



Machine learning phases and criticalities without using real data for trainingD.-R. Tan  and F.-J. Jiang *¹*Department of Physics, National Taiwan Normal University, 88, Sec. 4, Ting-Chou Road, Taipei 116, Taiwan*

(Received 14 August 2020; revised 2 December 2020; accepted 7 December 2020; published 28 December 2020)

We study the phase transitions of three-dimensional (3D) classical $O(3)$ model and two-dimensional (2D) classical XY model, as well as both the quantum phase transitions of 2D and 3D dimerized spin-1/2 antiferromagnets, using the technique of supervised neural network (NN). Moreover, unlike the conventional approaches commonly used in the literature, the training sets employed in our investigation are neither the theoretical nor the real configurations of the considered systems. Remarkably, with such an unconventional set up of the training stage in conjunction with some semiexperimental finite-size scaling formulas, the associated critical points determined by the NN method agree well with the established results in the literature. The outcomes obtained here imply that certain unconventional training strategies, like the one used in this study, are not only cost-effective in computation but are also applicable for a wide range of physical systems.

DOI: [10.1103/PhysRevB.102.224434](https://doi.org/10.1103/PhysRevB.102.224434)**I. INTRODUCTION**

The applications of machine learning (ML) methods and techniques to the studies of many-body systems have recently inspired the communities of physics, applied physics and physical chemistry. Moreover, many important and exciting achievements have been obtained using the ML approach in the last a few years [1–86]. Among these achievements, first principles calculations of properties of materials and analyzing the signals from colliders in high energy physics are two such notable examples. Yet another significant accomplishment is the success of investigating critical phenomena using both the supervised and unsupervised neural networks (NN).

By employing the dedicated convolutional neural network techniques (CNN) which can capture certain characteristics of the studied models, it has been demonstrated that the phase transitions associated with many classical and quantum systems, including the Ising model, the XY models, as well as the Hubbard model have been studied with various extent of satisfaction. Because of these numerous successful examples mentioned, it is optimistically believed that with the ideas of ML one may be able to uncover features of certain systems that cannot be obtained by the conventional methods. Even those days, seeking devoted ML techniques to surpass the success that the traditional approaches can reach is still vigorous.

The standard procedure, i.e., the most considered scheme, of investigating the phase transitions of physical models by supervised NN consists of three steps [22,77,80], namely, the training, the validation, and the testing stages. Among these three stages, the training is the most flexible one and various strategies have been used for this step [22,29,48,86]. Typically real configurations of the studied systems obtained from certain numerical methods are employed as the training sets. In addition, the training has been applied to various chosen temperatures T (relevant parameter) across the tran-

sition temperature T_c (critical point). This indicates that in principle T_c (or the critical point) should be known in advance before one can employ the NN techniques. Such a training approach has led to success in studying the critical phenomena associated with several many-body systems such as the Ising and the Hubbard models. Other schemes for which the locations of the critical points are not required are introduced as well [48,79,86,87]. For instance, the method of using the theoretical ground-state configurations in the ordered phase as the training sets are demonstrated to be valid for ferromagnetic and antiferromagnetic Potts models. Although the details of the confusion scheme outlined in Ref. [29] are different from the method of using the theoretical ground states as the training set, it is yet another approach which does not require the knowledge of the critical points and needs fewer degree of freedom to carry out the NN calculations. For the readers who are interested in the details of these training processes, see Refs. [22,29,48,77,80,86].

The strategy of considering the theoretical ground states in the ordered phase as the training sets requires only one training and the knowledge of the associated critical point(s) is not needed. This approach has been applied to both the ferromagnetic and the antiferromagnetic Potts models, and the obtained outcomes show that the idea is effective [48,86]. In particular, numerical evidence strongly suggests that with this method, the computational demanding for the training stage is tremendously reduced, and its applicability is broad.

Despite the NN results of estimating the critical points associated with the Potts models, using the method of considering the ground-state configurations in the ordered phase as the training sets, are impressive, an interesting question arises. Specifically, is this approach applicable for studying the zero temperature phase transitions of quantum spin systems, as well as the phase transitions of models with continuous variables such as the classical $O(3)$ model?

To answer the crucial question outlined in the previous paragraph, here we study the phase transitions of three-dimensional (3D) classical $O(3)$ (Heisenberg) model and

*fjjiang@ntnu.edu.tw

two-dimensional (2D) classical XY model, as well as the quantum phase transitions of 2D and 3D dimerized spin-1/2 antiferromagnets, using the simplest deep learning neural network, namely, the multilayer perceptron (MLP). In particular, unlike the conventional or the unconventional training procedures introduced previously, in this investigation, we have adopted an alternative strategy for the training. In particular, the training sets employed here belong to neither the theoretical nor the real configurations of the considered systems.

The motivation for using the simplest deep learning NN in the study is that whether a NN idea is valid or not should not depend on the detailed infrastructure of the built NN. Hence, a MLP made up of only three layers is employed here. One can definitely consider a more complicated (and dedicated as well) NN such as CNN for the associated investigations. This will be left for future work.

Remarkably, even using the extraordinary training sets mentioned above and some semiexperimental finite-size scaling (which will be introduced later), the constructed MLP can effectively detect the critical points of all the studied classical and quantum physical systems. The intriguing outcomes obtained here strongly suggest that the approach of investigating the targeted physical systems before employing any objects for the training, such as those done here and in Refs. [48,86], is not only cost-effective in computation, but also leads to accurate determination of the associated critical points. Finally, it is amazing that the simple procedure described here is not only valid for studying the phase transitions associated with spontaneous symmetry breaking (SSB), but also works for those related to topology.

This paper is organized as follows. In Sec. II, the studied microscopic models and the employed NN are described. In particular, the NN training sets and labels are introduced thoroughly. In Sec. III the resulting numerical results determined by applying the NN techniques are presented. Finally, Sec. IV concludes our investigation.

II. THE MICROSCOPIC MODELS AND OBSERVABLES

A. The 3D classical $O(3)$ (Heisenberg) and 2D classical XY model

The Hamiltonian $H_{O(3)}$ of the 3D classical $O(3)$ (Heisenberg) model on a cubical lattice considered in our study is given by

$$\beta H_{O(3)} = -\beta \sum_{\langle ij \rangle} \vec{s}_i \cdot \vec{s}_j, \quad (1)$$

where β is the inverse temperature and $\langle ij \rangle$ stands for the nearest neighboring sites i and j . In addition, in Eq. (1) \vec{s}_i is a unit vector belonging to a 3D sphere S^3 and is located at site i .

Starting with an extremely low temperature, as T rises, the classical $O(3)$ system will undergo a phase transition from an ordered phase, where majority of the unit vectors point toward the same direction, to a disordered phase for which these mentioned vectors are oriented randomly. Relevant observables used here to signal out the phenomenon of this phase transition are the first and the second Binder ratios (Q_1 and

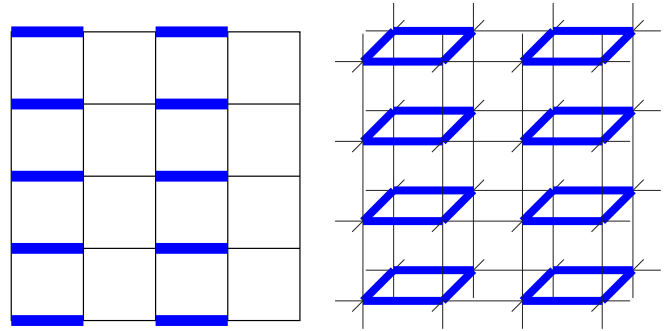


FIG. 1. The studied dimerized quantum antiferromagnetic Heisenberg models: 2D ladder (left) and 3D plaquette (right) models. The bold and thin bonds shown in both sub-figures represent J' and J couplings, respectively.

Q_2) defined by

$$Q_1 = \langle |m|^2 \rangle / \langle m^2 \rangle, \quad (2)$$

$$Q_2 = \langle m^2 \rangle^2 / \langle m^4 \rangle, \quad (3)$$

where $m = \frac{1}{L^3} \sum_i \vec{s}_i$ and L is the linear box size of the system [88].

The Hamiltonian of the 2D classical XY model on the square lattice has the same expression as $H_{O(3)}$, except that the corresponding unit vector \vec{s}_i at site i belongs to a (2D) circle instead of a 3D sphere.

B. The 2D and 3D dimerized quantum antiferromagnetic Heisenberg models

The 2D and 3D dimerized quantum antiferromagnetic Heisenberg model share a similar form of Hamiltonian given as

$$H = \sum_{\langle i,j \rangle} J_{ij} \vec{S}_i \cdot \vec{S}_j, \quad (4)$$

where again $\langle i, j \rangle$ stands for the nearest neighboring sites i and j , $J_{ij} > 0$ is the associated antiferromagnetic coupling (bond) connecting i and j , and \vec{S}_i is the spin-1/2 operator located at i . The cartoon representation, in particular the spatial arrangement of the antiferromagnetic couplings, of the studied models are shown in Fig. 1 (In this study, these quantum spin models will be called 2D ladder and 3D plaquette models if no confusion arises). From the figure one sees that, as the ratios J'/J (of both models) being tuned, quantum phase transitions from ordered to disordered states will take place in these models when $g := J'/J$ exceed certain values g_c . Relevant observables considered in our investigation for studying the quantum phase transitions are again the first and the second Binder ratios described above. For the studied spin-1/2 systems, Q_1 and Q_2 have the following definitions:

$$Q_1 = \langle |M_s|^2 \rangle / \langle M_s^2 \rangle, \quad (5)$$

$$Q_2 = \langle M_s^2 \rangle^2 / \langle M_s^4 \rangle, \quad (6)$$

$$M_s = \frac{1}{L^d} \sum_i (-1)^{i_1 + \dots + i_d} S_i^z, \quad (7)$$

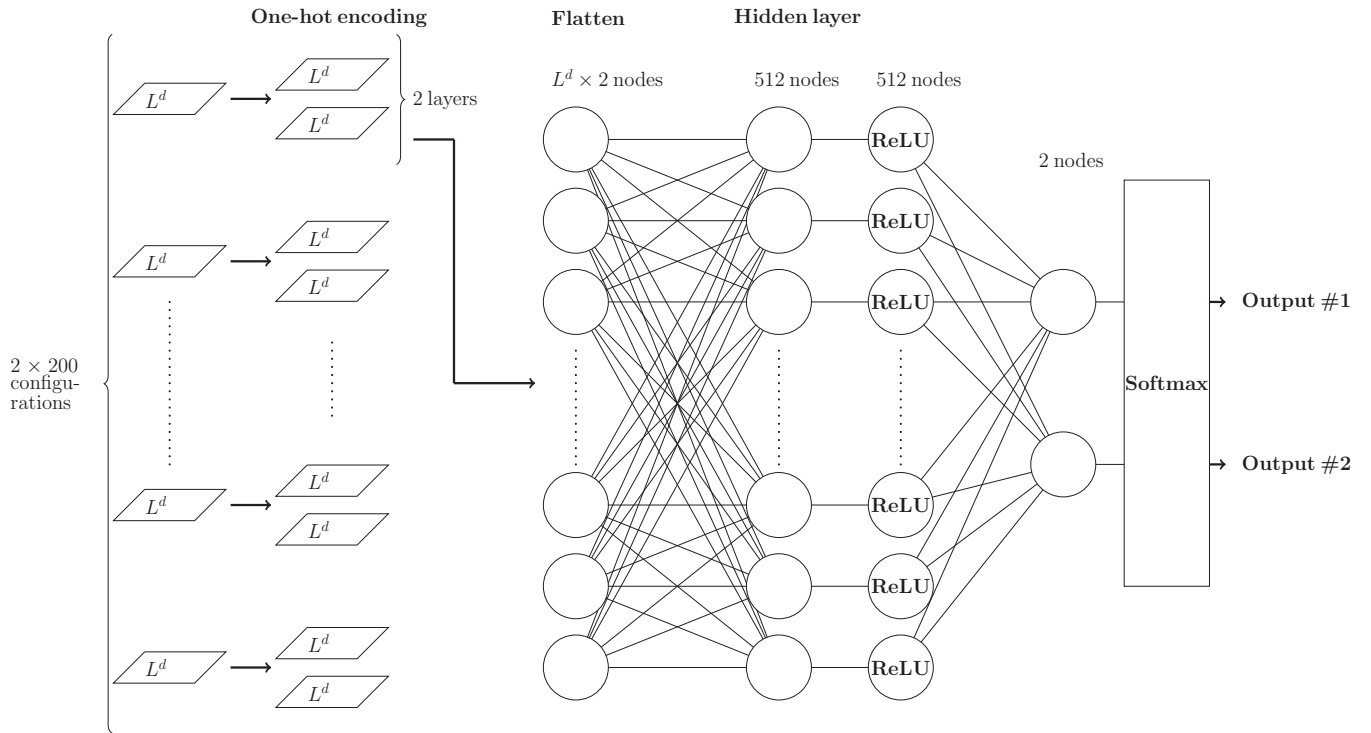


FIG. 2. The NN (MLP), which consists of one input layer, one hidden layer, and one output layer, used here and in Ref. [86]. In the figure d is the dimensionality of the considered system. In addition, the objects in the input layer are made up of 200 copies of only two configurations for all the studied models. Finally, there are 512 (or 1024) nodes in the hidden layer and each of these nodes is independently connected to every object in the input layer. Before each training object is connected to the nodes in the hidden layer, the steps of one-hot encoding and flatten are applied. The activation functions (ReLU and softmax) and where they are employed are demonstrated explicitly. For all the considered systems, the output layers consist of two elements.

where d is the dimensionality of the studied models.

These mentioned g_c of the quantum spin systems, as well as the T_c of the 3D classical $O(3)$ and 2D classical XY models introduced previously, have been calculated with high accuracy in the literature [89–93].

III. THE CONSTRUCTED SUPERVISED NEURAL NETWORKS

In this section, we will review the supervised NN, namely, the multilayer perceptron (MLP) used in our study. The employed training sets and the associated labels for the studied models will be described as well.

A. The built multilayer perceptron (MLP)

The MLP used in our investigation is already detailed in Ref. [86]. Specifically, using the NN library keras [94], we construct a supervised NN which consists of only one input layer, one hidden layer of 512 (or 1024) independent nodes, and one output layer.

The pictorial representation of the constructed MLP, including the steps of one-hot encoding (which is natural and somehow essential for a NN with multi-component outputs) and flatten is shown in Fig. 2. The motivation for such a construction will be detailed later when the training strategies for the studied four models are introduced. Here we would like to point out that the output of the built NN is a two-component

vector. The norm R of such a vector \vec{v} , which is defined by

$$R(\vec{v}) = \sqrt{v_1^2 + v_2^2}, \quad (8)$$

will be employed to investigate the considered phase transitions. As we will demonstrate later, by studying the T (g) dependence of R , the associated critical points T_c (g_c) can be estimated with high precision.

A typical NN investigation consists of three stages, namely, the training, the validation, and the testing stages. Here we will focus on describing the training procedures employed in this study since the ideas behind the other two stages are similar to that of the training stage.

First of all, 200 identical copies of two certain configurations, which will be introduced later, are used as the training set. Second of all, the steps of one-hot encoding and flatten are performed. Figure 2 contains the details of how these two steps are done. The main algorithm adopted in the NN calculations is the mini-batch. As a result, objects in the training set are permuted randomly before the optimization procedure is conducted. Therefore if one arbitrarily picks up two different batches, it is unlikely that they are identical in all aspects. Figure 3 is a cartoon representation of how the mini-batch algorithm works in reality.

Each node in the hidden layer is connected independently to every object resulting from the procedures of one-hot encoding and flatten. Such connections are named weights. Every node (of the hidden layer) is also linked to each element of

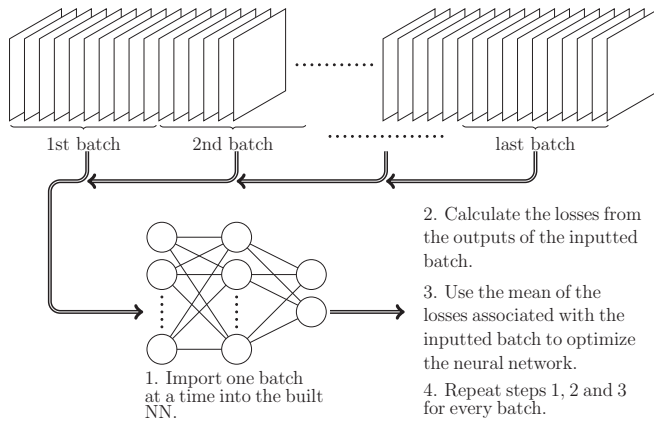


FIG. 3. The mini-batch algorithm employed in the NN calculations.

the output layer, after being activated by the ReLU function (see Figs. 2 and 4). These links are called weights as well. Finally, the softmax function σ ($\sigma(x_i) = \exp(x_i) / \sum_j \exp(x_j)$ for $i = 1, 2, 3, \dots$) is applied to the output layer so that every output vector $\vec{v} = (v_1, v_2)$ has the property of $v_1 + v_2 = 1$.

The Adam optimizer is employed to update the weights so that the loss function (defined later) will reach a minimum. The loss function used is the categorical crossentropy C which is given by

$$-\frac{1}{n} \sum_x \sum_j^2 y_j \ln a_j, \quad (9)$$

where n is the total number of objects contained in the input set (for the training stage, n is the number of objects included in each batch). a_j are the outcomes obtained after applying all the layers. Moreover, x and y appearing above are the training inputs and the corresponding designed outputs (labels), respectively. One cycle of training all batches is called an epoch. At least 800 epochs is used in our NN calculations. Finally, in our study the learning rate η is the default one, and L_2 regularization is applied as well to avoid overfitting.

For the four studied models, results calculated using 10 sets of random seeds are all taken into account when presenting the final outcomes. Specifically, each NN calculation leads to a value of R , and the R as well as its uncertainty quoted in this study are obtained from these 10 results of R . The motivation of considering 10 sets of random seeds is to examine whether the initial weights as well as other elements employed in the training stage have any nonnegligible influence on the final outcomes. We would also like to point out that in the testing stage, each of these 10 calculations uses the same set of configurations produced from the Monte Carlo simulations. Later we will come back to this and make a comment about it.

B. Training set and output labels for the 3D classical $O(3)$ and 2D classical XY models

Regarding the training sets employed in the calculations, instead of using real configurations obtained from simulations or the theoretical ground states in the ordered phase of the considered system, here we use a slightly different alternative.

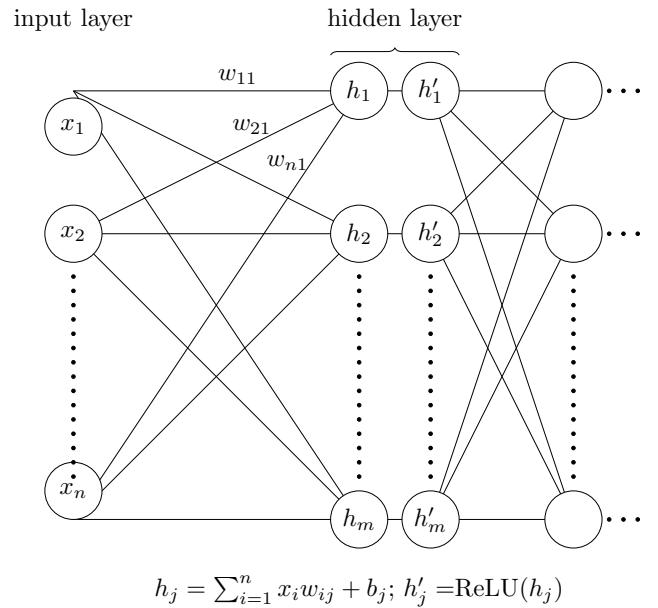
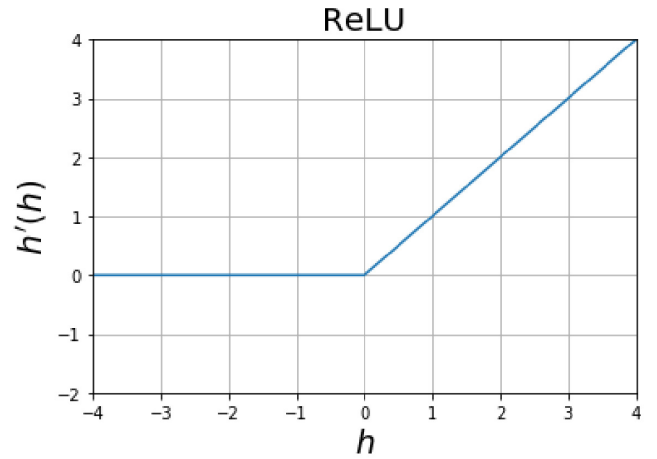


FIG. 4. The activation function ReLU.

Specifically, to train the NN on a L by L by L cubical lattice for the 3D classical $O(3)$ model, the training set consists of only two configurations. In addition, 0 is assigned to every site of one configuration and the other configuration is made up by giving each of its sites the value of 1. As a result, the output labels are the vectors of (1,0) and (0,1). Configurations and labels obtained using the same ideas as that of the 3D $O(3)$ model are employed for the 2D classical XY model. The motivation for considering such training sets will be explained in next subsection.

C. The expected output vectors for the 3D classical $O(3)$ and 2D classical XY models at various T

It should be pointed out that an $O(3)$ configuration is specified completely by the associated two parameters, namely, θ and ψ at each site of the underlying cubical lattice. At an extremely low temperature T , all the unit vectors of a $O(3)$ configuration point toward a particular direction. Under such a

circumstance, $\psi \bmod \pi$ is either 0 or 1 for every unit vector [of an $O(3)$ configuration]. The employed training set described in the previous subsection is motivated by this observation. As a result, the magnitude R of the output vector for a ground state $O(3)$ configuration is 1. When the temperature rises, one expects that R diminishes with T and for $T \geq T_c$, R takes its possible minimum value $1/\sqrt{2}$. Consequently, studying the magnitude of the NN output vectors as a function of T can reveal certain relevant information of T_c .

Here we would like to emphasize the fact that ψ rather than θ is considered in our investigation. This is because for any two given fixed values of ψ , their related arc length on the 3D unit sphere is the same. For θ , this is not the case. Therefore, with ψ one should arrive at more accurate outcomes.

The same scenario described above for the 3D classical $O(3)$ model applies to the 2D classical XY model as well.

D. Training set and output labels for the 2D and 3D quantum spin models

For the 2D and 3D dimerized quantum antiferromagnetic Heisenberg models investigated here, their associated classical ground-state configurations (in the ordered phase) are adopted as the training sets. Specifically, the training set for each of these two models consists of two configurations. Moreover, the spin value of every lattice site is either 1 or -1 and they are arranged alternatively. In other words, for a site which has a spin value 1 (-1), the spin values for all of its nearest neighbor sites are -1 (1). With such set up of the training sets, the employed output vectors should be (1,0) and (0,1) naturally.

We would like to emphasize the fact that the training sets considered for the studied 2D and 3D quantum spin models are not even among any of the possible ground-state configurations of these two systems.

E. The expected output vectors for the 2D and 3D dimerized quantum spin models at various g

Due to quantum fluctuations, it is not possible to assign any definite spin configurations for these investigated quantum spin models when $g = 1$ and $g > g_c$. Therefore, how the corresponding output vectors behave with respect to the dimerization strength g will be treated classically here. Consequently, R should be 1 and $1/\sqrt{2}$ for $g = 1$ and $g \geq g_c$, respectively. As we will demonstrate shortly, the R (magnitude) of the outputs associated with the NN studies of these quantum spin systems follow these rules (i.e., the values of R are 1 and $1/\sqrt{2}$ for $g = 1$ and $g > g_c$, respectively) in a satisfactory manner, hence lead to fairly good estimations of the critical points.

IV. THE NUMERICAL RESULTS

The configurations associated with the considered systems, namely, the 3D (classical) $O(3)$, the 2D classical XY , the 3D plaquette, as well as the 2D ladder models are generated by the Wolff and the stochastic series expansion (SSE) algorithms [92,95,96]. In addition, for each of the studied model, the corresponding configurations are recorded once in (at least) every two thousand Monte Carlo sweeps after the thermalization,

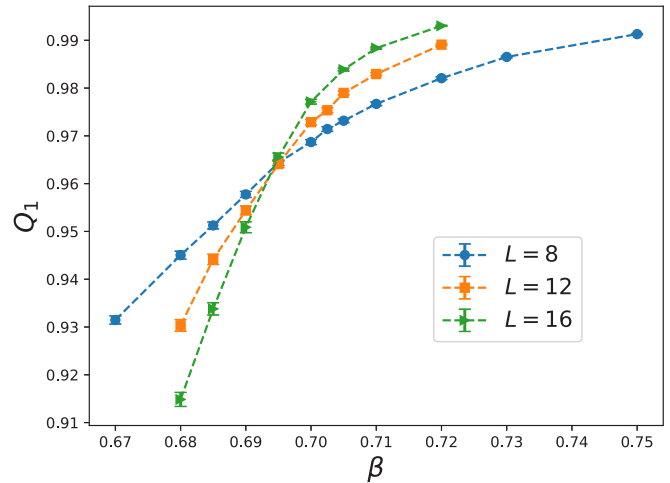


FIG. 5. Q_1 as functions of β for the 3D classical $O(3)$ model.

and at least one thousand configurations are produced. These spin compositions are then used for the calculations of NN. A semiexperimental finite-size scaling, which is adopted to estimate the critical points, will be introduced as well in this section.

A. Results of 3D classical $O(3)$ model

In Fig. 5, the observable first Binder ratio Q_1 are considered as functions of β for $L = 8, 12, 16$. As can be seen from the figure, the curves corresponding to various L intersect at a value of β close to the predicted critical point $\beta_c = 0.6929$ [89,90].

For a $O(3)$ configuration obtained from the simulation, all the S^3 vectors associated with it are converted to $\psi \bmod \pi$ and the resulting configuration is then fed into the trained NN, see Fig. 6 for a cartoon representation of such a conversion.

Density plots for one x - y layer of two $O(3)$ classical spin states as well as their corresponding $\psi \bmod \pi$ configurations are shown in Figs. 7 and 8. The outcomes presented in these figures are obtained at the inverse temperatures $\beta = 5.0$ (Fig. 7) and $\beta = 0.05$ (Fig. 8). Significant difference of ψ

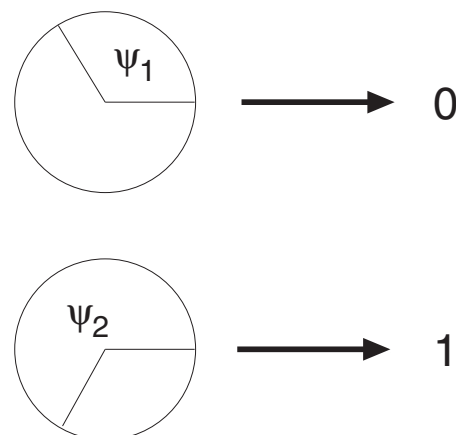


FIG. 6. Conversion of ψ associated with the $O(3)$ model into the used discrete variables 0 and 1.

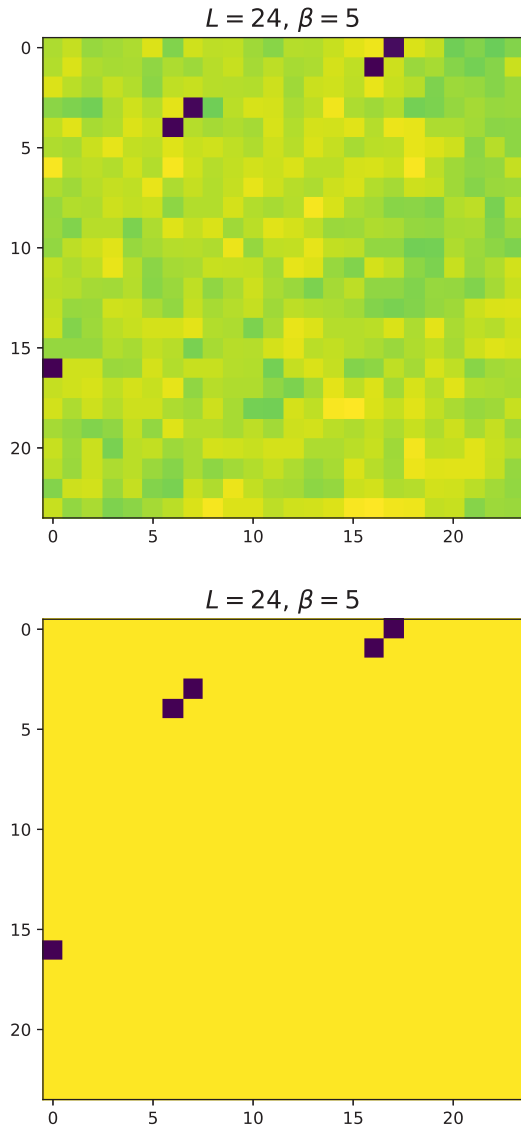


FIG. 7. Density plots of one x - y layer of one $O(3)$ classical spin state (top) as well as the corresponding $\psi \bmod \pi$ configuration (bottom). The results are obtained at $\beta = 5.0$.

mod π configurations as demonstrated in the figures implies that a phase transition occurs at an inverse temperature between $\beta = 5.0$ and $\beta = 0.05$.

It should be pointed out that due to fluctuations, occasionally configurations obtained at $\beta = 5.0$ will show the feature that appears in Fig. 8.

R as functions of β for $L = 8$ and $L = 20$ are shown in Fig. 9. While it is clear that both panels of Fig. 9 imply R change rapidly close to $\beta_c = 0.6929$, β_c cannot be calculated unambiguously when only the information of R is available.

If one assumes that R diminishes linearly with β in the critical region, then β_c can be approximately estimated by the intersection of the curves of R and $1/\sqrt{2} + 1 - R$. Such an idea has been used in Ref. [86] to calculate the T_c of the 3D 5-state ferromagnetic Potts model as well as the g_c of the 3D plaquette model (the latter will be studied in more detail here). Here we adopt a more appropriate approach for the determination of the considered critical points by taking into account the

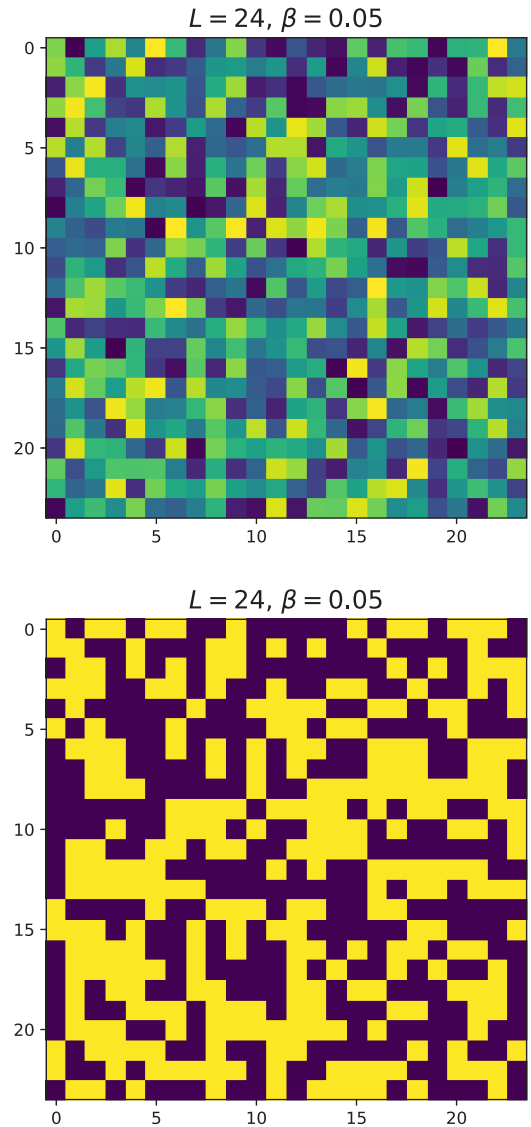


FIG. 8. Density plots of one x - y layer of one $O(3)$ classical spin state (top) as well as the corresponding $\psi \bmod \pi$ configuration (bottom). The results are obtained at $\beta = 0.05$.

deviation between the theoretical and the calculated R . Notice that the designation of the training set is based on the ground-state configurations of the considered models. In other words, they are indeed related to the associated ground-state configurations, but not directly. With such a set up, one would expect that in the testing stage, the R corresponding to configurations obtained at extremely low temperatures should reach the value of 1. In reality, calculation done at a very low temperature takes huge amount of computing resource. Hence, an overall shift Δ , which is the difference between 1 and the R associated with the largest β , is conducted [97]. With such a strategy $R + \Delta$ at the simulated lowest temperature take the value of 1. It should be noted that the simulated lowest temperature T_{low} should have the property that the R closed to T_{low} should (almost) saturate to a constant. Although such an approach will certainly introduce some systematic uncertainties to the determination of T_c (g_c), it turns out that the critical points estimated by this approach are in good agreement with the

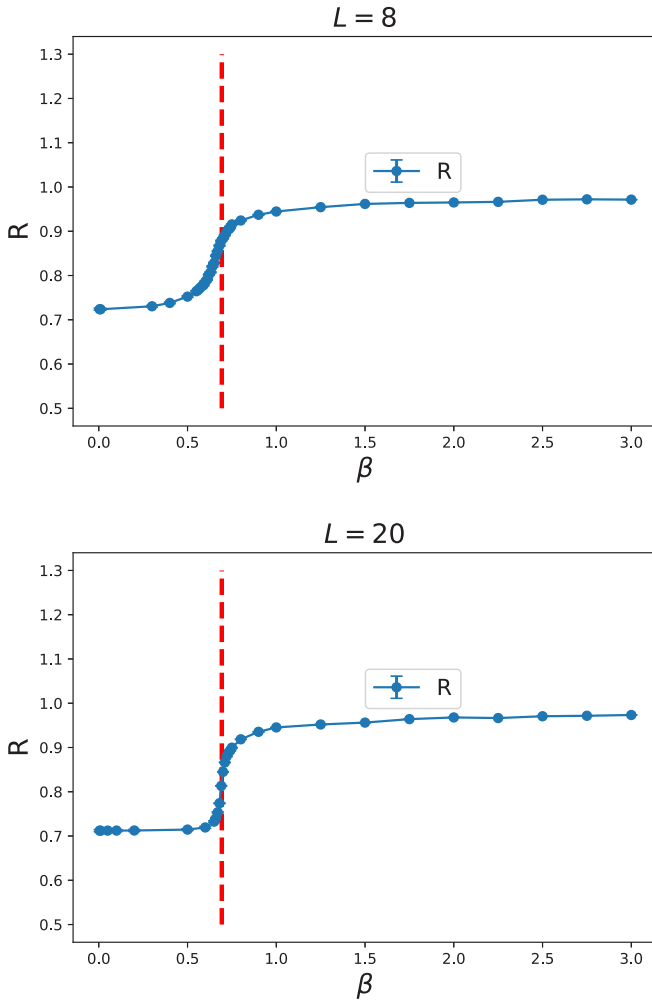


FIG. 9. R as functions of β for the 3D classical $O(3)$ model. The top and bottom panels are for $L = 8$ and $L = 20$, respectively.

known outcomes in the literature. Later we will come back to discuss the option of choosing a relevant temperature to calculate Δ .

Figures 10 and 11 demonstrate the associated curves made up of considering the data of $R + \Delta$ and $1/\sqrt{2} + 1 - R - \Delta$ as functions of β for $L = 4, 8, 12, 20, 24$. As can be seen from the figures, the intersections of these two curves for all the L (except the one of $L = 4$) are in good agreement with the theoretical prediction $\beta_c \sim 0.693$ (which are the vertical dashed lines in these figures). While for large L , the estimated values of β at which the mentioned two curves intersect are slightly away from $\beta_c = 0.6929$, the results shown in Figs. 10 and 11 indicate that the idea of estimating β_c by considering the intersection of the curves associated with $R + \Delta$ and $1/\sqrt{2} + 1 - R - \Delta$ is an effective approach. In particular, considering the simplicity of both the training procedure and the semiexperimental method of calculating β_c (for any finite L) employed in this study, the achievement reaches here for the determination of the T_c of the complicated 3D classical $O(3)$ model is remarkable.

The success of calculating the T_c of 3D classical $O(3)$ model through the idea of only considering $\psi \bmod \pi$ indicates

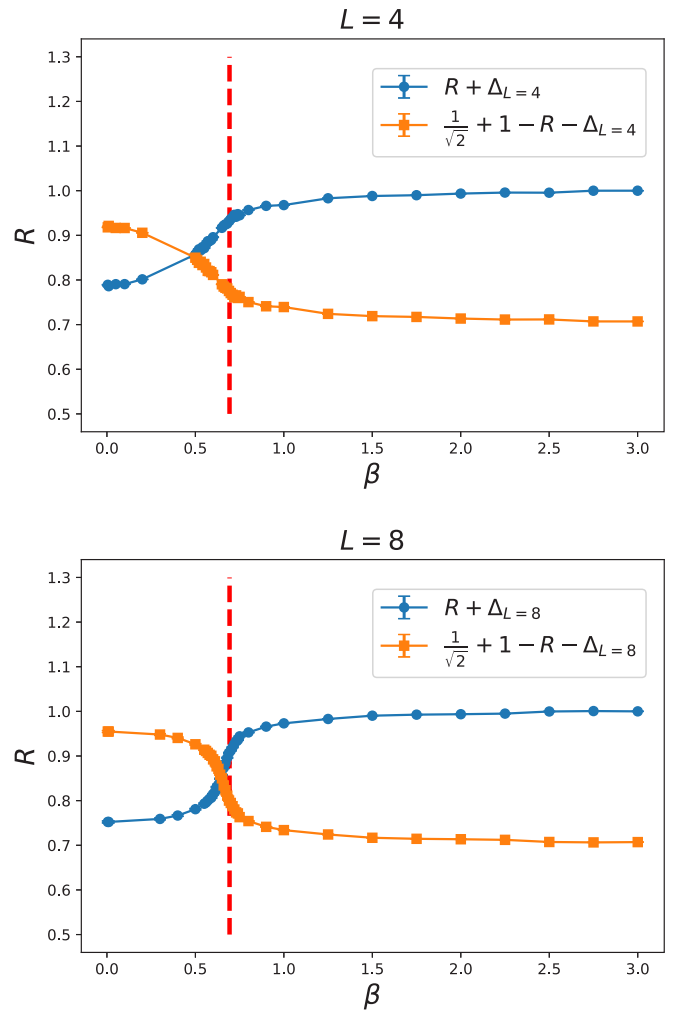


FIG. 10. $R + \Delta$ and $1/\sqrt{2} - \Delta + 1 - R$ as functions of β for the 3D classical $O(3)$ model. The top and bottom panels are for $L = 4$ and $L = 8$, respectively.

that partial information of the model is sufficient to estimate its associated critical point accurately.

It is worthy to mention that an equivalent method of calculating the critical point is to look for the temperature T_1 so that $R + \Delta$ takes the value of $(1 + 1/\sqrt{2})/2$ at T_1 . Then T_c is just T_1 . Figure 12 demonstrates the determination of T_c for the $O(3)$ model using this idea. In Fig. 12 the vertical dashed and horizontal solid lines represent the expected T_c and $(1 + 1/\sqrt{2})/2$. Clearly the intersection of R and $(1 + 1/\sqrt{2})/2$ matches T_c very well. This indicates that the described method indeed can be employed to determine T_c with high accuracy.

To calculate the critical points of the studied models with high precision using the intersections described above, one may apply certain forms of finite-size scaling to those crossing points. Based on the outcomes demonstrated in Figs. 10 and 11, it is clear that the R associated with the 3D classical $O(3)$ model receives mild finite-size effect. Apart from this, accurate determination of the crossing places in the relevant parameter space, particularly high precision estimated uncertainties for these crossing points, is needed to

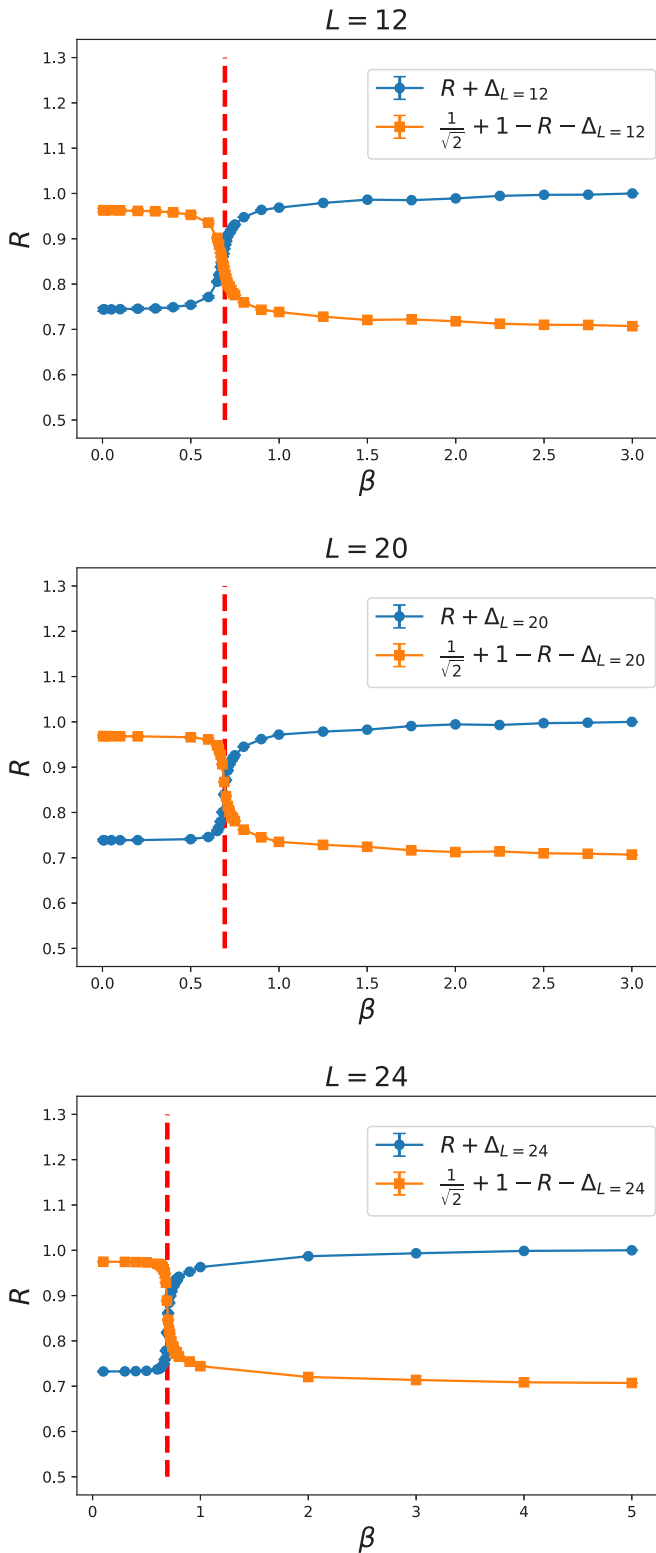


FIG. 11. $R + \Delta$ and $1/\sqrt{2} - \Delta + 1 - R$ as functions of β for the 3D classical $O(3)$ model. The top, middle, and bottom panels are for $L = 12$, $L = 20$, and $L = 24$, respectively.

carry out the fits. Hence, we postpone such an analysis to a latter subsection where 2D three-state ferromagnetic Potts model and 2D classical XY models on the square lattices are discussed.

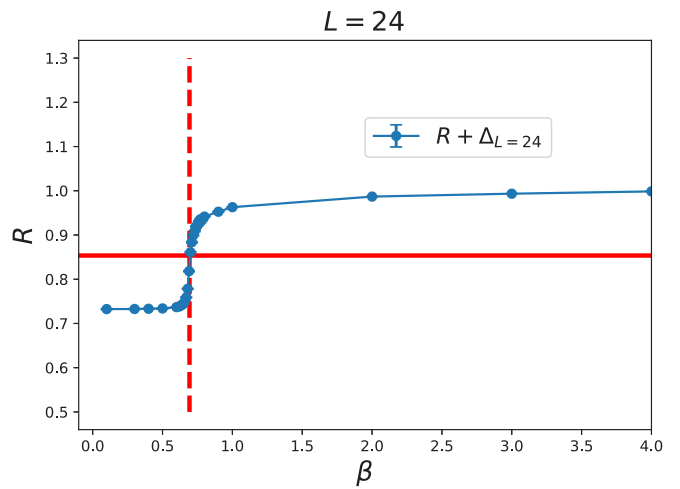


FIG. 12. $R + \Delta$ as functions of β for the 3D classical $O(3)$ model. The vertical dashed and horizontal solid lines represent the expected T_c and $(1 + 1/\sqrt{2})/2$.

B. Results of 2D quantum spin system

The first Binder ratio Q_1 close to g_c for the studied 2D dimerized spin-1/2 antiferromagnet (2D ladder model) are shown in Fig. 13. Similar to the case of 3D classical $O(3)$ model, various curves of large L tend to intersect at a value of g around 1.9. The estimated intersection $g \sim 1.9$ matches nicely with the known result $g_c = 1.90948(5)$ in the literature [92]. Of course, a better determination of g_c requires the performance of a dedicated finite-size scaling analysis.

For the 2D ladder model, the associated R as functions of g for $L = 24, 48$ are shown in Fig. 14. Moreover, by using the idea of estimating β_c for the 3D classical $O(3)$ model, the curves resulting from treating $R + \Delta$ and $1/\sqrt{2} + 1 - R - \Delta$ as functions of g are demonstrated in Fig. 15 ($L = 24, 32$) and 16 ($L = 48, 64$). The vertical dashed lines in these figures are the theoretical g_c . Here Δ is the difference between the theoretical and the calculated values of R at $g = 1$. As can be seen from the figures, when box size L increases, the g

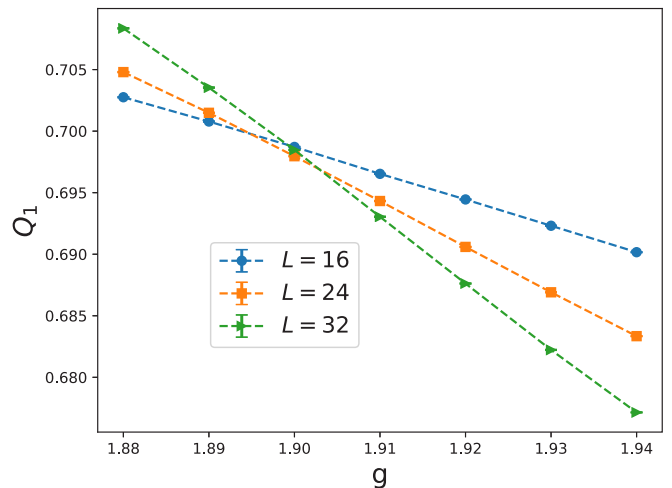


FIG. 13. Q_1 (of $L = 16, 24, 32$) as functions of g for the 2D dimerized quantum ladder model.

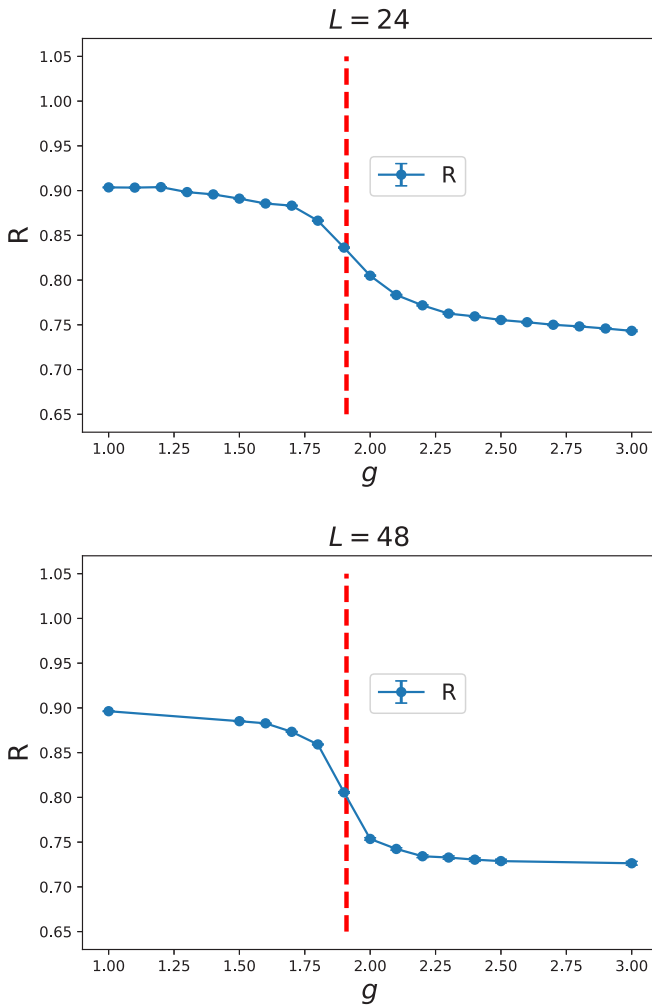


FIG. 14. R as functions of g for the 2D dimerized quantum ladder model. The top and bottom panels are for $L = 24$ and $L = 48$, respectively.

at which the mentioned two curves intersects gradually approaches the theoretical $g_c \sim 1.9$. One can apply certain kind of finite-size scaling formula to fit the intersection points in order to obtain an estimation of g_c . Here we do not make such an attempt. The validity of our approach for the determination of the critical point associated with a quantum phase transition will be confirmed in next subsection where the NN results of the 3D plaquette model is presented.

C. Results of 3D quantum spin system

The g_c of the 3D plaquette model studied in Ref. [93] can be determined by considering the Néel temperatures T_N of various g close to g_c . Specifically, if the logarithmic correction is not taken into account, then close to g_c , T_N can be described by $T_N \sim A|g - g_c|^c + B|g - g_c|^{2c}$, here A , B , and c are some constants. As a result, g_c can be calculated by fitting the data of T_N of various g to this form. The g_c estimated by this approach lies between 4.35 and 4.375; see Fig. 17. This obtained g_c will be used to examine the effectiveness of the NN method for calculating the g_c of the 3D plaquette model.

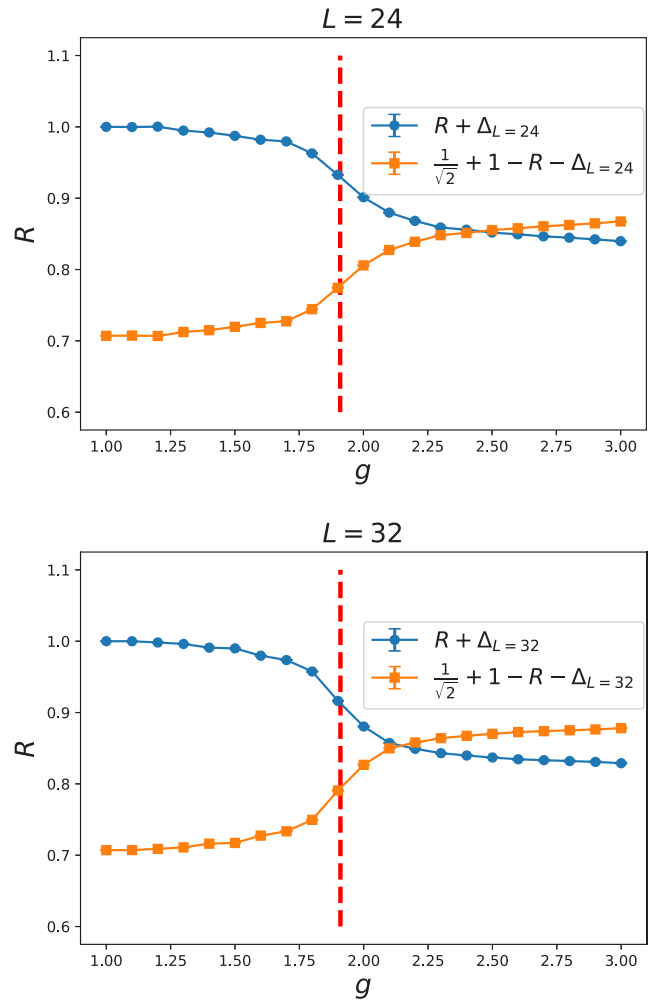


FIG. 15. $R + \Delta$ and $1/\sqrt{2} - \Delta + 1 - R$ as functions of g for the 2D dimerized quantum ladder model. The top and bottom panels are for $L = 24$ and $L = 32$, respectively.

R as functions of g for $L = 16$ and 32 for the 3D plaquette model are shown in Fig. 18. In addition, the curves resulting from considering $R + \Delta$ and $1/\sqrt{2} + 1 - R - \Delta$ as functions of g are demonstrated in Fig. 19 ($L = 16, 32$). The vertical dashed lines in these figures are 4.35 which is the estimated lower bound for g_c discussed in the previous paragraph. Here Δ is again the difference between the theoretical and the calculated values of R at $g = 1$.

Remarkably, just like what we have found for the 2D quantum ladder model, the results shown in the figure clearly reveal the message that our NN method is valid for 3D quantum spin model as well.

It is interesting to notice that the crossing points in both panels of Fig. 19 are slightly below the critical point calculated from T_N . We attribute this to the facts that the systematic influence of some tunable parameters of NN as well as certain corrections to the employed finite-size scaling method are not taken into account here.

Nevertheless, based on the outcomes associated with both the investigated 2D and 3D dimerized quantum antiferromagnetic Heisenberg models, it is beyond doubt that the NN

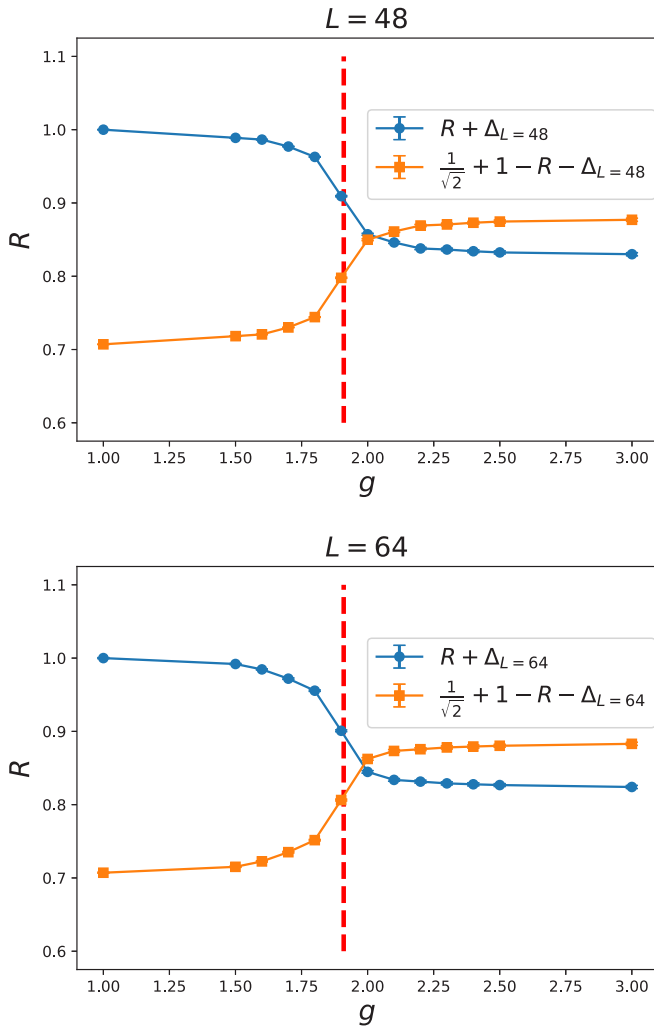


FIG. 16. $R + \Delta$ and $1/\sqrt{2} - \Delta + 1 - R$ as functions of g for the 2D dimerized quantum ladder model. The top and bottom panels are for $L = 48$ and $L = 64$, respectively.

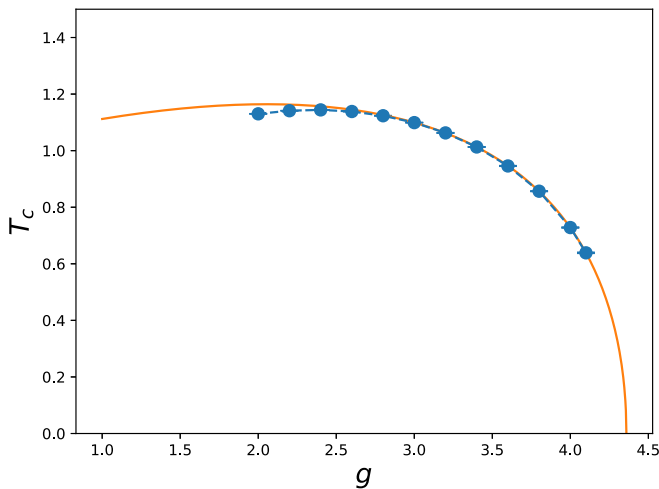


FIG. 17. T_N as a function of g for the 3D dimerized quantum plaquette model. The solid line shown in the figure is obtained by using the results from a fit.

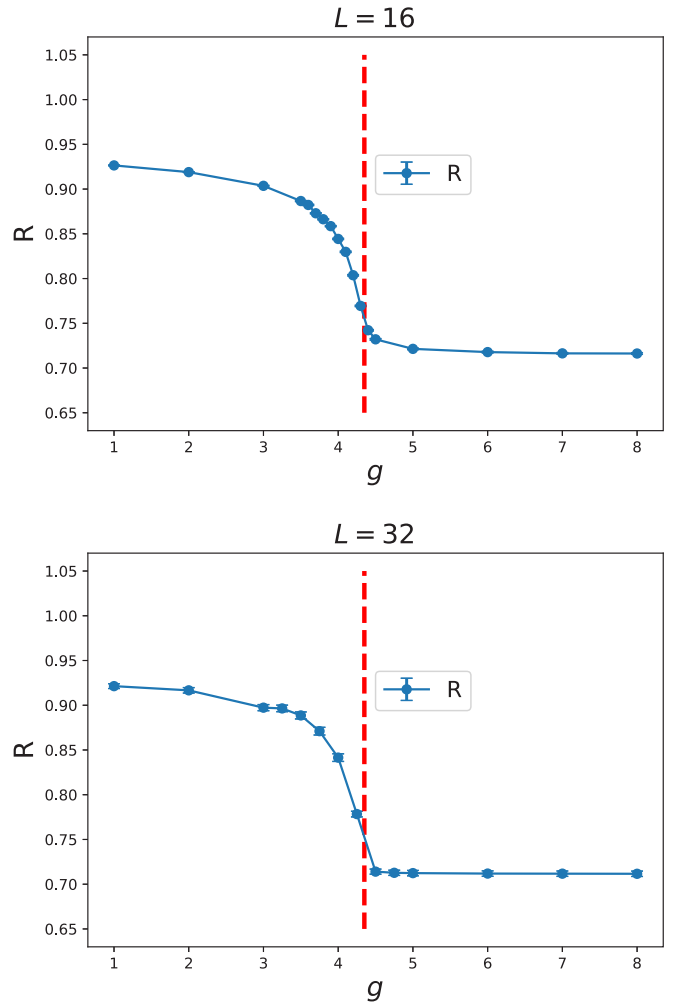


FIG. 18. R as functions of g for the 3D dimerized quantum plaquette model. The top and bottom panels are for $L = 16$ and $L = 32$, respectively.

approach employed here can be used to estimate the critical points of quantum phase transitions efficiently.

D. Verification of the semiexperimental finite-size scaling formulas: 2D three-state ferromagnetic Potts model and 2D classical XY model on the square lattices

1. 2D three-state ferromagnetic Potts model

In previous subsections, it is shown that the critical point can be obtained by considering the intersection of two curves made up of quantities associated with R . To obtain a high precision estimation for the critical point using the crossing points, one can apply certain expression of finite-size scaling to fit the data (of the crossing points). Here we use the data of the 2D three-state ferromagnetic Potts model on the square lattice available in Ref. [48] to carry out such an investigation. For each L , the data are obtained using a single set of relevant NN parameters. As a result, the quoted errors are associated with the Potts configurations themselves.

Figure 20 shows that data of $R + \Delta$ and $1/\sqrt{3} + 1 - R - \Delta$ as functions of T for various L for the 2D three-state ferromagnetic Potts model on the square lattice [48]. A fit

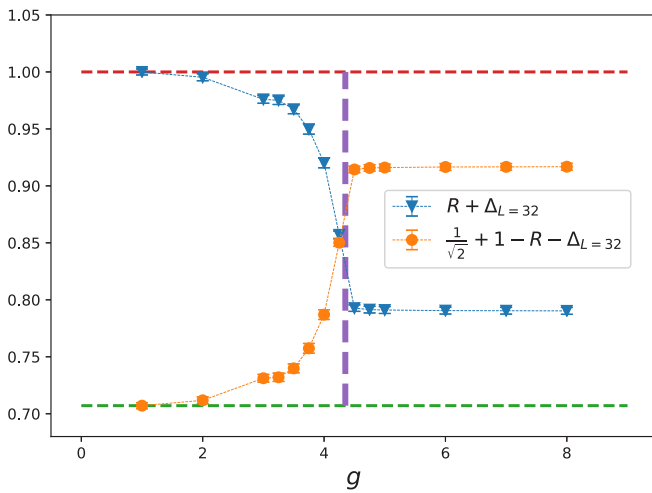
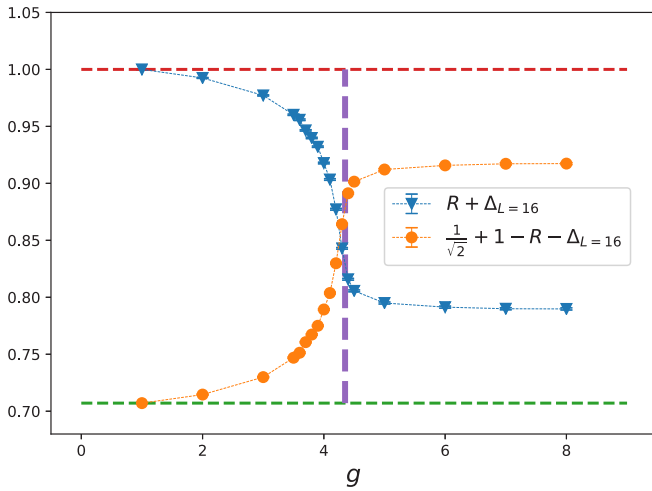


FIG. 19. $R + \Delta$ and $1/\sqrt{2} - \Delta + 1 - R$ as functions of g for the 3D dimerized quantum plaquette model. The top and bottom panels are for $L = 16$ and $L = 32$, respectively.

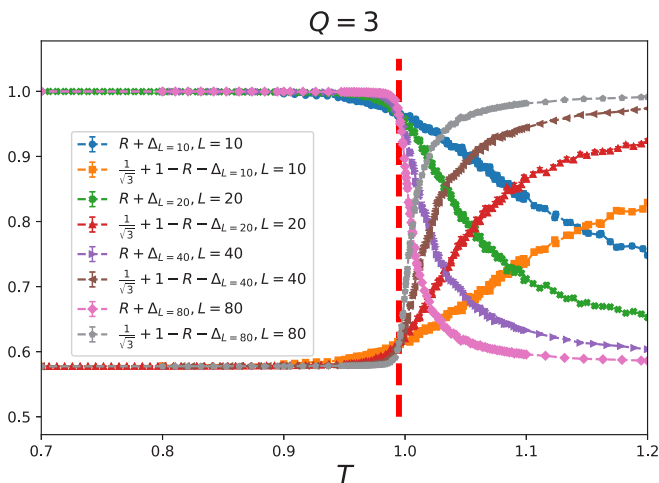


FIG. 20. $R + \Delta$ and $1/\sqrt{3} + 1 - R - \Delta$ as functions of T for various L for the 2D three-state ferromagnetic Potts model on the square lattice.

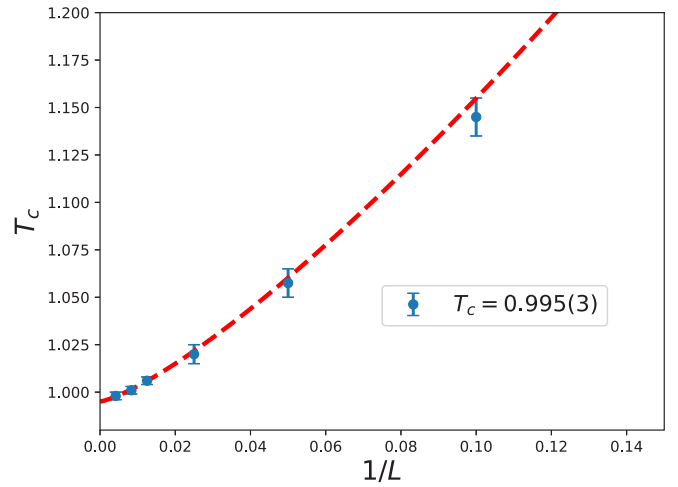


FIG. 21. Fit of the crossing points (of various finite L) to the ansatz $a + b/L^c$. The data are associated with the three-state ferromagnetic Potts model on the square lattice [48] and the dashed line in the figure is obtained by using the results of the fits.

of the form $a + b/L^c$, where a , b , and c are some to be determined constants (a is exactly the desired T_c), is used to fit the data of the crossing points obtained from $L = 10, 20, 40, 80, 120, 240$ (The data of $L = 120$ and 240 are not presented in Fig. 20). When carrying out the fits, Gaussian noises are considered to estimate the corresponding errors of the constants a , b , and c .

The fits lead to $a = 0.995(3)$ which matches quantitatively with the theoretical prediction $T_c \sim 0.99497$; see Fig. 21. Moreover, it is anticipated that the exponent c of the ansatz is given by $1/\nu$ (ignore the subleading exponents), where ν is the correlation length exponent. For three-state ferromagnetic Potts model, $\nu \sim 0.824$ [98]. In our calculation, we find $c \sim 1.29$ which is in reasonably good agreement with the expected $1/0.824 \sim 1.21$. All these results in turn confirm the validity of calculating the critical points using the NN approach presented in this study.

2. 2D classical XY model

The $R + \Delta$ and $1/\sqrt{2} + 1 - R - \Delta$ as functions of β for several L of the 2D classical XY model are demonstrated in Fig. 22. Similar to the analysis done for the 2D three-state ferromagnetic Potts model, we would like to calculate the crossing points for various L and use some kind of finite-size scaling to fit the obtained data so that one can determine the associated β_c (or T_c). After obtaining coarse estimations of the crossing points for various L from Fig. 22, more simulations are carried out to reach a better precision for these crossing points. These refined $\beta_c(L)$ are then fitted to the same formula (i.e., $a + b/L^c$) as that used for the three-state Potts model. We find that the obtained results are not satisfactory. This can be expected since the topological characteristics of the Kosterlitz-Thouless transition should reflect on R .

Motivated by the finite-size scaling formulas used in Refs. [52,91,99] for the 2D classical XY model, we use an ansatz of the form $a_1 + b_1/[\log(L)]^2$ (here a_1 is the β_c) to fit the newly obtained data of crossing points. The outcome

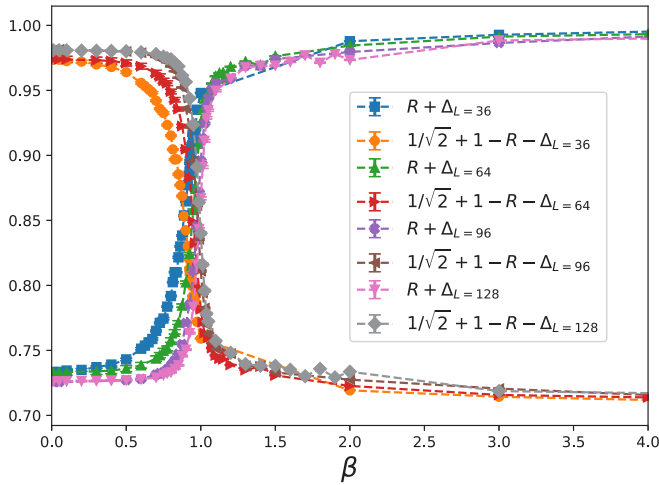


FIG. 22. $R + \Delta$ and $1/\sqrt{2} + 1 - R - \Delta$ as functions of β for various L for the 2D classical XY model on the square lattice.

is good and we find the β_c is given by $\beta_c = 1.119(7)$ (see both panels of Fig. 23) which match very well with the known result $\beta_c \sim 1.1199$ in the literature. It is intriguing that the simple NN procedure used here works for the phase transition(s) associated with topology as well.

Before ending this subsection, we would like to point out that in principle the supervised NN method is a optimization procedure. As a result, to obtain a more accurate estimation of the critical point in a (supervised) NN investigation, the systematic impact associated with the tunable parameters of a built NN, such as the number of epoch, batchsize, nodes in the hidden layers and so on, should be examined.

V. DISCUSSIONS AND CONCLUSIONS

In this study we investigate the phase transitions of 3D classical $O(3)$ model and 2D classical XY model, as well as the quantum phase transitions of both the 2D and the 3D dimerized spin-1/2 antiferromagnetic Heisenberg models using the simplest deep learning NN, namely, a MLP that is made up of only one input layer, one hidden layer, as well as one output layer.

In our investigation, the training set for each of the studied models consists of only two objects. In particular, none of the used training objects belongs to the theoretical or the real configurations of the considered physical systems.

Remarkably, with such an unconventional approach of carrying out the training processes in conjunction with certain semiexperimental finite-size scaling formulas, the resulting outcomes from the built MLP lead to very good estimations of the targeted critical points. The results reached here as well as that shown in Refs. [48,86] provide convincing evidence that the performance of some unconventional strategies, such as employing the theoretical ground-state configurations as the training sets, are impressive. Particularly, the simplicity of these approaches make them cost-effective in computation. It is amazing that the simple procedures used in Refs. [48,86] and here are not only valid for phase transitions associated with SSB, but also work for those related to topology.

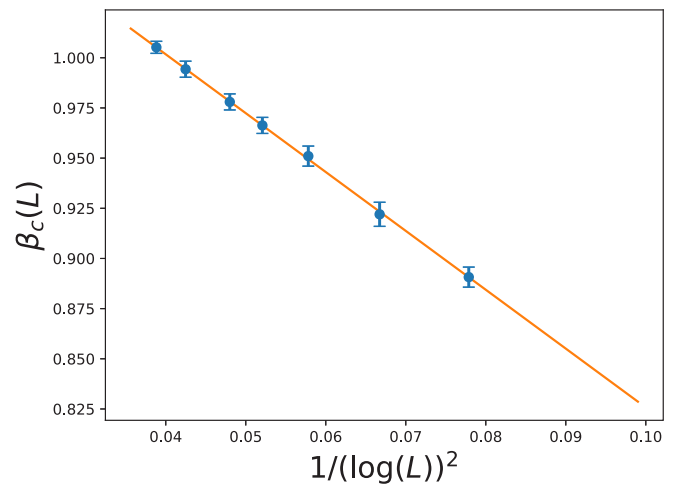
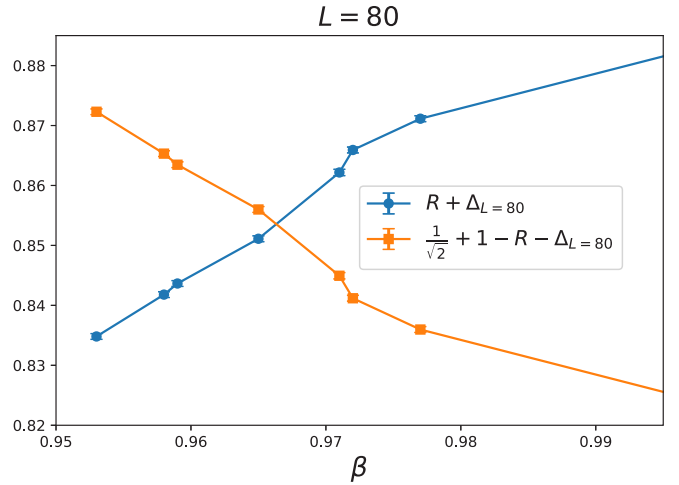


FIG. 23. (Top) Estimation of the crossing point for $L = 80$. (Bottom) Fit of $\beta_c(L)$ to the ansatz $a_1 + b_1/[\log(L)]^2$. The solid line is obtained using the results from the fit.

We would like to point out that for the 3D classical $O(3)$ model, the training set used here consists of two configurations (their elements are either all 1 or all 0). In principle, one can consider training set made up of three, four, or even five configurations following the same idea as that of two objects training set. To examine whether using the training sets, which constitute more than two objects, one can arrive at the same level of success as that shown in the previous section, we have performed three more NN calculations using $n = 3$, $n = 4$, and $n = 5$ training sets. Here n denotes the number of objects contained in the training set. Interestingly, the precision of the estimated T_c of the 3D classical $O(3)$ model obtained from these additional calculations is becoming slightly less satisfactory with n ; see Fig. 24 for a outcome related to $n = 4$ and $L = 20$. Intuitively, this can be understood as follows. Let us assume that initially all the unit vectors belong to a category of the classification scheme implemented in the training stage. Then any local fluctuation will have greater impact on the resulting NN outputs if the training set contains more types of objects. Despite this, it is beyond doubt that the outcomes associated with training sets consisting of only two configurations, including those from all the four studied

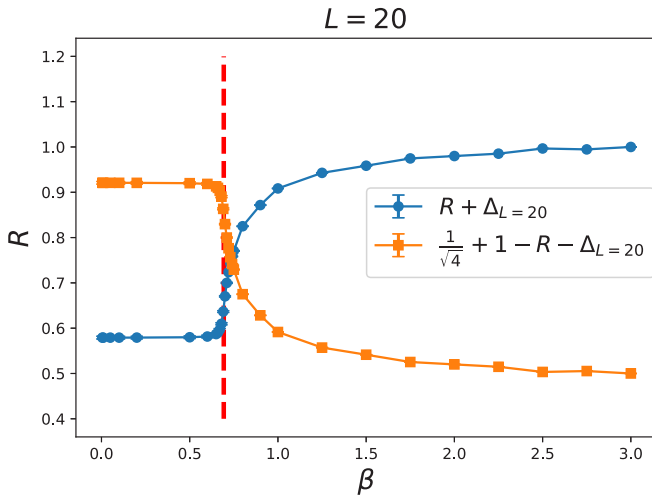


FIG. 24. $R + \Delta$ and $1/\sqrt{4} - \Delta + 1 - R$ as functions of β for the 3D classical $O(3)$ model. The results are obtained from the calculations which use four configurations as the training set.

models, strongly suggest the effectiveness of the approach presented in this study.

The NN results related to all the models considered here are obtained using 10 sets of random seeds with other parameters of NN being fixed in the calculations. It turns out that the influence on the NN outcomes due to the use of different random seeds is very mild. Moreover, for several L of the studied 3D $O(3)$ and 2D ladder models, we have performed analysis using only one of the 10 trained NNs. Some of the resulting outcomes are shown in Figs. 25 and 26 (the errors of the data shown in these new figures are associated with the configurations determined from QMC simulations). These new figures match nicely with that determined with 10 sets of random seeds. Apart from this, we have also carried out several calculations using various batchsize, epoch, and nodes in the hidden layer. These new calculations lead to very good agreement with that shown explicitly in this study as well, see Fig. 27 for one result from these new calculations. The additional investigations introduced in this paragraph imply that the tunable parameters of NN have very mild effects on the resulting outcomes of R for the considered models. Hence, the obtained conclusion here should be reliable. Of course, as already being pointed out before, considering other systematic impacts are required if a highly accurate estimation of the targeted critical point is desirable.

In the following, we would like to discuss a little bit more about the option of choosing a relevant temperature to calculate the overall shift Δ for R . We will consider the $O(3)$ model as a demonstration. Previously Δ was determined by the difference between 1 and the R associated with the lowest temperature. Due to the employed configurations in the training stage, this approach is indeed natural. However, considering the fact that simulations carried out at low temperatures require more computing power, an alternative to perform an overall shift of R (by Δ) can be conducted as follows. First, one notices that R should obtain the value of $1/\sqrt{2}$ at extremely high temperatures and the associated simulations are low-cost. Hence, instead of carrying out the

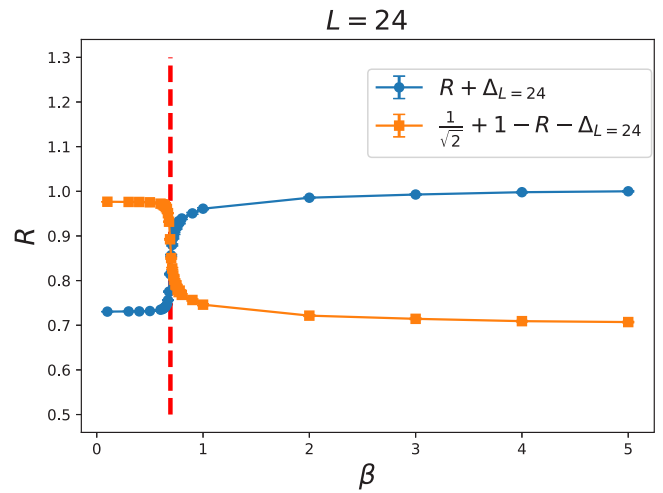
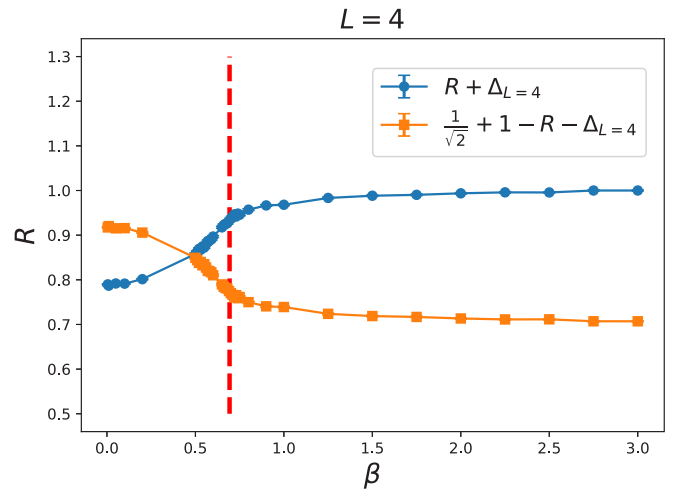


FIG. 25. $R + \Delta$ and $1/\sqrt{2} - \Delta + 1 - R$ as functions of β for the 3D classical $O(3)$ model. The data are obtained using only one set of random seeds. The top and bottom panels are for $L = 4$ and $L = 24$, respectively.

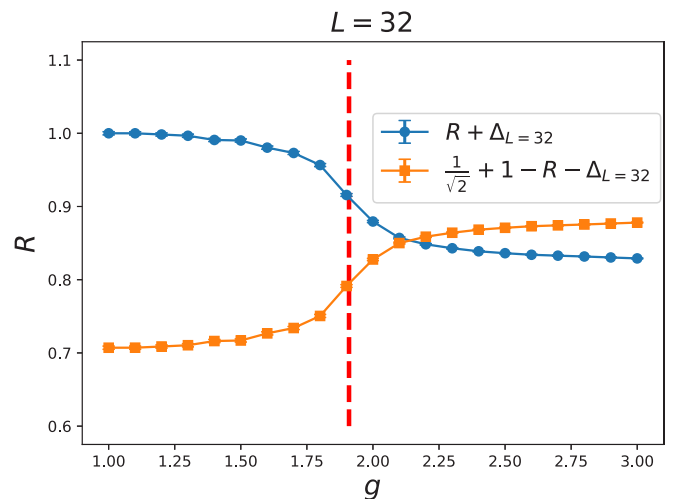


FIG. 26. $R + \Delta$ and $1/\sqrt{2} - \Delta + 1 - R$ as a function of g for the 2D ladder model. The data are obtained using only one set of random seeds.

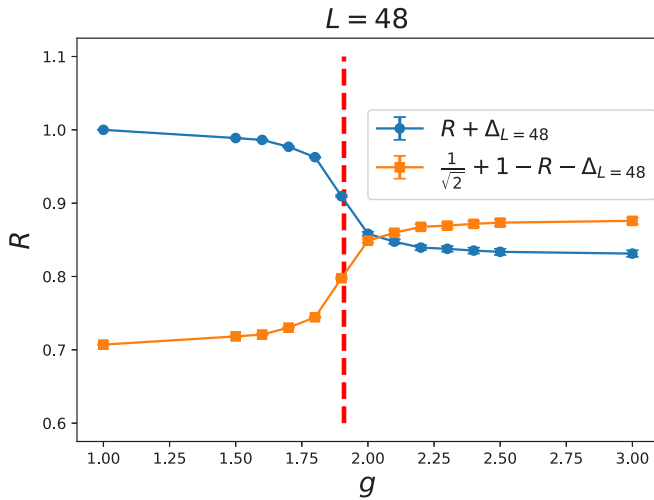


FIG. 27. $R + \Delta$ and $1/\sqrt{2} - \Delta + 1 - R$ as a function of g for the 2D ladder model. The data are obtained using different NN parameters from that shown in the previous subsection.

shift based on the result calculated at the lowest temperature, one can execute such an operation at the highest temperature T_h . With this new strategy, $R - \Delta$ is $1/\sqrt{2}$ at T_h . In Fig. 28, $R - \Delta$ and $1 + 1/\sqrt{2} - R + \Delta$ of the $O(3)$ model are shown as functions of β (Here Δ is the difference between $1/\sqrt{2}$ and the R obtained at T_h). The top and bottom panels in the figure are for $L = 12$ and $L = 24$, respectively. Based on the outcomes demonstrated in Fig. 28, the validity of this new approach is beyond doubt. Since one can simulate very high temperatures with ease, this new strategy of conducting the shift of R is cost-effective.

Although in this study we have focused on studying the phase transitions of several models, it is probable that simple NN approaches, similar to the one(s) considered here, are available for investigating other physical properties of many-body systems. We would like to emphasize the motivations for the series of our studies of applying the NN techniques to investigate the phase transitions of several physical systems, as shown in Refs. [48,86] and here. Conventionally, the application of a supervised NN to explore the critical phenomenon of a specific system has a caveat, namely, the knowledge of the critical point is required in advance before one can employ the methods of NN for the investigation. Hence, for systems with unknown critical points, it may not be easy to apply such standard NN procedures to the studies in a straightforward manner. The approaches considered in Refs. [48,86] and here definitely can take care of this issue, hence promote the use of NN methods in various fields of many-body systems. In particular, these unconventional methods are adequate for carrying out any NN investigations of examining whether certain proposed theories are relevant for a real and unexplored physical system.

Finally, it will be interesting to examine the accuracy in the determination of the critical points with the NN approach versus the traditional methods.

The NN approach is in principle a variation method and there are several adjustable parameters which can be tuned to reach a optimized performance for NN. In the mean-

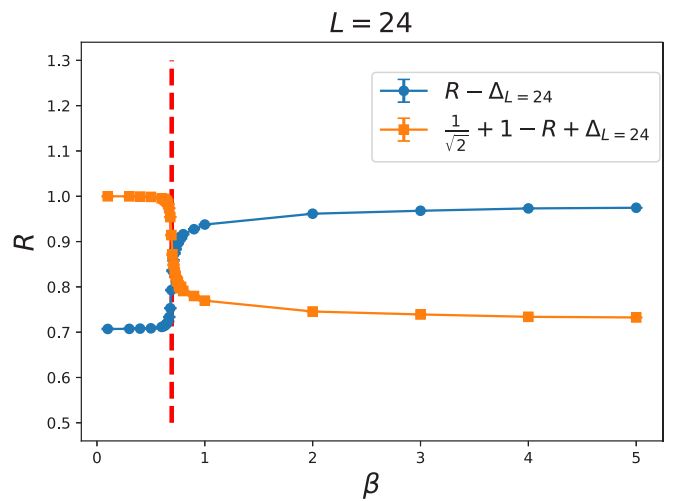
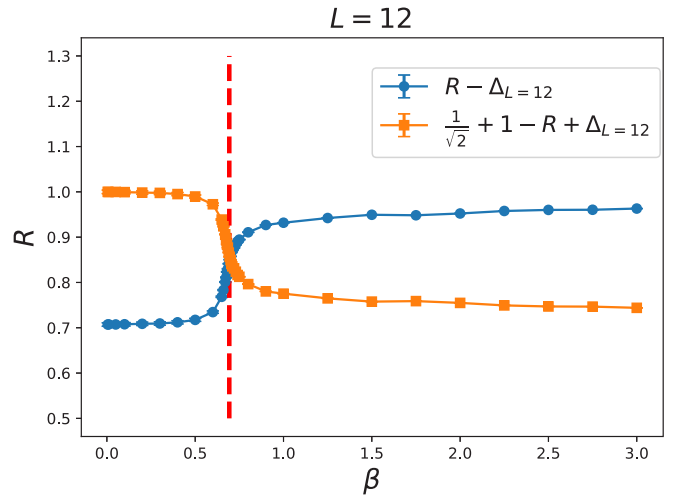


FIG. 28. $R - \Delta$ and $1/\sqrt{2} + \Delta + 1 - R$ as functions of β for the 3D classical $O(3)$ model. The top and bottom panels are for $L = 12$ and $L = 24$, respectively. The vertical dashed line is the expected T_c . The Δ used here is the difference between $1/\sqrt{2}$ and the R associated with the highest temperature.

time this feature of NN also introduce certain systematic uncertainties to the determination of the critical points, as being described previously. Apart from this, different optimizers and truncation strategies could bring systematic errors to the results as well. Traditional methods which involve variation principle may encounter similar situations as that of NN. From this point of view, Monte Carlo is an excellent approach for studying phase transitions if it does not suffer limitation due to the nature of the investigated systems.

Any theoretical methods eventually will confront the experimental data. Hence, considering the fact that uncertainties always come with experimental results, the NN approach is a good alternative to the traditional methods when studying phases and criticalities is concerned. Particularly, NN may well have the potential to surpass other approaches since the research of improving the performance and expanding the applicability of NN is vigorous and is on the fly.

ACKNOWLEDGMENT

Partial support from Ministry of Science and Technology of Taiwan is acknowledged.

-
- [1] M. Rupp, A. Tkatchenko, K.-R. Müller, and O. Anatole von Lilienfeld, *Phys. Rev. Lett.* **108**, 058301 (2012).
- [2] John C. Snyder, M. Rupp, K. Hansen, K.-R. Müller, and K. Burke, *Phys. Rev. Lett.* **108**, 253002 (2012).
- [3] G. Montavon, M. Rupp, V. Gobre, A. Vazquez-Mayagoitia, K. Hansen, A. Tkatchenko, K.-R. Müller, and O. Anatole von Lilienfeld, *New J. Phys.* **15**, 095003 (2013).
- [4] A. P. Bartók, R. Kondor, and G. Csányi, *Phys. Rev. B* **87**, 184115 (2013).
- [5] G. Pilania, C. Wang, Xun Jiang, S. Rajasekaran, and R. Ramprasad, *Sci. Rep.* **3**, 2810 (2013).
- [6] B. Meredig, A. Agrawal, S. Kirklin, J. E. Saal, J. W. Doak, A. Thompson, K. Zhang, A. Choudhary, and C. Wolverton, *Phys. Rev. B* **89**, 094104 (2014).
- [7] K. T. Schütt, H. Glawe, F. Brockherde, A. Sanna, K. R. Müller, and E. K. U. Gross, *Phys. Rev. B* **89**, 205118 (2014).
- [8] Z. Li, J. R. Kermode, and A. De Vita, *Phys. Rev. Lett.* **114**, 096405 (2015).
- [9] P. Baldi, P. Sadowski, and D. Whiteson, *Phys. Rev. Lett.* **114**, 111801 (2015).
- [10] V. Mnih, K. Kavukcuoglu, D. Silver, A. A. Rusu, J. Veness, M. G. Bellemare, A. Graves, M. Riedmiller, A. K. Fidjeland, G. Ostrovski, S. Petersen, C. Beattie, A. Sadik, I. Antonoglou, H. King, D. Kumaran, D. Wierstra, S. Legg, and D. Hassabis, *Nature* **518**, 529 (2015).
- [11] T. Lang, F. Flachsenberg, U. von Luxburg, and M. Rarey, *J. Chem. Inf. Model.* **56**, 12 (2016).
- [12] J. Lee, A. Seko, K. Shitara, K. Nakayama, and I. Tanaka, *Phys. Rev. B* **93**, 115104 (2016).
- [13] J. Searcy, L. Huang, M. A. Pleier, and J. Zhu, *Phys. Rev. D* **93**, 094033 (2016).
- [14] T. Castro, M. Quartin, and S. Benítez-Herrera, *Phys. Dark Univ.* **13**, 66 (2016).
- [15] P. Baldi, K. Cranmer, T. Faucett, P. Sadowski, and D. Whiteson, *Eur. Phys. J. C* **76**, 235 (2016).
- [16] P. Baldi, K. Bauer, C. Eng, P. Sadowski, and D. Whiteson, *Phys. Rev. D* **93**, 094034 (2016).
- [17] G. Torlai and R. G. Melko, *Phys. Rev. B* **94**, 165134 (2016).
- [18] L. Wang, *Phys. Rev. B* **94**, 195105 (2016).
- [19] M. A. Shandiz and R. Gauvin, *Comput. Mater. Sci.* **117**, 270 (2016).
- [20] T. Ohtsuki and T. Ohtsuki, *J. Phys. Soc. Jpn.* **85**, 123706 (2016).
- [21] B. Hoyle, *Astron. Comput.* **16**, 34 (2016).
- [22] J. Carrasquilla and R. G. Melko, *Nat. Phys.* **13**, 431 (2017).
- [23] G. Carleo and M. Troyer, *Science* **355**, 602 (2017).
- [24] Z. Wu, B. Ramsundar, E. N. Feinberg, J. Gomes, C. Geniesse, A. S. Pappu, K. Leswing, and V. Pande, *Chem. Sci.* **9**, 513 (2018).
- [25] P. Broecker, J. Carrasquilla, R. G. Melko, and S. Trebst, *Sci. Rep.* **7**, 8823 (2017).
- [26] K. Ch'ng, J. Carrasquilla, R. G. Melko, and E. Khatami, *Phys. Rev. X* **7**, 031038 (2017).
- [27] J. Barnard, E. N. Dawe, M. J. Dolan, and N. Rajcic, *Phys. Rev. D* **95**, 014018 (2017).
- [28] A. Tanaka and A. Tomiya, *J. Phys. Soc. Jpn.* **86**, 063001 (2017).
- [29] E. P. L. van Nieuwenburg, Y.-H. Liu, and S. D. Huber, *Nat. Phys.* **13**, 435 (2017)
- [30] A. Mott, J. Job, J. R. Vlimant, D. Lidar, and M. Spiropulu, *Nature* **550**, 375 (2017).
- [31] J. Liu, H. Shen, Y. Qi, Z. Y. Meng, and L. Fu, *Phys. Rev. B* **95**, 241104(R) (2017).
- [32] T. Tamura, M. Karasuyama, R. Kobayashi, R. Arakawa, Y. Shiihara, and I. Takeuchi, *Modell. Simul. Mater. Sci. Eng.* **25**, 075003 (2017).
- [33] J. Tubiana and R. Monasson, *Phys. Rev. Lett.* **118**, 138301 (2017).
- [34] X. Y. Xu, Y. Qi, J. Liu, L. Fu, and Z. Y. Meng, *Phys. Rev. B* **96**, 041119(R) (2017).
- [35] L. Huang and L. Wang, *Phys. Rev. B* **95**, 035105 (2017).
- [36] M. Zevin, S. Coughlin, S. Bahaadini, E. Besler, N. Rohani, S. Allen, M. Cabero, K. Crowston, A. Katsaggelos, S. Larson, T. K. Lee, C. Lintott, T. Littenberg, A. Lundgren, C. Østerlund, J. Smith, L. Trouille, and V. Kalogera, *Class. Quant. Grav.* **34**, 064003 (2017).
- [37] J. Liu, Y. Qi, Z. Y. Meng, and L. Fu, *Phys. Rev. B* **95**, 041101(R) (2017).
- [38] Y. Liu, T. Zhao, W. Ju, and S. Shi, *J. Materiomics* **3**, 159 (2017).
- [39] Q. Wei, R. G. Melko, and J. Z. Y. Chen, *Phys. Rev. E* **95**, 032504 (2017).
- [40] Y. Nagai, H. Shen, Y. Qi, J. Liu, and L. Fu, *Phys. Rev. B* **96**, 161102(R) (2017).
- [41] B. Kolb, L. C. Lentz, and A. M. Kolpak, *Sci. Rep.* **7**, 1192 (2017).
- [42] D.-L. Deng, X. Li, and S. Das Sarma, *Phys. Rev. B* **96**, 195145 (2017).
- [43] P. Ponte and R. G. Melko, *Phys. Rev. B* **96**, 205146 (2017).
- [44] G. Kasieczka, T. Plehn, M. Russell, and T. Schell, *J. High Energy Phys.* **05** (2017) 006.
- [45] Y. Zhang, R. G. Melko, and E.-A. Kim, *Phys. Rev. B* **96**, 245119 (2017).
- [46] Y. Zhang and E.-A. Kim, *Phys. Rev. Lett.* **118**, 216401 (2017)
- [47] W. Hu, R. R. P. Singh, and R. T. Scalettar, *Phys. Rev. E* **95**, 062122 (2017).
- [48] C.-D. Li, D.-R. Tan, and F.-J. Jiang, *Ann. Phys.*, **391**, 312 (2018).
- [49] K. Ch'ng, N. Vazquez, and E. Khatami, *Phys. Rev. E* **97**, 013306 (2018).
- [50] S. Lu, Q. Zhou, Y. Ouyang, Y. Guo, Q. Li and J. Wang, *Nat. Commun.* **9**, 3405 (2018).
- [51] D. George and E. A. Huerta, *Phys. Rev. D* **97**, 044039 (2018).
- [52] M. J. S. Beach, A. Golubeva, and R. G. Melko, *Phys. Rev. B* **97**, 045207 (2018).
- [53] L. G. Pang, K. Zhou, N. Su, H. Petersen, H. Stöcker, and X. N. Wang, *Nat. Commun.* **9**, 210 (2018).
- [54] P. E. Shanahan, D. Trewartha, and W. Detmold, *Phys. Rev. D* **97**, 094506 (2018).
- [55] N. Artrith, A. Urban, and G. Ceder, *J. Chem. Phys.* **148**, 241711 (2018).

- [56] K. T. Butler, D. W. Davies, H. Cartwright, O. Isayev, and A. Walsh, *Nature* **559**, 547 (2018).
- [57] A. P. Bartók, J. Kermode, N. Bernstein, and G. Csányi, *Phys. Rev. X* **8**, 041048 (2018).
- [58] P. Zhang, H. Shen, and H. Zhai, *Phys. Rev. Lett.* **120**, 066401 (2018).
- [59] J. Graser, S. K. Kauwe, and T. D. Sparks *Chem. Mater.* **30**, 3601 (2018).
- [60] A. Butter, G. Kasieczka, T. Plehn, and M. Russell, *SciPost Phys.* **5**, 028 (2018).
- [61] J. Gao, L.-F. Qiao, Z.-Q. Jiao, Y.-C. Ma, C.-Q. Hu, R.-J. Ren, A.-L. Yang, H. Tang, M.-H. Yung, and X.-M. Jin, *Phys. Rev. Lett.* **120**, 240501 (2018).
- [62] W. Zhang, J. Liu, and T.-C. Wei, *Phys. Rev. E* **99**, 032142 (2019).
- [63] J. Greitemann, K. Liu, and L. Pollet, *Phys. Rev. B* **99**, 060404(R) (2019).
- [64] J. Ren, L. Wu, J. M. Yang, and J. Zhao, *Nucl. Phys. B* **943**, 114613 (2019).
- [65] X.-Y. Dong, F. Pollmann, and X.-F. Zhang, *Phys. Rev. B* **99**, 121104(R) (2019).
- [66] R. Jinnouchi, F. Karsai, and G. Kresse, *Phys. Rev. B* **100**, 014105 (2019).
- [67] M. Cavaglia, K. Staats, and T. Gill, *Commun. Comput. Phys.* **25**, 963 (2019).
- [68] D. W. Davies, K. T. Butler, and A. Walsh, *Chem. Mater.* **31**, 7221 (2019).
- [69] G. P. Conangla, F. Ricci, M. T. Cuairan, A. W. Schell, N. Meyer, and R. Quidant, *Phys. Rev. Lett.* **122**, 223602 (2019).
- [70] B. Yoon, T. Bhattacharya, and R. Gupta, *Phys. Rev. D* **100**, 014504 (2019).
- [71] J. Fluri, T. Kacprzak, A. Lucchi, A. Refregier, A. Amara, T. Hofmann, and A. Schneider, *Phys. Rev. D* **100**, 063514 (2019).
- [72] A. Canabarro, F. F. Fanchini, A. L. Malvezzi, R. Pereira, and R. Chaves, *Phys. Rev. B* **100**, 045129 (2019).
- [73] L. Li, Y. You, S. Hu, Y. Shi, G. Zhao, C. Chen, Y. Wang, A. Stroppa, and W. Ren, *Appl. Phys. Lett.* **114**, 083102 (2019).
- [74] H. Chan, B. Narayanan, M. J. Cherukara, F. G. Sen, K. Sasikumar, S. K. Gray, M. K. Y. Chan, and S. K. R. S. Sankaranarayanan, *J. Phys. Chem. C*, **123**, 6941 (2019).
- [75] W. Lian, S.-T. Wang, S. Lu, Y. Huang, F. Wang, X. Yuan, W. Zhang, X. Ouyang, X. Wang, X. Huang, L. He, X. Chang, D.-L. Deng, and L. Duan, *Phys. Rev. Lett.* **122**, 210503 (2019).
- [76] L. Zhu, W. Zhang, J. Kou, and Y. Liu, *Phys. Fluids* **31**, 015105 (2019).
- [77] P. Mehta, M. Bukov, C.-H. Wang, A. G. R. Day, C. Richardson, C. K. Fisher, and D. J. Schwab, *Phys. Rep.* **810**, 1 (2019).
- [78] K. T. Schütt, M. Gastegger, A. Tkatchenko, K.-R. Müller, and R. J. Maurer, *Nat. Commun.* **10**, 5024 (2019).
- [79] J. F. Rodriguez-Nieva and M. S. Scheurer, *Nat. Phys.* **15**, 790 (2019).
- [80] G. Carleo, I. Cirac, K. Cranmer, L. Daudet, M. Schuld, N. Tishby, L. Vogt-Maranto, and L. Zdeborová, *Rev. Mod. Phys.* **91**, 045002 (2019).
- [81] T. Ohtsuki, and T. Mano, *J. Phys. Soc. Jpn.* **89**, 022001 (2020).
- [82] W. Hashimoto, Y. Tsuji, and K. Yoshizawa *J. Phys. Chem. C* **124**, 9958 (2020).
- [83] R. Jinnouchi, F. Karsai, and G. Kresse, *Phys. Rev. B* **101**, 060201(R) (2020).
- [84] A. J. Larkoski, I. Moulton, and B. Nachman, *Phys. Rept.* **841**, 1 (2020).
- [85] X. Han and S. A. Hartnoll, *Phys. Rev. X* **10**, 011069 (2020).
- [86] D.-R. Tan, C.-D. Li, W.-P. Zhu, and F.-J. Jiang, *New J. Phys.* **22**, 063016 (2020).
- [87] J. Singh, V. Arora, V. Gupta, and M. S. Scheurer, *arXiv:2006.11868*.
- [88] K. Binder, *Z. Phys. B* **43**, 119 (1981).
- [89] C. Holm and W. Janke, *Phys. Lett. A* **173**, 8 (1993).
- [90] M. Campostrini, M. Hasenbusch, A. Pelissetto, P. Rossi, and E. Vicari, *Phys. Rev. B* **65**, 144520 (2002).
- [91] M. Hasenbusch, *J. Phys. A* **38**, 5869 (2005).
- [92] A. W. Sandvik, *Computational Studies of Quantum Spin Systems*, AIP Conf. Proc. 1297, 135 (AIP, New York, 2010).
- [93] D.-R. Tan, C. D. Li, and F.-J. Jiang, *Phys. Rev. B* **97**, 094405 (2018).
- [94] <https://keras.io>.
- [95] U. Wolff, *Phys. Rev. Lett.* **62**, 361 (1989).
- [96] A. W. Sandvik, *Phys. Rev. B* **59**, R14157(R) (1999).
- [97] Based on the employed training sets, the correction Δ should be calculated at the R obtained with the largest β for the 3D classical $O(3)$ and 2D classical XY models. Similarly, the Δ associated with the studied quantum spin models are determined at the R corresponding to $g = 1$.
- [98] F. Y. Wu, *Rev. Mod. Phys.* **54**, 235 (1982).
- [99] Y.-D. Hsieh, Y.-J. Kao, and A. W. Sandvik, *J. Stat. Mech.* (2013) P09001.