Machine learning of frustrated classical spin models. I. Principal component analysis

Ce Wang[1] and Hui Zhai[1,2]
[1]*Institute for Advanced Study, Tsinghua University, Beijing 100084, China*
[2]*Collaborative Innovation Center of Quantum Matter, Beijing 100084, China*
(Received 17 August 2017; revised manuscript received 16 October 2017; published 26 October 2017)

This work aims at determining whether artificial intelligence can recognize a phase transition without prior human knowledge. If this were successful, it could be applied to, for instance, analyzing data from the quantum simulation of unsolved physical models. Toward this goal, we first need to apply the machine learning algorithm to well-understood models and see whether the outputs are consistent with our prior knowledge, which serves as the benchmark for this approach. In this work, we feed the computer data generated by the classical Monte Carlo simulation for the XY model in frustrated triangular and union jack lattices, which has two order parameters and exhibits two phase transitions. We show that the outputs of the principal component analysis agree very well with our understanding of different orders in different phases, and the temperature dependences of the major components detect the nature and the locations of the phase transitions. Our work offers promise for using machine learning techniques to study sophisticated statistical models, and our results can be further improved by using principal component analysis with kernel tricks and the neural network method.

DOI: 10.1103/PhysRevB.96.144432

I. INTRODUCTION

The most ambitious goal for applying machine learning to physics research is to discover new physics without prior human knowledge. For example, in the near future one practical use will be to analyze data from quantum simulation. The ultimate goal of quantum simulation is to perform simulating quantum computation on physical models that cannot be solved by a classical computer. That is to say, there is no reliable prior human knowledge of these models. One such example is the quantum simulation of the Fermi-Hubbard model [1]. The Fermi-Hubbard model lies at the heart of understanding many strongly correlated materials; however, it cannot be solved by reliable classical computational methods once the filling is away from half filling, and the exact diagonalization is limited to a very small size, making it an issue of debate for several centuries. Experiments on ultracold Fermi atoms in optical lattices can now simulate the Fermi-Hubbard model with a system size much larger than the computation capability of a classical computer [2]. The output of these quantum simulation experiments will be a huge amount of data measured at different temperatures and different parameters. Hence, how one can analyze these big data sets to extract useful, and sometimes latent, physical information without prior knowledge of the model has become a crucially important problem. From this perspective, artificial intelligence could be useful.

Toward this goal, we should first see whether the machine learning algorithm can reproduce the known results of well-understood models, which serves as an important benchmark for the machine learning approach. For this purpose, we focus on classical statistical models. For these models, the classical Monte Carlo simulation can generate many configurations at different temperatures and different parameter regimes, and these configurations are used as input data to feed the computer. Over many decades, physicists have developed the concept of an "order parameter," with which these configurations can be classified into different phases; however, a computer does not have a concept of an order parameter as prior knowledge. Nevertheless, through the algorithm developed in machine learning studies [3], a computer can also successfully group the configurations into different classes. We will show that the classification by the machine learning algorithm is consistent with the order-parameter classification of these models. Further, by adding information for the temperature associated with each configuration, the computer can also determine the critical temperature of the phase transition.

Recent studies along this line have investigated the Ising model and some other related models by applying various machine learning methods such as principal component analysis with or without kernel tricks and artificial neural networks [4–11]. The classical XY model with continuous variables on an unfrustrated square lattice has also been studied [8,10]. Some works have also studied quantum models of strongly correlated fermions and accelerated Monte Carlo simulations [12–19]. In this work we study the classical XY model in frustrated lattices [20–28], including both triangular and union jack lattices. The key advance is that these models are frustrated and exhibit two order parameters, which are a $U(1)$ order parameter and a chirality order parameter, and, consequently, two successive phase transitions as the temperature decreases [26–28]. We show that the machine learning method can successfully capture both. Our work represents an important step toward applying machine learning to sophisticated physical models.

II. MODEL AND METHOD

In this work we consider the XY model, whose Hamiltonian is given by

$$\mathcal{H} = J \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j), \tag{1}$$

where $\theta_i \in (0, 2\pi]$ is a classical variable defined at each site and $\langle ij \rangle$ denotes all the nearest-neighboring bonds. Below we will consider different lattice structures, including two-dimensional square, triangular, and union jack lattices,

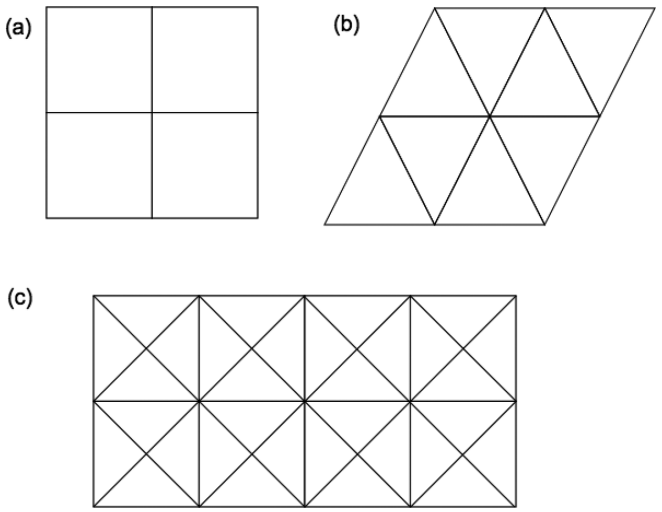

FIG. 1. Schematic of (a) square, (b) triangular, and (c) union jack lattices.

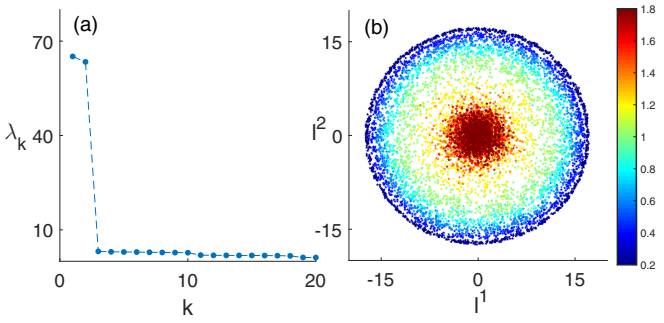

FIG. 2. (a) The eigenvalues of the $\mathcal{S}$ matrix for the XY model in a two-dimensional 18×18 square lattice. Two eigenvalues are significantly larger than the others. (b) The projection of all $N = 9000$ data sets onto the two-dimensional principal subspace $l_n = (l_n^1, l_n^2)$ $(n = 1, \ldots, N)$. The temperature of the data set ranges from $0.2J$ to $1.8J$, with $\Delta T = 0.2J$, and at each temperature 1000 data sets are taken from the classical Monte Carlo simulation. The color bar indicates the temperature at which the data are generated.

respectively, as shown in Fig. 1. We use classical Monte Carlo to generate equilibrium configurations, denoted by $\{x_n\}$ $(n = 1, \ldots, N)$, on a lattice with L total lattice sites and at different temperatures. N is the total number of data sets, and $N = N_T \times N_0$, where N_0 is the number of configurations generated at each given temperature and N_T is the number of different temperatures considered. At each site i, the spin is described by its orientation $(\cos\theta_i, \sin\theta_i)$, and in this case, each x_n is a vector of $2L$ dimensions, and for later convenience, they are organized as

$$x_n = (\cos\theta_1, \ldots, \cos\theta_L, \sin\theta_1, \ldots, \sin\theta_L). \tag{2}$$

Here $\{x_n\}$ $(n = 1, \ldots, N)$ will be the data and the only input that we feed to the computer. In our analysis below, in most cases we will treat all N data sets with different temperatures together. In some cases that we specify as temperature-resolved analyses, we treat each N_0 set of data with a given temperature individually.

The principal component analysis (PCA) method is to find out one or a few directions denoted by u, and the projection of $(x_n - \bar{x})$ to u can maximally distinguish all data sets, where $\bar{x} = 1/N \sum_n x_n$ is the average of the data sets [3]. This is equivalent to searching for a vector u such that the variance σ^2, defined as

$$\sigma^2 = \frac{1}{N} \sum_n [u^T (x_n - \bar{x})]^2, \tag{3}$$

will be maximized. This is further equivalent to finding out the eigenvector corresponding to the largest or the largest few eigenvalues of the matrix $\mathcal{S}$, which is defined as

$$\mathcal{S} = \frac{1}{N} \sum_n (x_n - \bar{x})(x_n - \bar{x})^T. \tag{4}$$

$\mathcal{S}$ is a $2L \times 2L$ dimensional matrix in this case, and it should have $2L$ eigenvalues, whose eigenvectors u_i $(i = 1, \ldots, 2L)$ form a complete basis set. If there exist W eigenvalues λ_k $(k = 1, \ldots, W)$ that are much larger than the rest λ_k $(k = W + 1, \ldots, 2L)$, their corresponding eigenvectors u_k $(k = 1, \ldots, W)$ are called the "principal component," and they form the "principal subspace." In this case, each vector x_n is characterized by its projection into the principal subspace, described by a W-dimension number $l_n \equiv (l_n^k) = (u_k^T x_n)$ $(k = 1, \ldots, W)$. In other words, each data set is now replaced by

$$x_n \approx \sum_{k=1}^{W} l_n^k u_k + \sum_{k=W+1}^{2L} (u_k^T \bar{x}) u_k. \tag{5}$$

Below we will apply this method to the XY model in square, triangular, and union jack lattices. We will analyze the outputs of the PCA algorithm, which are the largest few λ_k, their corresponding u_k, and the structure of l_n. We will show that these results are consistent with our understanding of the order parameters in different lattices.

III. SQUARE LATTICES

The PCA for the square-lattice XY model was reported in Ref. [10]. Here we review the results and analyze the results based on a simple toy model. This viewpoint will be useful for later discussion of triangular and union jack lattices, and the results for the square lattice also help us understand those of the union jack lattice. The results from the square lattice are shown in Fig. 2. Clearly, in Fig. 2(a) one can see that there are two eigenvalues that are much larger than all the rest, and their corresponding eigenvectors are denoted by u_1 and u_2. Projecting all x_n to (u_1, u_2), we plot all $l_n = (l_n^1, l_n^2)$ in Fig. 2(b). One can see that some data are concentrated near the center, and actually, they are data generated at high temperatures; other data form a circle, and they are generated at low temperatures. The question is whether and how these outputs are consistent with our prior knowledge of the XY model.

The insight for understanding this output can be obtained by considering the following toy model. Let us consider just two sites denoted by A and B, and the data will be

$$x_n = (\cos\theta_A, \cos\theta_B, \sin\theta_A, \sin\theta_B). \tag{6}$$

For simplicity, we assume that among all N data sets, p percent of the data mimic the lower-temperature case where an antiferromagnetic order is formed; therefore, we set $\theta_B = \theta_A + \pi$, with θ_A being uniformly distributed among $\in (0,2\pi]$, and the other $1-p$ percent of data mimic the higher-temperature case where θ_A and θ_B are uncorrelated and they are uniformly distributed among $\in (0,2\pi]$. Hence, it is straightforward to obtain, averaging over a sufficiently large number of data sets,

$$\mathcal{S} = (1-p)\mathcal{S}_h + p\mathcal{S}_l, \tag{7}$$

where $\mathcal{S}_h = \mathcal{I}/2$, and

$$\mathcal{S}_l = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda \end{pmatrix}, \tag{8}$$

where

$$\Lambda = \frac{1}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}. \tag{9}$$

$\mathcal{S}$ has two degenerate larger eigenvalues $\lambda_1 = \lambda_2 = (1+p)/2$, and the other two eigenvalues are $(1-p)/2$. The eigenvectors corresponding to λ_1 and λ_2 are

$$u_1 \propto (1,-1,0,0), \tag{10}$$

$$u_2 \propto (0,0,1,-1). \tag{11}$$

Projecting higher-temperature data into the (u_1,u_2) subspace results in $l_n \propto (\cos\theta_A - \cos\theta_B, \sin\theta_A - \sin\theta_B)$, and projecting lower-temperature data into the (u_1,u_2) subspace results in $l_n \propto (\cos\theta_A, \sin\theta_A)$.

This toy model can be straightforwardly changed to a full lattice model by considering A and B as a unit cell and using the translational symmetry to generate a full two-dimensional square lattice. We still consider two types of data: p percent of data simulate the low-temperature case with antiferromagnetic long-range order, where all A sublattice sites have the same spin angle θ and the B sublattice sites have the spin angle $\theta+\pi$, and the other $1-p$ percent data simulate the high-temperature case where the spin angles at all sites are not correlated at all. We note that this is still an oversimplified situation compared to real data, but we can obtain some insights because it is straightforward to construct the $\mathcal{S}$ matrix and find out all eigenvalues analytically. Interestingly, we find in this case there are two degenerate larger eigenvalues $[1+(L/2-1)p]/2$, which is much larger than the remaining $(1-p)/2$ for large L. The eigenvectors corresponding to the larger two eigenvalues are

$$u_1 \propto (1,-1,\ldots,1,-1,0,0,\ldots,0,0), \tag{12}$$

$$u_2 \propto (0,0,\ldots,0,0,1,-1,\ldots,1,-1). \tag{13}$$

Projecting higher-temperature data into this principal subspace, when the system is large enough, $\cos\theta$ or $\sin\theta$ at different sites will average each other out, and eventually, $l_n \approx (0,0)$. Projecting lower-temperature data into this principal subspace results in $l_n \propto (\cos\theta, \sin\theta)$, which form a circle. This basically shows that the main features in the outputs of the PCA algorithm shown in Fig. 2(b) capture the difference between the presence and absence of $U(1)$ spin order in lower- and higher-temperature phases.

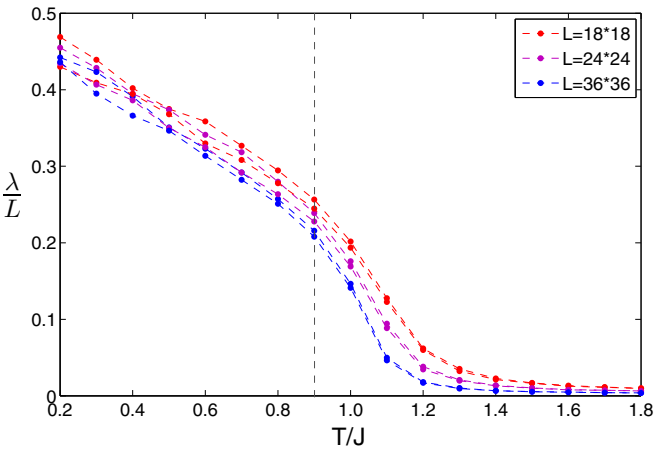

FIG. 3. Temperature-resolved PCA for the square-lattice XY model in $L = 18\times 18$, $L = 24\times 24$, and $L = 36\times 36$ lattices, respectively. We consider a temperature range from $0.2J$ to $1.8J$, with $\Delta T = 0.1J$, and at each fixed temperature, $N_0 = 1000$ data sets are input. The dashed line indicates the expected temperature for the KT transition in the square-lattice XY model.

From the above analysis, one can also see that if all the data are collected from high temperature, that is to say, when p is set to zero, there will be no major eigenvalues that are much larger than the others. Thus, we perform a temperature-resolved PCA, in which all data fed to the computer are generated at the same temperature, and then we can plot the normalized eigenvalues λ_1/L and λ_2/L as a function of the temperature of the input data. The result for the square lattice is shown in Fig. 3. On the other hand, our knowledge of statistical physics already tells us that this model displays a Kosterlitz-Thouless (KT) phase transition at $T_c \approx 0.9J$ [29], below which a quasi-long-range antiferromagnetic order is formed. Indeed, it is found that the two major normalized eigenvalues are insensitive to system size for temperature below T_c, and for temperature above T_c, the two major normalized eigenvalues quickly decrease as the system size increases.

IV. TRIANGULAR LATTICES

Now we consider the XY model in a triangular lattice. For the triangular lattice, we know that there are two different order parameters at low temperatures. One of them is the $U(1)$ spin order whose transition is a Kosterlitz-Thouless transition expected at $T_{c1} = 0.504J$ [27,28]. The other order is a Z_2 local chiral order defined at each triangle. This chiral order is shown in Fig. 4. The planar spins have a $2\pi/3$ angle between two neighboring sites, and at each triangle, the planar spins rotate either clockwise or counterclockwise, as shown in Fig. 4. This Z_2 transition takes place at $T_{c2} = 0.512J$ [27,28]. Since these two transitions are very close in this model, we do not have enough samples at the temperature window between these two transitions; therefore, the purpose of this model is to discuss the situation with two order parameters, but practically, it is considered a single transition. In order to discuss the situation with two separated transitions, for a triangular lattice we need

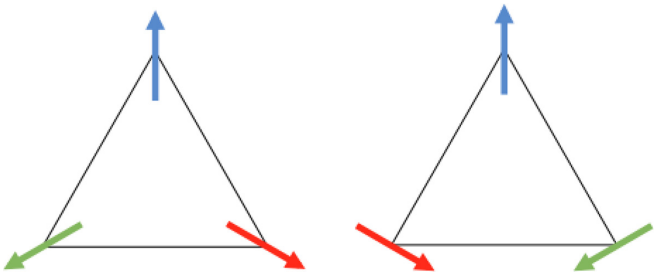

FIG. 4. Schematics of two types of chiral orders for the XY model in a triangular lattice.

a huge amount of data such that there will be enough samples between the two transitions. Instead, this physics is easier to elaborate with the union jack model in which the two transition temperatures are quite different.

A similar PCA is performed for data collected from the Monte Carlo simulation of nine different temperatures. Figure 5 shows that there are four eigenvalues that are much larger than the others. They are denoted $\lambda_1,\ldots,\lambda_4$, and their corresponding eigenvectors are $u_1,\ldots,u_4$, which form the principal subspace. Thus, each data set x_n is now characterized by a four-dimensional number $l_n = (l_n^1,l_n^2,l_n^3,l_n^4)$. In Fig. 6 we present these numbers for all data.

To understand these results, we perform a similar toy model analysis. Now let us consider three sites denoted by A, B, and C. Among all N, in $p/2$ percent of data, the planar spins rotate clockwise, and the spin angles at the three sites are $\theta,\theta+2\pi/3$, and $\theta+4\pi/3$, respectively; for another $p/2$ percent of data, the planar spins rotate counterclockwise, and the spin angles at the three sites are $\theta,\theta-2\pi/3$, and $\theta-4\pi/3$, respectively, which mimics two different types of chiral orders, as shown in Fig. 4. In the remaining $1-p$ percent of data, the spin angles at the three different sites are not correlated. In this case, it is straightforward to calculate

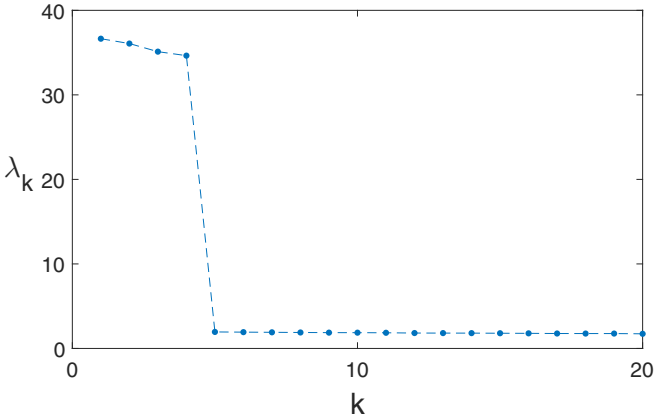

FIG. 5. The eigenvalues of the $\mathcal{S}$ matrix for the XY model in a two-dimensional 18×18 triangular lattice. Four eigenvalues are significantly larger than the others. The temperature of the data set ranges from $0.3J$ to $0.7J$, with $\Delta T = 0.05J$, and at each temperature $N_0 = 1000$ data sets are taken from the classical Monte Carlo simulation.

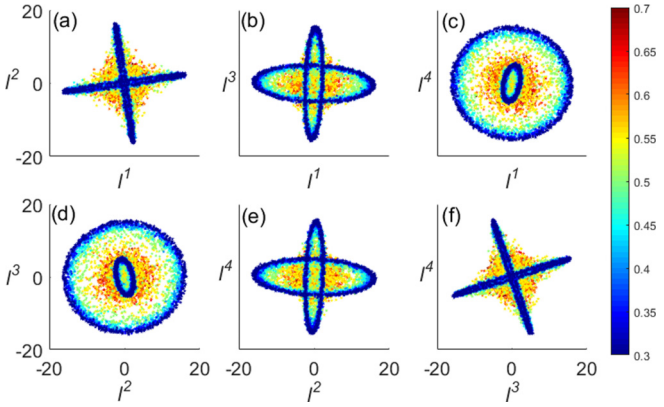

FIG. 6. The projection of data for the XY model in a triangular lattice to the principal subspace. The principal subspace is generated by data from all temperature. (a)–(f) For (l_n^1,l_n^2), (l_n^1,l_n^3), (l_n^1,l_n^4), (l_n^2,l_n^3), (l_n^2,l_n^4), and (l_n^3,l_n^4), respectively. The color bar indicates the temperature at which the data are generated.

that

$$\mathcal{S} = (1-p)\mathcal{S}_h + p\mathcal{S}_l, \tag{14}$$

where $\mathcal{S}_h = \mathcal{I}/2$,

$$\mathcal{S}_l = \begin{pmatrix} \Lambda & 0 \\ 0 & \Lambda \end{pmatrix}, \tag{15}$$

and

$$\Lambda = \frac{1}{2}\begin{pmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} \\ -\frac{1}{2} & 1 & -\frac{1}{2} \\ -\frac{1}{2} & -\frac{1}{2} & 1 \end{pmatrix}. \tag{16}$$

Note that two eigenvalues of $\mathcal{S}_l$ are $3/4$ and the other is zero; the corresponding eigenvectors for eigenvalues of $3/4$ can be written as

$$d_1 \propto \left(1,\cos\left(\frac{2\pi}{3}\right),\cos\left(\frac{4\pi}{3}\right)\right) \propto (2,-1,-1), \tag{17}$$

$$d_2 \propto \left(0,\sin\left(\frac{2\pi}{3}\right),\sin\left(\frac{4\pi}{3}\right)\right) \propto (0,1,-1). \tag{18}$$

Therefore, the matrix $\mathcal{S}$ has four degenerate eigenvalues as $1/2+p/4$, and the remaining two are $1/2-p/2$. Corresponding to the four degenerate eigenvalues, one way to choose the orthogonal eigenvectors is

$$u_1 = (d_1,d_2), \tag{19}$$

$$u_2 = (d_2,d_1), \tag{20}$$

$$u_3 = (d_1,-d_2), \tag{21}$$

$$u_4 = (-d_2,d_1). \tag{22}$$

Projecting the low-temperature data into this subspace, for samples with clockwise chirality, $l_n = (3\cos\theta,0,0,3\sin\theta)$; and for samples with counterclockwise chirality, $l_n = (0,3\sin\theta,3\cos\theta,0)$. Keeping in mind that θ is uniformly distributed among $(0,2\pi]$, plotting (l_n^1,l_n^2), (l_n^1,l_n^3), (l_n^2,l_n^4), and (l_n^3,l_n^4) for all data gives a cross, and data with different chiralities sit at different lines of the cross. When plotting (l_n^1,l_n^4), the clockwise chirality sample forms a circle, and

the counterclockwise chirality samples are concentrated at the center. The situation is reversed when plotting (l_n^2,l_n^3).

Because all four eigenvalues are degenerate, one can also make a rotation for the bases $u_1,\ldots,u_4$. For instance, one can choose

$$u_1' = \cos\alpha u_1 + \sin\alpha u_2, \tag{23}$$

$$u_2' = -\sin\alpha u_1 + \cos\alpha u_2, \tag{24}$$

$$u_3' = \cos\beta u_3 + \sin\beta u_4, \tag{25}$$

$$u_4' = -\sin\beta u_3 + \cos\beta u_4. \tag{26}$$

Under these bases, for samples with clockwise chirality,

$$l_n = 3(\cos\alpha\cos\theta, -\sin\alpha\cos\theta, \sin\beta\sin\theta, \cos\beta\sin\theta), \tag{27}$$

and for samples with counterclockwise chirality,

$$l_n = 3(\sin\alpha\sin\theta, \cos\alpha\sin\theta, \cos\beta\cos\theta, -\sin\beta\cos\theta). \tag{28}$$

In this case, plotting (l_n^1,l_n^2) and (l_n^3,l_n^4) gives rise to two tilted crosses, as shown in the low-temperature data (dark blue) in Figs. 6(a) and 6(f). Without loss of generality, we can assume both α and β belong to $(0,\pi/4]$, and plotting (l_n^1,l_n^3) and (l_n^2,l_n^4) results in two ellipses, where the long axes of one ellipse are along l^1 and those of the other are along l^2, as in Figs. 6(b) and 6(e). Plotting (l_n^1,l_n^4) and (l_n^2,l_n^3) also results in two ellipses, with one always being larger than the other, as in Figs. 6(c) and 6(d).

Here we also remark that the appearance of l_n is quite sensitive to the choice of bases. For instance, instead of using Eqs. (19)–(22), one can also choose the bases as

$$u_1 = (d_2,0), \tag{29}$$

$$u_2 = (0,d_2), \tag{30}$$

$$u_3 = (d_1,0), \tag{31}$$

$$u_4 = (0,d_1). \tag{32}$$

Under these bases, $l_n = \frac{3}{2}(-\sin\theta,\cos\theta,\cos\theta,\sin\theta)$ for low-temperature data with clockwise chirality, and $l_n = \frac{3}{2}(\sin\theta,-\cos\theta,\cos\theta,\sin\theta)$ for low-temperature data with counterclockwise chirality. In this case, (l_n^1,l_n^2), (l_n^1,l_n^3), (l_n^2,l_n^4), and (l_n^3,l_n^4) are all circles, and (l_n^1,l_n^4), (l_n^2,l_n^3) are two crosses at 45°. Furthermore, one can rotate the bases with the same transformation as in Eqs. (23)–(26), and under these rotated bases, for clockwise and counterclockwise samples, $l_n = \frac{3}{2}(-\sin(\theta-\alpha),\cos(\theta-\alpha),\cos(\theta+\beta),\sin(\theta+\beta))$, and $l_n = \frac{3}{2}(\sin(\theta-\alpha),-\cos(\theta-\alpha),\cos(\theta+\beta),\sin(\theta+\beta))$, respectively. In this case, (l_n^1,l_n^2) and (l_n^3,l_n^4) will remain two circles, and the other four combinations of (l_n^1,l_n^3), (l_n^1,l_n^4), (l_n^2,l_n^3), and (l_n^2,l_n^4) will all become two ellipses perpendicular to each other and orientated at 45° and 135°. We will come back to this point in the discussion of the union jack lattice.

In this toy model with only three sites, the eigenvalues of the four degenerate ones are not significantly larger than the other two; however, like in Sec. III, one can consider these three sites as a unit cell and generate a triangular lattice model. Similarly, we assume that in all data, the spin angles either are identical for all unit cells, simulating the low-temperature situation, or are not correlated at all, simulating the high-temperature situation. In this case, one can show that the four degenerate eigenvalues become $1/2 + p(L-2)/4$, and these four are much larger than the rest, which remain $(1-p)/2$. Similarly, when projecting the high-temperature data into the principal subspace, the contributions from different sites also average out, and one obtains $l_n \approx (0,0,0,0)$, as we can see in Fig. 6. In this way, we understand the output of the principal subspace and the projection of the data from both ordered and disordered phases into this subspace.

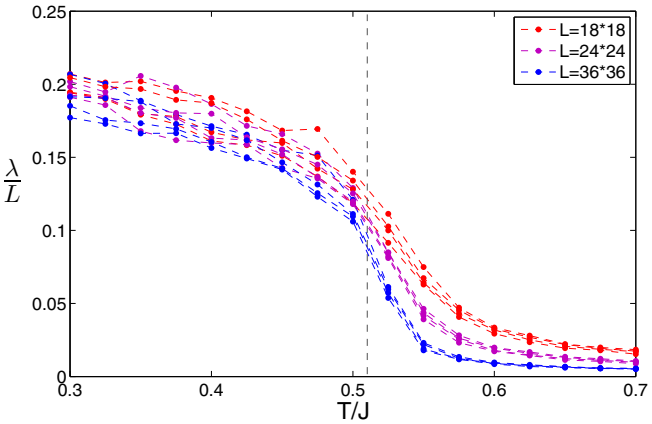

FIG. 7. Temperature-resolved PCA for the triangular-lattice XY model in $L = 18\times18$, $L = 24\times24$, and $L = 36\times36$ lattices. We consider a temperature range from $0.3J$ to $0.7J$, with $\Delta T = 0.025J$, and at each fixed temperature, $N_0 = 1000$ data sets are input. The dashed line indicates the expected temperature for the KT transition in the triangular-lattice XY model.

Before concluding this session, we also show in Fig. 7 a temperature-resolved PCA that reveals how the large principal component depends on temperature. Like in the square lattice case, below a certain temperature the four normalized major principal eigenvalues are not sensitive to system size, and above a certain temperature, they decrease to quite small ones as the system size increases. This transition temperature scale is consistent with the KT transition expected for the XY model in a triangular lattice.

V. UNION JACK LATTICE

Now we move to discussing the XY model in the union jack lattice [28]. The union jack lattice is made of a square lattice and an extra site at each center of the plaquette. The physics of the XY model in the union jack lattice is shown in Fig. 8. For $T > T_1$, it is a spin-disordered normal phase. In the range $T_1 > T > T_2$, the π antiferromagnetic spin order forms

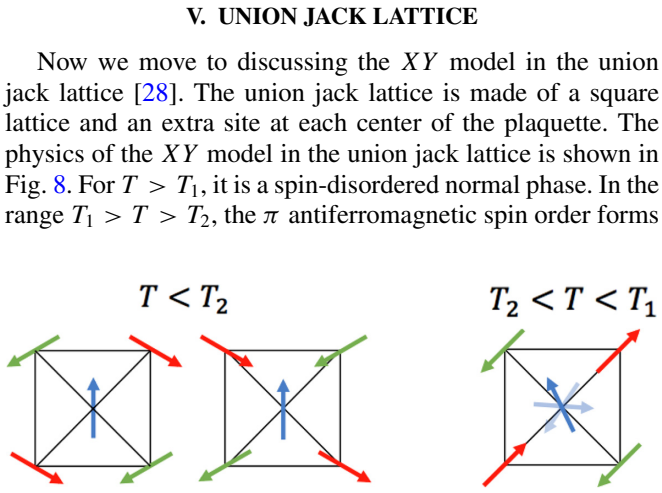

FIG. 8. Schematics of two types of chiral orders for the XY model in the union jack lattice at low temperature and an intermediate antiferromagnetic order at intermediate temperature.

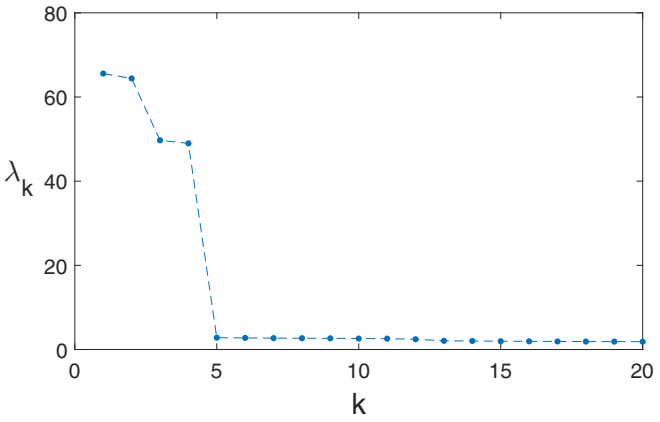

FIG. 9. The eigenvalues of the $\mathcal{S}$ matrix for the XY model in a two-dimensional 18×18 union jack lattice. Four eigenvalues are significantly larger than the others. The temperature of the data set ranges from $0.2J$ to $1.0J$, with $\Delta T = 0.1J$, and at each temperature $N_0 = 1000$ data sets are taken from the classical Monte Carlo simulation.

in the square lattice, as shown in Fig. 8(b). Further lowering the temperature to $T < T_2$, a $2\pi/3$ antiferromagnetic order is formed at each triangle, as shown in Fig. 8(a). Similar to the antiferromagnetic order in the triangular lattice, it has two opposite chiralities. The phase transition at T_1 is a Kosterlitz-Thouless type of $U(1)$ ordered transition, and the transition at T_2 is an Ising transition.

In the union jack lattice, $T_1 = 0.64J$ and $T_2 = 0.43J$ [28]. Since these two transition temperatures are well separated, we have enough data sampled at all three temperature regimes, i.e., $T > T_1$, $T_2 < T < T_1$, and $T < T_2$. Therefore, we use this model to address the issue of two phase transitions. We perform PCA for data collected from all three temperature regimes. The results are shown in Figs. 9 and 10. In Fig. 9 we can see that there are four eigenvalues that are considerably larger than the others. The corresponding four eigenvectors form the principal subspace. The projection of all data into this principal subspace is shown in Fig. 10.

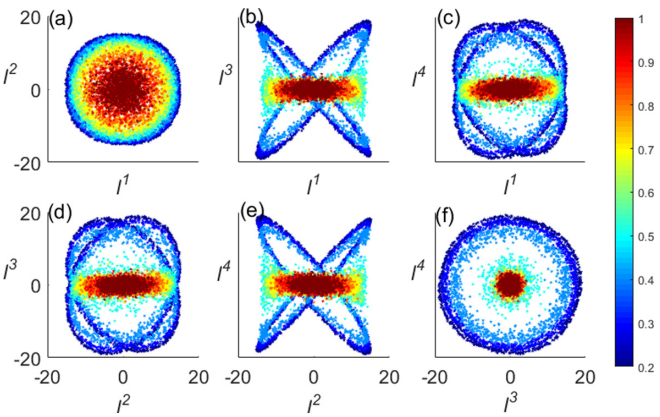

FIG. 10. The projection of data for the XY model in the union jack lattice to the principal subspace. (a)–(f) correspond to (l_n^1,l_n^2), (l_n^1,l_n^3), (l_n^1,l_n^4), (l_n^2,l_n^3), (l_n^2,l_n^4), and (l_n^3,l_n^4), respectively. The color bar indicates the temperatures at which the data are generated.

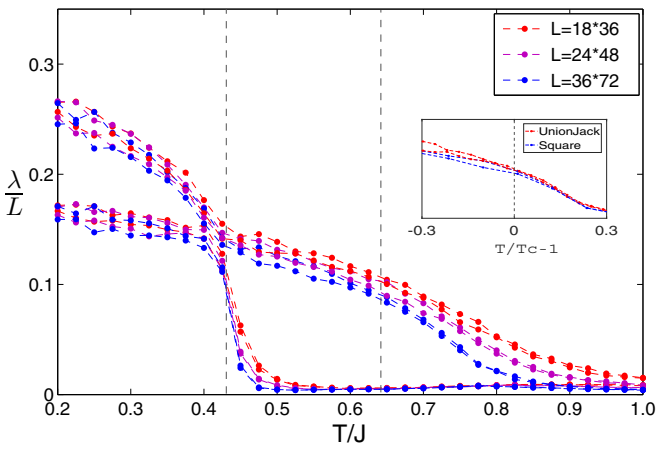

FIG. 11. Temperature-resolved PCA for the union-jack-lattice XY model in $L = 18 \times 36$, $L = 24 \times 48$, and $L = 36 \times 72$ lattices, respectively. We consider a temperature range from $0.2J$ to $1.0J$, with $\Delta T = 0.025J$, and at each fixed temperature, $N_0 = 1000$ data sets are input. The dashed line indicates the expected temperature for the Ising transition and the KT transition in the union-jack-lattice XY model. The inset shows a comparison of the critical behavior between the union jack lattice and square lattice; the eigenvalues are rescaled so that the data coincide at $T/T_c = 1$, and here we consider $L = 36 \times 72$ for the union jack lattice and $L = 36 \times 36$ for the square lattice.

In this case, the low-temperature ordered phase has four cites at each unit cell. As in the discussion in the above two sections, we can design a toy model with four sites to understand this behavior. We will not repeat a similar analysis here. In fact, for temperature below T_2, the situation is nearly the same as the low-temperature phase of the triangular lattice. Using the basis rotation of Eqs. (29)–(32), as we discussed in Sec. IV, two combinations, (l_n^1,l_n^2) and (l_n^3,l_n^4), behave as two circles, and the other four combinations of (l_n^1,l_n^3), (l_n^1,l_n^4), (l_n^2,l_n^3), and (l_n^2,l_n^4) are two ellipses perpendicular to each other. In fact, as one can see from Fig. 10, the data generated below T_2 behave in this way.

The reason we use the bases of Eqs. (29)–(32) to discuss the union jack lattice is worth emphasizing. It is an important observation that when there is only a π antiferromagnetic order in the square lattice as at intermediate temperature, as we discussed in Sec. III, there are two major components, and their eigenvalues are consistent with that of d_2. Hence, for data generated at intermediate temperature, $(l_n^3,l_n^4) \approx (0,0)$ because the site at the center of the plaquette is not ordered, (l_n^1,l_n^2) forms a circle as in the square lattice case, and (l_n^1,l_n^3), (l_n^1,l_n^4), (l_n^2,l_n^3), and (l_n^1,l_n^4) all behave as a one-dimensional line because they are a projection of the circle onto one of the axes. This is exactly what is found in Fig. 10.

In Fig. 11, we perform a temperature-resolved PCA. It is very clear that there are four major eigenvalues at the lowest temperature. Two of them vanish at a temperature $\sim 0.43J = T_2$, and the other two gradually vanish as temperature increases. As in the square and triangular lattice cases, the other two eigenvalues are not sensitive to system size below T_1, but above T_1 they decrease as the system size increases. Here we should stress that the different behaviors of these

eigenvalues at T_1 and T_2 are, in fact, quite physical. At T_2 the Ising transition is a second-order phase transition, and the eigenvalues indeed vanish pretty much close to the expected transition temperature. However, T_1 is a Kosterlitz-Thouless transition, and the transition takes place in a much smoother way. If we compare the behaviors of the principal eigenvalues near T_1 with the principal eigenvalues of the square lattice near its phase transition point by plotting them as a function of $T/T_c - 1$ ($T_c = 0.64J$ for the union jack lattice and $T_c = 0.9J$ for the square lattice), we find the data coinciding with each other by rescaling, as shown in the inset of Fig. 11. This result is consistent with the fact that the phase between T_2 and T_1 is actually the same as the low-temperature phase of the square lattice.

VI. OUTLOOK

In this work we utilized PCA analysis to classify the classical Monte Carlo data generated for the XY model in two-dimensional square, triangular, and union jack lattices, where the latter two are frustrated lattices and display two order parameters and the last one even has two well-separated phase transitions. Using simple toy models, we showed that the outputs of the PCA fully agree with our prior understanding of different orders in these models, and the temperature-resolved analysis of the principal components is also consistent with the critical temperature and the order of transitions.

Although this simple PCA is good enough to recognize different phases, the outputs are still too complicated to directly read out the order parameters. The physical reason is that some order parameters, such as the chirality, are a nonlinear function of the input function $\cos\theta$ or $\sin\theta$, and the kernel PCA can directly reveal these nonlinear order parameters [30]. Moreover, to reveal the Kosterlitz-Thouless transition more directly, one also needs some nonlinear order parameters such as superfluid density [31,32], and the kernel PCA can also help. Finally, we will use a neural network to find out the best kernel. These results will be published in subsequent publications.

ACKNOWLEDGMENTS

We thank Y. Deng for helpful discussions. This work is supported by NSFC Grant No. 11325418 and MOST under Grant No. 2016YFA0301600.

[1] A. Montorsi, *The Hubbard Model: A Reprint Volume* (World Scientific, Singapore, 1992).
[2] A. Mazurenko, C. S. Chiu, G. Ji, M. F. Parsons, M. Kanász-Nagy, R. Schmidt, F. Grusdt, E. Demler, D. Greif, and M. Greiner, Nature (London) **545**, 462 (2017).
[3] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2007).
[4] L. Wang, Phys. Rev. B **94**, 195105 (2016).
[5] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).
[6] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, Nat. Phys. **13**, 435 (2017).
[7] G. Torlai and R. G. Melko, Phys. Rev. B **94**, 165134 (2016).
[8] S. Wetzel, Phys. Rev. E **96**, 022140 (2017).
[9] P. Ponte and R. G. Melko, arXiv:1704.05848.
[10] W. J. Hu, R. Singh, and R. Scalettar, Phys. Rev. E **95**, 062122 (2017).
[11] S. Wetzel and M. Scherzer, arXiv:1705.05582.
[12] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Phys. Rev. X **7**, 031038 (2017).
[13] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Sci. Rep. **7**, 8823 (2017).
[14] J. Liu, H. Shen, Y. Qi, Z. Y. Meng, and L. Fu, Phys. Rev. B **95**, 241104(R) (2017).
[15] L. Huang and L. Wang, Phys. Rev. B **95**, 035105 (2017).
[16] L. Huang, Y.-f. Yang, and L. Wang, Phys. Rev. E **95**, 031301(R) (2017).
[17] X. Y. Xu, Y. Qi, J. Liu, L. Fu, and Z. Y. Meng, Phys. Rev. B **96**, 041119(R) (2017).
[18] L. Wang, arXiv:1702.08586.
[19] Y. Nagai, H. Shen, Y. Qi, J. Liu, and L. Fu, Phys. Rev. B **96**, 161102(R) (2017).
[20] J. Villain, J. Phys. C **10**, 1717 (1977).
[21] J. Villain, J. Phys. C **10**, 4793 (1977).
[22] D. H. Lee, J. D. Joannopoulos, J. W. Negele, and D. P. Landau, Phys. Rev. Lett. **52**, 433 (1984).
[23] S. Miyashita and H. Shiba, J. Phys. Soc. Jpn. **53**, 1145 (1984).
[24] S. Lee and K.-C. Lee, Phys. Rev. B **49**, 15184 (1994).
[25] S. E. Korshunov, Phys. Rev. Lett. **88**, 167007 (2002).
[26] M. Hasenbusch, A. Pelissetto, and E. Vicari, J. Stat. Mech. (2005) P12002.
[27] T. Obuchi and H. Kawamura, J. Phys. Soc. Jpn. **81**, 054003 (2012).
[28] J.-P. Lv, T. M. Garoni, and Y. J. Deng, Phys. Rev. B **87**, 024108 (2013).
[29] P. Olsson, Phys. Rev. B **52**, 4526 (1995).
[30] C. Wang and H. Zhai (unpublished).
[31] T. Ohta and D. Jasnow, Phys. Rev. B **20**, 139 (1979).
[32] H. Weber and P. Minnhagen, Phys. Rev. B **37**, 5986(R) (1988).