Berezinskii-Kosterlitz-Thouless transition from neural network flows

Kwai-Kong Ng,[1,*] Ching-Yu Huang [1,†] and Feng-Li Lin [2,†]

[1]*Department of Applied Physics, Tunghai University, Taichung 40704, Taiwan*
[2]*Department of Physics, National Taiwan Normal University, Taipei 11677, Taiwan*

We adopt the neural network (NN) flow method to study the Berezinskii-Kosterlitz-Thouless (BKT) phase transitions of the two-dimensional q-state clock model with $q \geqslant 4$. The NN flow consists of a sequence of the same units that proceed with the flow. This unit is a variational autoencoder trained by the data of Monte Carlo configurations in unsupervised learning. To gauge the difference among the ensembles of Monte Carlo configurations at different temperatures and the uniqueness of the ensemble of NN-flow states, we adopt the Jensen-Shannon divergence (JSD) as the information-distance measure "thermometer." This JSD thermometer compares the probability distribution functions of the mean spin value of two ensembles of states. Our results show that the NN flow will flow an arbitrary spin state to some state in a fixed-point ensemble of states. The corresponding JSD of the fixed-point ensemble takes a unique profile with peculiar features, which can help to identify the critical temperatures of BKT phase transitions of the underlying Monte Carlo configurations.

I. INTRODUCTION

Machine learning is a powerful tool for extracting relevant features from big data and then classifying the data according to these learned features. It has a wide range of applications to physical sciences [1–9]. When applied to physics study, the thermal states or quantum states of many-body systems are the natural arenas to exploit the power of machine learning in dealing with big data. Indeed, there have been many studies of identifying the phase transitions and critical states by machine learning, e.g. [10–27]. Most studies are based on the so-called supervised learning by training the neural network with the labeled data, e.g., labeling the thermal states by temperature and then using the trained machine to classify the unknown states and identify the phase transition point. There are two challenges to such studies. One is to discover new physics through machine learning, such as the new critical states. With supervised learning, the answers, e.g., critical states, are already encoded in the training set. The machine should be able to identify the target states without specifying them in advance in the training stage. This is the practice of unsupervised machine learning, by which the machine is trained with unlabeled data to find the posteriors of the data set encoding the information of the relevant features.

There are many approaches to uncovering the phase structures through unsupervised learning, e.g. [15–21]. Among them, there is an unconventional approach of adopting the unsupervised learning neural network to identify phase transitions, the so-called neural network (NN) flow [18–21], by mimicking the spirit of renormalization group (RG) flow. Instead of training the machine to classify the thermal states, to proceed with the NN flow, we first use the unlabeled data to train an unsupervised neural network such as autoencoder (AE) or variational autoencoder (VAE) imperfectly, i.e., without having high accuracy. We call this neural network the unit of NN flow. The AE structure with encoding and decoding parts bears a similar role to the single decimation (encoding) and rescaling (decoding) step in the RG flow. We then build up a machine of NN flow by connecting a sequence of the same flow units. The NN flows proceed by flowing the input states through the NN-flow machine to obtain the corresponding flow states. Surprisingly, the flow states will approach a fixed-point ensemble of states by which one can identify the phase transition by a temperature prediction network which plays the role of a thermometer. The validity of the NN flow-method has been demonstrated with its application to $q = 2, 3, 4$ Potts and clock spin models with various frameworks of unsupervised machine structure, such as the restricted Boltzmann machine (RBM) and (variational) autoencoder (AE/VAE) [21]. The success of NN flow demonstrates its similarity to the renormalization group flow (RG flow), which flows any state away from the critical state by sequential scaling or rescaling procedures, and the unit of NN flow plays a similar role of the scaling or rescaling operation. Despite that, the underlying mechanism for the success of NN flow is not completely clear.

Another challenge is to apply the machine-learning technique to identify the topological phase transition, such as the Berezinskii-Kosterlitz-Thouless (BKT) transition [28–31]. Unlike the usual Landau-Ginzburg-type phase transition, there is no local order parameter to characterize the BKT phase transition; thus, in some sense, the BKT phase transition is continuous, or so-called "topological." One may need to adopt nonlocal observables, such as the vortex condensation, to distinguish the BKT phase from the non-BKT ones. Regarding pattern recognition from the machine-learning point

*kkng@thu.edu.tw
†Corresponding authors: cyhuangphy@thu.edu.tw;
fengli.lin@gmail.com

of view, the relevant features of the BKT phase transition should be less evident than the Landau-Ginzburg ones. The BKT phase transition was first studied for the XY model. A family of its discrete versions called the q-state clock model with $q \geqslant 5$ can also display the BKT phase [32,33]. The BKT phase appears in the XY model whenever the temperature T is below a critical temperature. However, for the above clock models, the BKT phase appears when $T_1 \leqslant T \leqslant T_2$. Below T_1, the model is in the ferromagnetic (Z_q broken) phase, and above T_2, the paramagnetic (disordered) phase. Thus, there are the BKT phase transitions at both T_1 and T_2, which are not easy to identify precisely because of the lack of a local order parameter of the Landau-Ginzburg type even by numerical methods [34–39]. Despite that, it has been suggested there emerges an onset temperature T_{eu} for the $q > 4$ clock model, beyond which the thermodynamical observables of the discrete rotor models collapse into the ones of the continuous rotor model ($q = \infty$). Thus, thermodynamics in this regime becomes q-independent, which defines the so-called *extended universality* [40]. For $q \geqslant 8$ clock models, $T_2 > T_{eu}$ so that the phase transition at T_2 is BKT. On the other hand, for the cases of $q = 5, 6, 7$, $T_2 \leqslant T_{eu}$ and the phase transition at T_2 is not guaranteed to be BKT. It was claimed in [40] that the transition at T_2 for $q = 6, 7$ is non-BKT by some arguments, but in [32,34,41], it was showed otherwise.

Despite the intriguing feature of topological phase transitions and the difficulty of characterization, some machine-learning studies identify BKT phases for clock models based on the supervised learning scheme [42,43]. In this work we apply the NN-flow method to identify the BKT phase transitions for the clock models. In our previous work [21], we have studied the NN flows of the clock models with $q \leqslant 4$, of which the phase transitions are of the Landau-Ginzburg type. Regarding the $q \geqslant 5$ cases, we have also done some trials during the working period for [21], but we cannot identify the BKT phase transitions. There are two reasons for this failure. One is due to the inaccurate training data for the BKT states by our chosen Monte Carlo method. The other is due to an inaccurate machine-learning thermometer, which cannot give a precise temperature reading of the NN-flow configurations in the low-temperature regime because of the less relevant features in such a regime. In this work we have improved the implementation of NN flow for BKT phase transitions. We adopt a more accurate Monte Carlo scheme so that the simulated states can exhibit the BKT phase transitions and serve as the training data for the unit of NN flow. For characterizing the configurations, we give up the machine-learning thermometer, which is hard to train to be accurate enough for the low-temperature regime. Instead, we adopt the information-distance quantity, the so-called Jensen-Shannon divergence (JSD), which is a generalization of the Kullback-Leibler divergence, to measure the difference between two ensembles of thermal states at different temperatures. Therefore, we can use the Monte Carlo states of a given temperature as the gauge ensemble to compare with the ensemble of the NN-flow states.

In this way, we first see that the NN-flow ensembles obtained from the different temperature ensembles of Monte Carlo simulated configurations all yield the same information-distance measure profile. This implies that all the NN-flow states belong to the same ensemble of states, which we call the fixed-point ensemble (of states). Next, the JSD profile shows peculiar features at phase transitions to help identify critical temperatures $T_{1,2}$. For example, the information-distance measure shows a sharp drop near T_1 and a local minimum near T_2. Our results show that the NN flow can flow an arbitrary state to a state in the fixed-point ensemble, and the pattern of its JSD can be used to identify the critical temperatures of BKT phase transitions. Since the JSD shows a minimum near T_2, we may expect that the NN-flow states capture the essence of the critical states at T_2. However, this is not the case, as we will discuss later.

The remainder of this paper is organized as follows. In Sec. II we review the clock models and BKT phase transitions and describe our Monte Carlo method and the phase diagrams of the simulated results. In Sec. III we review the method of NN flow. In Sec. IV we introduce the information-distance "thermometer" for measuring the temperatures of thermal ensemble states. In Sec. V we show the results of the NN flows of the $q = 2, 4, 5, 6, 7, 8$ clock models on 20×20 and 40×40 lattices, including the patterns of JSDs and their features for identifying the (non-)BKT phase transitions. We also examine whether the fixed-point ensemble states bear the physical properties of the BKT phase. We conclude our paper in Sec. VI. In Appendix A we show more details on the phase diagrams of our Monte Carlo configurations. In Appendix B we show the typical schematic structure of the NN model for training the unit of NN flow and a typical training accuracy histogram.

II. MONTE CARLO SIMULATION OF Q-STATE CLOCK MODEL AND BEREZINSKII-KOSTERLITZ-THOULESS TRANSITION

We study the q-state clock model on the square lattice, which is described by the following Hamiltonian:

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

where the q-state spin on the site i is characterized by the Potts variable $\theta_i = 2\pi s_i / q$ with $s_i = 0, 1, \ldots, q-1$, and $\langle ij \rangle$ denotes the nearest neighbors that interact by the ferromagnetic coupling $J > 0$. The thermodynamic behaviors of the model and the associated phase transitions are encoded in the following partition function at finite temperature T:

$$Z_N = \sum_{\text{all possible } \theta_i, \theta_j} \exp \left(\frac{J}{k_B T} \sum_{\langle ij \rangle} \cos(\theta_i - \theta_j) \right). \tag{2}$$

In the following, we set $J = 1$ as the energy unit. The q-state clock model has been extensively investigated both analytically [32,33,41] and numerically [34–40]. The model is exactly solvable for $q = 2, 3, 4$ [44]. For $q = 2, 3$, they reduce to the Ising and the three-state Potts models, respectively; for $q = 4$, it can be mapped into a double copy of the Ising model. Therefore, their phase transitions are the second order of the Landau-Ginzburg types. As discussed before, for $q \geqslant 5$, the model displays the BKT phase sandwiched by a low-temperature ferromagnetic (Z_q) broken phase and a high-temperature paramagnetic (disordered) phase. Thus, we may
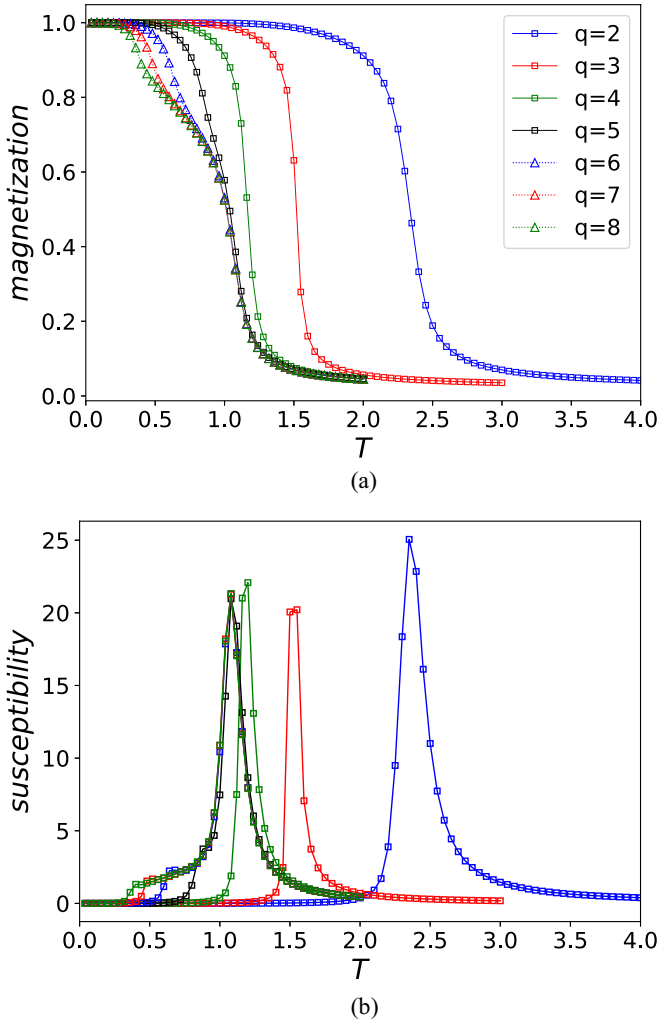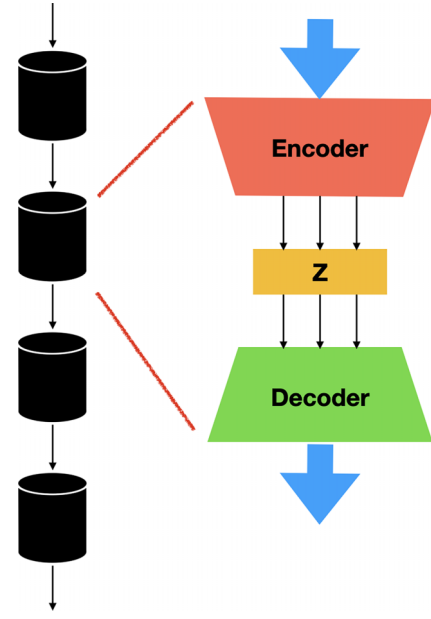
FIG. 2. Schematic structure of the NN flow (left) with the substructure of each unit, i.e., a (variational) AE (right). We consider VAE so that the hidden layer vector z is stochastic with Gaussian distributions.

FIG. 1. (a) Magnetization and (b) magnetic susceptibility for the q-state clock model are obtained from Monte Carlo simulations on a 40×40 lattice.

expect two continuous BKT transitions at temperature T_1 and T_2 ($> T_1$). In [45] the machine-learning technique and finite-size scaling are adopted to determine the critical temperatures of the $q = 8$ clock model, which are $T_1 = 0.410$ and $T_2 = 0.921$ in the thermodynamic limit. As $q \to \infty$, i.e., a planar rotor, the model is equivalent to the XY model for which the BKT phase extends throughout the low-temperature regime, that is, $T_1 = 0$. The BKT phase transitions are notoriously difficult to identify [34–40], as there is no Landau-Ginzburg-type order parameter to characterize them. This poses the challenge of identifying the BKT phase transitions in some novel ways and then motivates this work to resolve it by the NN flow. To proceed with the NN flow for identifying the BKT phase transitions of the clock models, we need to first produce the spin configurations as the training data to the neural network, i.e., the variational autoencoder (VAE) for the current work. We will implement the Monte Carlo method to simulate the spin configurations of the training set for various temperatures.

In the Monte Carlo simulation, it is known that, near the critical temperature, the single-flip updating approach will encounter the problem of critical slowing down, in which the autocorrelation time, i.e., the time for reaching independent configurations, becomes very large. To tackle this problem, Swendsen and Wang [46] proposed a cluster updating scheme in which connected sites of the same spins are grouped into a cluster according to some assigned probabilities. In this way, spins are flipped altogether inside the same cluster. Later Wolff [47] modified the cluster algorithm and extended it to multicomponent spins such that the spins are first projected into a randomly chosen direction before the Swednsen and Wang algorithm is carried out. We can obtain accurate Monte Carlo configurations for the NN flow by adopting the Wolff algorithm. We generate 1000 configurations for each temperature with 50 to 80 different temperatures for each q value. To investigate the size effect, simulations of both lattice sizes $L = 20$ and $L = 40$ are also carried out.

In Fig. 1 we show the magnetization and magnetic susceptibility of the clock models for $q = 2, 3, 4, 5, 6, 7, 8$, obtained from our Monte Carlo simulated configurations. The similar temperature profiles of energy and heat capacity can be found in Fig. 14 of Appendix A. For simplicity, we will focus only on Fig. 1 for discussion. We can see that these temperature profiles (or phase diagrams) move towards the low-temperature regime as q increases. The reason why our previous tests of the NN flow for $q \geqslant 5$ failed is partly because the resolution of our machine-trained thermometer is poor in the low-temperature regime. We can also see that for $q \geqslant 5$ models, there is a sharp peak near T_2 but only a minor jump near T_1. It is possible to determine the critical temperatures from these anomalies, in principle, but difficult. It may need highly accurate evaluations of critical points and precise size-scaling analysis to determine the BKT phase. Despite that, we will see that our information-distance thermometer will give clearer features near $T_{1,2}$.
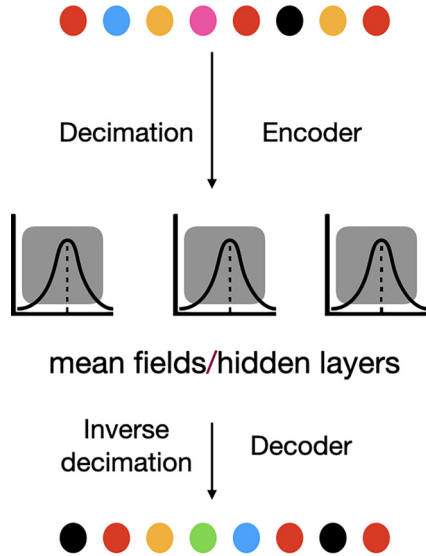
FIG. 3. Naive analog between RG flow and neural network flow. The decimation of the renormalization flow is in analogy to the encoder and its inverse to the decoder. Therefore, the stochastic hidden layer plays the role of the mean field in the renormalization flow. As a result, the different choices of the stochastic distribution for the hidden layer may correspond to the different mean fields and yield different fixed-point states.

Also, in Fig. 1 the lattice size is 40×40. Due to the topological nature of the BKT phase transitions, the critical temperatures T_1 and T_2 depend more sensitively on the lattice size than the Landau-Ginzburg ones; e.g., see [35] for the study of the finite-size effect of $q = 6$ clock model. Our goal here is not to find the true (theoretical) $T_{1,2}$ in the large size limit but to check whether the NN flow can be adopted to identify the $T_{1,2}$ of the Monte Carlo simulations, which we denote as $(T_{1,2})_{MC}$ or $(T_c)_{MC}$ to distinguish from the theoretical one denoted by $(T_c)_{\text{exact}}$. It will be much more challenging for machine learning to identify the critical temperatures using the Monte Carlo training set on the finite-size lattice. We will not specify the theoretical values of $T_{1,2}$ for simplicity.

Later we will examine the information-distance measure of the NN-flow states to the Monte Carlo ones. In this way we can examine the validity of the NN-flow method.

III. NEURAL NETWORK FLOW

In this section we review our setup for the NN flow, which was originally proposed in [18] (see also [19]) and then generalized in our previous work [21] by adopting not only the restricted Boltzmann machine (RBM) but also the autoencoder (AE) and the variational autoencoder (VAE) as the unit of NN flow. In this work we will adopt VAE since it yields more efficient and stable results from our own experiences.

The NN flow scheme is shown in the left part of Fig. 2, which consists of a sequential flow through the NN (i.e., VAE adopted in this work) units. The detail of the VAE unit is shown in the right part of Fig. 2. The procedure of the NN flow goes as follows. First, we train the VAE with Monte Carlo simulated configurations of the clock models. For each temperature T, we prepare about 2000 configurations as the training set, and the size of the temperature bin is 0.1 with $0 \leqslant T \leqslant 4$. The structure of VAE is very simple. It consists of an encoder and a decoder, as indicated in the right part of Fig. 2. The encoder consists of an input layer to which the input spin configuration is preprocessed by proper normalization as the common practice at the training stage of typical machine learning and a hidden/latent layer whose elements are random variables of unit Gaussian. One can also add an intermediate layer before the hidden layer. The intermediate and hidden layer sizes can be adjusted accordingly to optimize the training. The decoder then takes the hidden layer as the input, either or not going through an intermediate layer, and then gives the output layer of the same size as the input layer of the encoder. The mean square error between the input and output layers can be calculated as the penalty for optimizing the machine structure. Usually, we do not need to have very accurate training in the VAE to give some flexibility to flow the state. Otherwise, if the VAE is well trained, the output state will be perfectly the input state, and there will be no flow of states. The detailed machine structure of VAE and the training accuracy are almost the same as the one used in
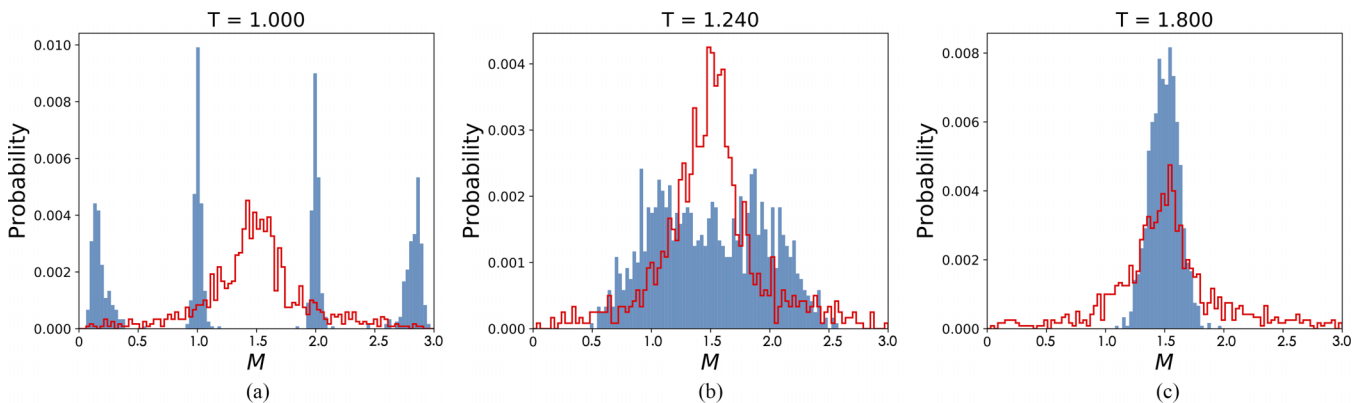

FIG. 4. Probability density functions (PDF) for the mean spin value M_n of $q = 4$ clock model on 20×20 lattice. Blue (shading): the PDFs for the temperature ensembles of Monte Carlo simulated configurations at (a) $T = 1$ (below T_c), (b) $T = 1.24$ (near T_c), and (c) $T = 1.8$ (above T_c). Red (line): the PDFs of the corresponding NN-flow ensemble after 100 iterations. It shows that the NN-flow states approach some states in a fixed-point ensemble of states.

[21]. For completeness, in Appendix B, we give the details of the hyperparameters for the typical NN model, along with the corresponding histogram of the training accuracy for training the unit of NN flow. Moreover, after some trials, we find that the convergence of the NN flow to the destined critical regime seems to be quite insensitive to the exact value of the training accuracy, as long as the accuracy is not high. This reflects that the unit of NN flow plays a similar role to the decimation or coarse-graining step in the RG flow, for which the loss of irrelevant information is expected.

In some sense, the VAE has a structure similar to a single step of renormalization group (RG) flow, as depicted in Fig. 3. The encoder compresses the input data by some truncation, similar to coarse graining by decimation. This procedure is used to approach a mean-field state, which in NN flow is the Gaussian hidden layer. The decoder then proceeds with the hidden layer resizes to the original size of the input layer, which is quite similar to performing the inverse decimation in the RG flow. It is interesting to examine how good this analog can be in the current study of the BKT phase transitions.

After training the VAE, we use it as the unit of NN flow. The NN flow then passes an arbitrary Monte Carlo configuration into a sequence of unit VAEs. All the VAEs have the same bias and weights determined at the training stage. The procedure of NN flow is similar to a sequence of decimation or inverse decimation steps in the RG flow to flow a UV state to the IR fixed-point state. Typically, after a few steps, the NN flow will also yield a fixed-point state. However, unlike the RG flow, we will see that the ensemble of fixed-point states bears no characteristic of criticality, such as power-law behaviors of correlation functions. Despite that, our information-distance thermometer can still indicate peculiar features near $T_{1,2}$.

We now motivate the need to construct a "thermometer" based on an information-distance measure to gauge the NN-flow states and to use it to identify the BKT phase transitions. For the Landau-Ginzburg type-phase transition, we can indeed calculate the order parameter such as the magnetization of the Ising model (or equivalently the two-state clock model [44]) and compare with the one by Monte Carlo method. Or we can measure the temperature of the NN-flow states by using a machine-learning thermometer. We can train the thermometer machine with the Monte Carlo simulated data set. In fact, this is what has been done in our previous work [21]. However, this kind of thermometer is not very accurate in low-temperature regimes. This is expected as the low-temperature phase is the ordered phase with fewer intrinsic features, thus more difficult to train. When studying the BKT phase, the above techniques become less useful because there is no local order parameter such as magnetization for the BKT phase transition, and especially the BKT critical temperatures tend to locate at the low-temperature regime. This is why we must adopt different techniques to characterize the NN-flow states to extract their BKT features. This will be discussed in the next section.

IV. INFORMATION-DISTANCE MEASURE AS THERMOMETER

Instead of training an additional machine-learning-type thermometer, we can, in fact, use the temperature ensembles
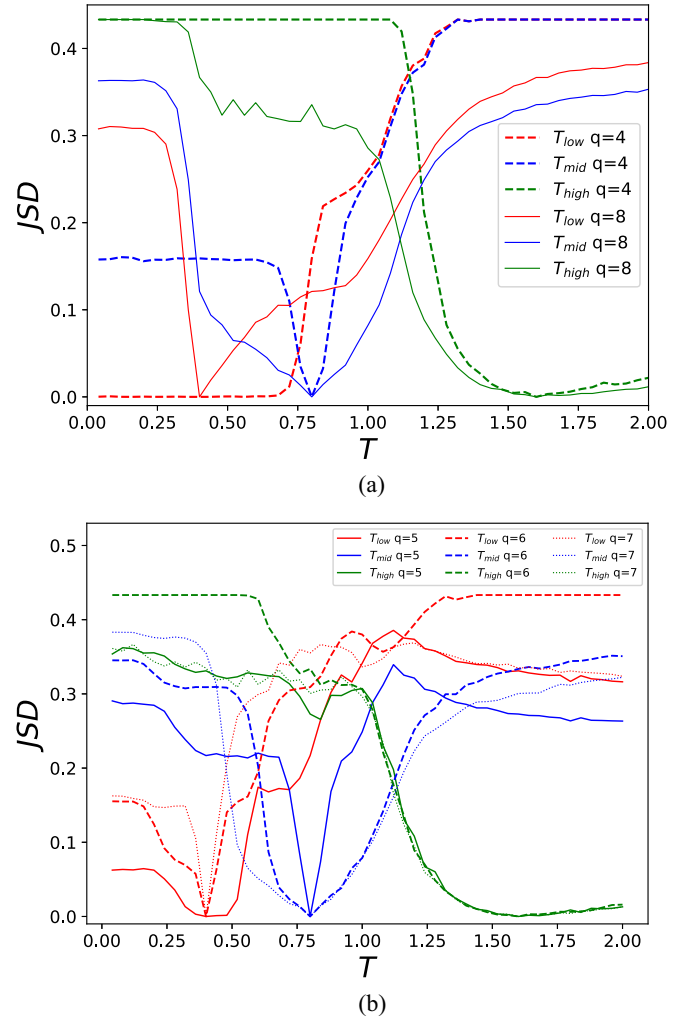

(a)

(b)

FIG. 5. Jensen-Shannon divergence $JSD[P_{T_p} || Q_T]$ of the (a) $q = 4$ (dashed line) and 8 (solid line) and (b) $q = 5$ (solid line), 6 (dashed line), and 7 (dotted line) clock models on a 40×40 lattice. Here T runs all the temperature bins in $0 < T < 4$, and T_p runs for three specific temperatures: $T_{\text{low}} = 0.4$ (red color), $T_{\text{mid}} = 0.8$ (blue color), and $T_{\text{high}} = 1.6$ (green color). The JSDs can serve as a good "thermometer" because the minimum of the JSDs sits right on the corresponding T_p.

of the Monte Carlo simulated configurations to gauge NN-flow states and their associated ensemble. In this sense we would like to measure the difference between two ensembles of thermal states. Thus, one can invoke some information-distance measures such as the root-mean-square distance, the Kullback-Leibler divergence (KLD) [48] or the Jensen-Shannon divergence (JSD) [49,50]. Given two (discrete) probability density functions (PDFs) denoted by P and Q, the KLD is given by

$$KL[P||Q] = -\sum_n P(n) \ln \frac{Q(n)}{P(n)}. \quad (3)$$

KLD is a very common information-distance measure, which is also known as relative entropy. However, it is not well defined if $Q(n)$ vanishes for some n. Therefore, a more
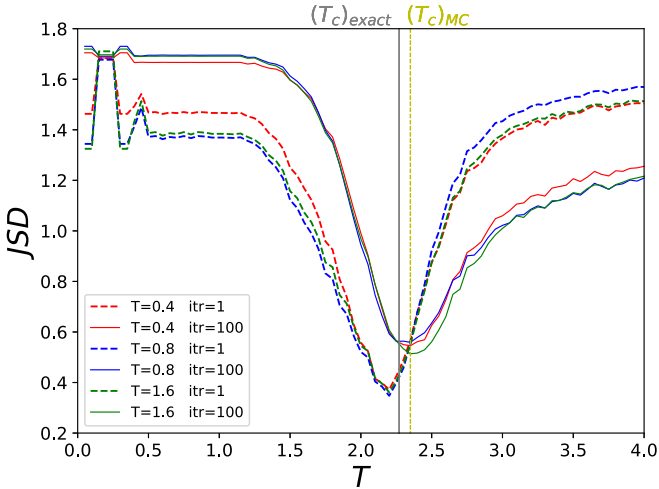
FIG. 6. NN-flow JSD$[P_{T_{\text{flow}}}||Q_T]$ of $q = 2$ clock model on a 20×20 lattice. $P_{T_{\text{flow}}}$ is the PDF of the mean spin value of NN-flow states after one iteration (dashes line) and 100 iterations (solid line) of the initial Monte Carlo ensemble at temperature T_{flow}, and Q_T is the reference PDF based on Monte Carlo ensemble at temperature T. For convenience, we choose $T_{\text{flow}} = T_{\text{low}}$ (red), T_{mid} (blue), and T_{high} (green) of Fig. 5. We see that there is a unique pattern for NN-flow JSDs with its minimum sitting inside the narrow window of $[(T_c)_{\text{exact}}, (T_c)_{MC}]$ with $(T_c)_{\text{exact}} = \frac{2}{\log(1+\sqrt{2})} = 2.27$ (vertical gray line), and $(T_c)_{MC} = 2.35$ (vertical light-yellow line) the theoretical and Monte Carlo critical temperatures, respectively.

well-defined measure is the JSD, which is given by

$$\text{JSD}[P||Q] = \tfrac{1}{2}\text{KL}[P||R] + \tfrac{1}{2}\text{KL}[Q||R], \quad (4)$$

where $R = \frac{1}{2}(P + Q)$. It is easy to see that JSD is well defined even though some of $P(n)$ or $Q(n)$ vanishes. Note that KL$[P||Q]$ and JSD$[P||Q]$ are positive definite and can measure the dissimilarity between the distributions P and Q. Thus, the JSD$[P||Q]$ becomes smaller if P and Q are more similar to each other and vanishes if $P = Q$.

Recall that temperature can also be defined statistically. Thus, our information-distance measure, in this sense, plays the role of the "thermometer." With the JSD as the chosen information-distance measure, the next question is what kind of PDFs we will choose so that the resultant JSD can serve as a good "thermometer." Since we would like to construct the PDFs for the ensemble of the thermal states, it is better to choose some specific property of the spin configurations and consider its PDF for a given ensemble.

A natural property for the spin configurations is the "mean spin value,"

$$M = \frac{1}{N}\sum_{i=1}^{N} s_i, \quad (5)$$

where N is the total number of the lattice sites, and $s_i = 0, 1, \ldots, q - 1$ is the spin value on the site i. Thus, M will be the random variable for an ensemble of a given temperature or its corresponding ensemble of the NN-flow states. We will further discretize M by proper rounding into bins denoted by M_n with n the integer index so that its associated PDF will be just denoted by $P(n)$ or $Q(n)$ as the ones used for the JSD.
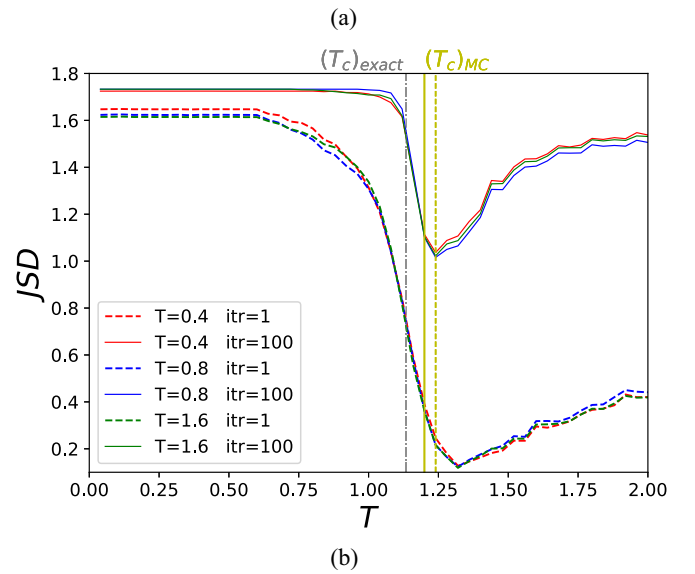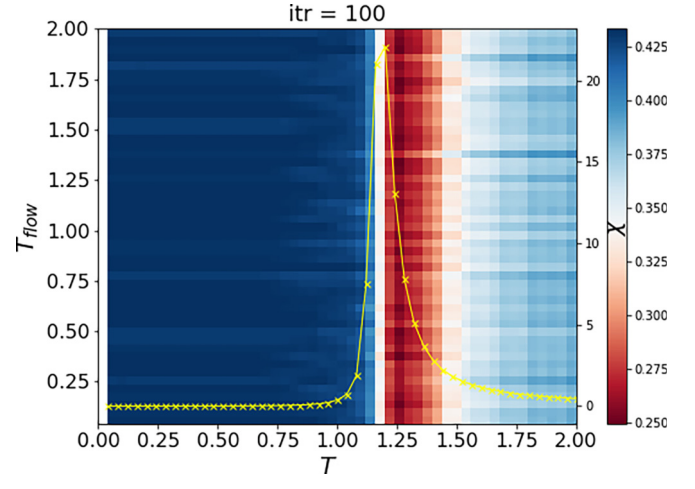

(a)

(b)

FIG. 7. (a) Color chart of JSD$[P_{T_{\text{flow}}}||Q_T]$ of T-T_{flow} plane for the $q = 4$ clock model on a 40×40 lattice for the ensemble of NN-flow states after 100 iteration steps. The corresponding diagram of magnetic susceptibility (yellow line with decorated crosses) is also superimposed, which shows a peak near $T = (T_c)_{\text{exact}}$. (b) NN-flow JSD$[P_{T_{\text{flow}}}||Q_T]$ for $T_{\text{flow}} = T_{\text{low}}$ (red), T_{mid} (blue), and T_{high} (green) of Fig. 5 on a 20×20 lattice (dashed line) and 40×40 lattice (solid line). We label $(T_c)_{\text{exact}} = \frac{1}{\log(1+\sqrt{2})} = 1.13$ by the vertical gray dash-dot line, and $(T_c)_{MC} = 1.24$ (20×20 lattice), 1.20 (40×40 lattice) with the vertical solid light-yellow lines. We see that the finite-size effect is relevant to improve the capability of using the minima of NN-flow JSDs to identify the critical temperature. Our results again show that NN flows can reach a fixed-point JSD pattern with its minimum located at the critical temperature.

Note that the mean spin value M is different from the usual physical magnetization of the state and cannot be served as an order parameter of the system.

In Fig. 4 we show the PDFs for M_n for the Monte Carlo configuration ensembles of the $q = 4$ clock model at three different temperatures (in blue shading): $T = 1$ (below T_c), $T = 1.24$ (near T_c), and $T = 1.8$ (above T_c), and also the PDFs of the corresponding ensemble states obtained by the NN flow (red line). We can see that the NN-flow PDFs
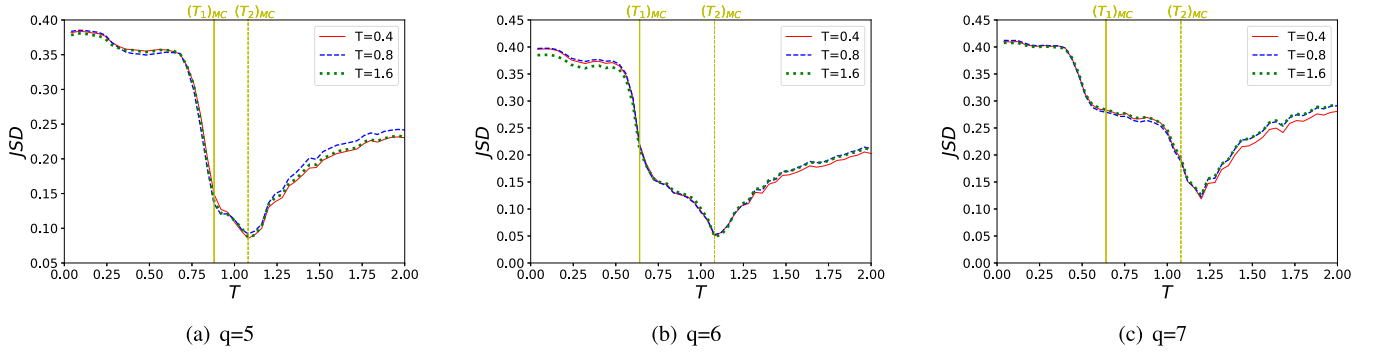
(a) q=5   (b) q=6   (c) q=7

FIG. 8. NN-flow JSD$[P_{T_{\text{flow}}}||Q_T]$ after 100 iterations for $T_{\text{flow}} = T_{\text{low}}$(red solid), T_{mid} (blue dashed), and T_{high} (green dotted) of Fig. 5 on a $40{\times}40$ lattice for the clock models of $q = $ (a) 5, (b) 6, and (c) 7. The critical temperatures $(T_1)_{MC}$ (yellow solid) and s $(T_1)_{MC}$ (yellow dashed) of the Monte Carlo simulation are indicated by the vertical light-yellow lines in each subfipanel. The universal JSDs have decreased sharply near the transition point $(T_1)_{MC}$ between the ordered phase and the BKT phase. Moreover, the JSD minima sit right at $T = (T_2)_{MC}$ in (a) and (b), but a bit higher than $T = (T_2)_{MC}$ in (c).

resemble each other. As we shall see later, this implies that the NN-flow states approach some states in a fixed-point ensemble of states. Also, as we can see from Fig. 4, because the mean spin value has no preferred directions, the PDFs are symmetric around the mean value, which is always 3/2 for all temperatures and is unrelated to the orderings.

To demonstrate the ability of the JSD "thermometer," in Fig. 5 we plot the JSD$[P_{T_p}||Q_T]$ for the $q = 4, 5, 6, 7, 8$ clock models on a $40{\times}40$ lattice. Here Q_T is the PDF of the mean spin value for the Monte Carlo ensemble at temperature T with T running across all the temperature bins with $0 < T < 4$, and P_{T_p} is the one for some specific temperature T_p. In Fig. 5 we choose three specific temperatures: $T_{\text{low}} = 0.4$, $T_{\text{mid}} = 0.8$, and $T_{\text{high}} = 1.6$ with the middle temperature T_{mid} in the BKT phase for the cases of $q > 4$. We see that in most cases, the JSD has its minimum sit right on $T = T_p$. This indicates that the JSD can serve as a good "thermometer" for the clock models considered in this work. The only exceptions are the JSDs in the low-temperature regime of the $q \leqslant 4$ cases, e.g., the one of $q = 4$ at T_{low} in Fig. 5, for which we find that the corresponding JSDs are almost zero in an extended regime at low temperature. It indicates that the thermal fluctuations in these low-temperature ordered phases are too small to be resolved by JSD.

V. THE RESULTS OF NN FLOW FOR BKT PHASE TRANSITIONS

In this section we will present the results of the NN flows for the $q = 2, 4, 5, 6, 7, 8$ clock models and discuss the properties of the NN-flow states.

A. $q = 2$

We first consider the $q = 2$ clock model on the $20{\times}20$ lattice as a benchmarking test to check if the JSD of mean spin value can be used as a kind of "thermometer" to gauge the ensembles of states. As mentioned, the $q = 2$ clock model is equivalent to the Ising model with a second-order phase transition from the low-temperature ordered phase to the high-temperature disordered one. We will consider the NN-flow JSD$[P_{T_{\text{flow}}}||Q_T]$ where $P_{T_{\text{flow}}}$ is the PDF of the mean spin value

of NN-flow states after 100 iterations of the initial Monte Carlo ensemble at temperature T_{flow}, and Q_T is again the Monte Carlo one at temperature T with T running over $0 < T < 4$. For convenience, we choose $T_{\text{flow}} = T_{\text{low}}$, T_{mid}, and T_{high} of Fig. 5. We expect that the NN-flow JSDs should show a unique pattern associated with a fixed-point ensemble of states at the critical temperature, thus being consistent with the results obtained in [21]. This is indeed the case as shown in Fig. 6, as we see that the minimum of the universal NN-flow JSD sits inside the narrow window of $[(T_c)_{\text{exact}}, (T_c)_{MC}]$ with $(T_c)_{\text{exact}}$ and $(T_c)_{MC}$ the theoretical and Monte Carlo critical temperatures, respectively. With the success of this benchmarking test, we can move on to the higher q clock models.

B. $q = 4$

The $q = 4$ clock model is the dividing line for the appearance of the BKT phase transitions. For the $q > 4$ clock models, the BKT phase appears to be characterized by some nonlocal order parameter, such as the condensation of the vortex/antivortex pair. However, we like to characterize the BKT phase transitions by the patterns of our JSD "thermometer" for the NN-flow ensemble of states after 100 iterations of the initial Monte Carlo ensemble. The result of the NN-flow JSD$[P_{T_{\text{flow}}}||Q_T]$ is shown in Fig. 7. The notation for $P_{T_{\text{flow}}}$ and Q_T and the three choices of $T_{\text{flow}} = T_{\text{flow}}$, T_{mid} and T_{high} in Fig. 7(b) are the same as for the $q = 2$ case. In Fig. 7(a) we plot the color chart for the distributions of JSD$[P_{T_{\text{flow}}}||Q_T]$ on the T-T_{flow} plane. This plot is for the $q = 4$ clock model on a $40{\times}40$ lattice. The color pattern of Fig. 7(a) is almost uniform along the y axis but varies along the x axis. This implies that all the NN-flow ensembles of different T_{flow} yield the same JSD profile; a universal fixed-point ensemble of the NN-flow states exists.

More interestingly, the color chart of Fig. 7(a) shows two discontinuities along the x axis. One is around $T = (T_c)_{\text{exact}} \approx 1.13$, which is the critical temperature of the $q = 4$ clock mode, and slightly deviates from the Monte Carlo critical temperature $(T_c)_{MC}$ as indicated by the peak of magnetic susceptibility (yellow line with decorated crosses). The second discontinuity is at around $T = 1.35$, beyond which the magnetic susceptibility starts to level off. In Fig. 7(b) we show
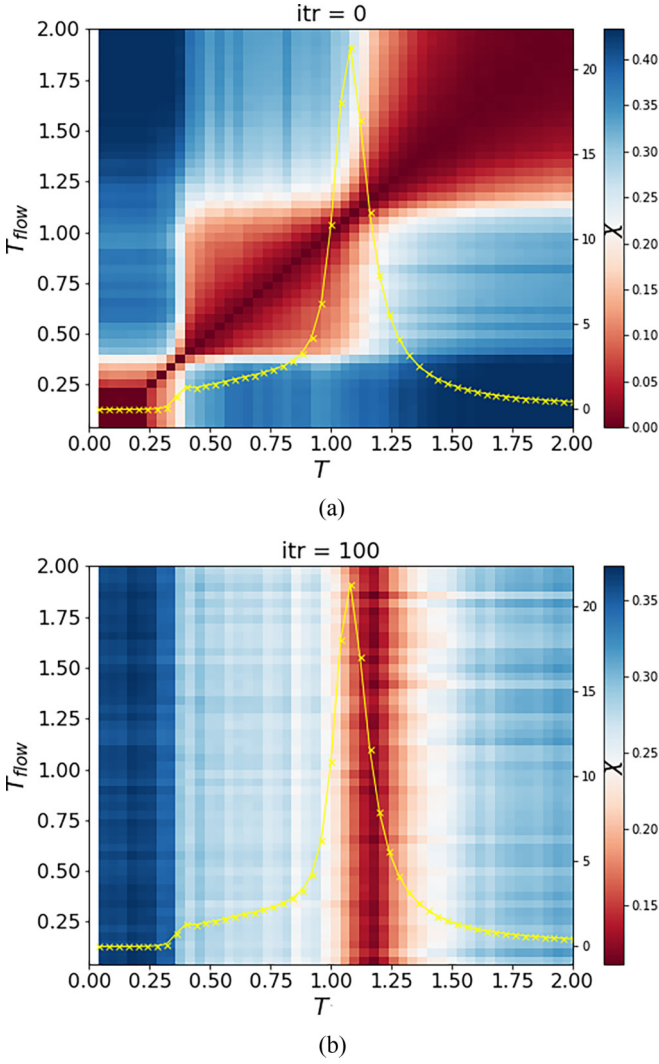
(a)

(b)

FIG. 9. (a) Color chart of JSD$[P_{T_{flow}}||Q_T]$ on the T-T_{flow} plane with $P_{T_{flow}}$ and Q_T the PDFs of the mean spin value for the Monte Carlo ensembles of the $q = 8$ clock model on a 40×40 lattice at temperatures T_{flow} and T, respectively. The three (red) blocks coincide with the ordered, BKT, and disordered phases, which are superimposed by the magnetic susceptibility, shown by the light-yellow line with decorated crosses. (b) The JSD color chart with the PDF $P_{T_{flow}}$ in (a) replaced by the NN-flow one after 100 iterations. The uniform color distribution along the y axis indicates a universal fixed-point ensemble, which can capture the Monte Carlo critical temperatures $(T_{1,2})_{MC} \approx 0.39, 1.06$ by its discontinuity along the x axis.

the NN-flow JSD$[P_{T_{flow}}||Q_T]$ for $T_{flow} = T_{low}$ (red), T_{mid} (blue), and T_{high} (green) on a 20×20 (dashed lines) and on a 40×40 (solid lines) lattice. Then, we can examine the finite-size effect by comparing these two subfigures. Again, we see that the NN flow drives the JSD curves of Monte Carlo ensembles of different temperatures to a universal JSD curve with its minimum sitting nearby the narrow window of $[(T_c)_{exact}, (T_c)_{MC}]$. This is consistent with the result of Fig. 7(a). We see that the minimum of the universal JSD on the 40×40 lattice is closer to $(T_c)_{MC}$ than that of the 20×20 lattice. This indicates the finite-size effect. Thus, after taking care of the finite-size effect, the universal NN-flow JSD can be used to identify the
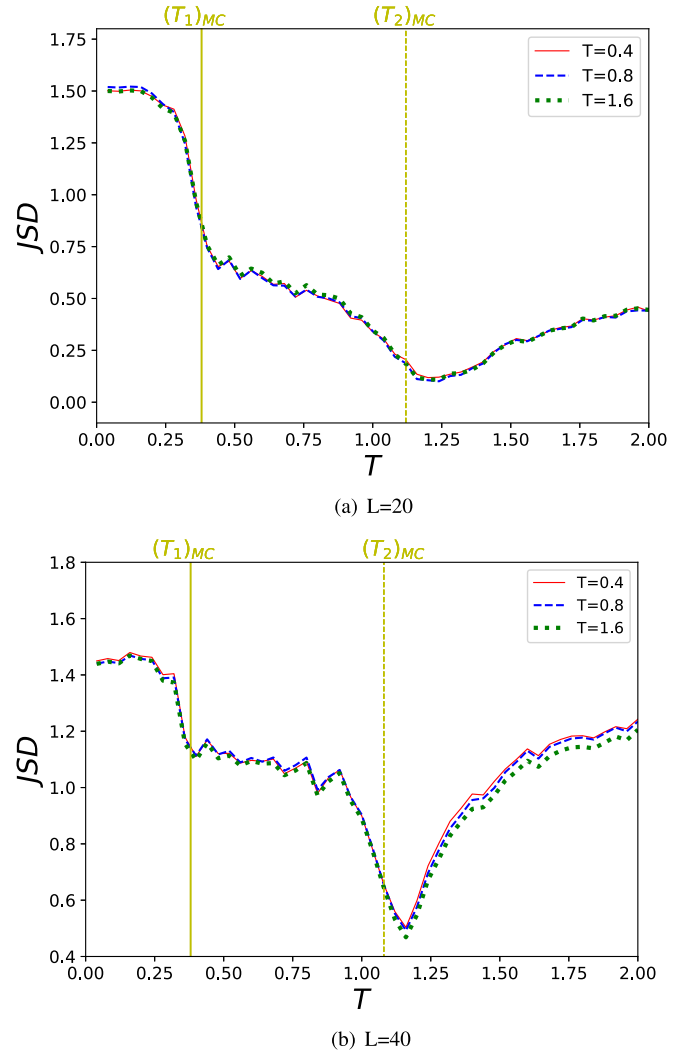

(a) L=20

(b) L=40

FIG. 10. NN-flow JSD$[P_{T_{flow}}||Q_T]$ after 100 iterations for $T_{flow} = T_{low}$ (red solid), T_{mid} (blue dashed), and T_{high} (green dotted) of Fig. 5 for the clock models of $q = 8$ (a) on a 20×20 lattice, and (b) on a 40×40 lattice. The results show that NN flow can yield a universal JSD to identify the BKT phase and the associated critical temperatures by taking account of the finite-size effect.

critical temperature of the $q = 4$ clock model as in the $q = 2$ case. Last, the NN-flow JSD of the 40×40 lattice also shows a sharp drop in a very small window between $[(T_c)_{exact}, (T_c)_{MC}]$. This can also be considered a peculiar feature to indicate the phase transition besides using its minimum.

C. $q = 5, 6, 7$

Now we will consider the clock models possibly exhibiting the BKT phase transitions, that is, $q > 4$. In this subsection we will consider $q = 5, 6, 7$ cases, and in the next subsection the $q = 8$ case. From both the analytical and numerical studies [32,33,35–39], it is argued that there is an extended BKT phase in a finite temperature interval $[T_1, T_2]$ for $q > 4$ cases. As discussed before, due to the emergence of T_{eu} for *extended universality*, there are discussions about the authentic BKT nature of the phase transition at T_2 for $q = 5, 6, 7$. Therefore
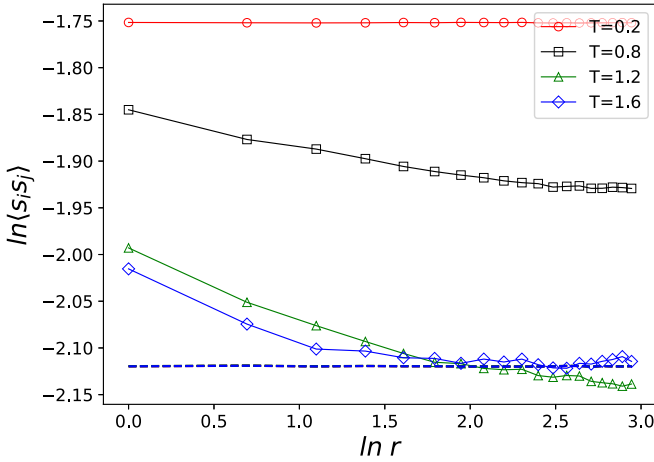
FIG. 11. Correlation of site spins, $\langle s_i s_j \rangle$ of the $q = 8$ clock model on a 40×40 lattice as a function of $r = |i - j|$ for the Monte Carlo ensemble at various temperatures and for the universal NN-flow ensemble after 100 iterations (dashed lines). The power-law behavior will be characterized by an approximately straight line of nonzero slope. The $T = 0.8$ one (black square line) represents the BKT phase and does show an approximate power-law behavior. However, the NN-flow one is a flat straight line, more like the low-temperature ordered phase (red circle line).

we will first discuss the result for $q = 5, 6, 7$ and leave the study of the $q = 8$ case in the next subsection.

Similar to the $q = 2, 4$ cases, the resultant NN-flow $\text{JSD}[P_{T_{\text{flow}}} || Q_T]$ for $q = 5, 6, 7$ on a 40×40 lattice for $T_{\text{flow}} = T_{\text{low}}$ (red solid), T_{mid} (blue dashed), and T_{high} (green dotted) of Fig. 5 is shown in Fig. 8. We again see that the NN-flow JSDs approach a universal pattern, which indicates a fixed-point ensemble of states. In Figs. 8(a) and 8(b) for the $q = 5, 6$ cases, respectively, we see that the JSD minima sit right on $T = (T_2)_{MC}$, which is the T_2 indicated by Monte Carlo simulation. However, for Fig. 8(c) for the $q = 7$ case, the minimum sits at the temperature slightly higher than $(T_2)_{MC}$.

Finally, in all three cases, as shown in Fig. 8, the JSD has dropped sharply near the transition point $T = (T_1)_{MC}$ and has a minimum close to $T = (T_2)_{MC}$, where $(T_{1,2})_{MC}$ are indicated by the vertical light-yellow lines in each panel. This indicates the capability of the NN flows to identify the BKT phase.

D. $q = 8$

Now we consider the NN flow of the $q = 8$ clock model on 20×20 and 40×40 lattices. This model has an extended BKT phase ranging from $(T_1)_{MC}$ to $(T_2)_{MC}$, for which the critical temperatures $(T_{1,2})_{MC}$ of the BKT phase transition are read from the typical phase diagrams of our Monte Carlo simulation, as shown in Fig. 1. Moreover, even there exists the collapse of thermodynamic observables, the phase transition at $(T_2)_{MC}$ is the BKT type for the $q = 8$ clock model.

We first plot the color charts of $\text{JSD}[P_{T_{\text{flow}}} || Q_T]$ on the T-T_{flow} plane. In Fig. 9(a) the PDFs $P_{T_{\text{flow}}}$ and Q_T are the ones of mean spin value for the Monte Carlo ensembles of states at temperatures T_{flow} and T, respectively. This can be seen as the autocorrelation between finite-temperature Monte Carlo ensembles. Interestingly, three red blocks appear, corresponding
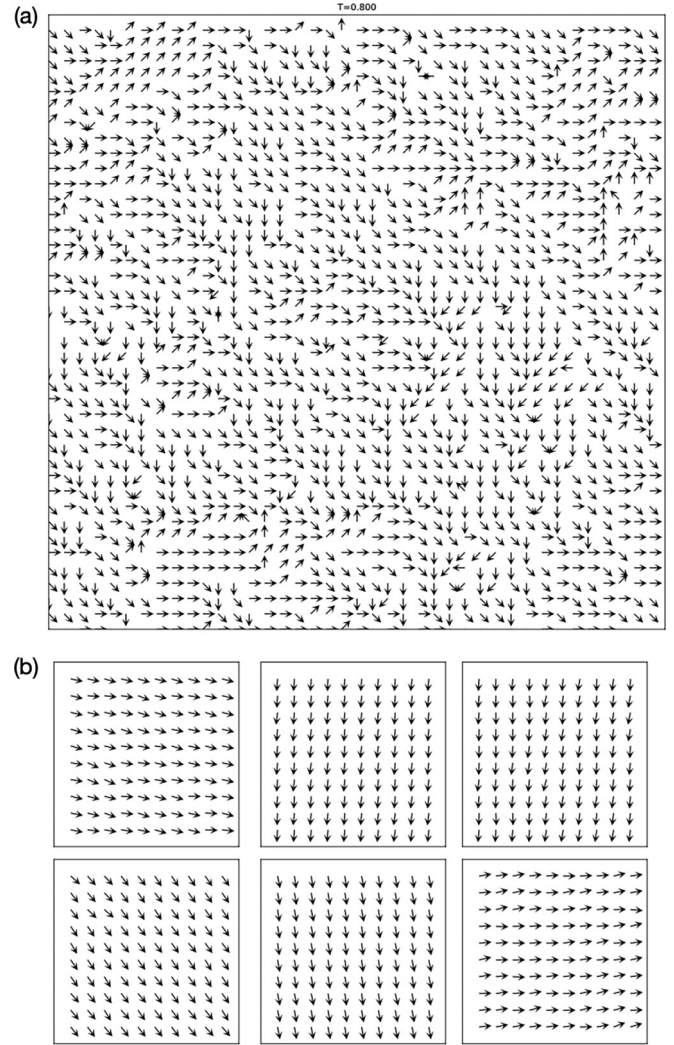

FIG. 12. (a) Picture of the orientations of the site spins for a chosen BKT Monte Carlo state of $q = 8$ clock model on a 40×40 lattice with temperature $T = 0.8$. This picture shows the typical vortex-antivortex condensation of the BKT phase. (b) The picture of the orientations of the site spins for some six chosen NN-flow states of the same model after 20 iterations. For simplicity, only part of the lattice is shown. Almost all the site spins in each NN-flow state point to the same direction.

to the ordered phase, the BKT phase, and the disordered phase of the $q = 8$ clock model. For comparison, in Fig. 9(b) we replace $P_{T_{\text{flow}}}$ in Fig. 9(a) by the PDF associated with the NN-flow ensembles after 100 iterations. As expected, the color chart is now uniform along the y axis, indicating a universal fixed-point ensemble of states after NN flow. More interestingly, we see a minor vertical jump near $(T_1)_{MC} \approx 0.39$, the lower critical temperature of the BKT phase transition. There is no such jump in Fig. 7(a) for the $q = 4$ case. On the other hand, there emerges a (red) stripe starting near the second BKT critical temperature $(T_2)_{MC} \approx 1.06$ as indicated by the peak of magnetic susceptibility, but ending when the magnetic susceptibility starts to level off. This kind of (red) stripe also appears in Fig. 7(a) for the $q = 4$ case.

Similarly to the cases of $q = 5, 6, 7$, we again show the NN-flow $\text{JSD}[P_{T_{\text{flow}}} || Q_T]$ of the $q = 8$ clock model for

$T_{flow} = T_{low}$ (red solid), T_{mid} (blue dashed), and T_{high} (green dotted) on a 20×20 lattice in Fig. 10(a) and on a 40×40 lattice in Fig. 10(b). The NN-flow JSDs again approach a unique pattern, indicating a universal fixed-point ensemble of states. Unlike the cases of $q = 4, 5, 6$ but similar to the case of $q = 7$, the minimum of universal JSD no longer locates at the critical Monte Carlo temperature $(T_2)_{MC}$ but at a higher one. Despite that, the increase of the lattice size does help to sharpen the minimum of the JSD and slightly move it toward $(T_2)_{MC}$. Note that the same effect also moves $(T_2)_{MC}$ to the lower value. By the trend, we expect that with a large enough lattice size, the minimum of the universal JSD could move closer to $(T_2)_{MC}$ and become sharper. We believe this should be the case for all BKT phase transitions from the BKT phase to the disordered phase; that is, it should hold for $q \geqslant 7$.

Regarding the BKT phase transition near $(T_1)_{MC}$, the JSDs do show a sharp drop, as in the cases of $q = 5, 6, 7$. Moreover, the increase of the lattice size also helps to pin down $(T_1)_{MC}$ more precisely by the endpoint of the sharp drop. However, we found some deviation of the NN flow results from the MC data for $q = 7, 8$, especially at T_2. One of the possible explanations is due to the emergence of *extended universality* discussed before. That is, as q increases, the phase transition at T_2 should be closer to the one of the continuous rotor model and harder for the JSD or other quantities to capture the exact phase transition point. We can then conclude that the NN flow

can yield a universal JSD pattern that can be used to identify the BKT phase and the associated critical temperatures once the finite-size effect is considered. Therefore, we can infer the extended regime of BKT between $(T_1)_{MC}$ and $(T_2)_{MC}$ from the peculiar pattern of JSD.

E. Properties of the NN-flow states

Our study demonstrates the ability of NN flows to identify BKT phase transitions. This then raises the question if the NN-flow ensemble fixed-point states can be the physical critical states for the BKT phase transition. Or does the ensemble of NN-flow fixed-point states bear any physical implications? In the following, we examine this issue.

To check this issue, we should first recall that our NN-flow states are no longer the q-spin states but with the site-spin values in the $[0,1]$ range. One can either use these continuous spin values to evaluate some physical quantities or try to discretize them to mimic the q-state spin. For the moment, we will use the continuous site-spin values. The first thing to check is to evaluate the spatial correlation function of the site spins and see if it exhibits power-law-like behaviors, which is the signature of critical states. The results for some chosen states of the $q = 8$ clock model on a 40×40 lattice are shown in Fig. 11. This is a log-log plot, so the power-law behavior should be represented by an approximately straight line of the nonzero slope, such as the BKT one at $T = 0.8$ shown by
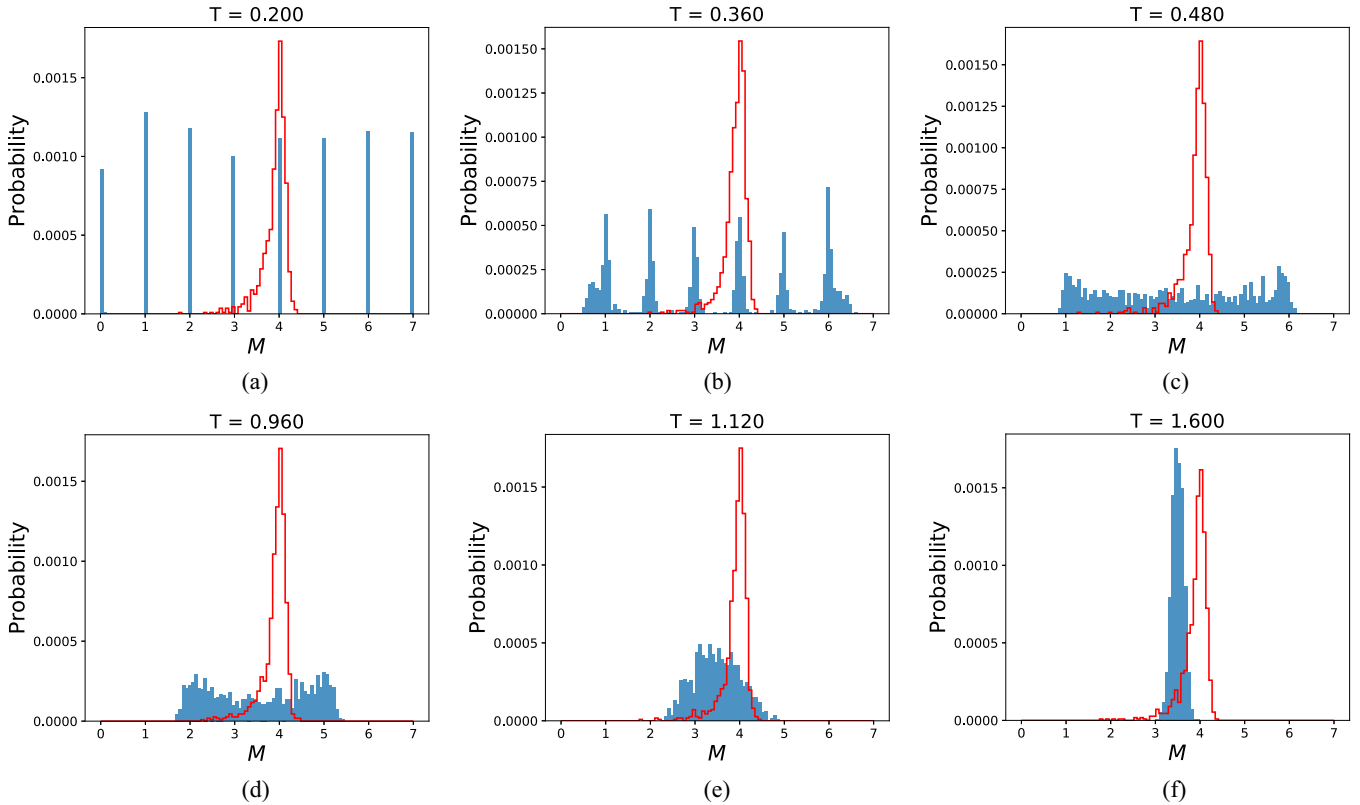

FIG. 13. Probability density functions (PDF) for the mean spin value M_n of $q = 8$ clock model on 40×40 lattice. Blue: the PDFs for the temperature ensembles of Monte Carlo simulated configurations at (a) $T = 0.2$ (ordered phase), (b) $T = 0.36$ [ordered phase, near $(T_1)_{MC}$], (c) $T = 0.48$ [BKT phase, near $(T_1)_{MC}$], (d) $T = 0.96$ [BKT phase, near $(T_2)_{MC}$], (e) $T = 1.12$ [disordered phase, near $(T_2)_{MC}$], and (f) $T = 1.6$ (disordered phase). Red: the PDFs of the corresponding NN-flow ensemble after 100 iterations, which are almost the same for all temperatures. The results show that although the flow PDFs do not resemble the ones at $(T_2)_{MC}$, it still yields the maximal overlap and thus the minimal JSD to help indicate the phase transition.

the black line. However, we see that the NN-flow one (the dashed line) is just a flat straight line, which is more like one of the low-temperature ordered phases such as the red-line one. This indicates that the NN-flow states are more like the low-temperature ordered states instead of the BKT ones. Such a conclusion remains even when we discretize the site spins of the NN-flow fixed-point states. This implies that the NN flow states do not show the features of BKT states, though the NN-flow JSDs can show a universal pattern that can capture the phase diagram of the BKT phase transitions.

We can explore the above conclusion more directly by showing the orientations of all the site spins, namely, the picture of $0 \leqslant \theta_i = 2\pi s_i/q \leqslant 2\pi$ for all i. In Fig. 12(a) we show a picture of the orientations of site spins for a typical BKT state of the $q = 8$ clock model at $T = 0.8$, for which we have verified its spin-spin correlation obeys an approximate power-law behavior. We see that the site spins slowly change their orientations either clockwise or anticlockwise in the local domains, similar to the vortex and antivortex structure. The clockwise domains will usually encounter the anticlockwise domain after some extent, expressing the pattern of vortex-antivortex condensation as expected for a BKT state. On the other hand, in Fig. 12(a) we show the pictures of orientations of site spins for six NN-flow states of the $q = 8$ clock model on a 40×40 lattice. Each of them has all the site spins oriented in the same direction, which is consistent with the result of the flat straight line shown in Fig. 11. For simplicity, we show only part of the lattice. However, different NN-flow states pick up different orientations for the whole lattice. This seems to imply that each NN-flow state is in a particular spontaneous-symmetry-broken state. Despite that, the statistics of the orientation or the mean spin value from the ensemble of NN-flow states can yield a universal and peculiar pattern by which the BKT phase transitions can be identified. Therefore, even though the NN flow can yield a fixed-point ensemble of states, the component states bear no feature of BKT states so we cannot use this universal ensemble to evaluate the critical exponents of the phase transitions. This is the essential difference of the NN flow from the RG flow. That is, the statistics of NN-flow ensemble states can help identify the BKT phase, but the states carry no feature of the mean-field states. Therefore, we can take the analog picture of Fig. 3 only in the literal sense.

F. Why can JSD be used to detect the transition?

To clarify why JSD of mean spin value can be used to indicate the phase transition even though the NN-flow states carry no vortex/antivortex features of the critical states, we now examine the probability distributions (PDF) for details. In Fig. 13 we choose to display the PDFs of the mean spin value M_n of $q = 8$ clock model on 40×40 lattice at six different temperatures (in blue): $T = 0.2$, $T = 0.36$ (below T_1), $T = 0.48$, $T = 0.96$ (between T_1 and T_2), $T = 1.12$, and $T = 1.6$ (above T_2), which cover three phases. We can see that the pattern of PDF changes from phase to phase. They can be compared to the PDFs of the corresponding ensemble states obtained by the NN flow (in red), which remains almost the same in all three phases.
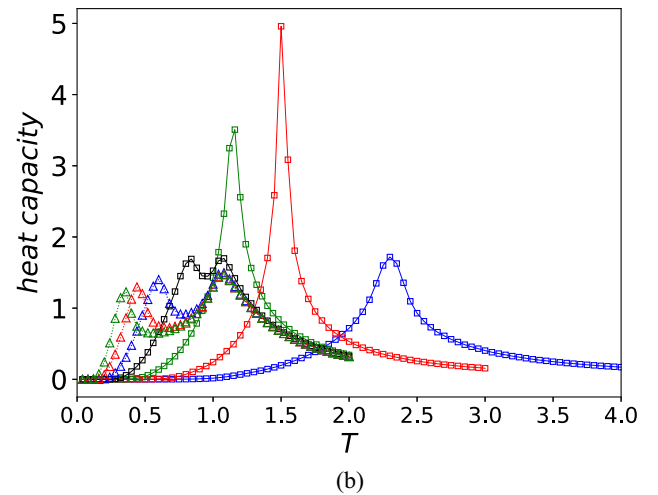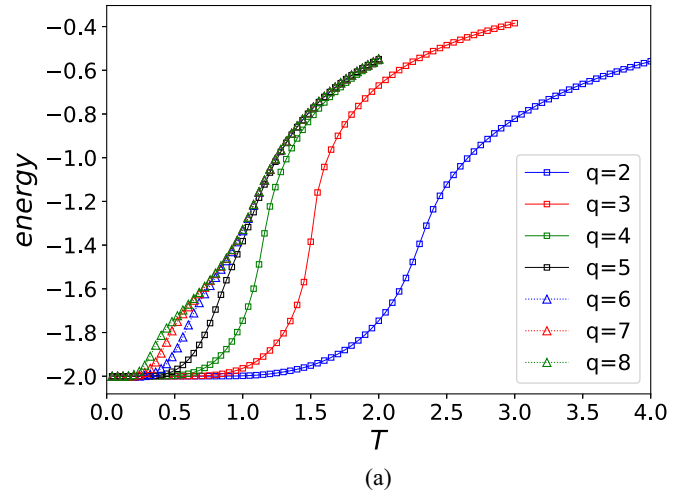

(a)

(b)

FIG. 14. (a) Energy and (b) heat capacity for $q = 2, 3, 4, 5, 6, 7, 8$ clock model is obtained from the Monte Carlo configurations on a 40×40 lattice.

By comparing the PDFs of the unflow states to the flow ones, we can see some interesting features and understand why JSD is minimal near the critical temperature, i.e., $(T_2)_{MC}$. The key point is that the overall shape of the flow PDFs is similar to the one of the disordered phase away from $(T_2)_{MC}$ [Fig. 13(f)] but with a shifted central position. This shift makes the value of the associated JSD larger than the ones near $(T_2)_{MC}$ [Figs. 13(d) and 13(e)]. This is in contrast to the case of $q = 4$ clock model as shown in Fig. 4. In that case, the flow PDFs do not resemble the one above T_c quite well, but with about the same central positions [Fig. 4(c)]. Despite that, we see that the flow PDFs in both $q = 4$ and $q = 8$ cases have the maximal overlap with the unflow PDFs at the critical temperatures so that the corresponding JSDs are minimal. Therefore, although the flow states of the BKT phases do not carry the topological features, the PDFs of their ensemble can still indicate the phase transition miraculously.

Finally, we shall mention that we have also studied the PDFs and thus of JSDs of energy or physical spins, but find that they cannot capture the BKT phase transition as the mean spin value can. However, we do not show these null results as there is no interesting implication to be drawn.
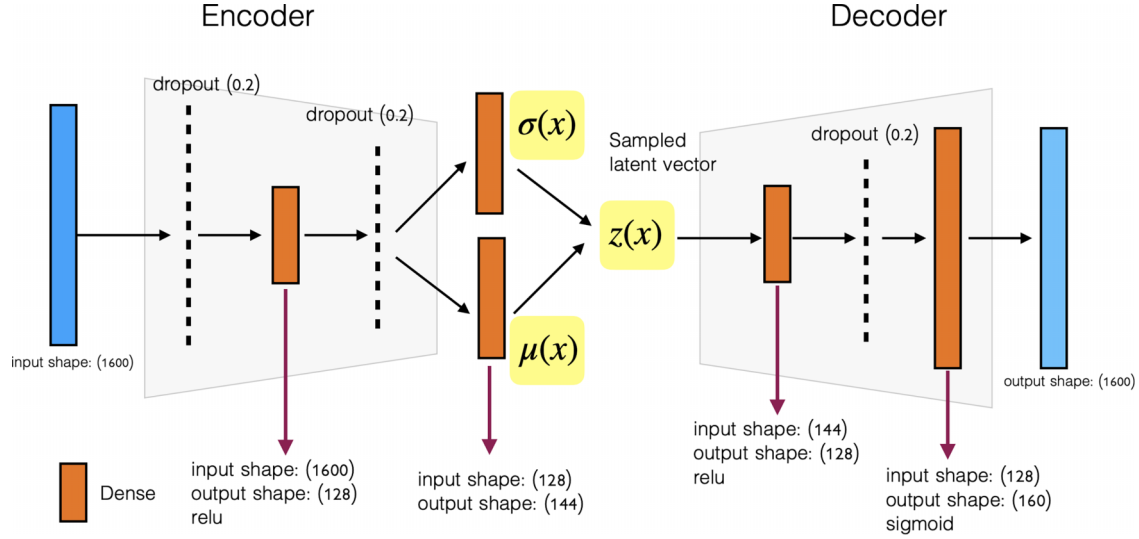
FIG. 15. A typical schematic architecture of a variational autoencoder for training the unit of NN flow with $L = 40$. The number inside the brackets indicates the dimension of each layer. In the middle are the hidden layers of VAE with dimensions indicated as 128 and 144, denoted as h_1 and h_2, respectively, in Fig. 15.

The deeper reason behind this is unclear and requires further investigation.

VI. CONCLUSION

In this paper we extend the neural flow (NN flow) based on the unsupervised machine learning of the variational autoencoder to study the BKT phase transitions of the q-state clock models. Due to the intricate nature of the BKT phase, we employ an information-distance measure, the Jensen-Shannon divergence (JSD), as the "thermometer" to gauge the NN-flow states with input Monte Carlo states. We find that the minimum of the JSD can pinpoint the temperatures of the ensembles far more precisely than the usual machine-learning thermometer. With the help of this JSD thermometer, we find that the NN-flow states from different finite-temperature Monte Carlo states form a fixed-point ensemble, which can capture the essence of the phase diagrams, such as the BKT phase and the associated critical temperatures of phase transitions.

However, the NN-flow states do not bear the feature of BKT states, but each looks more similar to the spontaneously-symmetry-broken states of the low-temperature phase. This

implies that the NN flow is not an analog to the RG flow. Despite that, the statistics of the mean spin value for evaluating the JSDs can indeed capture the essence of the BKT phase diagrams.

Our results demonstrate the power of machine learning in the study of topological phase transitions. The realization of the phase diagrams by the NN flow is quite different from the conventional way via coarse graining or RG flow. It is interesting to explore the deeper statistical structure of the NN-flow ensemble of states to uncover the mystery of the powerful NN flow method. After the demonstration of this work and the earlier ones [18–21], it is also crucial to look for the application of NN flow to other physical arenas with the similar critical phenomena so that NN flow can help to predict the locations of the phase transitions.
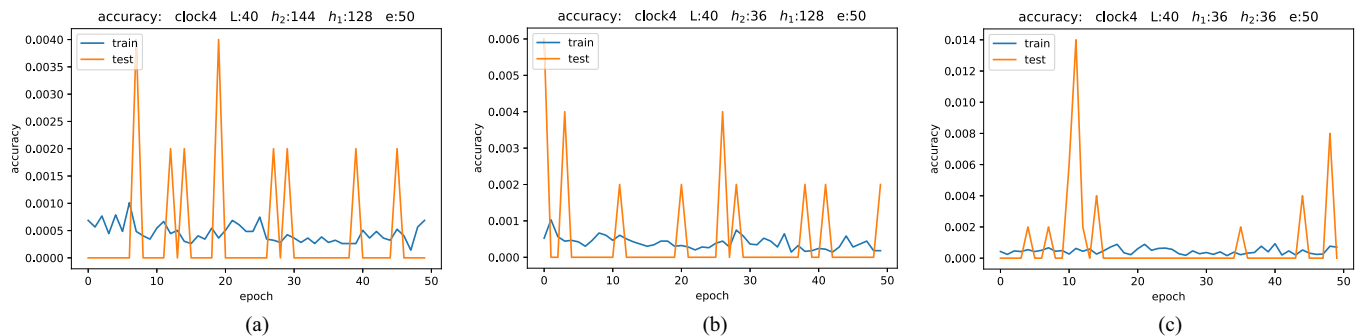
FIG. 16. Typical histograms of training accuracy for the NN model with different parameters, i.e., the dimensions of two hidden layers of VAE, h_1 and h_2 in Fig. 15: (a) $h_1 = 128$ and $h_2 = 144$; (b) $h_1 = 128$ and $h_2 = 36$; (c) $h_1 = 36$ and $h_2 = 36$.

APPENDIX A: DETAILS ON THE PHASE DIAGRAMS OF MONTE CARLO SIMULATED CONFIGURATIONS

For completeness and as complementary to Fig. 1, we show in this Appendix the phase diagrams of the $q = 2, 3, 4, 5, 6, 7, 8$ clock model based on the temperature profiles of (a) energy and (b) heat capacity. These results are obtained from the Monte Carlo configurations of 2000 samples for each temperature bin and are shown in Fig. 14. From the results, we find that it is more difficult to identify T_1's of the BKT phase transitions than T_2. The latter is signaled by a sharp rise or drop (magnetization and energy) or a peak (magnetic susceptibility and energy). The former, however, shows only a small jump or change.

APPENDIX B: SCHEMATIC STRUCTURE OF NN-FLOW UNIT AND TRAINING ACCURACY, AND THE RESULTANT INSENSITIVE JSD FEATURE

In this Appendix we demonstrate that the results of NN flow are quite independent of the hyperparameters of the training models as long as the training accuracy is low.

We take the $q = 4$ clock model as the example, and first display the details of a typical NN model for the NN-flow unit in Fig. 15, of which the dimensions of two hidden layers of the VAE are $h_1 = 128$ and $h_2 = 144$. The resultant training accuracy of this model is shown in Fig. 16(a), which is less than 1%. By changing the hyperparameters $h_{1,2}$, the training accuracy changes accordingly, as shown in Figs. 16(b) and 16(c), which are also lower than 2%.
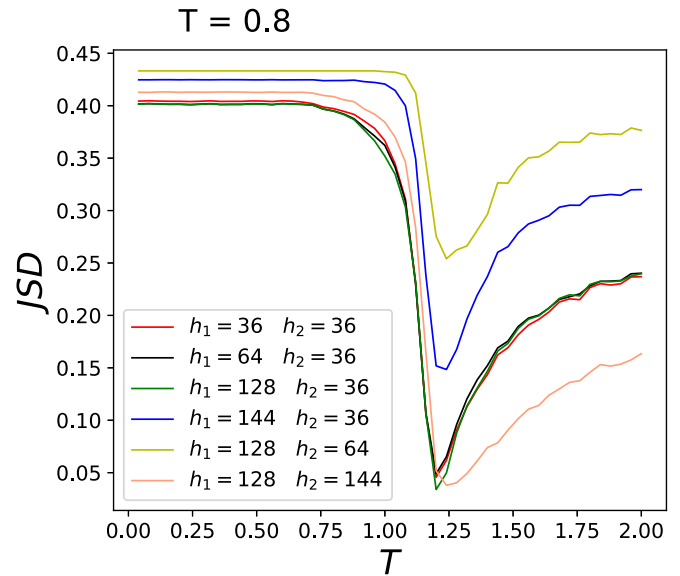

FIG. 17. NN-flow JSD$[P_{T_{\mathrm{flow}}}||Q_T]$ of the $q = 4$ clock model on a 40×40 lattice for the ensemble of NN-flow states after 100 iteration steps with different hyperparameters. Our results again show that NN flows can reach a fixed-point JSD pattern with its minimum located at the critical temperature. The minimum points seem to be independent of the hyperparameter of NN.

Finally, we show the resultant JSDs with different hyperparameters h_1 and h_2 in Fig. 17. We see that the patterns of JSD are quite independent of the hyperparameters, especially with the minima at the critical temperatures, although the overall scales are different. Therefore, we can conclude that identifying the phase transitions by NN flow is quite insensitive to the hyperparameters as long as the training accuracy is low. The low training accuracy ensures the input states will flow by the NN units to the ensemble of fixed-point states. Otherwise, the highly accurate VAE will reproduce the input state.

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

[2] M. R. Hush, Machine learning for quantum physics, Science **355**, 580 (2017).

[3] Z. Cai and J. Liu, Approximating quantum many-body wave functions using artificial neural networks, Phys. Rev. B **97**, 035116 (2018).

[4] G. Torlai, G. Mazzola, J. Carrasquilla, M. Troyer, R. Melko, and G. Carleo, Neural-network quantum state tomography, Nat. Phys. **14**, 447 (2018).

[5] P. Baldi, P. Sadowski, and D. Whiteson, Searching for exotic particles in high-energy physics with deep learning, Nat. Commun. **5**, 4308 (2014).

[6] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, Machine learning quantum phases of matter beyond the fermion sign problem, Sci. Rep. **7**, 8823 (2017).

[7] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, Machine learning and the physical sciences, Rev. Mod. Phys. **91**, 045002 (2019).

[8] C.-H. Liao and F.-L. Lin, Deep generative models of gravitational waveforms via conditional autoencoder, Phys. Rev. D **103**, 124051 (2021).

[9] H.-S. Kuo and F.-L. Lin, Conditional noise deep learning for parameter estimation of gravitational wave events, Phys. Rev. D **105**, 044016 (2022).

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

[11] Y. Zhang, R. G. Melko, and E.-A. Kim, Machine learning z_2 quantum spin liquids with quasiparticle statistics, Phys. Rev. B **96**, 245119 (2017).

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

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

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

[15] E. van Nieuwenburg, E. Bairey, and G. Refael, Learning phase transitions from dynamics, Phys. Rev. B **98**, 060301(R) (2018).

[16] C. Alexandrou, A. Athenodorou, C. Chrysostomou, and S. Paul, The critical temperature of the 2D-Ising model through Deep Learning Autoencoders, Eur. Phys. J. B **93**, 226 (2020).

[17] N. Käming, A. Dawid, K. Kottmann, M. Lewenstein, K. Sengstock, A. Dauphin, and C. Weitenberg, Unsupervised machine learning of topological phase transitions from experimental data, Mach. Learn.: Sci. Technol. **2**, 035037 (2021).

[18] S. Iso, S. Shiba, and S. Yokoo, Scale-invariant feature extraction of neural network and renormalization group flow, Phys. Rev. E **97**, 053304 (2018).

[19] S. S. Funai and D. Giataganas, Thermodynamics and feature extraction by machine learning, Phys. Rev. Res. **2**, 033415 (2020).

[20] E. De Mello Koch, R. De Mello Koch, and L. Cheng, Is deep learning a renormalization group flow? IEEE Access **8**, 106487 (2020).

[21] D. Giataganas, C.-Y. Huang, and F.-L. Lin, Neural network flows of low q-state Potts and clock models, New J. Phys. **24**, 043040 (2022).

[22] P. Suchsland and S. Wessel, Parameter diagnostics of phases and phase transition learning by neural networks, Phys. Rev. B **97**, 174435 (2018).

[23] S. S. Schoenholz, E. D. Cubuk, D. M. Sussman, E. Kaxiras, and A. J. Liu, A structural approach to relaxation in glassy liquids, Nat. Phys. **12**, 469 (2016).

[24] T. Ohtsuki and T. Mano, Drawing phase diagrams of random quantum systems by deep learning the wave functions, J. Phys. Soc. Jpn. **89**, 022001 (2020).

[25] P. Ronhovde, S. Chakrabarty, D. Hu, M. Sahu, K. F. Kelton, N. A. Mauro, K. K. Sahu, and Z. Nussinov, Detecting hidden spatial and spatio-temporal structures in glasses and complex physical systems by multiresolution network clustering, arXiv:1102.1519.

[26] K. A. Nicoli, S. Nakajima, N. Strodthoff, W. Samek, K.-R. Müller, and P. Kessel, Asymptotically unbiased estimation of physical observables with neural samplers, Phys. Rev. E **101**, 023304 (2020).

[27] P. Mehta, M. Bukov, C.-H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, A high-bias, low-variance introduction to Machine Learning for physicists, Phys. Rep. **810**, 1 (2019).

[28] J. Kosterlitz and D. Thouless, Ordering, metastability and phase transitions in two-dimensional systems, J. Phys. C **6**, 1181 (1973).

[29] V. L. Berezinsky, 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).

[30] V. L. Berezinsky, Destruction of long-range order in one-dimensional and two-dimensional systems possessing a continuous symmetry group. II. Quantum systems, Sov. Phys. JETP **34**, 610 (1972).

[31] J. M. Kosterlitz, The critical properties of the two-dimensional xy model, J. Phys. C **7**, 1046 (1974).

[32] J. V. Jose, L. P. Kadanoff, S. Kirkpatrick, and D. R. Nelson, Renormalization, vortices, and symmetry breaking perturbations on the two-dimensional planar model, Phys. Rev. B **16**, 1217 (1977).

[33] G. Ortiz, E. Cobanera, and Z. Nussinov, Dualities and the phase diagram of the p-clock model, Nucl. Phys. B **854**, 780 (2012).

[34] M. S. S. Challa and D. P. Landau, Critical behavior of the six-state clock model in two dimensions, Phys. Rev. B **33**, 437 (1986).

[35] Y. Tomita and Y. Okabe, Probability-changing cluster algorithm for two-dimensional XY and clock models, Phys. Rev. B **65**, 184405 (2002).

[36] R. Krčmár, A. Gendiar, and T. Nishino, Entanglement-entropy study of phase transitions in six-state clock model, Acta Physica Polonica A **137**, 598 (2020).

[37] C. Chatelain, DMRG study of the Berezinskii-Kosterlitz-Thouless transitions of the 2D five-state clock model, J. Stat. Mech. (2014) P11022.

[38] L. Vanderstraeten, B. Vanhecke, A. M. Läuchli, and F. Verstraete, Approaching the Kosterlitz-Thouless transition for the classical XY model with tensor networks, Phys. Rev. E **100**, 062136 (2019).

[39] Z.-Q. Li, L.-P. Yang, Z. Y. Xie, H.-H. Tu, H.-J. Liao, and T. Xiang, Critical properties of the two-dimensional q-state clock model, Phys. Rev. E **101**, 060105(R) (2020).

[40] C. M. Lapilli, P. Pfeifer, and C. Wexler, Universality Away from Critical Points in Two-Dimensional Phase Transitions, Phys. Rev. Lett. **96**, 140603 (2006).

[41] S. Elitzur, R. B. Pearson, and J. Shigemitsu, Phase structure of discrete abelian spin and gauge systems, Phys. Rev. D **19**, 3698 (1979).

[42] K. Shiina, H. Mori, Y. Okabe, and H. K. Lee, Machine-learning studies on spin models, Sci. Rep. **10**, 2177 (2020).

[43] Y. Tomita, K. Shiina, Y. Okabe, and H. K. Lee, Machine-learning study using improved correlation configuration and application to quantum Monte Carlo simulation, Phys. Rev. E **102**, 021302(R) (2020).

[44] F. Y. Wu, The Potts model, Rev. Mod. Phys. **54**, 235 (1982).

[45] Y. Miyajima, Y. Murata, Y. Tanaka, and M. Mochizuki, Machine learning detection of Berezinskii-Kosterlitz-Thouless transitions in q-state clock models, Phys. Rev. B **104**, 075114 (2021).

[46] R. H. Swendsen and J.-S. Wang, Nonuniversal Critical Dynamics in Monte Carlo Simulations, Phys. Rev. Lett. **58**, 86 (1987).

[47] U. Wolff, Collective Monte Carlo Updating for Spin Systems, Phys. Rev. Lett. **62**, 361 (1989).

[48] F. Pérez-Cruz, Kullback-Leibler divergence estimation of continuous distributions, in *Proceedings of the 2008 IEEE International Symposium on Information Theory* (IEEE, New York, 2008), pp. 1666–1670.

[49] J. Lin, Divergence measures based on the Shannon entropy, IEEE Trans. Inf. Theory **37**, 145 (1991).

[50] F. Nielsen, On the Jensen–Shannon symmetrization of distances relying on abstract means, Entropy **21**, 485 (2019).