ij}^\alpha$  of the two-color AT model on a  $8 \times 8$  square lattice for (a)  $K_2/T = 0.1$  and  $K_4/T = 0.1$  in the paramagnetic phase, and (c)  $K_2/T = 0.1$  and  $K_4/T = 1$  in the product phase. They are color coded such that blue stands for color index  $\alpha = 1$  and red for  $\alpha = 2$ . Comparison of the corresponding coupling coefficients  $w_{ij}^1$  versus  $w_{ij}^2$  for the two colors are shown in (b) for the paramagnetic phase and in (d) for the product phase.

phases. After fitting the histograms by a Gaussian distribution function  $p(w) = p_0 e^{-(w-\mu)^2/2\zeta^2}$ , we find  $\mu = -0.0049$  and  $\zeta = 0.069$  for the paramagnetic phase and  $\mu = -0.021$  and  $\zeta = 0.096$  for the product phase. In both cases the deviation of the peak from the origin is insignificant. The width of the two bells are similar; in the product phase the width is about 39% larger. In the earlier RBM learning of the 2D Ising model, Torlai and Melko<sup>3</sup> observed a sharp contrast in the histogram of the couplings between the ferromagnetic and paramagnetic phases. This is understandable because strong interlayer couplings are necessary to generate the long-range spin correlation in the ferromagnetic phase, while weak interlayer couplings imply that the visible spins behave like independent entities. We generated spin configurations in the ferromagnetic phase of the AT model by the Wolff algorithm. Due to the presence of the color index, more hidden nodes are needed in the training than those in the Ising case for the same lattice size. Qualitatively, we confirm that the width of the coupling strength distribution can be regarded as an order parameter to distinguish the ferromagnetic phase and the paramagnetic phase. Unfortunately, as we showed above, this would be an unreliable indicator for identifying the product phase.

The existence of the product order also prompts us to explore the correlation between the couplings connecting the same hidden spin and the two visible spins with different colors on the same lattice site. We plot the corresponding  $w_{ij}^1$  versus  $w_{ij}^2$  in Fig. 3(b) for the paramagnetic phase and in Fig. 3(d) for the product phase. In both cases, the data scat-

ters around the origin with no significant enhancement in any of the four quadrants. Once again, we cannot distinguish the paramagnetic and the product phases according to the RBM parameters. Is it, then, possible to define a machine-learning motivated order parameter based on the RBM parameters for the product phase?

## B. Non-ergodicity and the product order

Let us digress a moment to the ferromagnetic case. The conventional order parameter is the magnetization, whose emergence is related to the breaking of the  $Z_2$  symmetry of Ising spins. Below the critical temperature, magnetization can take on two signs. In the thermodynamic limit, ergodicity is broken and the magnetization is either positive or negative. In a finite system, however, ergodicity may not be broken in a Monte Carlo simulation, especially with the implementation of various cluster algorithms. Whether the ergodicity is broken or not will not affect the study of phases and phase transitions as long as we take proper care of the sign of the order parameter. How important, then, is the ergodicity broken in the RBM study? With these considerations in mind, we now return to the product phase to show that ergodicity breaking indeed holds a crucial role in the identification of an appropriate order parameter.

In the following we train the RBM with ergodicity-breaking data in a  $4 \times 4$  lattice for  $K_2/T = 0.1$  and  $K_4/T = 2$ . We note that deeper in the product phase spin configurations generated by Metropolis sampling are ergodicity broken, unless the simulations time is sufficiently long. Figure 4 shows the local fields  $a_i^\alpha$  and the histogram of the coupling coefficients  $w_{ij}^\alpha$  in the product phase. The local fields for color-1 (blue) spins are predominantly negative (13 out of 16), while those for color-2 (red) are predominantly positive (also 13 out of 16). Furthermore, they are so correlated that  $a_i^1 a_i^2 < 0$  for every lattice site  $i$ . Intriguingly, the histogram of  $w_{ij}^\alpha$  develops two broad but separate lumps on the two sides of the origin. We intentionally plot the histogram of  $w_{ij}^\alpha$  for individual colors. Obviously, the two lumps do not correspond to the two colors, even though the majority of  $w_{ij}^\alpha$  connecting blue spins clusters with negative values, while the majority of those connecting red spins clusters with positive values. The broken color symmetry among  $a_i^\alpha$  or among  $w_{ij}^\alpha$  motivates us to further explore the product of these parameters. Figure 5 shows the histogram of  $w_{ij}^1 w_{ij}^2$  and  $a_i^\alpha w_{ij}^\alpha$ . We find that *almost all* (above 99.9%)  $w_{ij}^1 w_{ij}^2$  are negative and 96.8% of  $a_i^\alpha w_{ij}^\alpha$  are positive, which means that the signs of  $a_i^\alpha$  and  $w_{ij}^\alpha$  associating with visible spins on the same lattice site  $i$  are well correlated. This again is not hard to understand if one explores the effective energy in Eq. 5. In the product phase with  $\langle \sigma_i^1 \sigma_i^2 \rangle < 0$ , negative  $w_{ij}^1 w_{ij}^2$  can lead to negatively large  $\sum_{i,\alpha} \sigma_i^\alpha w_{ij}^\alpha$ , hence the sum in the second term on the righthand side of Eq. 5 is positively large. Positive  $a_i^\alpha w_{ij}^\alpha$  also imply that  $a_i^\alpha \sigma_i^\alpha$  in the first term are negative. As a result, the effective energy is large but negative. In other words, configurations with  $\sigma_i^1 \sigma_i^2 < 0$  have larger weight than those with  $\sigma_i^1 \sigma_i^2 > 0$ , hence the product order develops. We would like to point out that an independent sampling and

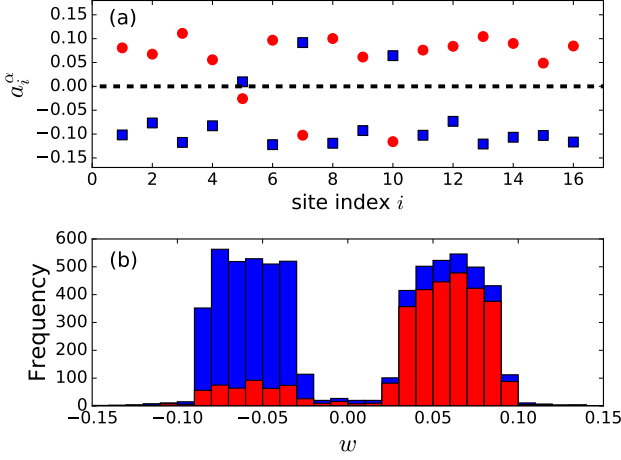


FIG. 4: RBM parameters obtained from the training of the two-color AT model on a  $4 \times 4$  square lattice with  $K_2/T = 0.1$  and  $K_4/T = 2$  in the product phase. (a) The local fields  $a_i^1$  and  $a_i^2$  for the two-color Ising spins on any site  $i$  are of opposite signs. (b) The histogram of the coupling coefficients  $w_{ij}^\alpha$  exhibits two lumps, one on each side of the origin. The figure is color coded such that blue stands for color index  $\alpha = 1$  and red for  $\alpha = 2$ .

the consequent training may as well result in the dominance of positive  $w_{ij}^1 w_{ij}^2$ , which corresponds to the product phase with  $\langle \sigma_i^1 \sigma_i^2 \rangle > 0$ .

Interestingly, even though more than 99.9% of the product  $w_{ij}^1 w_{ij}^2$  are negative, only 15.60% of  $w_{ij}^1$  are positive. Neither is the latter percentage close to 0 or 1 as the translational invariance may suggest, nor is it approaching 0.5 in a random fashion. In fact, the number is, as it should, close to the percentage of  $a_i^1$  being negative (3 out of 16). It turns out this number is training-dependent. We point out that in the product order phase there is an emergent local symmetry:  $\sigma_i^1 \leftrightarrow \sigma_i^2$ , which leaves the order parameter invariant. The RBM can accommodate this symmetry by the joint transformation:  $a_i^1 \leftrightarrow a_i^2$  and  $w_{ij}^1 \leftrightarrow w_{ij}^2$  for all  $j$ . As a result, the polarization of  $a_i^\alpha$  or  $w_{ij}^\alpha$  cannot serve as an indicator of the product phase. Instead, the polarization of the product  $w_{ij}^1 w_{ij}^2$  or  $a_i^\alpha w_{ij}^\alpha$  can. We can use, e.g., the product  $w_{ij}^1 w_{ij}^2$  to design a machine-learning motivated order parameter

$$\Gamma \equiv \frac{1}{NM} \sum_{i=1}^N \sum_{j=1}^M \text{sgn}(w_{ij}^1 w_{ij}^2), \quad (14)$$

where  $\text{sgn}(x)$  is the sign function. In the product phase we expect  $|\Gamma| \approx 1$ , while in the disordered paramagnetic phase  $\Gamma = O(1/\sqrt{N})$  due to fluctuations.

### C. Breaking the product symmetry in ergodic samples

To further demonstrate the feasibility of  $\Gamma$  as an order parameter for the product order, we consider the following setup. We start from the  $8 \times 8$  lattice with  $K_2/T = 0.1$  and

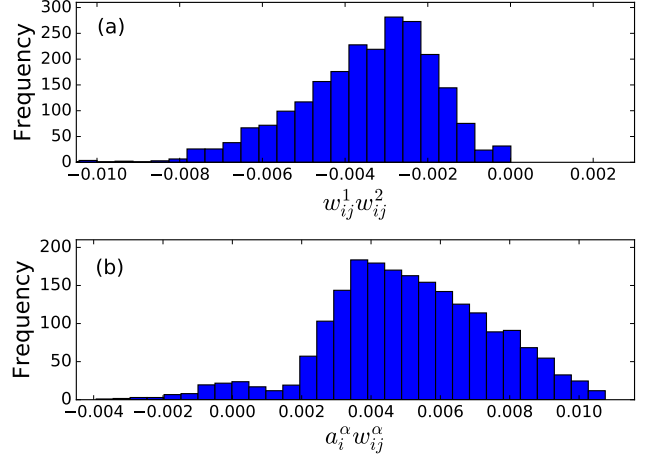


FIG. 5: Histograms of the products  $w_{ij}^1 w_{ij}^2$  and  $a_i^\alpha w_{ij}^\alpha$  resulting from the training in the product phase, as illustrated in Fig. 4. (a) More than 99.9% of  $w_{ij}^1 w_{ij}^2$  are negative. (b) 96.8% of  $a_i^\alpha w_{ij}^\alpha$  are positive.

$K_4/T = 1$  explored in Sec. III A. With preserved ergodicity in data from the simulation, we showed there that the RBM parameters do not show significant difference from those for a paramagnetic state. In particular, there is no nontrivial color pattern in the parameter, so one expects  $\Gamma = 0$ . To be consistent, we also have a vanishing conventional product order parameter

$$\langle \sigma^1 \sigma^2 \rangle \equiv \frac{1}{N_s} \sum_{j=1}^{N_s} \left( \frac{1}{N} \sum_{i=1}^N \sigma_i^1 \sigma_i^2 \right)_{\text{jth configuration}}, \quad (15)$$

where  $N_s$  is the number of spin configurations fed into the RBM training. We now introduce an ergodicity breaking term  $h_4 \sum_i \sigma_i^1 \sigma_i^2$  to the original AT Hamiltonian Eq. (1). Depending on the sign of  $h_4$ ,  $\langle \sigma^1 \sigma^2 \rangle$  polarizes accordingly. For  $-0.2 \leq h_4 \leq 0.2$ , we perform the RBM training to the spin configurations obtained from sufficiently long simulations, such that ergodicity is not broken.

Figure 6(a) shows the training result for  $h_4 = 0.2$ , at which the spin configurations are totally polarized to  $\langle \sigma_i^1 \sigma_i^2 \rangle = -1$  while spins of individual colors remain disordered. Interestingly, however, the histogram of  $w_{ij}^\alpha$  in Fig. 6(a) has no qualitative difference from the case in the absence of  $h_4$  [Fig. 3(c)]. However, we find that almost all products  $w_{ij}^1 w_{ij}^2$  are negative. More impressively, as shown in Fig. 6(b), we find strong correlation between  $w_{ij}^1$  and  $w_{ij}^2$ , in sharp contrast to Fig. 3(d). The symmetry breaking field  $h_4 = 0.2$  completely orders the couplings of the spins on the same lattice site to any hidden spin, even though there is no apparent order in the coupling coefficients for any individual color. This provides further evidence to the validity of Eq. (14) as the order parameter for the product phase.

We further vary  $h_4$  from -0.2 to 0.2 and compare the machine-learning motivated order parameter  $\Gamma$  and the conventional order parameter  $\langle \sigma^1 \sigma^2 \rangle$  in Fig. 7. We find that when the  $Z_2$  symmetry in  $\langle \sigma^1 \sigma^2 \rangle$  is broken by the external field that



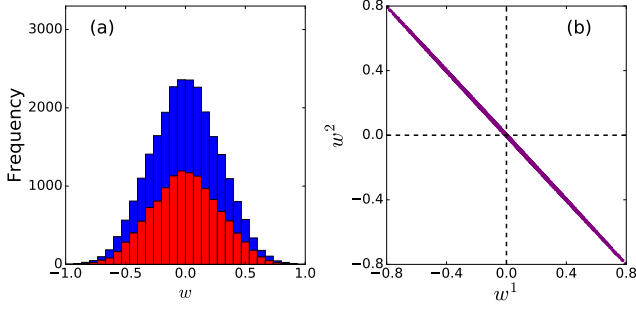


FIG. 6: (a) Histogram of the coupling coefficients  $w_{ij}^\alpha$  of the two-color AT model on a  $8 \times 8$  square lattice for  $K_2/T = 0.1$  and  $K_4/T = 1$  in the product phase with an additional symmetry breaking field  $h_4 \sum_i \sigma_i^1 \sigma_i^2$ , where  $h_4 = 0.2$ . They are color coded such that blue stands for color index  $\alpha = 1$  and red for  $\alpha = 2$ . (b) Comparison of the corresponding coupling coefficients  $w_{ij}^1$  versus  $w_{ij}^2$ . The straight line with slope -1 is consistent with  $\langle \sigma_i^1 \sigma_i^2 \rangle = -1$  for sufficiently large  $h_4 > 0$ .

couples to the local order parameter, the RBM study renders  $\Gamma$  in *quantitative* agreement with  $\langle \sigma^1 \sigma^2 \rangle$ . Due to the sign function we choose,  $\Gamma$  properly saturates to  $\pm 1$  at sufficiently large  $h_4$ , just as  $\langle \sigma^1 \sigma^2 \rangle$  does.

#### D. Generalization to the AT model with more color

We have now identified  $\Gamma$ , obtained from the products  $w_{ij}^1 w_{ij}^2$ , to be a legitimate order parameter in the RBM learning of the product phase in the two-color AT model. This prompts, however, interesting further questions. How do we generalize the results to the  $N_c$ -color AT model for  $N_c \geq 3$ ? Is the number of lumps determined by the number of colors  $N_c$ ? Can the product  $w_{ij}^\alpha w_{ij}^\beta$  (and  $a_i^\alpha w_{ij}^\alpha$ ) continue to serve as the indicators of the product phase? To answer these questions, we also study the  $N_c = 3$  case as a self-consistent check. We plot the local fields  $a_i^\alpha$  and the histogram of the coupling coefficients  $w_{ij}^\alpha$  resulting from the RBM training of a  $4 \times 4$  lattice for  $K_2/T = 0.1$  and  $K_4/T = 1$  in the product phase in Fig. 8, where all the training parameters are identical to the two-color training case. Ergodicity is chosen to be broken, again, in the data of spin configurations. We find that  $a_i^\alpha$  for color index  $\alpha = 1$  and 3 have the positive sign and  $a_i^2$  always assume the opposite sign. Even though the sign for each color may not be the same, we find  $a_i^1 a_i^2 < 0$ ,  $a_i^2 a_i^3 < 0$ , and  $a_i^3 a_i^1 > 0$  for every lattice site index  $i$ . Correspondingly, we find  $w_{ij}^1 w_{ij}^2 < 0$ ,  $w_{ij}^2 w_{ij}^3 < 0$ ,  $w_{ij}^3 w_{ij}^1 > 0$ , and  $a_i^\alpha w_{ij}^\alpha > 0$ . These results, as we argued in the two-color case, are consistent with the product order  $\langle \sigma_i^1 \sigma_i^2 \rangle < 0$ ,  $\langle \sigma_i^2 \sigma_i^3 \rangle < 0$ , and  $\langle \sigma_i^3 \sigma_i^1 \rangle > 0$ .

The combination of the training results in the product phase of the two-color and the three-color AT models suggests that the product phase in a generic  $N_c$ -color AT model features two lumps in the histogram of  $w_{ij}^\alpha$ , one on each side of the origin, when ergodicity in the spin configurations is broken. The sign of either  $w_{ij}^\alpha$  or  $a_i^\alpha$  depends on the history of training and the

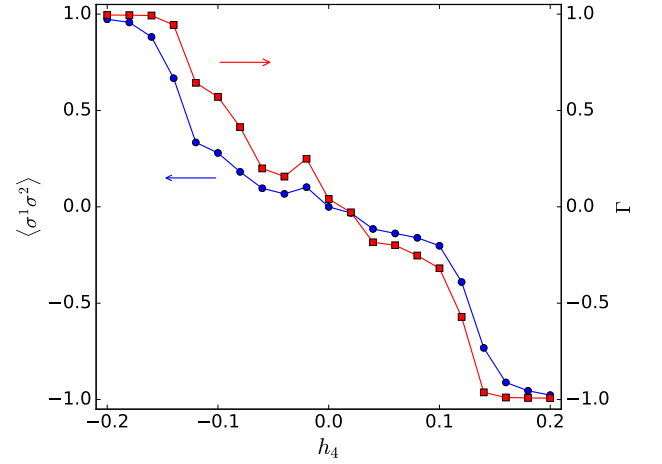


FIG. 7: Evolution of the machine-learning order parameters  $\Gamma$  [Eq. (14)] and the conventional order parameter  $\langle \sigma^1 \sigma^2 \rangle$  in the two-color AT model on a  $8 \times 8$  square lattice for  $K_2/T = 0.1$  and  $K_4/T = 1$  in the product phase with an additional symmetry breaking field  $h_4 \sum_i \sigma_i^1 \sigma_i^2$ , where  $h_4$  varies from -0.2 to 0.2.

site index  $i$ , but the sign of any one of the products  $w_{ij}^\alpha w_{ij}^\beta$ ,  $a_i^\alpha a_j^\beta$ , and  $a_i^\alpha w_{ij}^\alpha$  does not depend on the site index  $i$  or  $j$  of the visible or hidden Ising spin deep in the product phase. In other words, the sign of any of the products has translational invariance.

#### IV. SUMMARY AND DISCUSSION

We have trained the RBM to study the 2D two-color AT model and identified corresponding patterns in the RBM parameters for ferromagnetic, paramagnetic, and product phases. The ferromagnetic phase is characterized by a broad distribution of interlayer couplings  $w_{ij}^\alpha$  and polarized external fields  $a_i^\alpha$ , which are consistent with the spontaneous  $Z_2$  symmetry breaking in the Ising spins. The paramagnetic phase is characterized by a relatively narrower distribution of the coupling coefficients  $w_{ij}^\alpha$  with a zero mean and unpolarized external field  $a_i^\alpha$ , which imply that the  $Z_2$  symmetry of the Ising spins are conserved. The nontrivial product phase in the AT model is disordered with regard to spins of a single color, i.e.  $\langle \sigma_i^\alpha \rangle = 0$ , but ordered with regard to the product of the two spins on the same site, i.e.  $\langle \sigma_i^\alpha \sigma_i^\beta \rangle \neq 0$ . In the product phase, we find that the products  $w_{ij}^1 w_{ij}^2$  are polarized and can be used to construct an order parameter which quantitatively mimics  $\langle \sigma_i^1 \sigma_i^2 \rangle$ . The polarization of either  $w_{ij}^\alpha$  or  $a_i^\alpha$  depends on training and cannot serve as an indicator of the product phase. These results can be straightforwardly generalized to the generic  $N_c$ -color AT model.

The RBM learning of the Boltzmann distribution is an unsupervised learning. We only feed in the spin configurations by Monte Carlo simulation. We do not provide any knowledge from our physical understanding of the phases and corresponding phase transitions. Through the distribution and

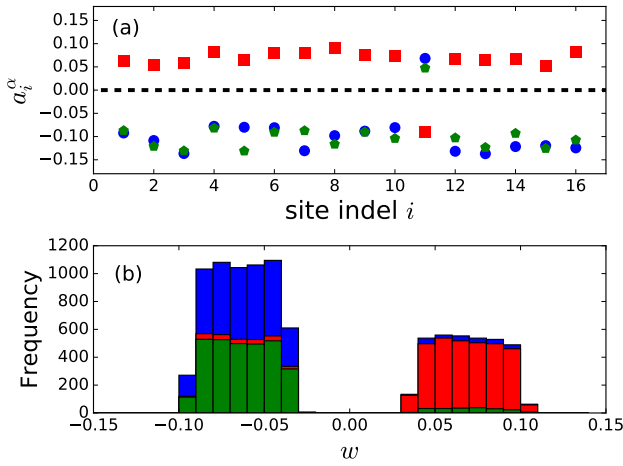


FIG. 8: RBM parameters obtained from the training of the three-color AT model on a  $4 \times 4$  square lattice with  $K_2/T = 0.1$  and  $K_4/T = 1$  in the product phase. The figure is color coded such that blue stands for color index  $\alpha = 1$ , red for  $\alpha = 2$ , and green for  $\alpha = 3$ . (a) The local fields for blue and green spins are of the same sign, which is opposite to that of the red spins on the same site. (b) The histogram of the coupling coefficients  $w_{ij}^\alpha$  exhibits two lumps, one on each side of the origin. The left lump is larger and formed mostly by edges connecting blue and green spins to hidden spins. The right lump is smaller and formed mostly by edges connecting red spins to hidden spins.

polarization of the parameters or the products of parameters, the RBM can distinguish disordered, partially ordered, and ordered phases. As demonstrated in the product phase, the information not only provides hints on what order is established in the corresponding phase, but also facilitates the construction of order parameter. This is an attractive application of various machine learning schemes, in particular, in poorly-studied models or models without local order parameters, as

evident in other machine learning studies.<sup>13</sup>

The success in the identification of the product phase, as well as of the ferromagnetic phase and the paramagnetic phase, allure us to explore beyond training translationally invariant classical models. One direction is to introduce disorder. The numerically obtained RBM parameters do not respect translational invariance. However, the sign of the parameters (e.g.,  $a_i^\alpha$  in the ferromagnetic phase) or that of the products of the parameters (e.g.,  $w_{ij}^1 w_{ij}^2$  in the product phase) respects translational invariance and can be used to replace local order parameters as the indicator of the corresponding phase. It would be interesting to ask to what extent the results can be generalized in the disordered AT model. The disordered model also features a continuous phase transition rounded by disorder from a first-order phase transition and the emerging critical behavior is shown to be in the clean 2D Ising universality class, accompanied by universal logarithmic corrections.<sup>37</sup> Whether it is possible to distinguish a continuous transition from a first-order one in machine learning and how to identify the universality class of a continuous transition are interesting questions. Another interesting direction is to learn topological quantum systems, whose topological properties are robust against disorder. The Kitaev model is one such example, which has simple and translationally invariant solutions when one represents the quantum many-body ground state by an RBM.<sup>28</sup> To what extent numerical training can identify the topology in such systems is of great interest.<sup>11</sup>

## V. ACKNOWLEDGEMENTS

This work is supported by the National Basic Research Program of China through Project No. 2015CB921101 and the National Natural Science Foundation of China through Grant No. 11674282.

- <sup>1</sup> I. Goodfellow, Y. Bengio and A. Courville, Deep Learning (MIT Press 2016).
- <sup>2</sup> J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
- <sup>3</sup> G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).
- <sup>4</sup> L. Wang, Phys. Rev. B **94**, 195105 (2016).
- <sup>5</sup> E. P. L. van Nieuwenburg, Y.-H. Liu and S. D. Huber, Nat. Phys. **13**, 435 (2017).
- <sup>6</sup> Y.-H. Liu and E. P. L. van Nieuwenburg, arXiv:1706.08111v1.
- <sup>7</sup> A. Tanaka and A. Tomiya, arXiv:1609.09807v1.
- <sup>8</sup> C.-D. Li, R.-R. Tan and F.-J. Jiang, arXiv:1703.02369v1.
- <sup>9</sup> K. Ch'ng, J. Carrasquilla, R. G. Melko and E. Khatami, arXiv:1609.02552v1.
- <sup>10</sup> F. Schindler, N. Regnault and T. Neupert, arXiv:1704.01578v1.
- <sup>11</sup> Y. Zhang and E.-A. Kim, Phys. Rev. Lett. **118**, 216401 (2017).
- <sup>12</sup> Y. Zhang, R. G. Melko and E.-A. Kim, arXiv:1705.01947v1.
- <sup>13</sup> T. Ohtsuki and T. Ohtsuki, J. Phys. Soc. Jpn. **85**, 123706 (2016).
- <sup>14</sup> P. Broecker, J. Carrasquilla, R. G. Melko and S. Trebst, arXiv:1608.07848v1.
- <sup>15</sup> J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, Phys. Rev. B **95**, 041101 (2017).
- <sup>16</sup> J. Liu, H. Shen, Y. Qi, Z. Y. Meng, and L. Fu, arXiv:1611.09364.
- <sup>17</sup> L. Huang and L. Wang, Phys. Rev. B **95**, 035105 (2017).
- <sup>18</sup> L. Huang, Y.-F. Yang, and L. Wang, Phys. Rev. E **95**, 031301 (2017).
- <sup>19</sup> X. Y. Xu, Y. Qi, J. Liu, L. Fu, and Z. Y. Meng, arXiv:1612.03804.
- <sup>20</sup> Y. Nagai, H. Shen, Y. Qi, J. Liu, and L. Fu, arXiv:1705.06724.
- <sup>21</sup> H. W. Lin and M. Tegmark, arXiv:1608.08225v2.
- <sup>22</sup> P. Mehta and D. J. Schwab, arXiv:1410.3831v1.
- <sup>23</sup> M. Koch-Janusz and Z. Ringel, arXiv:1704.06279v1.
- <sup>24</sup> M. E. J. Newman and G. T. Barkema, Monte Carlo Methods in Statistical Physics (Oxford University Press, 2001).
- <sup>25</sup> G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko and G. Carleo, arXiv:1703.05334v1.
- <sup>26</sup> G. Carleo and M. Troyer, Science **355**, 602, (2017).
- <sup>27</sup> Z. Cai, arXiv:1704.05148v1.
- <sup>28</sup> D.-L. Deng, X. Li and S. Das Sarma, arXiv:1609.09060.
- <sup>29</sup> D.-L. Deng, X. Li and S. Das Sarma, Phys. Rev. X **7**, 021021 (2017).

- <sup>30</sup> Y. Huang and J. E. Moore, arXiv:1701.06246.
- <sup>31</sup> J. Chen, S. Cheng, H. Xie, L. Wang and T. Xiang, arXiv:1701.04831.
- <sup>32</sup> X. Gao and L.-M Duan, arXiv:1701.05039.
- <sup>33</sup> M. H. Amin, E. Andriyash, J. Rolfe, B. Kilchytskyy and R. Melko, arXiv:1601.02036v1.
- <sup>34</sup> A. Morningstar and R. G. Melko, arXiv:1708.04622.
- <sup>35</sup> R. Baxter, Exactly Solved Models in Statistical Mechanics (Academic Press, New York, 1982).
- <sup>36</sup> D. Bollé and P. Kozłowski, J. Phys. A: Math. Gen. **31**, 6319 (1998).
- <sup>37</sup> Q. Zhu, X. Wan, R. Narayanan, J. A. Hoyos and T. Vojta, Phys. Rev. B **91**, 224201 (2015).
- <sup>38</sup> G. Hinton, Neural Comput. **14**, 1771 (2002). See, also, the appendix of Ref. 3.
- <sup>39</sup> L. Wang, arXiv:1702.08586v1.
- <sup>40</sup> N. Le Roux and Y. Bengio, Neural Comput. **20**, 1631 (2008).
- <sup>41</sup> G. Montufar and N. Ay, Neural Comput. **23**, 1306 (2011).