

Machine learning of the XY model on a spherical Fibonacci latticeChen-Hui Song^{1,2}, Qu-Cheng Gao,¹ Xu-Yang Hou,¹ Xin Wang,¹ Zheng Zhou,¹
Yan He³, Hao Guo^{1,*} and Chih-Chun Chien⁴¹*Department of Physics, Southeast University, Jiulonghu Campus, Nanjing 211189, China*²*Tsinghua Shenzhen International Graduate School, Tsinghua University, Shenzhen 518055, China*³*College of physics, Sichuan University, Chengdu, Sichuan 610064, China*⁴*Department of physics, University of California, Merced, California 95343, USA*

(Received 15 December 2021; accepted 9 March 2022; published 1 April 2022)

The XY model on a spherical surface, inspired by recently realized atomic gases trapped in a spherical shell, is analyzed here. Instead of a traditional latitude-longitude lattice, we introduce the much more homogeneous Fibonacci lattice and use classical Monte Carlo (MC) simulations to generate spin configurations. The results clearly show that topological defects, in the form of vortices, must exist in the stable configuration on a sphere but vanish in a plane due to a mathematical theorem. Using these spin configurations as training samples, a graph-convolutional network-based method is implemented to recognize different phases and successfully predict the phase-transition temperature. We also apply the density-based spatial clustering of applications with noise, a powerful machine-learning algorithm, to monitor the path of two vortices with topological charges on the sphere during MC simulations. Our results provide reliable predictions for quantum simulators using polaritons or future space-based experiments on ultracold atoms in microgravity.

DOI: [10.1103/PhysRevResearch.4.023005](https://doi.org/10.1103/PhysRevResearch.4.023005)**I. INTRODUCTION**

The XY model is a paradigm in statistical physics with important implications for condensed matter systems [1–7]. The two-dimensional (2D) XY model has inspired research on topological defects and unconventional phase transitions [8–10]. During the last half century, there have been many studies on various aspects of XY model [2,3,5–7,11–19]. Experimentally, it can be emulated by ultracold atoms in optical lattices [20] or in polariton systems [21]. Recently, the fast development of space-based technique has stimulated experimental efforts to confine ultracold atoms on a spherical surface in microgravity [22–30], making a systematic analysis of the XY models on a spherical lattice an intriguing quest.

In contrast with the planar case, however, there is no arbitrarily large and exactly uniform lattice on a sphere, which prevents a direct analytical study on such models. Nevertheless, one may gain insights by the aid of powerful tools in processing big data, such as machine learning. Recently, this approach has achieved tremendous and unexpected progresses in physics [31–45]. Even the simplest three-layer fully connected neutral network (FCN) can well recognize different phases of the Ising model [33]. For topological phase transitions, the FCN does not work equally well since it cannot

effectively capture the local spatial information. An alternative tool is the convolutional neural network (CNN), which has proved its power in many situations [18,46]. However, CNN is designed for image recognition [47] and is not applicable to lattices with no resemblance to a picturelike structure, including the lattices on a sphere and irregular lattices. In the following, a different tool is introduced to handle those irregular lattices.

Here we focus on the spherical XY model. The traditional latitude-longitude lattice is far from uniform, so we first introduce the Fibonacci lattice, which is basically the most uniform lattice on a spherical surface. With the help of Monte Carlo (MC) simulations, we find various spin configurations of the spherical XY model. Multiple vortices onset even at ultralow temperatures, and the net topological charge is always 2 due to the topology of the 2D sphere S^2 . We further introduce the powerful graph-convolutional network (GCN) method to overcome the disadvantages of the traditional CNN. Using the samples of the spin configurations from MC simulations, the GCN can effectively classify different phases and successfully predict the Berezinskii-Kosterlitz-Thouless (BKT) phase-transition temperature. The BKT transition [8,48] is not a typical phase transition that has different values of the order parameter across the transition. Moreover, it is impractical to determine the BKT transition by counting the vortices [49]. Therefore, determining the critical BKT temperature is an important but daunting task, and our work offers a solution to this challenge. Moreover, the results also confirm topological properties on 2D plane and sphere. While both Fibonacci lattice and GCN are promising tools for investigating physics on complex geometries, they have not been broadly introduced in condensed-matter systems. The MC annealing of a quantum

*guohao.ph@seu.edu.cn.

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

system is equivalent to the spontaneous relaxation dynamics in some sense. To capture the dynamics of topological defects, we introduce the density-based spatial clustering of applications with noise (DBSCAN), which shows the merging of vortices with different topological charges during MC annealing. Our methods offer useful tools and insights for spherically confined ultracold atoms or polaritons in engineered structures.

The rest of the paper is organized as follows. In Sec. II, we give an introduction to the spherical Fibonacci lattice and discuss its properties. Section III presents spin configurations of the spherical XY model by using MC simulations. We then use these results as the training samples for the GCN and predict the BKT transition temperature in Sec. IV. Next, we discuss the vortex dynamics during MC annealing in Sec. V. Finally, Sec. VI concludes our work. Some details are summarized in the Appendix.

II. THE SPHERICAL XY MODEL

The 2D classical XY model on a square lattice is described by

$$H = -J \sum_{(i,j)} \mathbf{s}_i \cdot \mathbf{s}_j = -J \sum_{(i,j)} \cos(\theta_i - \theta_j), \quad (1)$$

where J is the coupling constant, \mathbf{s}_i denotes the spin with angle θ_i at site i , and the sum is taken over all pairs of nearest-neighbor spins. At zero temperature, the system stays at the ground state in which all spins are aligned in the same direction. There are also excitations with nontrivial topology, such as vortices and anti-vortices, which are bounded below the BKT transition temperature T_c but proliferate above it. An interesting question is: What happens when all spins are confined on a topologically nontrivial surface like S^2 ?

A map of the planar XY model onto a spherical one is nontrivial since the 2D plane is topologically different from a spherical surface. On a square lattice, the ordinary XY model described by Eq. (1) is isotropic and the lattice site is evenly distributed. However, a spherical surface lacks a completely isotropic lattice of arbitrarily large size, which prevents a direct analytical study of the XY model on a sphere. The traditional latitude-longitude lattice is highly inhomogeneous with the lattice sites near the north or south pole much denser than anywhere else. If a spherical lattice is totally isotropic, its sites must be located at the vertices of a regular polyhedron. However, there are only five different types of regular polyhedrons. The regular dodecahedron has the largest number of vertices at 20, which is still too small for a systematic study of the XY model toward the thermodynamic limit. For a lattice with a large number of sites, we have to seek an alternative allowing the area occupied by each site to be almost identical. Fortunately, there exists such a spherical lattice, called the Fibonacci lattice [50], where the i -th site on a sphere of radius R is defined by

$$\begin{aligned} x_i &= \sqrt{R^2 - z_i^2} \cos(2i\pi\phi), & y_i &= \sqrt{R^2 - z_i^2} \sin(2i\pi\phi), \\ z_i &= R \left(\frac{2i-1}{N} - 1 \right), \end{aligned} \quad (2)$$

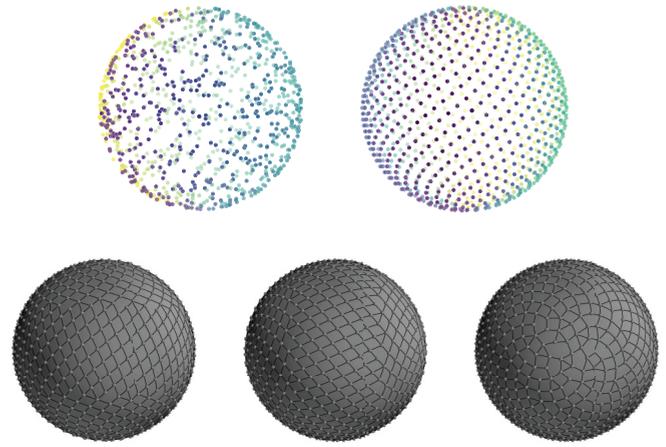


FIG. 1. (Top) Random lattice (left) and Fibonacci lattice (right) on a spherical surface. (Bottom) Perspectives of a $N = 1000$ Fibonacci lattice from different directions.

for $i = 1, 2, \dots, N$. Here N is the total number of sites, and $\phi = \frac{\sqrt{5}-1}{2}$ is the golden ratio. Figure 1 shows the comparison between a random and a Fibonacci lattices on a sphere in the top panels, both of which have $N = 1000$ sites. For each site, the points within its neighborhood of radius $r_c = 0.1298R$ are referred to as its neighbors. Apparently, the site distribution of the latter is much more uniform.

Although the site density of the Fibonacci lattices is quite uniform, this does not mean it is locally identical a 2D square lattice. To show the difference, we connect all pairs of neighbors of the Fibonacci lattice in the top panel of Fig. 1 and present the result in the bottom panels. Obviously, it is not a uniform lattice as some sites have three or five neighbors. The traditional CNN is for translationally invariant systems and is designed for images. Thus it can be applied to regular lattice structures that has a picturelike format, which is illustrated in Appendix A. Since the Fibonacci lattice has neither translational invariance nor picturelike structure, the CNN is not applicable.

To demonstrate the uniformity, we consider a unit sphere sprinkled with the above-mentioned Fibonacci lattice ($N = 1000$ and $r_c/R = 0.1298$). There are 850 sites with four neighbors, 76 sites with three neighbors, and 74 sites with five neighbors. Altogether, there are $850 \times 4 + 76 \times 3 + 74 \times 5 = 3998$ neighbors, and the average distance between neighboring sites and its standard deviation are $0.11455R$ and $0.00816R$, respectively. Basically, this is the best approximation on a sphere of a 2D square lattice, where each site has exactly four nearest neighbors.

III. SPIN CONFIGURATIONS

When mapping the XY model onto a sphere, we introduce a cutoff range of the interaction, such that only spins within r_c of each other interact. The Hamiltonian is further approximated by

$$H = -J \sum_{\langle i,j \rangle} e^{-\alpha r_{ij}^2} \mathbf{s}_i \cdot \mathbf{s}_j, \quad (3)$$

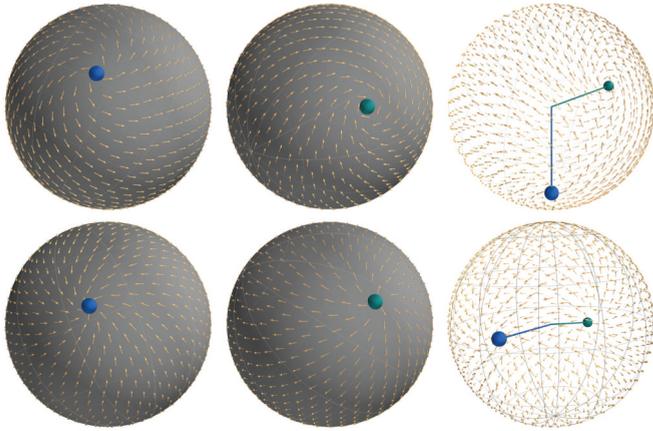


FIG. 2. MC simulations of the XY model on a spherical Fibonacci lattice after a complete annealing. Top row: (Left) A vortex on one side. (Middle) Another vortex on the opposite side. (Right) Perspective drawing. Here $r_c/R = 0.1298$. The angle between the two vortices with respect to the origin is 131.5° . Lower row: (Left) A vortex on one side. (Middle) Another vortex on the opposite side. (Right) Perspective drawing. Here $r_c/R = 0.4$. The angle between the two vortices is 174.3° .

where a Gaussian-type interaction is implemented since the separation between different pairs of sites is not a constant.

We set $N = 1000$, $\alpha = 10.0/R^2$, $T/J = 5.0 \times 10^{-4}$ and $r_c/R = 0.1298$ and use standard MC techniques to obtain samples of spin configurations. The spins are randomly oriented initially and then evolve under the MC annealing. Multiple vortices appear during the annealing, and some pairs of vortices merge or annihilate as the simulation continues. We plot the stable spin configuration in Fig. 2, showing only two unpaired vortices survive. Moreover, the distance between the two vortices tends to increase with the interaction range r_c during the MC annealing. For $r_c/R = 0.1298$, the angle between the two vortices with respect to the origin is about 131.5° . For $r_c/R = 0.4$, the two vortices eventually reside on two nearly opposite sites (not necessarily the two poles) on the sphere, and the angle between them is 174.3° . The relation between the vortex-separation and r_c is because r_c indicates the local scale of order. As it increases, the vortex-pair size will increase correspondingly.

The topological charge of a vortex is $n = \frac{1}{2\pi} \oint \nabla\theta \cdot d\mathbf{l}$. The charges of all vortices in Fig. 2 are found to be $+1$, which explains why the vortices move away from each other because of the repulsion between them. Since all spins are confined on the tangent plane at each site, they actually belong to a tangent vector field X on S^2 . Thus the result agrees with the Poincaré-Hopf theorem [51] that

$$\sum_i \text{ind}_{x_i}(X) = \chi(S^2). \tag{4}$$

Here $x_i \in S^2$ denotes the zeros of the vector field X , $\text{ind}_{x_i}(X)$ is the index of X at x_i , and $\chi(S^2)$ is the Euler characteristic and equal to 2 for S^2 . A zero at x_i actually corresponds to the center of a vortex. Thus the theorem states that the net topological charge of a spherical XY model is always 2, independent of temperature. Hence the vortices must exist on a sphere even at



FIG. 3. MC simulations of the XY model on a spherical Fibonacci lattice after a complete annealing. Here $N = 3000$, $T/J = 5.0 \times 10^{-4}$, and $r_c/R = 0.081$. The spin configuration is presented from three different directions to show all vortices.

zero temperature, essentially different from the planar square-lattice XY mode without any vortex in its stable ground state. Our simulations on the Fibonacci lattice also exhibit other types of stable spin configurations as the parameters change. For example, the case with $N = 3000$, $T/J = 5.0 \times 10^{-4}$ and $r_c/R = 0.081$ shows a final stable configuration with eight vortices, five of which have charge $+1$ while the other three have charge -1 , which is shown in Fig. 3. This may be understood as follows. The $N = 3000$ lattice becomes much denser than the $N = 1000$ case. Thus the effective distance between the vortices, or the sites on the shortest path between them, becomes larger and the interactions felt by the vortices become relatively smaller, which may prevent the merging of vortices as the annealing goes on. Moreover, these results agree with those on a spherical random lattice [52,53].

IV. PHASE-TRANSITION TEMPERATURE

To determine the phase-transition temperature, a powerful tool is the recently developed paradigm, the machine-learning method, which, exemplified by the FCN and CNN, has become a powerful tool for determining phase-transition temperatures in physical systems [31–35,39,42,43]. Even the simplest three-layer FCN can well recognize different phases of the Ising model [33]. However, for topological phase transitions, FCN does not work equally well since it cannot effectively capture the local spatial information. An alternative tool is the CNN, which relies on the convolutional kernels to extract local features from the input, which usually requires the information distribution to have spatial translation invariance. The Fibonacci lattice, however, is only approximately uniform.

Here we propose a more versatile method based on the GCN [54]. The GCN can effectively aggregate the local spatial information, thereby achieving a classification of vertices and graphs. Figure 4 shows a flow chart to classify different phases of the spherical XY model by the GCN. Here we give a brief introduction of the procedure. For a $N = 1000$ spherical Fibonacci lattice, we set $r_c = 0.1298R$, so most sites have four neighbors. We then map the lattice to a graph \mathcal{G} , and all information of \mathcal{G} is stored in the degree matrix \mathbf{D} and adjacency matrix \mathcal{A} , as shown in Fig. 4(a). The convolution is performed by using the Laplacian matrix $\mathbf{L} = \mathbf{D} - \mathcal{A}$. The input features are specified by the feature matrix $\mathbf{X} = (s_1^T, \dots, s_N^T)^T \in \mathbb{R}^{N \times 3}$, where $s_i = (x_i, y_i, z_i)$ is the spin at site i . Next, we apply the random-walk normalized Laplacian

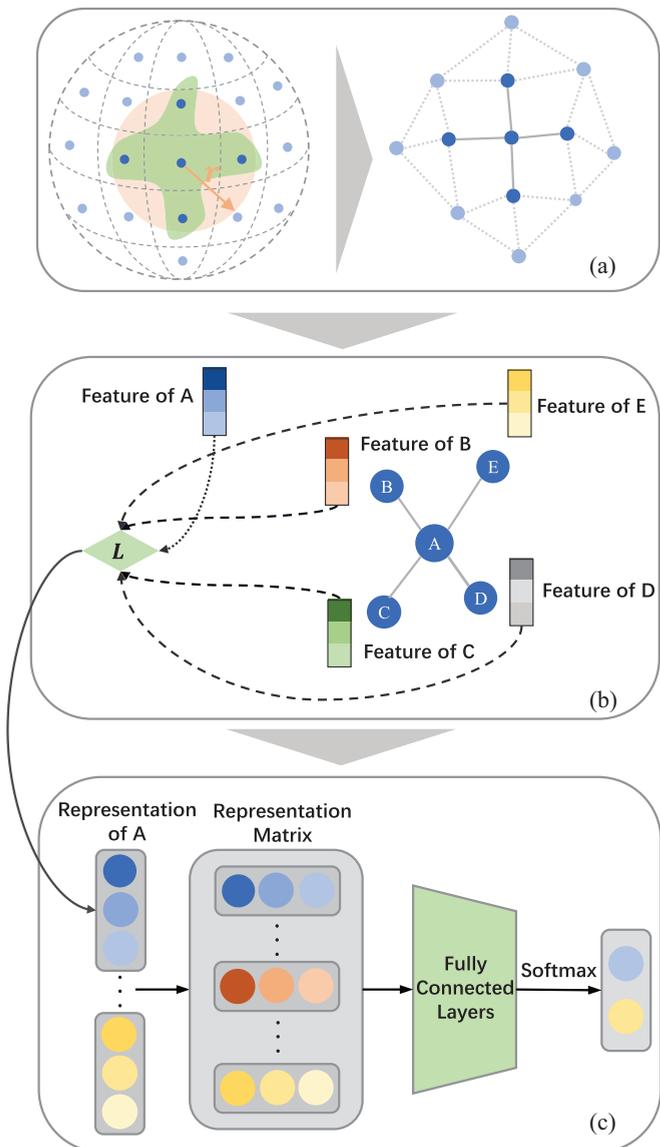


FIG. 4. Flow charts for phase classification of the XY model on a sphere based on the GCN. Panel (a) shows how to extract a graph structure from a spherical Fibonacci lattice. Panel (b) indicates that the input features are aggregated by the Laplacian matrix. Panel (c) shows that the classification probabilities are finally obtained by combining a softmax layer and the fully connected layer.

$L^m = D^{-1}L$ to aggregate the features [see Fig. 4(b)]:

$$H = \text{ReLu}(L^m X W_h + b_h), \quad (5)$$

where ReLu is the activation function, $W_h \in \mathbb{R}^{3 \times 1}$ is the weight, and $b_h \in \mathbb{R}^{N \times 1}$ is the bias. Finally, by using the fully connected layer and the *softmax* function to aggregate the hidden layer output $H \in \mathbb{R}^{N \times 1}$, we generate the classification confidences of the ordered phase (with probability p_o) and disordered phase (with probability p_d), as shown in Fig. 4(c). More details can be found in Appendix B. The criterion $p_o = p_d$ is adopted to identify the critical temperature T_c . Above T_c , the vortex number will proliferate but in a gradual fashion, as shown in Ref. [49] for the planar XY model, so a simple counting of the vortices is unreliable for determining T_c .

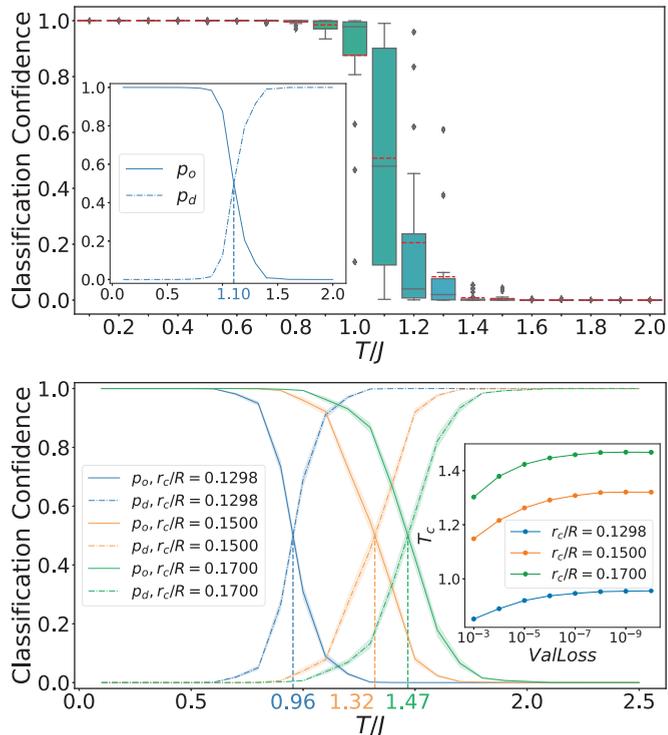


FIG. 5. (Top panel) Application of the GCN to the XY model on a 32×32 square lattice, where p_o is plotted vs temperature. Here the lower and top edges of each rectangle respectively label the lower and higher quartiles of the calculated p_o , and the black-solid and red-dashed lines denote the corresponding median and average values. The inset shows how to determine the critical temperature. (Bottom panel) Critical temperature of the XY model on a spherical Fibonacci lattice. The blue-dashed, brown dot-dashed, and green solid curves correspond to $r_c/R = 0.1298, 0.15, \text{ and } 0.17$, respectively; p_o and p_d are obtained from the average values. Here the inset shows the change in T_c with respect to loss function values of GCN on validation sets, which indicates the convergence of our estimation of T_c .

Using the results from MC simulations as the test sets for the GCN after training, we can identify phases and phase transitions. The top panel of Figure 5 shows our test of the GCN by a 32×32 square-lattice XY model. The estimation of T_c ($= 1.10J$) is very close to the result $1.08J$ from the renormalization group [55]. The bottom panel of Fig. 5 shows the GCN prediction of the spherical XY model with $N = 1000$ sites. The classification confidence is plotted as a function of temperature for $r_c/R = 0.1298, 0.15, \text{ and } 0.17$. The critical temperatures are determined by locating where $p_o = p_d$, which are $T_c/J = 0.96, 1.32, \text{ and } 1.47$, respectively. The translucent region of each curve indicates the 95% confidence intervals. To ensure the numerical accuracy, we also show the change in T_c with respect to the loss function values of the GCN on the validation sets in the inset of the bottom panel. The trends clearly indicate the convergence of our estimations. Moreover, our results show that T_c increases with r_c . Similar inferences also hold for planar-lattice XY models. A larger r_c means more neighbors are involved in the interactions. Thus more energetic thermal fluctuations are required to unbind vortices, resulting in higher T_c . On a 2D

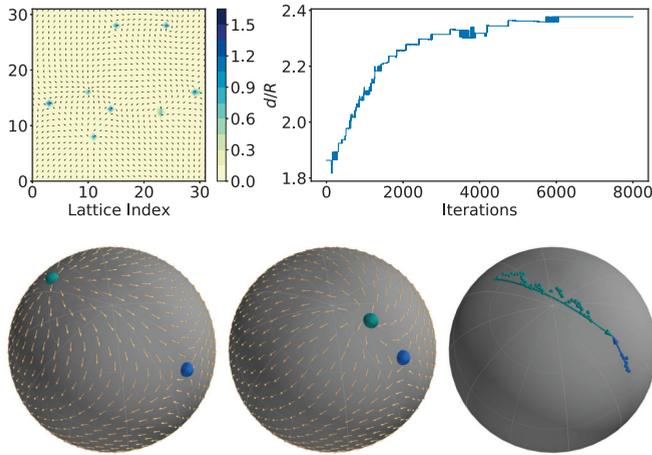


FIG. 6. (Top row) Tracking of vortices of a 2D square XY model by the GCN (left), and the distance between two vortices as a function of MC steps for a spherical XY model by the DBSCAN (right), in which one iteration is equivalent to 10,000 steps. (Bottom row) Snapshots of two annihilating vortices (left and middle) and their trajectories (right).

plane, T_c of the XY model on a triangular lattice with six neighbors [56] is 226% higher than that on a square lattice. For more details, please refer to Appendix B. The spherical Fibonacci lattice thus demonstrates a continuous tuning of the number of neighbors and its influence on T_c .

V. VORTEX DYNAMICS

We also investigate the motion of vortices on the sphere during a MC annealing. The number of MC steps is related to the real time via a certain function $f: n = f(t)$ [57]. Hence the MC simulation, which brings a system to equilibrium from a nonequilibrium starting point, is equivalent to the spontaneous relaxation dynamics governed by the Langevin equation [57]. Physically, the trajectory describes the time-dependent evolution of the vortex on a sphere while $n = f(t)$ can be thought of as a reparametrization of the evolution path. In Ref. [27], vortex dynamics of a Bose-Einstein condensate on a spherical shell is discussed from the view of an effective theory. To trace out the motion of vortices, we need to locate the vortices. Interestingly, we found that the hidden layer \mathbf{H} of the GCN, given by Eq. (5), quantifies the local disorder at site i since H_i is negligible at the sites where there is no vortex but finite at the sites where the vortices reside (details can be found in Appendix C). This may come from the fact that \mathbf{H} carries the feature information of the input data. Specifically, we set a threshold $\Delta = \frac{\gamma}{N} \sum_i H_i$, where γ is an adjustable parameter. By filtering out the coordinates with $H_i > \Delta$ at every 10,000 steps, the vortices can be located. We then use a well-known machine-learning method, the DBSCAN [58], to group these coordinates by the vortices that they belong to, which finally gives the dynamical picture of the vortices on a sphere.

As a test of this algorithm, we first use it to track vortices of a 32×32 square-lattice XY model and present the result in Fig. 6 (top-left panel). Evidently it works very well. We then apply it to the spherical XY model with $N = 1000$, $r_c/R = 0.225$ at $T/J = 5.0 \times 10^{-4}$ and visualize the results

in Fig. 6. Initially, there are multiple vortices, as discussed before. Eventually, only two of them survive in the final state, as others merge during the MC annealing. In this situation, the system shows a very clear merging path. In the top row (right panel), we present the distance between the two remaining vortices as a function of the MC step, which indicates that the positions of vortices become stable. Here the distance is evaluated along the great circle (geodesic curve) connecting the two vortices on the sphere. In the bottom row, we present the merging process of one pair of vortices with different topological charges during the MC annealing. The left panel shows an early vortex configuration, in which the two vortices are far from each other, and the middle panel shows an intermediate state, in which the two vortices move close to each other. The right panel presents their merging path. The whole figure tells that during the MC annealing, some vortices merge, and the remaining vortices of the same charge tend to stay away from each other due to the repulsion, which is similar to the behaviors of two equally charged particles moving on a sphere.

VI. CONCLUSION

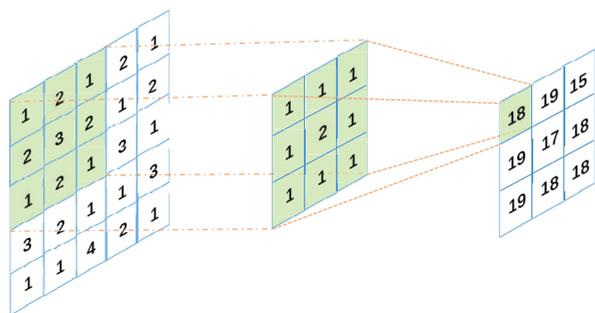
We have presented distinct features of the XY model on a spherical surface, inspired by cold-atom experiments in microgravity environments, such as the International Space Station. Using the Fibonacci lattice, MC simulations and machine-learning techniques like the GCN, our results predict the vortex distributions and phase-transition temperatures. Moreover, the DBSCAN clustering method allows us to visualize the evolution paths of the vortices during a MC annealing. The Fibonacci lattice, with its uniformity and tunable neighborhoods, can be generalized to other models on a sphere. Other than irregular lattices, the GCN has potential applications for multiparameter systems by using aggregation methods beyond the Laplacian matrix. The methods and analyses shed light on future experiments of engineered systems, including cold atoms and polaritons, with spherical geometries.

ACKNOWLEDGMENTS

H.G. was supported by the National Natural Science Foundation of China (Grant No. 12074064). C.C.C. was supported by the National Science Foundation under Grant No. PHY-2011360. C.H.S. acknowledges Shenzhen International Graduate School, Tsinghua University for providing computational resource and technical support that make a great contribution to this work.

APPENDIX A: A COMPARISON BETWEEN CNN AND GCN

It is crucial to introduce the GCN in this work since the traditional CNN cannot be applied to the irregular spherical lattices. The CNN was first implemented [47] in the field of computer vision. Since it can effectively capture the local features of images and enhance the image-sensing ability of computers, the CNN has been widely applied during the last thirty years. The spin configuration in the square-lattice XY model has explicit local order, which can be captured by the



Output[0][0]=1 × 1 + 2 × 1+1×1+2×1+3×2+2× 1+1×1+2×1+1×1=18

FIG. 7. Diagrammatic sketch of the principle of the convolution of pictures.

CNN. Hence the CNN has been applied to study the Ising model, XY model, and others in physics [18,33,59,60]. In the main text, we have cited some references to show the high-precision estimation of the critical temperature of the Ising model by the CNN.

However, the tradition CNN is designed for images. The convolution kernel (or filter) is like a sliding window. It slides back and forth with a given step length when placed over an input image, as indicated by Fig. 7. The characteristic pixel (destination pixel) will be obtained by performing the convolution on the source pixels, and accordingly the local features of the input (source) image can be extracted in this way (see Fig. 7). In other words, the CNN can only be applied to picturelike structures. In related topics of physics, only the regular lattices correspond to this kind of structure, since the lattice site can be thought of as a pixel and the spin-value (including the direction and angle) is accordingly the pixel values. This is why most machine-learning papers in this field are focused on the square lattice, a simple regular lattice. There are, however, many lattices that do not have picturelike structures, especially those without translational invariance. For example, in the studies of frustrated Ising models, researchers may face some complicated lattices like the Cairo pentagonal lattice [61] or the 3D garnet lattice [62], as shown in Fig. 8. Due to the complexity of these lattices, the traditional CNN does not work in these situations. Thus a new framework beyond the CNN has to be introduced. For lattices with translational invariance, such as the triangular and honeycomb lattices, it might be possible to change the convolution kernel to perform a “modified” CNN. How-

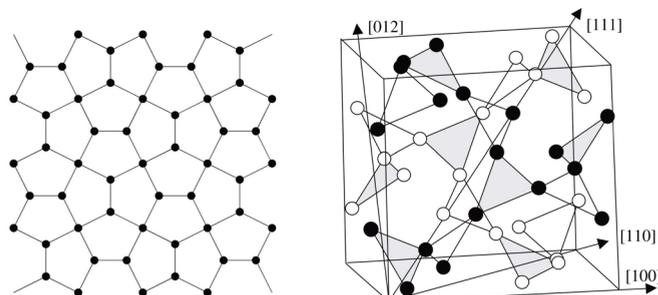


FIG. 8. (a) Cairo pentagonal lattice [61]. (b) Garnet lattice [62].

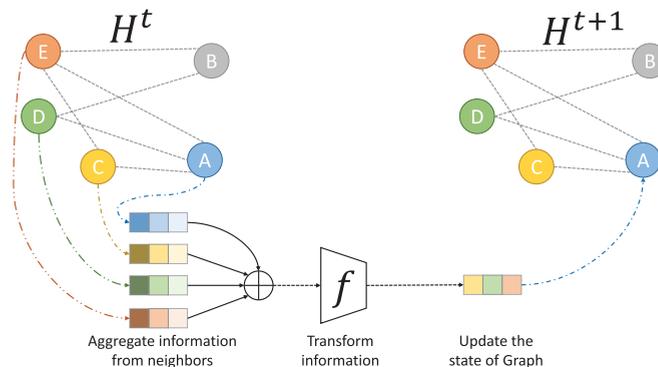


FIG. 9. A diagrammatic sketch of the GCN.

ever, as will be shown below, the GCN is the most versatile alternative.

The key of the convolution is to aggregate the information from nearby pixels (sites). Thus a natural idea to generalize the traditional CNN is to extend its ingredients in order to match arbitrary local topological structures. The GCN is proposed exactly under this guidance. We can map an arbitrary lattice into a graph, which not only preserves the neighbor structure but also breaks the limitation on information networks required by the traditional CNN. Thus GCN has a much broader range of applications. A diagrammatic sketch of the key idea of the GCN is shown in Fig. 9.

As an example, the spherical lattice does not have a picturelike structure due to its local variations; thus the traditional CNN is not applicable. In our paper, we first set a suitable r_c to determine the neighbors of each lattice site. Thus the topological structure of the whole lattice can be simply represented by a graph. There are already many aggregation methods in mathematics to capture the information from graphs. Here we find that by using the simplest random-walk Laplacian matrix and the FCN, it is possible to give results that are promising for the spherical XY models.

We want to emphasize that the GCN provides a unified machine-learning framework and can be applied to any kind of lattice, even including the irregular lattices. For various lattice models in physics (such as the Ising model, the XY model, and the Heisenberg models), once their topological structures are determined by certain graphs, we only need to choose suitable methods to perform information aggregations. Thus the GCN is much more powerful and has broad generalizations. Since the CNN is a special case of the GCN as the regular lattices can also be expressed by simple graphs, those problems studied by the CNN can also be studied by the GCN, as shown in Fig. 10. In our paper, we apply the GCN to both square and spherical lattices and indeed obtain excellent results. For example, our estimation of the critical temperature for the square lattice agrees very well with other theoretical or numerical predictions.

The GCN is expected to stimulate wide research interests. We envision some of its future applications, including physics of irregular lattices, crystals with defects, and other geometrical or topological problems.

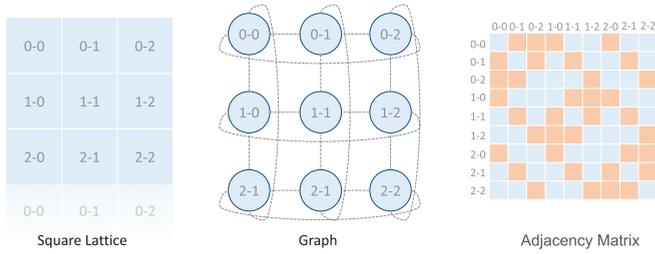


FIG. 10. The mapping from a 2D square lattice to a graph.

APPENDIX B: DETAILS OF GCN

1. Formulation

The GCN is a powerful machine-learning method which can effectively classify different phases of many-body systems confined on irregular lattices, including the spherical Fibonacci lattice.

Our goal is to distinguish the ordered phase at low temperatures and the disordered phase at high temperatures of the XY model on a spherical surface by using the GCN. Specifically, we are going to obtain a classification function $\hat{Y} = F(X; \Theta)$ by data training, which maps the feature matrix X containing all the spin information into the label space $\hat{Y} = (p_o, p_d)$. Here p_o and p_d , respectively, give the classification confidences of the ordered and disordered phases. The training process is to adjust the set Θ of parameters such that the classification result \hat{Y} is as close to the real label Y as possible.

2. Preliminaries

Graph is a powerful data structure for processing relational information, which can effectively realize the clustering of neighbor information. A graph \mathcal{G} comprises a set \mathcal{V} of vertices and a set \mathcal{E} of edges. For two vertices $v_i, v_j \in \mathcal{V}$, if there exists an edge $e_{ij} \in \mathcal{E}$, v_i and v_j are said to be connected by e_{ij} in the graph \mathcal{G} . The number of vertices, to which a vertex v_i is linked is defined as the degree d_i of v_i . Explicitly, \mathcal{G} can be represented by the adjacency matrix \mathcal{A} of which the entries are given by $\mathcal{A}_{ij} = 1$ if $e_{ij} \in \mathcal{E}$, or 0 otherwise. The feature of each vertex v_i can be expressed by a d -dimensional vector \mathbf{x}_i . Thus for all vertices, their features can be represented by a matrix $X \in \mathbb{R}^{n \times d}$, where n is the degree of \mathcal{G} .

The GCN is a generalization of the traditional CNN, which can be used to classify vertices and graphs. There are many ways to realize the GCN, which may be mainly classified into spatial-based and spectral-based convolutions. Both can be realized with the help of the Laplacian matrix L , defined by $L = D - \mathcal{A}$, where \mathcal{A} is the aforementioned adjacency matrix, and D is a diagonal matrix called the degree matrix, of which the diagonal elements are the degrees of the corresponding vertices. The definition of the Laplacian matrix is not unique, and the frequently used definitions include the symmetrically normalized Laplacian $L^{\text{sym}} = D^{-\frac{1}{2}} L D^{-\frac{1}{2}}$ and the random-walk normalized Laplacian $L^{\text{rm}} = D^{-1} L$.

The spatial-based GCN can be expressed by

$$H^{l+1} = \sigma(LH^l W^l), \tag{B1}$$

where W is the set of learnable parameters, H^l is the feature of the l -th layer (note that H^0 represents the input feature), and $\sigma(\cdot)$ is the activation function. In this method, the feature of each vertex is obtained by aggregating the information of its neighbors.

The spectral-based GCN is realized by first taking the spectral decomposition of the Laplacian matrix (i.e., the Fourier transformation) and then performing the associated inverse transformation. It can be expressed by

$$H^{l+1} = \sigma(U g_\theta(\Lambda) U^T H^l), \tag{B2}$$

where U is the matrix comprised by the eigenvectors of L , Λ is the diagonal matrix comprised by the eigenvalues of L , and the convolutional kernel $g_\theta(\Lambda)$ represents the set of learnable parameters. A widely adopted choice of g_θ is the Chebyshev polynomial, which can reduce the number of steps of eigen-decomposition and is thus more suitable for large-scale graph networks.

3. Methodology of phase classification

Here we illustrate how to apply the GCN to recognize different phases of the XY model on a sphere, schematically illustrated by the flow charts in Fig. 4. The convolutional kernel of the traditional CNN can capture the local spatial information, which corresponds to the short-range order of the XY models on a square lattice at low temperatures. This is why the traditional CNN works well in those situations. Note that the key point in phase recognition is to effectively aggregate the local spatial information. For the XY model, the first step is to establish the neighbor relation of each spin, which is guaranteed by the homogeneity of the square lattice. However, no such homogeneous lattice exists on a spherical surface, and the Fibonacci lattice is only an approximate uniform lattice without translation invariance. Thus the traditional CNN does not apply here. Fortunately, the GCN provides an ideal structure to characterize the neighbor relation of irregular lattices.

For a spherical Fibonacci lattice, all the sites naturally comprise a set of vertices \mathcal{V} . Moreover, by choosing a critical radius r_c such that $e_{ij} \in \mathcal{E}$ if $\|\mathbf{r}_i - \mathbf{r}_j\| < r_c$, a graph \mathcal{G} can be constructed. To better capture the local features, the value of r_c must be carefully selected, such that each site is only linked to the nearby sites. For example, when $N = 1000$, the (almost) best choice is $r_c/R = 0.1298$, shown in Fig. 4(a). For the XY model on a sphere, the spin at site i is denoted by a unit vector $\mathbf{s}_i = (x_i, y_i, z_i)$ with $i = 1, 2, \dots, N$. Thus the feature matrix is given by $X = (\mathbf{s}_1^T, \dots, \mathbf{s}_N^T)^T \in \mathbb{R}^{N \times 3}$. In this work, we use the random-walk normalized Laplacian $L^{\text{rm}} = D^{-1} L$ to aggregate the features and apply the ReLu function as the activation function, as shown in Fig. 4(b). Explicitly,

$$H = \text{ReLu}(L^{\text{rm}} X W_h + \mathbf{b}_h), \tag{B3}$$

where the weight $W_h \in \mathbb{R}^{3 \times 1}$ and the bias $\mathbf{b}_h \in \mathbb{R}^{N \times 1}$ are both learnable parameters, and $H \in \mathbb{R}^{N \times 1}$ is the feature representation. We further use the fully connected layer and the softmax function to aggregate the representation H , as demonstrated in Fig. 4(c), which can also be expressed by

$$\hat{Y} = \text{softmax}(H^T W_p + \mathbf{b}_p). \tag{B4}$$

Here $\mathbf{W}_p \in \mathbb{R}^{N \times 2}$ and $\mathbf{b}_p \in \mathbb{R}^2$ are both learnable parameters, and the output $\hat{\mathbf{Y}} \in \mathbb{R}^2$ gives the classification confidences that the input features, respectively, correspond to the ordered or disordered phase.

4. Training and prediction

a. General settings

The trial data come from MC simulations using the common Metropolis-Hasting algorithm and local spin update. As an example, we generate a $N = 1000$ Fibonacci lattice on a spherical surface of radius R . On the tangent plane of each lattice site, a randomly oriented spin is assigned initially. When mapping the lattice into a graph, we choose $r_c/R = 0.1298$ and use the cross entropy function as the loss function.

b. Data settings

At first, we give a very rough estimation of the critical temperature T_c by several runs of MC simulations, which show $T_c/J \sim 1.0$. Then, at temperatures far below T_c , we perform MC simulations over the temperature range $[0.001, 0.005]J$ with an increment of $0.001J$ and get a data set of the spin distributions in the low-temperature ordered phase. We repeat the simulations over the same temperature range for 100 times and get 100 sampling sets. Similarly, at a temperature far above T_c , we perform MC simulations over the temperature range $[101.0, 105.0]J$ with an increment of $1.0J$ and get a data set of the spin distributions in the high-temperature disordered phase. We also repeat the simulations to get 100 sampling sets. Finally, we split them into the training set, the validation set, and the test set, which have, respectively, an 80, a 10, and a 10 sampling set for the ordered and disordered phases.

c. Critical temperature

A prediction of the critical temperature can be achieved if the GCN model is trained sufficiently. We sample the data in a temperature range, in which T_c is included, and input those features into the GCN model. The output gives the confidences that the input features correspond to the ordered (p_o) and disordered (p_d) phases. Obviously, $p_o + p_d = 1$. We extract the critical temperature as the location when $p_o = p_d$. In order to suppress the errors, we carry out multiple simulations and classifications with the same set of parameters and average the corresponding classification confidences. For example, we execute our algorithm in the range $[0.1, 2.5]J$ with an increment of $0.1J$ and repeat the procedure 500 times to average the confidences for $r_c/R = 0.1298, 0.15, \text{ and } 0.17$, respectively.

d. Numerical results

When testing the GCN model with the samples from the test sets, it gives the correct classification results of $\geq 99.9\%$ classification confidences and 100% accuracy, even with only one convolutional layer.

5. Details of the results

We outline the numerical results in Figure 5. We first test our GCN model by using the data from the XY model on a

32×32 square lattice. We start with an arbitrary spin configuration and use MC simulations to give a spin sample; then we use the sample as the test set for the GCN and find an estimation of the classification confidences p_o and p_d . To obtain low-temperature data samples, we perform MC simulations over the temperature range $[0.001, 0.005]J$ with an increment of $0.001J$ and repeat the simulations for 100 times. For high-temperature data samples, we perform MC simulations over the temperature range $[101.0, 105.0]J$ with an increment of $1.0J$ and also repeat the simulations for 500 times. We then divide each of the 100 sampling sets into the training sets of 80 each, validation sets of 10 each, and test sets of 10 each. After the GCN model is trained with those samples, we apply it to the temperature range $[0.1, 2.0]J$ with an increment of $0.1J$ and repeat the procedure 500 times at each temperature to give 500 sets of raw data. In the top panel of Figure 5, the lower and upper edges of each rectangle, respectively, label the lower and higher quartiles of the data set, and the black-solid and red-dashed lines denote the corresponding median and average values (we only show the estimation of p_o). We use the average values for the estimated p_o and p_d . The critical temperature is determined by $p_o = p_d$. The estimation of $T_c = 1.10J$ is very close to the result ($T_c = 1.08J$) by the renormalization group calculation [55]. There are also very few statistically abnormal data labeled by the black dots, which basically have no influence on the final result. Note here we consider a finite-size system. Strictly speaking, T_c is not the Kosterlitz-Thouless transition temperature T_{KT} in the thermodynamic limit, but the effective transition temperature at which the correlation length ξ is comparable to the system size ($L = 32$) [55]. T_c can be inferred from T_{KT} via $T_c \approx T_{KT} + \frac{\pi^2}{c(\ln L)^2}$ where c is a constant [55]. Above T_c , the vortex number will proliferate. Is it possible to determine T_c simply by counting the vortex number as the temperature varies? For the ordinary XY model, the answer is negative on a square lattice as shown in Ref. [49].

In the bottom panel, we apply the GCN to the XY model on a spherical Fibonacci lattice. Similarly, we repeat the procedure of the GCN for 100 times and then average those raw data. We plot the classification confidences vs temperature for $r_c/R = 0.1298, 0.15, \text{ and } 0.17$, and the corresponding critical temperatures are $T_c/J = 0.96, 1.32, \text{ and } 1.47$, respectively. To ensure the accuracy of our numerics, we also show the change in T_c with respect to the loss function values of GCN on validation sets in the inset of the bottom panel of Fig. 5. The trends clearly indicate the convergence of our estimations. The increase of T_c with r_c is also reasonable since a larger r_c means more neighbors are involved in the interactions. Therefore more energetic thermal fluctuations are needed to unbind the vortices, signifying higher T_c . We emphasize that similar inferences also hold for XY models on a planar lattice. For XY models on a large square lattice ($L \rightarrow \infty$), each site has four neighbors and $T_{KT}/J \approx 0.898$ [63]. While for XY models on a large triangular lattice, each site have six neighbors and $T_{KT}/J \approx 2.93$ [56]. (Ref. [56] actually found $\beta_{KT} \approx 0.683$ with $J = \frac{1}{2}$, implying $T_{KT}/J \approx 2.93$.) Hence they demonstrate the same influence of the number of neighbors on the transition temperature. The numerical results indicate the SGD optimizer works very well while others, like the Adam, do not produce accurate results.

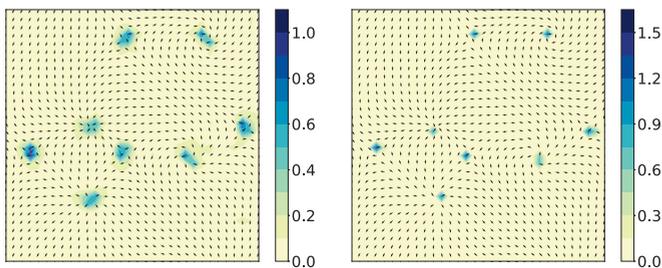


FIG. 11. (Left) Detection of vortices by using the “local disorder” ξ_i in the XY model on a 2D square lattice, where $r_c/a = 1.0$, and $T/J = 0.005$. (Right) Detection of vortices by using the element H_i of the hidden layer \mathbf{H} in the same XY model.

APPENDIX C: DETAILS OF THE VORTEX DYNAMICS

As pointed out in the main text, the dynamics of the vortices on a spherical surface can be traced by a well-known machine-learning method, the DBSCAN [58]. We present the details of the method here. Previous discussions have shown that the GCN has superior performance on recognizing the ordered and disordered phases even with only one convolutional layer. In the main text, we pointed out that the hidden layer \mathbf{H} , which is given by Eq. (5), can actually be used as an indicator of the local vortices. Here we present a qualitative explanation by a direct comparison.

For a one-convolutional-layer GCN, the Laplacian matrix realizes the feature aggregation of the neighbors for each site, i.e.,

$$\mathbf{R} = \mathbf{L}^m \mathbf{X} = (\mathbf{r}_1^T, \dots, \mathbf{r}_N^T)^T. \quad (\text{C1})$$

Here $\mathbf{r}_i = \frac{1}{|\mathcal{E}_i|} \sum_{j \in \mathcal{E}_i} (\mathbf{s}_j - \mathbf{s}_i)$, \mathcal{E}_i is the set of neighbors of site i , and $|\mathcal{E}_i|$ is its cardinal number. Note the spins change slowly at the sites away from any vortex. Thus $\|\mathbf{r}_i\|$ is close to zero. On the contrary, the sites around a vortex have drastic changes of the spins, and $\|\mathbf{r}_i\|$ must be finite. Hence we can introduce a suitable “observable:” $\xi_i = \|\mathbf{r}_i\|$ as the “local disorder” at site i to probe the local order. Specifically, we set a threshold $\Delta = \gamma \|\bar{\mathbf{r}}\|$, where $\|\bar{\mathbf{r}}\| = \frac{1}{N} \sum_i \|\mathbf{r}_i\|$ is the average local disorder, and γ is an adjustable parameter. As a check of this method, we apply it to the XY model on a 32×32 square lattice and present the result in the left panel of Fig. 11. Evidently, all the vortices are accurately by their local disorder.

Interestingly, we find that the detection of vortices can also be directly probed by the hidden layer \mathbf{H} of the GCN instead of introducing the artificial “local disorder.” As we pointed out before, \mathbf{H} is also the feature representation of the input data, which is given by Eq. (5). The i -th element H_i of \mathbf{H} plays a similar role as the local disorder. Similarly, we can also introduce $\Delta = \frac{\gamma}{N} \sum_i H_i$ and detect the vortices in the same way. In the right panel of Fig. 11, we show the results of probing the vortices in the same XY model. A direct comparison shows that \mathbf{H} indeed captures all the vortices effectively. This also provides an example about how the GCN classifies different objects by learning local features.

We then apply the method to the XY model on a spherical Fibonacci lattice. After identifying the vortices, their evolution paths can further be traced out by a data-clustering algorithm. Specifically, we first filter out the coordinates such

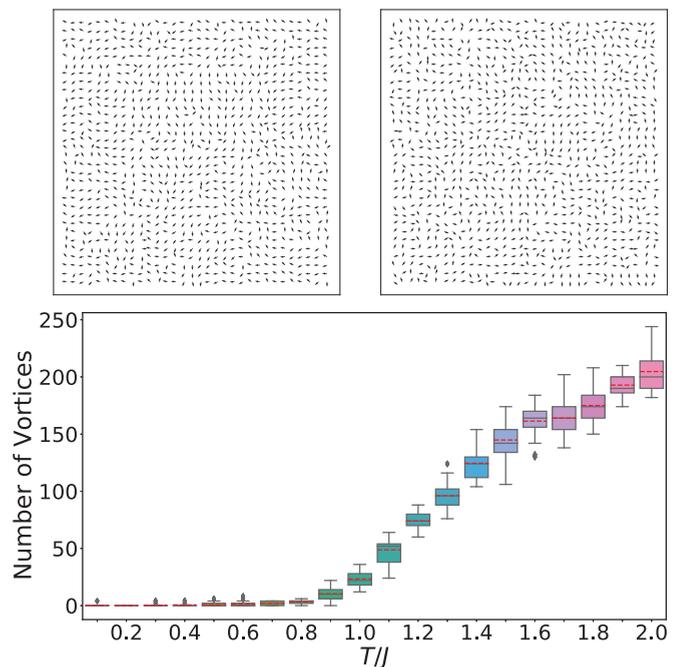


FIG. 12. (Top) Snapshots of the equilibrium spin configurations of a 2D 32×32 square-lattice XY model at $T = 1.0J$ (left) and $1.2J$ (right). Here $T_c = 1.10J$ in this model. (Bottom) The vortex number vs T/J of the same model.

that $\xi_i > \Delta$ at every 10000 steps. Those data should belong to the sites hosting the vortices. We label the set of those coordinates by P and use the DBSCAN to group those coordinates by the vortices:

$$\{P_1, \dots, P_k\} = \text{DBSCAN}(\text{Minpts}, r, P). \quad (\text{C2})$$

Here Minpts and r are two required parameters, k is the number of different vortices, and $P_i (i \in \{1, \dots, k\})$ is the classified coordinate set, each of which belongs to a single vortex. The position of a vortex is given by

$$\mathbf{x}_i^v = \underset{\|\mathbf{x}\|=R}{\text{argmin}} \sum_{\mathbf{x}^p \in P_i} \|\mathbf{x} - \mathbf{x}^p\|^2. \quad (\text{C3})$$

With these tools, we finally obtain the dynamics of the vortices on a spherical surface.

APPENDIX D: T_c AND VORTEX NUMBER

When the temperature is above the BKT transition temperature T_c , the vortex number of the XY models will proliferate. However, Ref. [49] shows that the vortex number increases smoothly for the standard XY model on a plane; therefore it is not possible to determine T_c by checking the vortex numbers. To confirm the result, we consider a 32×32 square-lattice XY model and run MC simulations to obtain equilibrium results. At high temperatures, the spin configuration is indeed very disordered, and we present an example in the top panels in Figure 12. The complex patterns make it difficult to find any significant difference between the spin configurations just below (left) and above (right) $T_c = 1.10J$ (the numerical estimation in Appendix B 5). In the bottom panel, we plot the vortex number vs T , and there is no abrupt change of

the vortex numbers, especially T_c . We mention that Ref. [49] also considered an extended square-lattice XY model with a modified potential

$$V_{ij} = 2J \left[1 - \left(\cos^2 \frac{\theta_i - \theta_j}{2} \right)^{p^2} \right], \quad J > 0. \quad (\text{D1})$$

When $p^2 = 1$, this model reduces to the ordinary 2D XY model. Only when p^2 is large enough (like $p^2 = 50$), the vortex number may increase abruptly at the critical temperature. According to these results, we expect that the method by plotting vortex number vs T may not work as an indicator to determine the BKT temperature of the standard XY model. Although it is hard to count the vortex number directly from the figure, we found that the GCN can still recognize those spin patterns very well.

-
- [1] E. Lieb, T. Schultz, and D. Mattis, Two soluble models of an antiferromagnetic chain, *Ann. Phys.* **16**, 407 (1961).
- [2] S. Katsura, T. Horiguchi, and M. Suzuki, Dynamical properties of the isotropic XY model, *Physica* **46**, 67 (1970).
- [3] L. J. de Jongh and A. R. Miedema, Experiments on simple magnetic model systems, *Adv. Phys.* **23**, 1 (1974).
- [4] D. D. Betts, Critical properties of the XY model, *Physica B+C* **86–88**, 556 (1977).
- [5] J. Tobochnik and G. V. Chester, Monte Carlo study of the planar spin model, *Phys. Rev. B* **20**, 3761 (1979).
- [6] P. Olsson, Effective vortex interaction in the two-dimensional XY model, *Phys. Rev. B* **46**, 14598 (1992).
- [7] L. Delfini, S. Lepri, and R. Liv, A simulation study of energy transport in the Hamiltonian XY model, *J. Stat. Mech.: Theory Exp.* (2005) P05006.
- [8] J. M. Kosterlitz and D. J. Thouless, Metastability and phase transitions in two-dimensional systems, *J. Phys. C: Solid State Phys.* **6**, 1181 (1973).
- [9] J. M. Kosterlitz, The critical properties of the two-dimensional XY model, *J. Phys. C: Solid State Phys.* **7**, 1046 (1974).
- [10] J. M. Kosterlitz, Kosterlitz–Thouless physics: A review of key issues, *Rep. Prog. Phys.* **79**, 026001 (2016).
- [11] R. Gupta and C. F. Baillie, Critical behavior of the two-dimensional XY model, *Phys. Rev. B* **45**, 2883 (1992).
- [12] P. Olsson, Monte Carlo analysis of the two-dimensional XY model. I. Self-consistent boundary conditions, *Phys. Rev. B* **52**, 4511 (1995).
- [13] P. Olsson, Monte Carlo analysis of the two-dimensional XY model. II. Comparison with the Kosterlitz renormalization-group equations, *Phys. Rev. B* **52**, 4526 (1995).
- [14] N. Nagaosa, *Quantum Field Theory in Condensed Matter Physics* (Springer and Beijing World Publishing Corporation, Beijing, 2010).
- [15] A. Altland and B. Simons, *Condensed Matter Field Theory* (Cambridge University Press and Beijing World Publishing Corporation, Beijing, 2015).
- [16] X. G. Wen, *Quantum Field Theory of Many-Body Systems* (Oxford University Press, Oxford, 2004).
- [17] A. Kushima, X. Lin, J. Li, J. Eapen, J. C. Mauro, X. F. Qian, P. Diep, and S. Yip, Computing the viscosity of supercooled liquids, *J. Chem. Phys.* **130**, 224504 (2009).
- [18] W. Zhang, J. Liu, and T.-C. Wei, Machine learning of phase transitions in the percolation and XY models, *Phys. Rev. E* **99**, 032142 (2019).
- [19] J. Wang, W. Zhang, T. Hua, and T.-C. Wei, Unsupervised learning of topological phase transitions using the Calinski-Harabaz index, *Phys. Rev. Research* **3**, 013074 (2021).
- [20] J. Struck, M. Weinberg, C. Ölschläger, P. Windpassinger, J. Simonet, K. Sengstock, R. Höppner, P. Hauke, A. Eckardt, L. M., and L. Mathey, Engineering Ising-XY spin-models in a triangular lattice using tunable artificial gauge fields, *Nat. Phys.* **9**, 738 (2013).
- [21] N. G. Berlo, M. Silva, K. Kalinin, A. Askitopoulos, J. D. Töpfer, P. Cilibrizzi, W. Langbein, and P. G. Lagoudakis, Realizing the classical XY Hamiltonian in polariton simulators, *Nat. Mater.* **16**, 1120 (2017).
- [22] N. S. Möller, F. E. A. dos Santos, V. S. Bagnato, and A. Pelster, Bose–Einstein condensation on curved manifolds, *New J. Phys.* **22**, 063059 (2020).
- [23] A. Tononi, F. Cinti, and L. Salasnich, Quantum Bubbles in Microgravity, *Phys. Rev. Lett.* **125**, 010402 (2020).
- [24] A. Tononi, A. Pelster, and L. Salasnich, Topological superfluid transition in bubble-trapped condensates (2021), *Phys. Rev. Research* **4**, 013122 (2022).
- [25] A. Tononi and L. Salasnich, Bose-Einstein Condensation on the Surface of a Sphere, *Phys. Rev. Lett.* **123**, 160403 (2019).
- [26] N. Lundblad, R. A. Carollo, C. Lannert, M. J. Gold, X. Jiang, D. Paseltiner, N. Sergay, and D. C. Aveline, Shell potentials for microgravity Bose-Einstein condensates, *npj Microgravity* **5**, 30 (2019).
- [27] S. J. Bereta, M. A. Caracanhas, and A. L. Fetter, Superfluid vortex dynamics on a spherical film, *Phys. Rev. A* **103**, 053306 (2021).
- [28] T. van Zoest, N. Gaaloul, Y. Singh, H. Ahlers, W. Herr, S. T. Seidel, W. Ertmer, E. Rasel, M. Eckart, and E. Kajari *et al.*, Space-borne Bose-Einstein condensation in microgravity, *Science* **328**, 1540 (2010).
- [29] D. Becker, M. D. Lachmann, S. T. Seidel, H. Ahlers, A. N. Dinkelaker, J. Grosse, O. Hellmig, H. Müntinga, V. Schkolnik, T. Wendrich *et al.*, Space-borne Bose–Einstein condensation for precision interferometry, *Nature (London)* **562**, 391 (2018).
- [30] R. Carollo, D. Aveline, B. Rhyno, S. Vishveshwara, C. Lannert, J. Murphree, E. Elliott, J. Williams, R. Thompson, and N. Lundblad, Observation of ultracold atomic bubbles in orbital microgravity, (2021), [arXiv:2108.05880](https://arxiv.org/abs/2108.05880).
- [31] L. Wang, Discovering phase transitions with unsupervised learning, *Phys. Rev. B* **94**, 195105 (2016).
- [32] G. Carleo and M. Troyer, Solving the quantum many-body problem with artificial neural networks, *Science* **355**, 602 (2017).
- [33] J. Carrasquilla and R. G. Melko, Machine learning phases of matter, *Nat. Phys.* **13**, 431 (2017).

- [34] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, Machine Learning Phases of Strongly Correlated Fermions, *Phys. Rev. X* **7**, 031038 (2017).
- [35] D. L. Deng, X. Li, and S. DasSarma, Quantum Entanglement in Neural Network States, *Phys. Rev. X* **7**, 021021 (2017).
- [36] C. Wang and H. Zhai, Unsupervised learning of frustrated classical spin models I: Principle component analysis, *Phys. Rev. B* **96**, 144432 (2017).
- [37] D. L. Deng, X. Li, and S. DasSarma, Exact machine learning topological states, *Phys. Rev. B* **96**, 195145 (2017).
- [38] F. Schindler, N. Regnault, and T. Neupert, Probing many-body localization with neural networks, *Phys. Rev. B* **95**, 245134 (2017).
- [39] S. J. Wetzel, Unsupervised learning of phase transitions: From principal component analysis to variational autoencoders, *Phys. Rev. E* **96**, 022140 (2017).
- [40] 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).
- [41] 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).
- [42] P. Zhang, H. Shen, and H. Zhai, Machine Learning Topological Invariants with Neural Networks, *Phys. Rev. Lett.* **120**, 066401 (2018).
- [43] J. Venderley, V. Khemani, and E. A. Kim, Machine Learning Out-of-Equilibrium Phases of Matter, *Phys. Rev. Lett.* **120**, 257204 (2018).
- [44] A. Sheverdin, F. Monticone, and C. Valagiannopoulos, Photonic Inverse Design with Neural Networks: The Case of Invisibility in the Visible, *Phys. Rev. Appl.* **14**, 024054 (2020).
- [45] C. Valagiannopoulos, Predicting the quantum texture from transmission probabilities, *J. Appl. Phys.* **127**, 174301 (2020).
- [46] M. J. S. Beach, A. Golubeva, and R. G. Melko, Machine learning vortices at the Kosterlitz-Thouless transition, *Phys. Rev. B* **97**, 045207 (2018).
- [47] S. Albawi, T. A. Mohammed, and S. Al-Zawi, Understanding of a convolutional neural network, *2017 International Conference on Engineering and Technology (ICET)* (IEEE, 2017), pp. 1–6.
- [48] V. L. Berezinskii, Destruction of long-range order in one-dimensional and two-dimensional systems having a continuous symmetry group I. Classical systems, *Sov. Phys. JETP* **32**, 493 (1971).
- [49] J. E. Van Himbergen, From Continuous to First-Order Transition in a Simple XY Model, *Phys. Rev. Lett.* **53**, 5 (1984).
- [50] R. Dixon, *Mathographics* (Dover Publications, New York, 1997).
- [51] J. W. Milnor, *Topology from the Differentiable Viewpoint* (Princeton University Press, Princeton, 1997).
- [52] R. L. B. Selinger, A. Konya, A. Traveset, and J. V. Selinger, Monte Carlo studies of the XY model on two-dimensional curved surfaces, *J. Phys. Chem. B* **115**, 13989 (2011).
- [53] L. S. Hirst, A. Ossowski, M. Fraser, J. Geng, J. V. Selinger, and R. L. B. Selinger, Morphology transition in lipid vesicles due to in-plane order and topological defects, *Proc. Natl. Acad. Sci. U.S.A.* **110**, 3242 (2013).
- [54] J. Zhou, G. Cui, S. Hu, Z. Zhang, C. Yang, Z. Liu, L. Wang, C. Li, and M. Sun, Graph neural networks: A review of methods and applications, *AI Open* **1**, 57 (2020).
- [55] S. T. Bramwell and P. C. W. Holdsworth, Magnetization and universal sub-critical behaviour in two-dimensional XY magnets, *J. Phys.: Condens. Matter* **5**, L53 (1993).
- [56] P. Butera and M. Comi, High temperature study of the Kosterlitz-Thouless phase transition in the XY model on the triangular lattice, *Phys. Rev. B* **50**, 3052 (1994).
- [57] P. Meakin, H. Metiu, R. G. Petschek, and D. J. Scalapino, The simulation of spinodal decomposition in two dimensions: A comparison of Monte Carlo and Langevin dynamics, *J. Chem. Phys.* **79**, 1948 (1983).
- [58] M. Ester, H.-P. Kriegel, J. Sander, and X. Xu, A density-based algorithm for discovering clusters in large spatial databases with noise, in *Proceedings of the 2nd International Conference on Knowledge Discovery and Data Mining (KDD-96)*, edited by E. Simoudis, J. Han, and U. Fayyad (AAAI Press, 1996), pp. 226–231.
- [59] S. Efthymiou, M. J. S. Beach, and R. G. Melko, Super-resolving the Ising model with convolutional neural networks, *Phys. Rev. B* **99**, 075113 (2019).
- [60] R. Zhang, B. Wei, D. Zhang, J.-J. Zhu, and K. Chang, Few-shot machine learning in the three dimensional Ising model, *Phys. Rev. B* **99**, 094427 (2019).
- [61] M. Rojas, O. Rojas, and S. M. de Souza, Frustrated Ising model on the Cairo pentagonal lattice, *Phys. Rev. E* **86**, 051116 (2012).
- [62] T. Yoshioka, A. Koga, and N. Kawakami, Frustrated Ising Model on the Garnet Lattice, *J. Phys. Soc. Jpn.* **73**, 1805 (2004).
- [63] R. Gupta, J. DeLapp, G. G. Batrouni, G. C. Fox, C. F. Baillie, and J. Apostolakis, Phase Transition in the 2D XY Model, *Phys. Rev. Lett.* **61**, 1996 (1988).