Rapid Communications

Machine-learning study using improved correlation configuration and application to quantum Monte Carlo simulation

Yusuke Tomita,[1,*] Kenta Shiina,[2,3,†] Yutaka Okabe [2,‡] and Hwee Kuan Lee[3,4,5,6,§]

[1]*College of Engineering, Shibaura Institute of Technology, Saitama 330-8570, Japan*

[2]*Department of Physics, Tokyo Metropolitan University, Hachioji, Tokyo 192-0397, Japan*

[3]*Bioinformatics Institute, Agency for Science, Technology and Research (A*STAR), 30 Biopolis Street, No. 07-01 Matrix, 138671, Singapore*

[4]*School of Computing, National University of Singapore, 13 Computing Drive, 117417, Singapore*

[5]*Singapore Eye Research Institute (SERI), 11 Third Hospital Avenue, 168751, Singapore*

[6]*Image and Pervasive Access Laboratory (IPAL), 1 Fusionopolis Way, No. 21-01 Connexis (South Tower), 138632, Singapore*

We use the Fortuin-Kasteleyn representation-based improved estimator of the correlation configuration as an alternative to the ordinary correlation configuration in the machine-learning study of the phase classification of spin models. The phases of classical spin models are classified using the improved estimators, and the method is also applied to the quantum Monte Carlo simulation using the loop algorithm. We analyze the Berezinskii-Kosterlitz-Thouless (BKT) transition of the spin-1/2 quantum XY model on the square lattice. We classify the BKT phase and the paramagnetic phase of the quantum XY model using the machine-learning approach. We show that the classification of the quantum XY model can be performed by using the training data of the classical XY model.

Remarkable developments in machine-learning-based techniques have been made in the past decade, which have impacted many areas in industry including automated driving, healthcare, etc. At the same time, the potential of machine learning for fundamental research has gained increasing interest. Statistical physics is one such scientific discipline [1].

Carrasquilla and Melko [2] used a technique of supervised learning to propose a paradigm that is complementary to the conventional approach of studying interacting spin systems. By using large data sets of spin configurations, they classified and identified a high-temperature paramagnetic phase and a low-temperature ferromagnetic phase. It was similar to image classification using machine learning. They demonstrated the use of neural networks for the study of the two-dimensional (2D) Ising model and an Ising lattice gauge theory.

Shiina *et al.* [3] reported a machine-learning study on phase transitions. The configuration of a long-range spatial correlation was treated instead of the spin configuration itself. By doing so, a similar treatment was provided to various spin models including multicomponent systems and systems with a vector order parameter. Not only were the second-order and the first-order transitions studied, but the Berezinskii-Kosterlitz-Thouless (BKT) transition [4–7] was studied as well. The disordered and the ordered phases, along with the BKT-type topological phase, were successfully classified.

Cluster algorithms [8,9] have been used to overcome slow dynamics in the Monte Carlo simulation. Swendsen and Wang (SW) [8] applied the Fortuin-Kasteleyn (FK) [10,11] representation to identify clusters of spins. The single-cluster variant of the cluster algorithm was proposed by Wolff [9]. Wolff also proposed the idea of an embedded cluster formalism [9,12,13] to treat vector spin models. By projecting a vector spin onto a randomly chosen unit vector, the Ising degrees of freedom are picked up. Then, a cluster spin flip can be performed with the FK cluster. A further advantage of cluster algorithms is that they lead to so-called improved estimators [13] which are designed to reduce statistical errors. In calculating spin correlations, only the spin pair belonging to the same FK cluster should be considered. The feature of manifesting spin correlations in a spin configuration is utilized in the probability-changing cluster algorithm, a self-adapted algorithm to tune the critical point automatically [14].

Evertz *et al.* [15,16] presented another type of cluster algorithm, which is called a loop algorithm. In treating vertex models, closed paths of bonds are flipped. Constraints at the vertices are automatically satisfied. The loop algorithm was applied to quantum spin systems in the worldline representation [17–19]. The improvements accomplished on the quantum Monte Carlo simulation were largely due to the global update, in which configurations are updated in units of some nonlocal clusters. By using the loop algorithm, nondiagonal quantities can be measured.

In this study, we consider an improved estimator for the correlation configuration in a cluster representation. We use the machine-learning method of Shiina *et al.* [3] for the classification of phases using the improved correlation configuration. Then, we apply this technique to quantum spin

---

*ytomita@shibaura-it.ac.jp

†16879316kenta@gmail.com

‡okabe@phys.se.tmu.ac.jp

§leehk@bii.a-star.edu.sg

(a) spin   **(b)** correlation   **(c)** improved

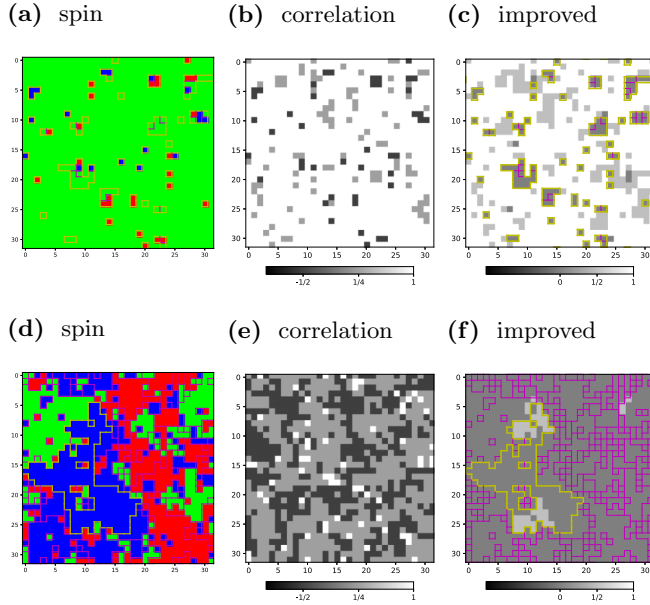**(d)** spin   **(e)** correlation   **(f)** improved

FIG. 1. Examples of the spin configuration $\{s_i\}$ [(a), (d)], correlation configuration $\{g_i(L/2)\}$ [(b), (e)], and improved correlation configuration $\{\tilde{g}_i(L/2)\}$ [(c), (f)] of the 2D three-state Potts model. The upper figures [(a)–(c)] are snapshots at the low temperature, $T = 0.9$, and the lower figures [(d)–(f)] are those at the high temperature, $T = 1.06$. The borders of FK clusters for the spin configuration are drawn by lines. They are copied in an improved correlation configuration.

systems. As an example, we show the results of the spin-1/2 XY model on the square lattice. This model exhibits the BKT transition [20].

We consider the configuration of a spin correlation with a distance of half of the system size, $L/2$. We note that this type of correlation function was used along with the generalized scheme for the probability-changing cluster algorithm [21]. For the q-state Potts model (including the Ising model), the correlation between two spins becomes 1 for the same spin pair, whereas it becomes $-1/(q-1)$ for a pair of different states. In the improved estimator for the cluster representation, the correlation becomes 1 for the spin pair belonging to the same FK cluster, whereas it becomes 0 for spins of different clusters. When the embedded algorithm for continuous spins is used, the projection of spins onto a randomly chosen reflection axis is made. We denote the site-dependent correlation configuration as $g_i(L/2)$. For the actual calculation, we treat the average value of the x direction and the y direction, that is,

$$g_i(L/2) = (g[s_{x_i,y_i}, s_{x_i+L/2,y_i}] + g[s_{x_i,y_i}, s_{x_i,y_i+L/2}])/2, \quad (1)$$

where $g[s, s']$ denotes a spin-spin correlation between a spin pair s and s'.

In Fig. 1, we show examples of the spin configuration $\{s_i\}$, correlation configuration $\{g_i(L/2)\}$, and an improved correlation configuration $\{\tilde{g}_i(L/2)\}$ of the 2D three-state Potts model. The spin configuration is generated by the Monte Carlo simulation, and the correlation configuration and the improved correlation configurations are calculated from the spin configuration. The upper figures are snapshots at the low temperature, $T = 0.9$, and the lower figures are those at the high temperature, $T = 1.06$. Temperatures are measured in units of the coupling J. We note that the exact second-order transition temperature T_c for this model is known as $1/\ln(1 + \sqrt{3}) \approx 0.995$.

Spins are displayed in one of three colors, red, green, or blue. The ordinary correlation takes a value of 1, $-1/2$, or $+1/4$, whereas the improved correlation takes a value of 1, 0, or $+1/2$. Both correlations from $+1$ to -1 are mapped in grayscale from 255 (white) to 0 (black). The permutation of three-state spins yields an essentially identical configuration, and the correlation configurations are invariant under the permutation. The borders of FK clusters for the spin configuration are drawn by lines in Figs. 1(a) and 1(d). They are copied in an improved correlation configuration. The border of the largest cluster is drawn as a thick yellow line for convenience.

At high temperatures, the spin configurations and the correlation configurations are randomly distributed, and the fluctuation of these quantities gives the susceptibility. In the improved correlation, the cancellation among different FK clusters is automatically satisfied. Figures 1(e) and 1(f) show the difference between the two correlation configurations. While the ordinary correlation configuration in Fig. 1(e) fluctuates in space, a couple of brighter areas in the largest cluster show the improved correlation in Fig. 1(f).

For convenience, we provide animations of the spin configuration, the correlation configuration, and the improved correlation configuration for the 2D Ising model (.mp4 files) in the Supplemental Material [22]. The animations at various temperatures are compared at the low temperature ($T = 2.1$), at $T_c = 2.269$, and at the high temperature ($T = 2.4$). The system sizes are $L = 32$ and 64.

We use the same technique of supervised learning as Shiina *et al.* [3] for the classification of the phases of the spin systems. We consider a fully connected neural network implemented with a standard TENSORFLOW library [23] using the 100-hidden-unit model to classify the ordered, the BKT, and the disordered phases. For the input layer, we use the improved correlation configurations $\{\tilde{g}_i(L/2)\}$. We have used a cross-entropy cost function supplemented with an $L2$ regularization term. The neural networks were trained using the Adam method [24].

We first analyzed the 2D three-state Potts model. The output layer averaged over a test set as a function of T for the 2D three-state Potts model is shown in Fig. 2(a). The probabilities of predicting the phases, disordered or ordered, are plotted for each temperature. The system sizes are $L = 32, 48$, and 64. The samples of T within the ranges $0.85 \leqslant T \leqslant 0.94$ and $1.06 \leqslant T \leqslant 1.15$ were used for the training data. We have not used the samples close to T_c for the training data because we assumed the situation that the exact T_c is not known. For a whole temperature range, around 35 000 training data sets are used, and we use 500 test data sets for each temperature. Ten independent calculations were performed to provide an error analysis. This figure corresponds to Fig. 2(a) of Ref. [3], and we again observe that the neural network could successfully classify the disordered and ordered phases using the improved correlation configuration. In the inset of Fig. 2(a), we show a comparison of the results of the improved correlation (the present study) and those of the previous study [3] of the
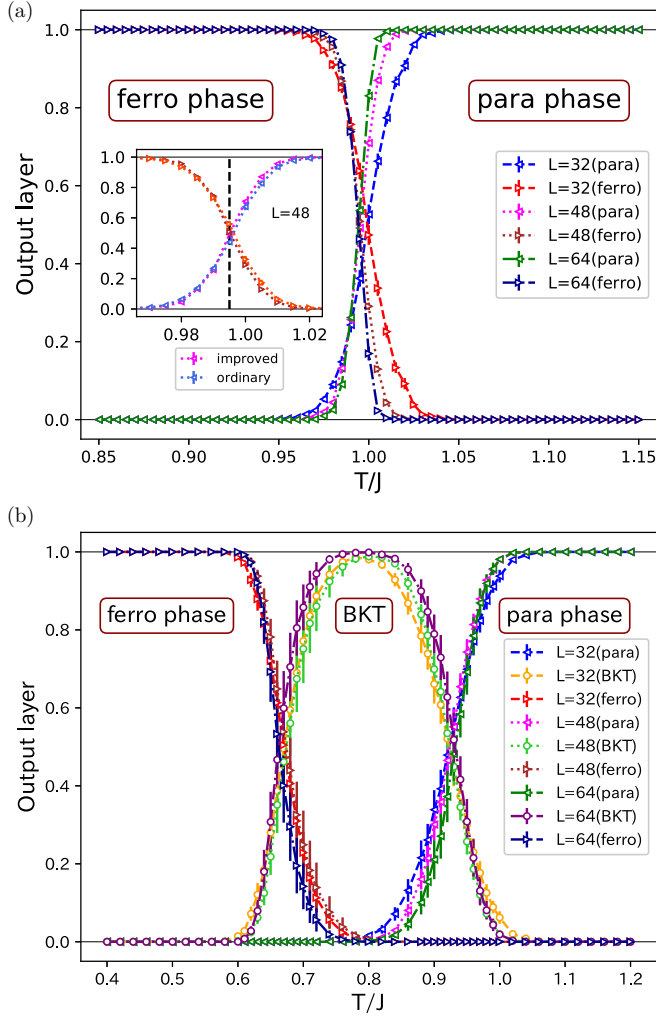
FIG. 2. (a) The output layer averaged over a test set as a function of T for the 2D three-state Potts model. The system sizes are $L = 32$, 48, and 64. The samples of T within the ranges $0.85 \leqslant T \leqslant 0.94$ and $1.06 \leqslant T \leqslant 1.15$ are used for the training data. In the inset, the comparison is made between the results of improved correlation and those of ordinary correlation in the case of $L = 48$. The exact T_c, $1/\ln(1 + \sqrt{3}) \approx 0.995$, is shown as a dashed line for convenience. (b) The same plot for the 2D six-state clock model. The system sizes are the same. The samples of T within the ranges $0.4 \leqslant T \leqslant 0.64$, $0.77 \leqslant T \leqslant 0.83$, and $0.96 \leqslant T \leqslant 1.2$ are used for the training data.

ordinary correlation in the case of $L = 48$. We used the same conditions for both the training data and test data of the improved and ordinary correlations produced from the same spin configurations. The point where the probabilities of predicting two phases are 50% is slightly closer to the exact critical temperature, shown as the dashed line in the inset, for the improved correlation, but the difference is small. The advantage of the improved estimator appears at high enough temperatures [compare Fig. 1(f) with Fig. 1(e)].

We next consider the 2D six-state clock model. Because of the discreteness, there are two transitions. One is a higher BKT transition, T_2, between the disordered and BKT phases, and the other is a lower transition, T_1, between the BKT and ordered phases. The output layer averaged over a test set as a function of T for the 2D six-state clock model is shown

in Fig. 2(b). The system sizes are $L = 32$, 48, and 64. The samples of T within the ranges $0.4 \leqslant T \leqslant 0.64$, $0.77 \leqslant T \leqslant 0.83$, and $0.96 \leqslant T \leqslant 1.2$ were used for the low-temperature, midrange temperature, and high-temperature training data, respectively. The recent numerical estimates of T_1 and T_2 for the six-state clock model are 0.701(5) and 0.898(5), respectively [25]. This figure corresponds to Fig. 4(a) of Ref. [3], and the present figure again shows the successful classification into the three phases.

We have classified the phases of transitions by means of the machine-learning approach by Shiina *et al.* [3] using the improved correlation configuration. There is no appreciable difference in accuracy between the use of the correlation configuration and that of the improved correlation configuration. The result indicates that the machine-learning-based phase classification is robust; that is, the phase classification does not discriminate the improved correlation configuration from the ordinary one.

Many applications of the loop updating method have been done for quantum systems. Here, we consider the quantum spin-$1/2$ XY model in two dimensions, which clearly demonstrated the utility of the loop algorithm [20]. The Hamiltonian is written as

$$H = -J \sum_{\langle i,j \rangle} \left(\hat{S}_i^x \hat{S}_j^x + \hat{S}_i^y \hat{S}_j^y \right). \tag{2}$$

Here, the spin operators $\hat{S}^{x,y}$ are one-half of the Pauli matrices $\sigma^{x,y}$. The summation is taken over the nearest-neighbor pairs. This model exhibits the BKT transition at around $T = 0.342$ [20].

We performed the quantum Monte Carlo simulation using the loop algorithm, and calculated the spatial correlation with a distance of $L/2$. A D-dimensional quantum system can be treated as a $(D + 1)$-dimensional classical system with an extra dimension of imaginary time. In calculating the spatial correlation, the summation over the imaginary-time axis is taken. The $\hat{S}^x$ component of the correlation function is calculated as [26]

$$\tilde{g}_{\boldsymbol{r}_i, \boldsymbol{r}_j}^x = \frac{4}{\beta^2} \int_0^\beta \int_0^\beta \hat{S}^x(\boldsymbol{r}_i, \tau_1) \hat{S}^x(\boldsymbol{r}_j, \tau_2) d\tau_1 d\tau_2$$

$$= \frac{1}{\beta^2} \int_0^\beta \int_0^\beta \delta_\ell(\boldsymbol{r}_i, \tau_1; \boldsymbol{r}_j, \tau_2) d\tau_1 d\tau_2, \tag{3}$$

where $\delta_\ell(\cdot)$ is the function that returns 1 if the loop of the position $\boldsymbol{r}_i$ and the time τ_1 and that of the position $\boldsymbol{r}_j$ and the time τ_2 belong to the same loop, whereas it returns 0 otherwise. Due to the $O(2)$ symmetry of the model, the $\hat{S}^y$ component of the correlation function is exactly the same as the $\hat{S}^x$ component [26]. A factor of 4 is introduced for the comparison of the spin-$1/2$ system with the classical model. We checked our calculation by the consistency with the precise calculations at $T = 0$ [27,28].

It is instructive to compare the correlation configurations of the quantum XY model and the classical XY model. In Fig. 3, examples of the snapshots of $\{\tilde{g}_i(L/2)\}$ of two models are displayed. At high temperatures above T_{BKT} [Figs. 3(b) and 3(d)], both improved configurations represent the behavior of a finite correlation length. At low temperatures below T_{BKT} [Figs. 3(a) and 3(c)], they are different from the high-temperature

(a) quantum $T = 0.24$

(b) quantum $T = 0.44$
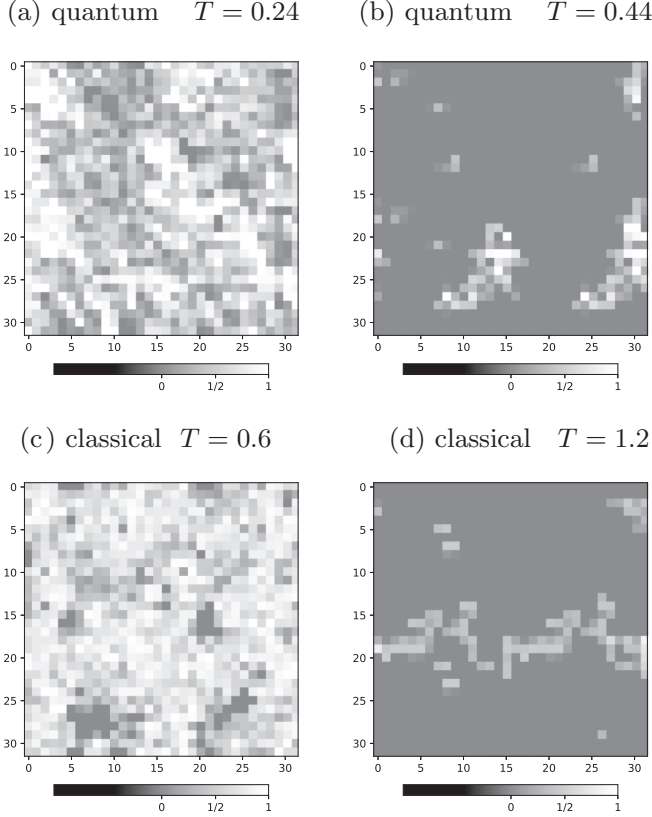
(c) classical $T = 0.6$

(d) classical $T = 1.2$

FIG. 3. Comparison of snapshots of $\{\tilde{g}_i(L/2)\}$ of the 2D quantum [(a), (b)] and classical [(c), (d)] XY models. Examples of snapshots below T_{BKT} [(a), (c)] and those above T_{BKT} [(b), (d)] are displayed.

configurations and at the same time they are different from the behavior of the ordered state, which was shown in Fig. 1(c). (Note that the precise estimate of the BKT temperature of the classical XY model is $T_{\mathrm{BKT}} = 0.8929$ [29].)

The classification of the BKT and paramagnetic phases of the spin-1/2 XY model using the machine-learning technique is shown in Fig. 4(a). The samples of T within the ranges $0.16 \leqslant T \leqslant 0.32$ and $0.38 \leqslant T \leqslant 0.54$ were used for the BKT-temperature and high-temperature training data, respectively. If we estimate the value of T_{BKT} from the point that the probabilities of predicting two phases are 50%, this temperature becomes around $T = 0.40$. It is slightly higher than the precise estimate for the infinite system, $T_{\mathrm{BKT}} = 0.342$ [20], although this temperature becomes lower as the system size increases. We also tested the classification of the quantum XY model using the training data of the classical model. For the classical model, not only the classical XY model (plane rotator) but also the anisotropic Heisenberg model with the XY interaction was treated. This anisotropic Heisenberg model has an out-of-plane fluctuation and the BKT transition temperature is slightly lowered at around $T_{\mathrm{BKT}} = 0.70$ [30,31]. In Fig. 4(b), we show the result of the classification of the quantum XY model using the training data of the classical XY model (plane rotator). We reproduced the BKT transition of the quantum XY model. The same conclusion was obtained when using the anisotropic Heisenberg model as the training data. The classification into two phases is slightly sharper for
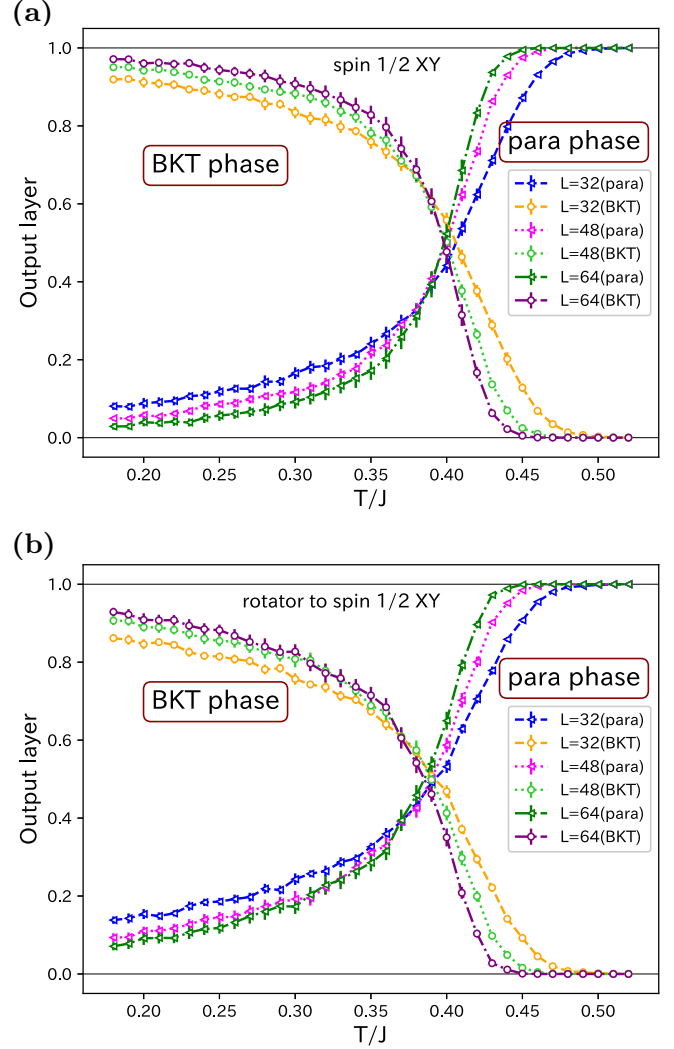

FIG. 4. (a) The output layer averaged over a test set as a function of T for the 2D spin-1/2 XY model. The system sizes are $L = 32$, 48, and 64. The samples of T within the ranges $0.16 \leqslant T \leqslant 0.32$ and $0.38 \leqslant T \leqslant 0.54$ are used for the training data. (b) The classification of the quantum XY model using the training data of the classical XY model (plane rotator). The samples of T within the ranges $0.50 \leqslant T \leqslant 0.84$ and $0.96 \leqslant T \leqslant 1.30$ for the classical XY model are used for the training data.

the anisotropic Heisenberg model than the classical XY model (plane rotator). The opposite direction, using the training data of the quantum model in the classification of the classical models, was also successful.

To summarize, we have proposed a method to use the improved estimator of the correlation configuration in the machine-learning study of the phase classification of spin models. For the classical spin systems, we have demonstrated the machine-learning studies of the 2D three-state Potts model (the second-order transition) and the 2D six-state clock model (the BKT transition). The results were compared with those of a previous study [3] using the ordinary correlation instead of the improved correlation. The method was also applied to the quantum Monte Carlo simulation using the loop algorithm. We treated the spin-1/2 quantum XY model, and

analyzed the BKT transition of the model. We emphasize that the classification scheme based on the training data of the classical XY model can be used for the phase classification of the quantum model. It indicates the universality of the phase transition, and at the same time, the generalized feature of the phase classification based on the machine learning. We also point out the effectiveness of the improved estimators in the loop algorithm to bridge classical and quantum Monte Carlo simulations.

We have opened a door to using the improved estimators for the machine-learning study of quantum systems. It is not trivial whether loop clusters in quantum spin systems can be identified with FK clusters in classical spin systems [32]. In this study, we clarified that the phase classification using machine learning does not discriminate between loop clusters and FK clusters. The BKT transition of the present study is a thermal phase transition. The investigation of a quantum phase transition at $T = 0$ will be interesting. For future studies, we may list several models for spin and charge degrees of freedom with loop algorithms. Examples are several quantum spin models, strongly correlated electron models, hard-core boson models, optical lattices, etc.

Another direction of future study is related to the inverse renormalization group approach [33]. Efthymiou *et al.* [34] have proposed a method to increase the size of the lattice spin configuration using superresolution, deep convolutional neural networks. At high temperatures, however, there is a problem that the noise is largely random and difficult to learn. The present improved correlation configuration could reduce this difficulty at high temperatures.

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

[2] J. Carrasquilla and R. G. Melko, Nat. Phys. **13**, 431 (2017).

[3] K. Shiina, H. Mori, Y. Okabe, and H. K. Lee, Sci. Rep. **10**, 2177 (2020).

[4] V. L. Berezinskii, Sov. Phys. JETP **32**, 493 (1971).

[5] V. L. Berezinskii, Sov. Phys. JETP **34**, 610 (1972).

[6] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).

[7] J. M. Kosterlitz, J. Phys. C **7**, 1046 (1974).

[8] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **58**, 86 (1987).

[9] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989).

[10] P. Kasteleyn and C. Fortuin, J. Phys. Soc. Jpn. Suppl. **26**, 11 (1969).

[11] C. Fortuin and P. Kasteleyn, Physica **57**, 536 (1972).

[12] U. Wolff, Nucl. Phys. B **322**, 759 (1989).

[13] U. Wolff, Nucl. Phys. B **334**, 581 (1990).

[14] Y. Tomita and Y. Okabe, Phys. Rev. Lett. **86**, 572 (2001).

[15] H. G. Evertz, G. Lana, and M. Marcu, Phys. Rev. Lett. **70**, 875 (1993).

[16] H. G. Evertz, Adv. Phys. **52**, 1 (2003).

[17] B. B. Beard and U.-J. Wiese, Phys. Rev. Lett. **77**, 5130 (1996).

[18] N. Kawashima and K. Harada, J. Phys. Soc. Jpn. **73**, 1379 (2004).

[19] J. Gubernatis, N. Kawashima, and P. Werner, *Quantum Monte Carlo Methods: Algorithms for Lattice Models* (Cambridge University Press, Cambridge, UK, 2016).

[20] K. Harada and N. Kawashima, Phys. Rev. B **55**, R11949 (1997).

[21] Y. Tomita and Y. Okabe, Phys. Rev. B **66**, 180401(R) (2002).

[22] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.102.021302 for animations of the spin configuration, the correlation configuration, and the improved correlation configuration for the 2D Ising model (.mp4 files).

[23] M. Abadi *et al.*, https://tensorflow.org, arXiv:1603.04467.

[24] D. P. Kingma and J. Ba, arXiv:1412.6980.

[25] T. Surungan, S. Masuda, Y. Komura, and Y. Okabe, J. Phys. A: Math. Theor. **52**, 275002 (2019).

[26] R. Brower, S. Chandrasekharan, and U.-J. Wiese, Physica A: Stat. Mech. Appl. **261**, 520 (1998).

[27] A. W. Sandvik and C. J. Hamer, Phys. Rev. B **60**, 6588 (1999).

[28] H.-Q. Lin, J. S. Flynn, and D. D. Betts, Phys. Rev. B **64**, 214411 (2001).

[29] M. Hasenbusch, J. Phys. A: Math. Gen. **38**, 5869 (2005).

[30] H. G. Evertz and D. P. Landau, Phys. Rev. B **54**, 12302 (1996).

[31] T. Figueiredo, J. Rocha, and B. Costa, Physica A: Stat. Mech. Appl. **488**, 121 (2017).

[32] Y. Nonomura and Y. Tomita, Phys. Rev. E **101**, 032105 (2020).

[33] D. Ron, R. H. Swendsen, and A. Brandt, Phys. Rev. Lett. **89**, 275701 (2002).

[34] S. Efthymiou, M. J. S. Beach, and R. G. Melko, Phys. Rev. B **99**, 075113 (2019).