Improved dynamical scaling analysis using the kernel method for nonequilibrium relaxation

Yuki Echinaka and Yukiyasu Ozeki
Department of Engineering Science, Graduate School of Informatics and Engineering,
The University of Electro-Communications, 1-5-1 Chofugaoka, Chofu-shi, Tokyo 182-8585, Japan
(Received 10 August 2016; published 21 October 2016)

The dynamical scaling analysis for the Kosterlitz-Thouless transition in the nonequilibrium relaxation method is improved by the use of Bayesian statistics and the kernel method. This allows data to be fitted to a scaling function without using any parametric model function, which makes the results more reliable and reproducible and enables automatic and faster parameter estimation. Applying this method, the bootstrap method is introduced and a numerical discrimination for the transition type is proposed.

DOI: 10.1103/PhysRevE.94.043312

I. INTRODUCTION

The transition into the Kosterlitz-Thouless (KT) phase [1–4] can be found in models such as the XY and q-state clock models in two dimensions and it corresponds to experimentally observed phenomena such as in the super-fluid film of ^{4}He. In the KT phase, there is no spontaneous magnetization, but the correlation length always diverges. Though theoretical studies for the KT transition have been reported, there are some difficulties in its numerical analysis. In KT transitions, the correlation length diverges exponentially as the system approaches the transition temperature, unlike the usual second-order transitions in which it diverges according to a power law. Therefore, numerical simulations require very long computation times to reach the system equilibrium near the KT transition temperature, making it difficult to conduct analysis using equilibrium simulations.

This difficulty has been overcome by the nonequilibrium relaxation (NER) method [5] combined with dynamical scaling analysis [6]. The NER method is an efficient numerical technique for analyzing equilibrium phase transitions. It provides the critical temperature and critical exponents accurately for second-order transition systems and has been used successfully to study problems including frustrated and/or random systems. In NER analysis, the equilibration step is not necessary. The simulation only proceeds up to the point at which the asymptotic behavior indicates the properties of the equilibrium state. Thus, one can analyze larger systems than with equilibrium simulations. The NER method has been extended beyond second-order transitions, including the KT transition and the first-order transition [7–20].

Regarding the KT transition, NER analysis is an efficient tool because of the dynamical scaling for the relaxation of the order parameter. To estimate the KT transition temperature, one may calculate the relaxation of the order parameter for several temperatures and fit the data to a dynamical scaling formula. As the equilibration step is not necessary and the treatment is systematic, irrespective of the detail of interactions, it has been used successfully to study various problems including frustrated and/or random systems [5,18]. Despite such results, the method still faces some problems (see the next section for details). (i) In the fitting procedure, a model of the scaling function is necessary. (ii) Although simulations are necessary to calculate the relaxation of magnetization near the transition temperature, it is difficult to obtain data showing the critical crossover in the regime close to the transition point. (iii) It remains problematic to estimate the error bars.

In the present study, we propose an improved dynamical scaling analysis for NER data. The Bayesian inference and the kernel method, which have already been applied to finite-size scaling analysis for static data [21], are used to solve the above-mentioned problems in the conventional dynamical scaling analysis. In the improved method, one can fit data to a scaling function without using any parametric model function. The result does not depend on the scaling method, making it more reliable and reproducible. This makes the fitting task simple and efficient and realizes automatic and faster parameter estimation. Further, we propose applications suitable for our method; the bootstrap method is introduced to improve the precision and obtain error bars for the estimated parameters and a numerical discrimination is presented for the transition type.

The remainder of this paper is organized as follows. The NER method and its dynamical scaling analysis are explained in Sec. II. In Sec. III the improved dynamical scaling analysis is introduced and demonstrated. Section IV emphasizes the advantages of our method and proposes an application to the bootstrap method. In Sec. V another application, the discrimination of the transition type, is proposed. Section VI gives a summary of this study and our results.

II. NONEQUILIBRIUM RELAXATION ANALYSIS

The NER analysis for an equilibrium phase transition is based on the relaxation of the order parameter $m(t)$, which estimates the transition temperature, and a dynamical exponent. In the case of ferromagnetism, it is the magnetization from the all-aligned state. It is expected that $m(t)$ decays to zero exponentially in the paramagnetic (PM) phase and to the value of the spontaneous magnetization $m_{\rm eq}$ in the ferromagnetic (FM) phase. The algebraic decay appears at the critical temperature. The asymptotic behavior is summarized as

$$m(t) \sim \begin{cases} \exp(-t/\tau), & T > T_c \\ t^{-\lambda}, & T = T_c \\ m_{\rm eq} & T < T_c, \end{cases} \tag{2.1}$$

where τ is the relaxation time and λ is a dynamical exponent. To distinguish the asymptotic relaxation behavior, it is

convenient to introduce the following logarithmic derivative of $m(t)$:

$$\lambda(t) = -\frac{d \ln m(t)}{d \ln t}. \tag{2.2}$$

The function $\lambda(t)$ is called the local exponent for the dynamical exponent λ. In practice, we estimate the transition temperature T_c from the temperature at which $\lambda(t)$ clearly approaches zero as the lower bound and that at which $\lambda(t)$ shows diverging tendency as the upper bound. Note that these estimated lower and upper bounds of T_c are reliable, because the estimated error bar defined by the difference between them does not include a statistical uncertainty.

This NER analysis of the order parameter can be applied to the case of KT transition systems, although some difficulties arise. The asymptotic behavior of the order parameter is expected to be

$$m(t) \sim \begin{cases} \exp(-t/\tau), & T > T_{\mathrm{KT}} \\ t^{-\lambda(T)}, & T \leqslant T_{\mathrm{KT}}. \end{cases} \tag{2.3}$$

The dynamical exponent $\lambda(T)$, the asymptotic power of the relaxation, is defined as completely in the KT phase and depends on the temperature. One cannot estimate the lower bound of T_{KT}, only the upper bound. Thus, it is difficult to estimate a reliable error bar.

To overcome this difficulty in applying the NER method to transition temperature estimates in KT transition systems, we introduce a dynamical scaling analysis based on the following natural scaling form [18]:

$$m(t,T) = \tau^{-\lambda} \Psi(t/\tau), \tag{2.4}$$

where $m(t,T)$ is the relaxation of magnetization obtained by the NER method and λ is a dynamical exponent. The relaxation time τ depends on the temperature and is expected to diverge at the KT transition temperature with the asymptotic form

$$\tau(T) = a \exp\left(\frac{b}{\sqrt{T - T_{\mathrm{KT}}}}\right), \tag{2.5}$$

similarly to the correlation length. Note that the same dynamical scaling can be applied to the second-order transition, where a typical power-law form

$$\tau(T) = a|T - T_c|^{-b} \tag{2.6}$$

is substituted for the asymptotic form of the relaxation time.

To estimate T_{KT} we can calculate $m(t,T)$ in a sufficient interval of Monte Carlo Steps (MCSs) for several values of T and fit the data to the above formula. Let us use the label i for all data points $m(t_i,T_i)$. The corresponding relaxation time is also dependent on i, i.e., τ_i, which should be identical for those with the same temperature, i.e., $\tau_i = \tau_j$ when $T_i = T_j$. If one assumes the scaling law, all data points converted as

$$X_i \equiv t_i/\tau_i, \tag{2.7}$$

$$Y_i \equiv \tau_i^{\lambda} m(t_i,T_i), \tag{2.8}$$

$$E_i \equiv \tau_i^{\lambda} \delta m(t_i,T_i) \tag{2.9}$$

should collapse according to a scaling function as

$$Y_i = \Psi(X_i), \tag{2.10}$$

where $\delta m(t_i,T_i)$ is the statistical error of $m(t_i,T_i)$ estimated in the simulation and E_i is that of Y_i. Estimating the critical exponent and corresponding transition temperature is known as scaling analysis (dynamical scaling).

As mentioned in the Introduction, although this method has been used successfully to study problems such as frustrated and/or random systems, it faces certain problems. (i) A model function for $\Psi(x)$ is necessary for the fitting procedure. This presents the problem of overfitting. (ii) We have used least-squares fitting or some standard nonlinear fitting method. Thus, simulations are needed to calculate the relaxation of magnetization near the transition temperature and observe the critical crossover from the critical power-law relaxation to the off-critical exponential one. The closer to the transition point the simulation is performed at, the longer MCS we need to reach the critical crossover. As the NER analysis uses data from large lattices in which the finite-size effect does not appear in the observed MCS, it is difficult to obtain data showing the critical crossover in the regime close to the transition point. (iii) It remains a problem to estimate the error bars.

III. IMPROVEMENT OF DYNAMICAL SCALING

In this section we present an improved dynamical scaling for NER data and resolve the difficulties encountered by the previous method. In the present method we use the kernel method with Bayesian statistics, which have been used for finite-size scaling analysis [21,22] and are known to be efficient. The basic formalism for the dynamical scaling is similar to that for the finite-size scaling case. We explain this in detail in the following, because the parameter t (the MCSs) appears in the dynamical analysis, unlike the static scaling case.

Let us use the vector form $\vec{Y} \equiv \{Y_1, Y_2, \ldots\}$. The conditional probability of observing the values of the scaling function Ψ by $Y_1, Y_2, \ldots$ with conditions of physical parameters $\vec{\theta}_p$ such as the transition temperature and/or the critical exponents is denoted by $P(\vec{Y}|\Psi,\vec{\theta}_p)$ and is assumed to be a multivariate Gaussian distribution with mean vector $\vec{\Psi}$ and covariance matrix $\mathcal{E}$ given by

$$P(\vec{Y}|\Psi,\vec{\theta}_p) \equiv \mathcal{N}(\vec{Y}|\vec{\Psi},\mathcal{E}), \tag{3.1}$$

where

$$(\vec{Y})_i \equiv Y_i, \tag{3.2}$$

$$(\vec{\Psi})_i \equiv \Psi(X_i), \tag{3.3}$$

$$(\mathcal{E})_{ij} \equiv E_i^2 \delta_{ij}, \tag{3.4}$$

and

$$\mathcal{N}(\vec{Y}|\vec{\mu},\Sigma) \equiv \frac{1}{\sqrt{|2\pi\,\Sigma|}} \exp\left(-\frac{1}{2}(\vec{Y} - \vec{\mu})^t \Sigma^{-1} (\vec{Y} - \vec{\mu})\right). \tag{3.5}$$

Introducing a statistical model $P(\Psi|\vec{\theta}_h)$ for the scaling function, where $\vec{\theta}_h$ contains the control parameters (hyper parameters), the conditional probability of $\vec{Y}$ for $\vec{\theta}_p$ and $\vec{\theta}_h$

is defined as

$$P(\vec{Y}|\vec{\theta}_p,\vec{\theta}_h) \equiv \int P(\vec{Y}|\Psi,\vec{\theta}_p)P(\Psi|\vec{\theta}_h)d\Psi. \tag{3.6}$$

In Bayes's theorem, the conditional probability of $\vec{\theta}_p$ and $\vec{\theta}_h$ for $\vec{Y}$ is written as

$$P(\vec{\theta}_p,\vec{\theta}_h|\vec{Y}) = P(\vec{Y}|\vec{\theta}_p,\vec{\theta}_h)P(\vec{\theta}_p,\vec{\theta}_h)/P(\vec{Y}), \tag{3.7}$$

where $P(\vec{\theta}_p,\vec{\theta}_h)$ denotes the prior distribution for $\vec{\theta}_p$ and $\vec{\theta}_h$ and $P(\vec{Y})$ denotes the prior distribution of $\vec{Y}$. In Bayesian statistics, the conditional probability $P(\vec{\theta}_p,\vec{\theta}_h|\vec{Y})$ is called a posterior estimator, by which a posterior probability of $\vec{\theta}_p$ and $\vec{\theta}_h$ can be estimated for a given $\vec{Y}$. The resulting parameter values are estimated by maximizing this conditional probability, known as the maximum *a posteriori* (MAP) estimate. In general, it is difficult to obtain any information on the prior distribution $P(\vec{\theta}_p,\vec{\theta}_h)$. Thus, we assume that it is uniform. In addition, $P(\vec{Y})$ can be recognized as a constant to normalize the posterior probability, allowing Eq. (3.7) to be rewritten as

$$P(\vec{\theta}_p,\vec{\theta}_h|\vec{Y}) \propto P(\vec{Y}|\vec{\theta}_p,\vec{\theta}_h). \tag{3.8}$$

In this formalism, the MAP estimator is identical to the maximization of the likelihood function

$$L(\vec{\theta}_p,\vec{\theta}_h) = P(\vec{Y}|\vec{\theta}_p,\vec{\theta}_h), \tag{3.9}$$

which is called the maximum likelihood (ML) estimator. In practice, the log-likelihood function $\ln L(\vec{\theta}_p,\vec{\theta}_h)$ is used in the ML estimator. From Eqs. (3.1)–(3.6), as the posterior probability is expressed by a multivariate Gaussian distribution, the log-likelihood function can be written as

$$\ln L(\vec{\theta}_p,\vec{\theta}_h) \equiv -\tfrac{1}{2}\ln|2\pi\,\Sigma| - \tfrac{1}{2}(\vec{Y}-\vec{\Psi})^t\Sigma^{-1}(\vec{Y}-\vec{\Psi}), \tag{3.10}$$

where the covariance matrix Σ is obtained by an inner product of model functions of the scaling function. In the previous method, various model functions can be tested. Instead, the present approach applies the kernel method and considers the scaling function Ψ to be nonparametric. Then, Ψ is expressed by a linear combination of kernel functions $K(X_i,X_j)$ and the covariance matrix Σ is assumed to be

$$\Sigma = \mathcal{E} + \Sigma', \tag{3.11}$$

$$(\Sigma')_{ij} = K(X_i,X_j). \tag{3.12}$$

The Bayesian inference based on Eqs. (3.10)–(3.12) is called a Gauss process regression, where the scaling function is given by

$$\Psi(X) \equiv \vec{k}^t\Sigma^{-1}\vec{Y}, \tag{3.13}$$

$$(\vec{k})_i \equiv K(X_i,X). \tag{3.14}$$

Following the static case [21], we use the Gauss kernel function

$$K(X_i,X_j) = \theta_0^2\delta_{ij} + K_G(X_i,X_j), \tag{3.15}$$

$$K_G(X_i,X_j) \equiv \theta_1^2\exp\left(-\frac{(X_i-X_j)^2}{2\theta_2^2}\right), \tag{3.16}$$

where θ_0, θ_1, and θ_2 are hyper parameters. Then the covariance matrix Σ in Eq. (3.10) is obtained by

$$(\Sigma)_{ij} = \begin{cases} \theta_1^2\exp\left(-\frac{(X_i-X_j)^2}{2\theta_2^2}\right), & i \neq j \\ \theta_0^2+\theta_1^2+E_i^2, & i=j. \end{cases} \tag{3.17}$$

To maximize the log-likelihood function practically, we use the conjugate gradient algorithm [23]. This algorithm requires the derivative of the log-likelihood function

$$\begin{aligned}\frac{\partial \ln L(\vec{\theta}_p,\vec{\theta}_h)}{\partial\theta} &= -\frac{1}{2}\mathrm{Tr}\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\theta}\right) - (\vec{Y}-\vec{\Psi})^t\Sigma^{-1}\frac{\partial(\vec{Y}-\vec{\Psi})}{\partial\theta} \\ &\quad -\frac{1}{2}(\vec{Y}-\vec{\Psi})^t\Sigma^{-1}\frac{\partial\Sigma}{\partial\theta}\Sigma^{-1}(\vec{Y}-\vec{\Psi}),\end{aligned} \tag{3.18}$$

where the iterations are performed by modifying the parameters along the direction of steepest descent. In this study, we perform dynamical scaling for the logarithms of X_i and Y_i as

$$\begin{aligned} X_i' &\equiv \ln X_i = \ln t_i - \left(\ln a + \frac{b}{\sqrt{T_i - T_{\mathrm{KT}}}}\right), \\ Y_i' &\equiv \ln Y_i = \ln m(t_i,T_i) + \lambda\left(\ln a + \frac{b}{\sqrt{T_i - T_{\mathrm{KT}}}}\right), \\ E_i' &\equiv \ln\left(1+\frac{E_i}{Y_i}\right) \approx \frac{E_i}{Y_i} = \frac{\delta m(t_i,T_i)}{m(t_i,T_i)}. \end{aligned} \tag{3.19}$$

We apply dynamical scaling analysis using the present method with previously calculated data from the XY model in two dimensions [5]: The Hamiltonian is given by

$$\mathcal{H} = -J\sum_{\langle ij\rangle}\cos(\phi_i-\phi_j), \tag{3.20}$$

where $\langle ij\rangle$ runs over all nearest-neighboring pairs on a square lattice. We use

$$m(t) = N^{-1}\sum_i \cos\phi_i(t) \tag{3.21}$$

as the dynamical order parameter at time t, which is usually averaged over independent samples. Hereafter, we measure the temperature in the unit of J/k_{B}. The calculations were carried out on a 1001×1000 square lattice with skew boundary conditions up to an observation time of 1.5×10^5 MCSs. Approximately 256 independent samples were taken for statistical averaging at each temperature. In the NER analysis of KT transitions, we usually prepare relaxation data for about 10^3-10^5 MCS points for each temperature. Note that the proposed method requires the inverse matrix to be calculated; the dimension of the matrix N_{m} is the total number of input data points. As the calculation time is proportional to N_{m}^3, we do not use all of the data points and, practically, we sample about 10^3 points from the data. We use data for 12 temperatures, as shown in Fig. 1, and sample 100 points for each temperature so as to give equal intervals of $\ln t$ (the X axis). To minimize the log-likelihood function (3.10), the parameters T_{KT}, λ, a, b, θ_0, θ_1, and θ_2 appearing in Eqs. (2.4),

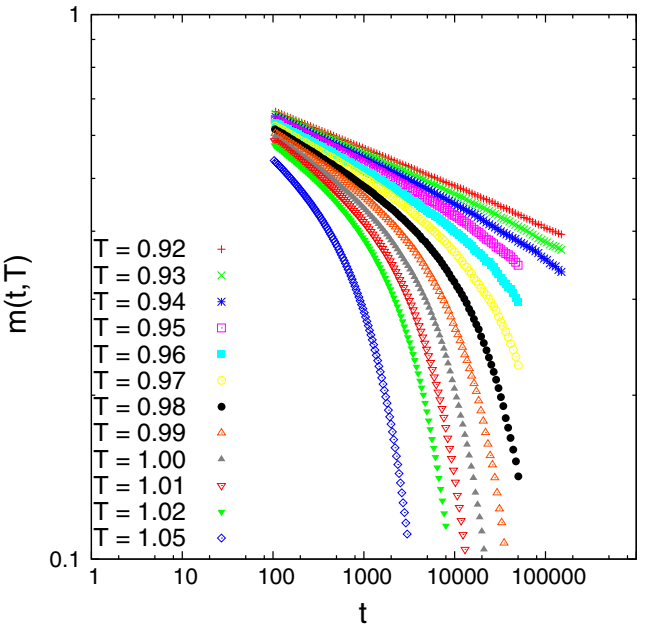

FIG. 1. Relaxation of magnetization for the XY model in two dimensions for $T = 0.92,0.93,0.94,0.95,0.96,0.97,0.98,0.99,1.00,1.01,1.02,1.05$ plotted on a double-logarithmic scale. For each temperature, 100 data points are chosen so as to give equal intervals on the X axis.

(2.5) and (3.16) are optimized. The results are shown in Fig. 2 with $T_{\mathrm{KT}} = 0.8955$ and $\lambda = 0.069$, which are consistent with the estimates of $T_{\mathrm{KT}} = 0.894(4)$ and $\lambda = 0.068(6)$ given by the previous method [18]. The error bars are not shown here, but will be introduced later to demonstrate the advantage of our method.

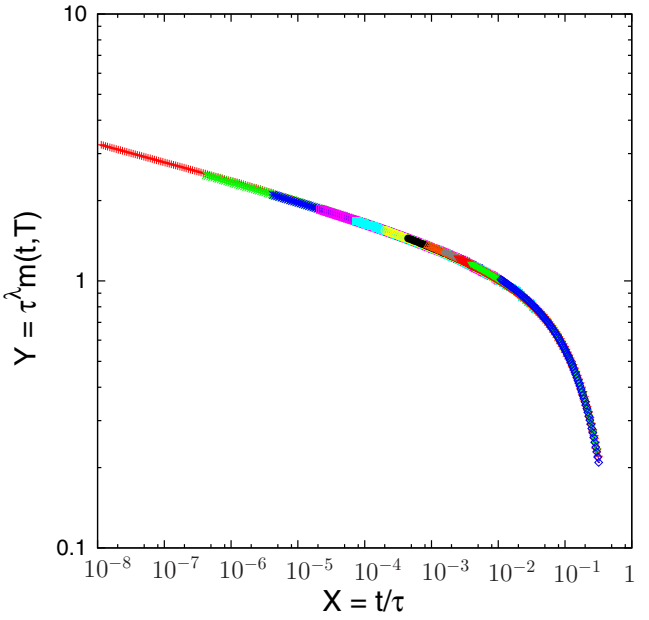

FIG. 2. Scaling plot for the data in Fig. 1 with $T_{\mathrm{KT}} = 0.8955$ and $\lambda = 0.069$.

IV. ADVANTAGES OF THE IMPROVED DYNAMICAL SCALING ANALYSIS

The proposed method for dynamical scaling analysis has two advantages over the previous approach. First, by means of the kernel method for constructing the scaling function, there is no uncertainty regarding the suitability of various scaling functions. This reduces the workload of a trial-and-error approach. Next, by means of the conjugate gradient method for minimizing the likelihood function, the scaling function can be optimized almost automatically and proceeds much faster. Note that this automation is another result of using the kernel method to guarantee a suitable trial function. These advantages enhance the effectiveness of dynamical scaling analysis.

A. Automation and efficiency

In the previous method, relaxation data were fitted for every two adjacent temperatures and the relative relaxation time was estimated because of the slower optimization. This means that we needed to calculate the relaxation so that the scaling function Ψ for the two temperatures overlapped. As the relaxation time increases rapidly in KT systems, it can be difficult to add data for temperatures close to the transition point, where the relaxation time is too large to overlap with those of adjacent temperatures. The faster optimization process in the present method allows us to fit relaxation data for all analyzed temperatures simultaneously, which provides a reliable scaling function over these temperatures and the adjacent regime. Thus, temperatures for which the relaxation data have no overlap with other analyzed temperatures can be used for the dynamical scaling.

As an example, in Fig. 3, we add data for $T = 0.90,0.91$, which have no overlap with the original data in Fig. 1. The

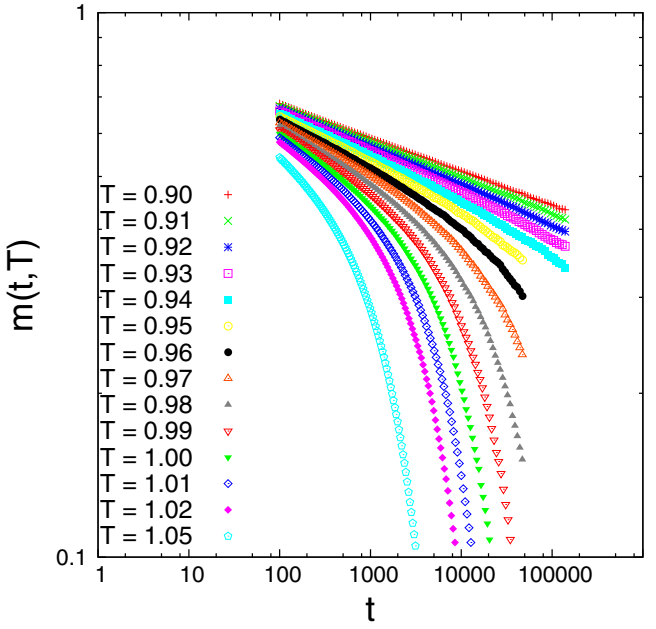

FIG. 3. Relaxation of magnetization for the XY model in two dimensions including $T = 0.91,0.90$. The data points for $T = 0.90$ have no overlap with those for other temperatures.

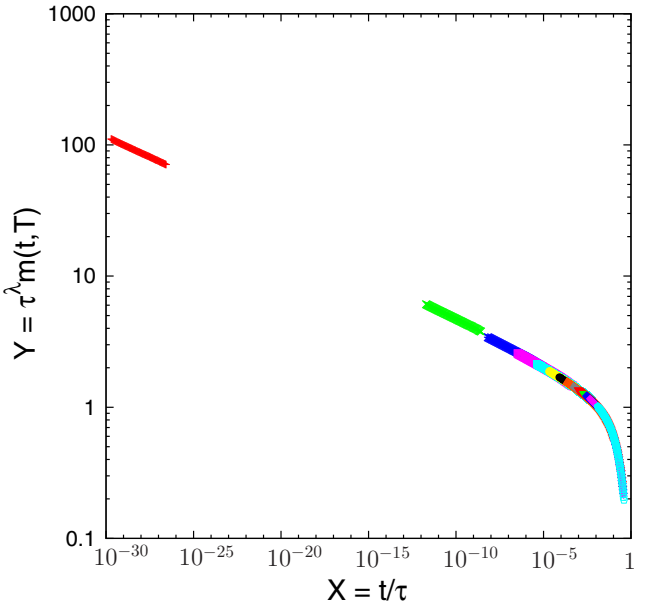

FIG. 4. Scaling plot for the data in Fig. 3 with $T_{\rm KT} = 0.8977$ and $\lambda = 0.070$. The isolated sector corresponds to the data for $T = 0.90$.

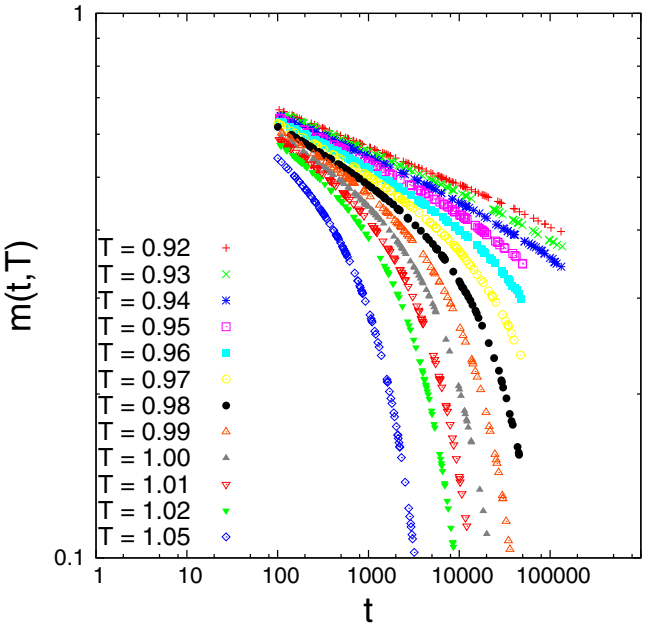

FIG. 5. Example of random sampling for the relaxation of magnetization for the XY model in two dimensions. For each temperature, 100 data points are chosen at random along the X axis.

resulting scaling fit given by the present method is shown in Fig. 4. There is no overlap between the data for $T = 0.90$ and those for other temperatures. Regardless, we obtain $T_{\rm KT} = 0.8977$ and $\lambda = 0.070$, which agree with the values obtained by the previous method.

B. Bootstrap method

One remarkable feature of dynamical scaling in NER analysis is the huge number of data points required to fit the scaling. In the usual finite-size scaling analysis, one observes the temperature dependence of some physical quantities for several different-size systems. Thus, the number of data points is the number of temperatures multiplied by the number of sizes, which is $O(10) \sim O(10^2)$. In NER analysis, one observes the relaxation of some physical quantities for several temperatures. Thus, the number of data points is the number of temperatures multiplied by the maximum number of observation steps, which is $O(10^3) \sim O(10^6)$. This feature has not yet been exploited. The efficiency and automation of the improved dynamical scaling process can be employed in the following bootstrap method [24,25], which systematically obtains precise estimates and their error bars. This provides a clear criterion for the accuracy of the result.

We perform multiple dynamical scaling analyses for randomly sampled data and compute various statistics from the measured results. Thereby, it is possible to improve the precision of the estimates and estimate the error. The relaxation data used in the analysis for Figs. 1 and 2 amount to 1.5×10^5 points for each temperature. These data points can be divided into various samples. In Sec. III the samples were divided into 100 points for each temperature so as to give equal intervals of $\ln t$. Here we prepare a sample by randomly choosing 100 points along the $\ln t$ axis for each temperature and then perform dynamical scaling for each sample.

We use data for 12 temperature values, as shown in Fig. 1. In Fig. 5 we show an example of the sampled data. Applying the present dynamical scaling to these data, we obtain the optimized scaling function in Fig. 6 and estimate $T_{\rm KT}$ and λ for this sample. Repeating this process 100 times, we can compute the average values and their statistical errors for parameters such as the transition temperature and critical exponent. The resulting values of the bootstrap method are obtained as $T_{\rm KT} = 0.8954(4)$ and $\lambda = 0.0690(2)$. The precision is one digit better than that given by the conventional method,

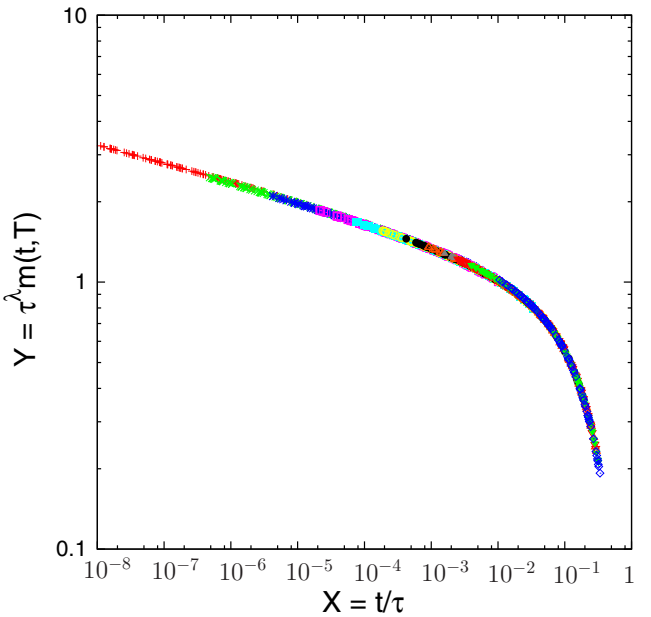

FIG. 6. Scaling plot for the data in Fig. 5.

$T_{\mathrm{KT}} = 0.894(4)$ and $\lambda = 0.068(6)$. Note that the estimated errors of the bootstrap method in the present study are not the confidence intervals usually shown in NER analysis, but are instead intervals of high reliability that can be estimated from the prepared data.

V. DISCRIMINATION OF TRANSITION TYPE

We propose another application for the improved dynamical scaling with automation and faster performance, namely, the discrimination of phase transition features from the relaxation of the order parameter. In some theoretical systems, there have been controversies regarding the discrimination between second-order and KT transitions, e.g., the five-state clock model in two dimensions [26,27] and the RP^2 model in two dimensions [28–32]. Numerical simulations are rarely able to solve such controversies. Here we examine some methods to discriminate between these transitions using the improved dynamical scaling. The basic idea is to compare the fit of the dynamical scaling with the asymptotic forms of τ between Eq. (2.5) for the KT transition and Eq. (2.6) for the second-order transition. We propose two methods. To check their efficiency, we apply them to the Ising model in two dimensions as a typical second-order transition system and the XY model in two dimensions as a typical KT transition system.

A. Consistency of estimated transition temperature

For a controversial system, as in previous sections, we can calculate the relaxation of the order parameter for temperatures close to the expected transition point. Then, for both transition types, dynamical scaling is performed under the assumption that the asymptotic function of τ is given by Eq. (2.5) or (2.6). The transition temperature and some exponents are estimated for each type. If the assumption is incorrect, there will generally be some inconsistency between the relaxation data close to the transition point. These inconsistencies allow us to identify the transition type.

As an example, consider the Ising model in two dimensions. The Hamiltonian is given by

$$\mathcal{H} = -J \sum_{\langle ij \rangle} S_i S_j, \tag{5.1}$$

where S_i takes ± 1 and $\langle ij \rangle$ runs over all nearest-neighbor pairs on a square lattice. We use

$$m(t) = N^{-1} \sum_i S_i(t) \tag{5.2}$$

as the dynamical order parameter at time t. The valid transition is the second-order type. Calculations were carried out on a 1501×1500 square lattice with skew boundary conditions up to an observation time of 10^5 MCSs. About 1024 independent samples were taken for statistical averaging at $T = 2.268$ and 2.270–2.279, where the exact transition temperature is given by $T_{\mathrm{c}} = 2.269\ldots$. In Fig. 7 we show relaxation data for $T = 2.268$ and 2.270. Note the upward trend for $T = 2.268$, indicating $T < T_{\mathrm{c}}$, and the downward trend for $T = 2.270$, indicating $T > T_{\mathrm{c}}$. In the standard NER analysis for the FM transition, we would conclude that the transition temperature was located between these two temperatures. In Fig. 8 we show

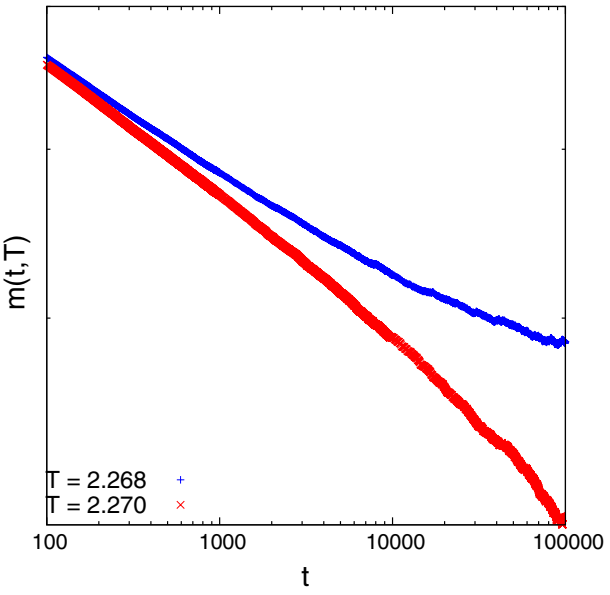

FIG. 7. Relaxation of magnetization for the Ising model in two dimensions for $T = 2.268$ and 2.270 with a double-logarithmic scale. Note the upward trend for $T = 2.268$ and the downward trend for 2.270.

relaxation data for $T = 2.270$–2.279, where the data points were sampled following Sec. III. Two different dynamical scalings were performed for these data. For the assumption of the second-order transition, Eq. (2.6) was used to find $T_{\mathrm{c}} = 2.269\,23$, which is consistent with the observation in Fig. 7. For the assumption of the KT transition, Eq. (2.5) was used to obtain $T_{\mathrm{KT}} = 2.267\,24$, which contradicts the behavior of the data at $T = 2.268$. This provides the conclusion that the second-order transition is more suitable. Hence,

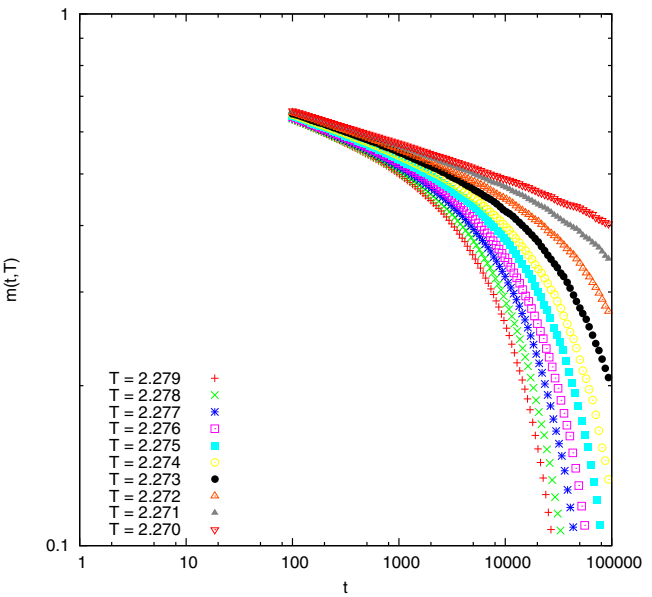

FIG. 8. Relaxation of magnetization for the Ising model in two dimensions for $T = 2.270, 2.271, 2.272, 2.273, 2.274, 2.275, 2.276, 2.277, 2.278, 2.279$ with a double-logarithmic scale.

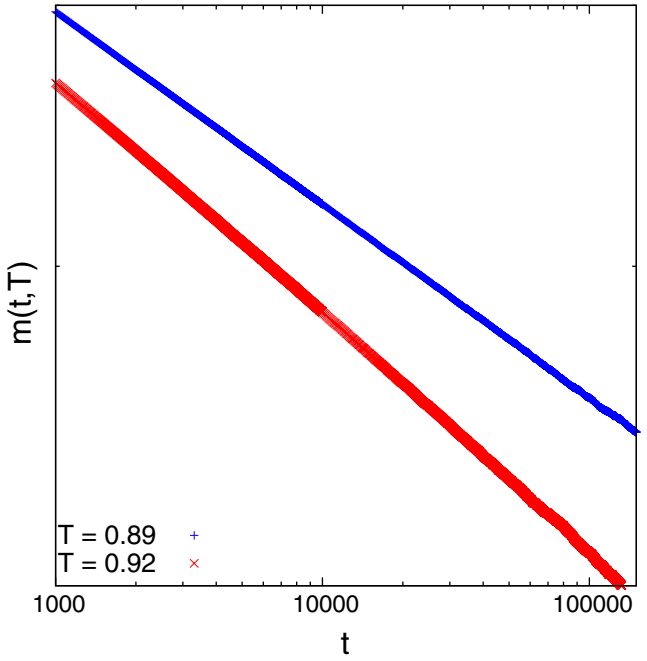

FIG. 9. Relaxation of magnetization for the XY model in two dimensions for $T = 0.89$ and 0.92 with a double-logarithmic scale. The trend is weak and hard to observe, but the data indicate a downward trend for $T = 0.92$ and an almost straight line for 0.89.

dynamical scaling reveals the correct transition temperature if the asymptotic function of τ is assumed to be the second-order transition case, as expected.

Next we examine the XY model in two dimensions, where the valid transition is the KT type. Using the data in Fig. 1, we conducted a simulation at $T = 0.89$, close to the KT transition temperature. In Fig. 9 we show relaxation data for $T = 0.89$ and 0.92. One can observe the downward trend for $T = 0.92$, indicating $T > T_{\mathrm{KT}}$. Note that, while this trend is very weak, it can be found by analyzing the data carefully. The power-law behavior observed for $T = 0.89$ does not give any bound, but suggests that $T \sim T_{\mathrm{KT}}$. In Sec. III we estimated $T_{KT} = 0.8955$ using dynamical scaling for Fig. 1 under the assumption of the KT transition, Eq. (2.5). Using the assumption of a second-order transition, Eq. (2.6) gives $T_{\mathrm{c}} = 0.9295$, which contradicts the behavior of the data at $T = 0.92$ in Fig. 9. This leads to the conclusion that the KT transition is more suitable. We show that dynamical scaling reveals a correct transition temperature if the asymptotic function of τ is assumed to be the KT transition case, as expected.

We have demonstrated the proposed method for systems exhibiting KT and second-order transitions. In both cases we obtained estimates for the transition temperature that are consistent with the relaxation data close to the transition point when the correct asymptotic form of τ was used. This provides a means of discriminating between transition types, although it is somehow difficult to observe the local trend in relaxation behavior.

B. Dynamical scaling with no assumption for τ

Let us propose another method for discriminating between transition types. In the present dynamical scaling scheme, we fit the relaxation data by introducing the values of τ for all temperatures as the fitting parameters, rather than using the asymptotic form of τ. If, for example, the temperatures of the simulated data are $T_1, T_2, \ldots, T_M$, the corresponding τ values, i.e., $\tau_1, \tau_2, \ldots, \tau_M$, are introduced and optimized by a similar scaling process. This provides the temperature dependence of τ for the chosen temperatures without any assumption regarding its asymptotic form. Then we have three different dynamical scaling results for the function $\tau(T)$: (i) the assumption of second-order type using Eq. (2.6), (ii) the assumption of KT type using Eq. (2.5), and (iii) no assumption of the transition type. We can compare these to determine the most suitable transition type.

Note that, even if the assumption of one particular transition type is correct and the assumed asymptotic function is suitable, the result cannot give a better fit than that without any assumption. Thus, one can discriminate the transition type by checking how close the result with some assumption is to that without the assumption. Such a comparison could take many forms; we use the residual for the logarithm of τ, defined by

$$r = \frac{1}{M}\sum_{m=1}^{M}\{\ln \tau(T_m) - \ln \tau_m\}^2, \tag{5.3}$$

where $\tau(T_m)$ is estimated through the assumption (i) or (ii) using Eq. (2.6) or (2.5), respectively, and τ_m is estimated through no assumption, i.e., (iii).

We present examples for both the Ising and XY models in two dimensions. For the Ising model, we used the relaxation data in Fig. 8. The τ values obtained by the three types of dynamical scaling are plotted in Fig. 10. At a glance, the

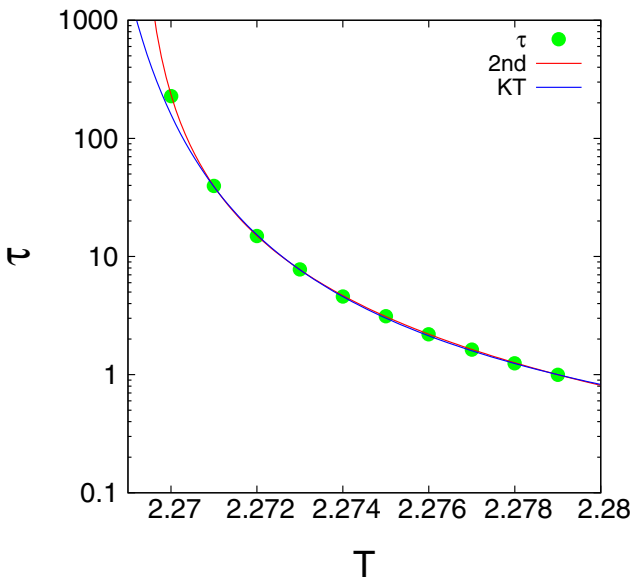

FIG. 10. Estimated relaxation time τ for the Ising model in two dimensions. Three different dynamical scaling results are plotted for the function $\tau(T)$ under the assumption of a second-order transition, KT transition, and without assumption (closed circles).

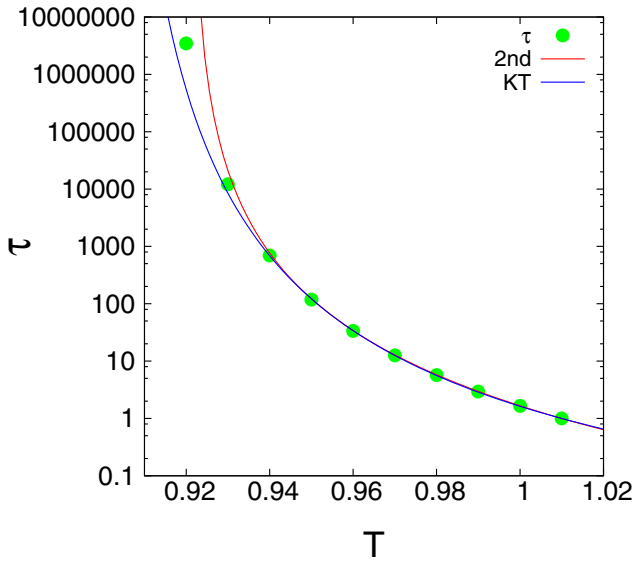

FIG. 11. Estimated relaxation time τ for the XY model in two dimensions. Three different dynamical scaling results are plotted for the function $\tau(T)$, under the assumption of a second-order transition, KT transition, and without assumption (closed circles).

curves of τ for both assumptions are almost consistent with the values of τ estimated by the scaling without assumptions. For a more precise comparison, we used the residual defined in Eq. (5.3). This gave $r_{2nd} = 1.3 \times 10^{-4}$ for the assumption of the second-order transition and $r_{KT} = 9.5 \times 10^{-3}$ for the assumption of the KT transition. The residual for the correct assumption, i.e., the second-order transition, is suitable and much smaller than that for the KT type.

For the XY model, we used the relaxation data in Fig. 1. The τ values obtained by the three types of dynamical scaling are plotted in Fig. 11. With the exception of the point at the lowest temperature ($T = 0.92$), the curves of τ for both assumptions are almost consistent with the values of τ estimated by the scaling without assumptions. For a more precise comparison, we used the residual defined in Eq. (5.3). This gave $r_{2nd} = 5.8 \times 10^{-2}$ for the assumption of the second-order transition and $r_{KT} = 9.0 \times 10^{-3}$ for the assumption of the KT transition. The residual for the correct assumption, i.e., the KT transition, is suitable and several times smaller than that for the second-order type. Furthermore, as can be seen in Fig. 11, the curve for the second-order transition is diverging at temperatures greater than $T = 0.92$, which is inconsistent with the observation of the finite value of τ in the PM phase and supports the identification of the correct KT transition.

VI. SUMMARY

Using Bayesian statistics and the kernel method, which are known to be efficient for finite-size scaling analysis with static data [21], we have described an improved dynamical scaling analysis scheme for the Kosterlitz-Thouless transition in the nonequilibrium relaxation method. Data can be fitted to a scaling function without using any parametric model function. The results do not depend on the scaling method and are more reliable and reproducible. This makes the task of fitting simpler and more efficient and enables a faster, automated parameter estimation procedure. We presented two applications for this method. The bootstrap method was introduced to improve the precision and obtain error bars for the estimated parameters. The numerical discrimination of second-order and KT transitions was also described. This method can be used to solve controversial problems concerning the discrimination between second-order and KT transitions such as the five-state clock model in two dimensions [26,27] and the RP^2 model in two dimensions [28–32].

ACKNOWLEDGMENTS

This work was supported by JSPS KAKENHI Grant No. 15K05205. The authors would like to thank K. Harada and K. Hukushima for fruitful discussions and comments. The authors are also grateful to the Supercomputer Center at the Institute for Solid State Physics, University of Tokyo, for the use of their facilities.

[1] V. L. Berezinskii, Sov. Phys. JETP **32**, 493 (1971).
[2] V. L. Berezinskii, Sov. Phys. JETP **34**, 610 (1971).
[3] J. M. Kosterlitz and D. J. Thouless, J. Phys. C **6**, 1181 (1973).
[4] J. M. Kosterlitz, J. Phys. C **7**, 1046 (1974).
[5] Y. Ozeki and N. Ito, J. Phys. A **40**, R149 (2007).
[6] Z. B. Li, L. Schülke, and B. Zheng, Phys. Rev. Lett. **74**, 3396 (1995).
[7] N. Ito, Physica A **192**, 604 (1993).
[8] N. Ito, Physica A **196**, 591 (1993).
[9] N. Ito, K. Hukushima, K. Ogawa, and Y. Ozeki, J. Phys. Soc. Jpn. **69**, 1931 (2000).
[10] Y. Ozeki, N. Ito, and K. Ogawa, *Activity Report of the Super Computer Center in the Institute of Solid State Physics of the University of Tokyo* (University of Tokyo, Tokyo, 2000), p. 37.
[11] Y. Ozeki and N. Ito, J. Phys. A **31**, 5451 (1998).
[12] N. Ito, Y. Ozeki, and H. Kitatani, J. Phys. Soc. Jpn. **68**, 803 (1999).
[13] K. Ogawa and Y. Ozeki, J. Phys. Soc. Jpn. **69**, 2808 (2000).
[14] Y. Ozeki and N. Ito, Phys. Rev. B **64**, 024416 (2001).
[15] Y. Ozeki, N. Ito, and K. Ogawa, J. Phys. Soc. Jpn. **70**, 3471 (2001).
[16] H. J. Luo, L. Schülke, and B. Zheng, Phys. Rev. Lett. **81**, 180 (1998).
[17] H. J. Luo, L. Schülke, and B. Zheng, Phys. Rev. E **57**, 1327 (1998).
[18] Y. Ozeki, K. Ogawa, and N. Ito, Phys. Rev. E **67**, 026702 (2003).
[19] Y. Ozeki and N. Ito, Phys. Rev. B **68**, 054414 (2003).
[20] Y. Ozeki, K. Kasono, N. Ito, and S. Miyashita, Physica A **321**, 271 (2003).
[21] K. Harada, Phys. Rev. E **84**, 056704 (2011).

[22] C. M. Bishop, *Pattern Recognition and Machine Learning* (Springer, New York, 2006).
[23] E. Polak, *Computational Methods in Optimization* (Academic, New York, 1971).
[24] B. Efron, Ann. Stat. **7**, 1 (1979).
[25] B. Efron and R. Tibshirani, *An Introduction to the Bootstrap* (Chapman and Hall, London, 1993).
[26] S. K. Baek and P. Minnhagen, Phys. Rev. E **82**, 031102 (2010).
[27] Y. Kumano, K. Hukushima, Y. Tomita, and M. Oshikawa, Phys. Rev. B **88**, 104427 (2013).
[28] H. Kunz and G. Zumbach, Phys. Rev. B **46**, 662 (1992).
[29] E. Mondal and S. K. Roy, Phys. Lett. A **312**, 397 (2003).
[30] R. Paredes V., A. I. Fariñas-Sánchez, and R. Botet, Phys. Rev. E **78**, 051706 (2008).
[31] A. I. Fariñas-Sánchez, R. Botet, B. Berche, and R. Paredes, Condens. Matter Phys. **13**, 13601 (2010).
[32] Y. Tomita, Phys. Rev. E **90**, 032109 (2014).