Temperature scaling in nonequilibrium relaxation in three-dimensional Heisenberg model in the Swendsen-Wang and Metropolis algorithms

Yoshihiko Nonomura [*]

International Center for Materials Nanoarchitectonics, National Institute for Materials Science, Tsukuba, Ibaraki 305-0044, Japan

Yusuke Tomita[†]

College of Engineering, Shibaura Institute of Technology, Saitama 337-8570, Japan

Recently the present authors proposed the nonequilibrium-to-equilibrium scaling (NE-ES) scheme for the critical Monte Carlo relaxation process [Nonomura, J. Phys. Soc. Jpn. **83**, 113001 (2014)], which scales relaxation data in the whole simulation-time regions regardless of functional forms, namely, both for the stretched-exponential critical relaxation in cluster algorithms and for the power-law critical relaxation in local-update algorithms. In the present study, we generalize this scheme to off-critical relaxation process and scale relaxation data for various temperatures in the whole simulation-time regions. This proposal of the off-critical scaling in cluster algorithms cannot be described by the dynamical finite-size scaling theory based on the power-law critical relaxation. As an example, we investigate the three-dimensional classical Heisenberg model previously analyzed with the NE-ES [Nonomura and Tomita, Phys. Rev. E **93**, 012101 (2016)] in the Swendsen-Wang and Metropolis algorithms.

I. INTRODUCTION

The nonequilibrium relaxation (NER) method is one of the improved Monte Carlo schemes to study phase transitions against the critical slowing down. In general, basic formulation of the NER method is based on the critical relaxation, and off-critical behaviors are described by scaling analyses. In local-update algorithms, the critical relaxation is characterized by the power-law behavior of physical quantities, and the critical point is determined as the most probable point to exhibit such a behavior [1]. This NER behavior is derived from the dynamical finite-size scaling (DFSS) theory [2,3], and the off-critical scaling behavior is also derived from it.

Recently the present authors revealed that the critical NER behaviors in cluster algorithms [4,5] are described by the stretched-exponential simulation-time dependence in various classical spin systems [6–8] and in a quantum phase transition [9]. Although the critical point can be determined from such early-time relaxation behaviors, more precise estimation is possible from the nonequilibrium-to-equilibrium scaling (NE-ES) [6,8,9], which connects the early-time and equilibrium behaviors smoothly. In addition to these numerical findings, the present authors derived this relaxation formula phenomenologically in the Ising models in the Swendsen-Wang (SW) algorithm [10].

Although the DFSS is not defined in cluster algorithms, in the present article we generalize the NE-ES to the off-critical region and confirm this "temperature scaling" in the three-dimensional (3D) classical Heisenberg model in the SW

algorithm, which we analyzed precisely with the NE-ES [8]. Here we also show that this extended formalism is applicable even to local-update algorithms.

The outline of the present article is as follows: In Sec. II we briefly summarize the model and Monte Carlo method used in the present article and review the NER method, the DFSS, and the NE-ES. In Sec. III we derive the temperature scaling in cluster and local-update algorithms and compare the formula with the one obtained from the DFSS. In Sec. IV we numerically confirm the temperature scaling with the magnetic susceptibility in the 3D classical Heisenberg model. As typical cluster and local-update algorithms, the SW and Metropolis [11] ones are utilized. In the Metropolis algorithm, the conventional scaling analysis based on the DFSS is also made for comparison. In Sec. V these results are compared with each other and with the previous numerical results, and we propose a general framework to investigate critical phenomena efficiently by combining the present scheme and the NE-ES. The above descriptions are summarized in Sec. VI. In the Appendix, similar analyses of the absolute value of magnetization are summarized.

II. MODEL AND METHOD

In the present article, the 3D classical ferromagnetic Heisenberg model on a simple cubic lattice described by the Hamiltonian

$$\mathcal{H} = -J \sum_{\langle ij \rangle \in \text{n.n.}} \vec{S}_i \cdot \vec{S}_j, \quad |\vec{S}_i| = 1 \tag{1}$$

with summation over all the nearest-neighbor bonds and $J > 0$, is simulated with the SW-type cluster algorithm in which all the spin clusters are flipped with 50% probability at each

---

[*]nonomura.yoshihiko@nims.go.jp

[†]ytomita@shibaura-it.ac.jp

Monte Carlo step (MCS). Although the original SW algorithm [4] can be applied only to the Potts model [12], vector spin models such as the Heisenberg model can be treated by constructing spin clusters with respect to the Ising element of vector spins projected onto a randomly chosen direction at each MCS [5].

At the critical point T_c, all the physical quantities can be treated with the NER scheme. However, in the off-critical region, the situation changes drastically. The spontaneous magnetization is vanishing above T_c and its temperature dependence can be analyzed only for $T \leqslant T_c$. Although the absolute value of it shows a diverging behavior for $T > T_c$, such a behavior is nothing but that of the square root of the magnetic susceptibility. While the magnetic susceptibility shows diverging behaviors in both sides of T_c, such a behavior is observed after subtracting the contribution from the spontaneous magnetization for $T < T_c$. Critical exponents of the susceptibility and magnetization are different, and NER analysis of a quantity including two critical exponents is quite complicated. Moreover, discontinuity of relaxation behaviors below and above T_c results in the restriction of initial states in the NER process. That is, NER started from the perfectly ordered state (corresponding to the configuration at $T = 0$) can be applied only for $T \leqslant T_c$, and that from the perfectly disordered states (one of the configurations at $T = \infty$) for $T \geqslant T_c$.

To summarize the above arguments, the spontaneous magnetization can be analyzed from the perfectly ordered state for $T \leqslant T_c$, and the magnetic susceptibility from the perfectly disordered states for $T \geqslant T_c$. Although other physical quantities can also be treated in principle, those derived from the temperature derivative (i.e., correlation with energy, e.g., the specific heat) show larger fluctuations, and the correlation length is evaluated indirectly (from the scale dependence of the correlation function or from the wave-number dependence of the magnetic susceptibility), and therefore they are not preferred for precise estimation. The scaled critical exponents β/ν and γ/ν can be evaluated from the NE-ES, and the bare exponent γ from the temperature scaling of the magnetic susceptibility as will be seen later. All the critical exponents can be obtained from these three exponents through the scaling relations. Although the bare exponent β can also be estimated from the temperature scaling of the absolute value of magnetization, it is not as accurate as γ. Details will be explained in the Appendix.

Next, established scaling formulas are briefly reviewed. The DFSS for a quantity Q is expressed as [2]

$$Q(t, L; T) \sim L^{x_Q/\nu} f[L/\xi(T), t/\tau(T)], \qquad (2)$$

with the simulation time t, linear size L, critical exponent x_Q defined in $Q(\infty, \infty; T) \sim (T - T_c)^{-x_Q}$ for $T \to T_c$, scaling function f, correlation length $\xi(T) \sim (T - T_c)^{-\nu}$, and correlation time $\tau(T) \sim (T - T_c)^{-z\nu}$ in local-update algorithms. Assuming equivalence of the functional form of f with respect to t and L, these two parameters are related with each other as $L \sim t^{1/z}$, or

$$Q(t, T) \sim t^{x_Q/(z\nu)} f[t^{1/(z\nu)}(T - T_c)] \qquad (3)$$

for a fixed system size [1]. From this formula, the critical point T_c can be evaluated from the power-law

simulation-time dependence of $Q(t, T_c)$, and an off-critical scaling $t^{-x_Q/(z\nu)}Q(t, T)$ versus $t^{1/(z\nu)}(T - T_c)$ is derived.

Such a formula does not hold in cluster algorithms, because the stretched-exponential critical relaxation is not consistent with the power-law size dependence. Then, the NE-ES is derived from the critical simulation-time dependence, $Q(t; T_c) \sim \exp(ct^\sigma)$ (in the NER from the perfectly-disordered states), and equilibrium size dependence at T_c, $Q(L; T_c) \sim L^{x_Q/\nu}$. Combining these formulas, we have $L^{-x_Q/\nu}Q(t, L; T_c) \sim \exp(ct^\sigma - \ln L^{x_Q/\nu})$ [6], or in a more general form corresponding to Eq. (3),

$$Q(t, L; T_c) \sim L^{x_Q/\nu} f_{sc}(ct^\sigma - \ln L^{x_Q/\nu}), \qquad (4)$$

with a scaling function f_{sc} on the NE-ES. This scaling form has been confirmed in classical spin systems [6,8] and in a quantum phase transition [9].

III. TEMPERATURE SCALING

Similarly to the NE-ES, the temperature scaling in cluster algorithms is derived from the onset and equilibrium behaviors. Namely, from the initial-time critical relaxation $Q(t; T_c) \sim \exp(ct^\sigma)$ and the temperature dependence in equilibrium $Q(\infty, T) \sim (T - T_c)^{-x_Q}$, we have $Q(t, T)(T - T_c)^{x_Q} \sim \exp[ct^\sigma + \ln(T - T_c)^{x_Q}]$, or

$$Q(t, T) \sim (T - T_c)^{-x_Q} f_{tsc}[ct^\sigma + \ln(T - T_c)^{x_Q}], \qquad (5)$$

with a scaling function f_{tsc} on the temperature scaling. Although the above derivation seems more nontrivial than that of the NE-ES, usage of the initial-time critical-relaxation formula can be justified in comparison with the off-critical scaling (3), which consists of the initial-time dependence at T_c and its modification by a scaling function with temperature dependence.

The above derivation is also possible in local-update algorithms. From the initial-time critical relaxation $Q(t, T_c) \sim t^{x_Q/(z\nu)}$ and the temperature dependence in equilibrium $Q(\infty, T) \sim (T - T_c)^{-x_Q}$, we result in $Q(t, T)(T - T_c)^{x_Q} \sim [t^{1/(z\nu)}(T - T_c)]^{x_Q}$, or

$$Q(t, T) \sim (T - T_c)^{-x_Q} f_{tsc}[t^{1/(z\nu)}(T - T_c)]. \qquad (6)$$

In comparison with the conventional off-critical scaling form (3), the prefactor of the scaling function is changed from $t^{x_Q/(z\nu)}$ to $(T - T_c)^{-x_Q}$ in the present formalism.

IV. NUMERICAL RESULTS

A. Swendsen-Wang algorithm

First, we verify the temperature scaling in cluster algorithms (5) with the Swendsen-Wang (SW) algorithm. Here we concentrate on the magnetic susceptibility, i.e., $Q(t, T) \to \chi(t, T)$ and $x_Q \to \gamma$ in Eq. (5). In our previous article to investigate the 3D classical Heisenberg model with the NE-ES based on the SW algorithm [8], the maximum system size was $L = 560$. Here we also take $L = 560$ and 225 Monte Carlo steps (MCS) and average 4×10^4 random-number sequences (RNS). The raw data for various temperatures (from $T = 1.4436 J/k_B$ to $1.4520 J/k_B$) are shown in Fig. 1, together with the data at the most probable value of the critical point, $T_c = 1.442987 J/k_B$ [8]. At $t = 225$MCS, χ for $T =$
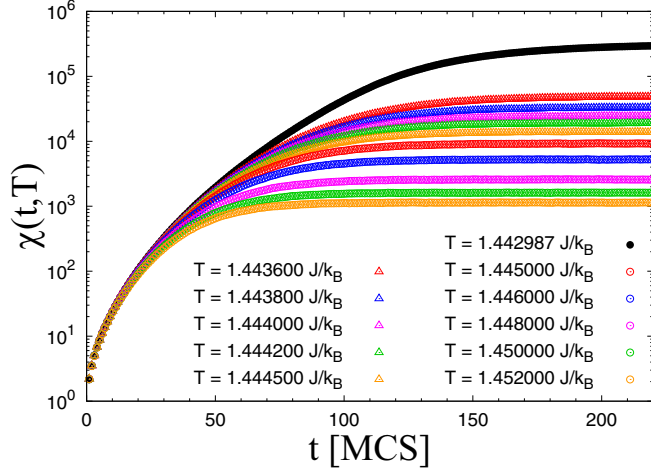
FIG. 1. Simulation-time dependence of the magnetic susceptibility for $L = 560$ at T_c and various temperatures used for the temperature scaling in the SW algorithm. The susceptibility decreases monotonically as the temperature increases.

$1.4436 \, J/k_B$ is about $1/6$ of that at $T = T_c$, while that at $T = 1.4520 \, J/k_B$ is about $1/40$ of that at $T = 1.4436 \, J/k_B$. Although the range of temperature for scaling does not seem so wide, that of χ is actually wide enough. In general, the temperature range of scaling is determined by the system size in the vicinity of T_c, and by the temperature itself far from T_c. Although the present formulation is based on the diverging behavior $\chi(t = \infty, L = \infty, T) \sim (T - T_c)^{-\gamma}$ for $T \to T_c$, the actual finite-size behavior is saturated with $\chi(t = \infty, L, T_c) \sim L^{\gamma/\nu}$, and the range of scaling near T_c increases as L increases. On the other hand, as temperature becomes away from T_c, the weight of the correction terms to scaling increases independently of L.
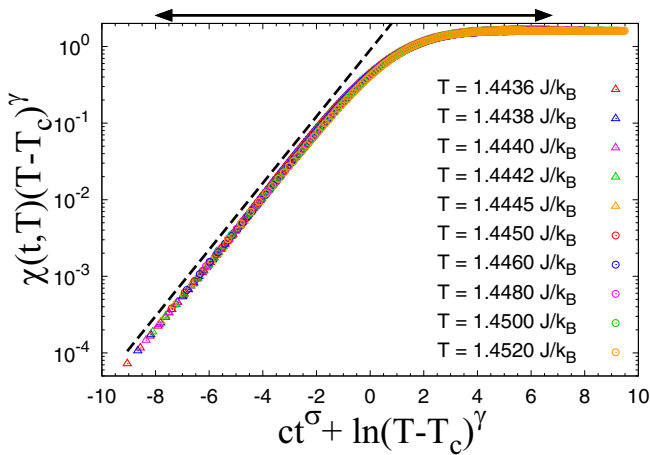

FIG. 2. Temperature scaling plot of the data in Fig. 1 using $T_c = 1.442987 \, J/k_B$ and $\sigma = 0.47$ [8] with $\gamma = 1.3945(19)$ and $c = 1.2595(43)$ in a semilog scale. The arrow specifies the range of data used for the fitting, and the dashed line corresponds to a simple exponential curve as a guide for eyes. Here the data for $t = 1$ MCS are not included.
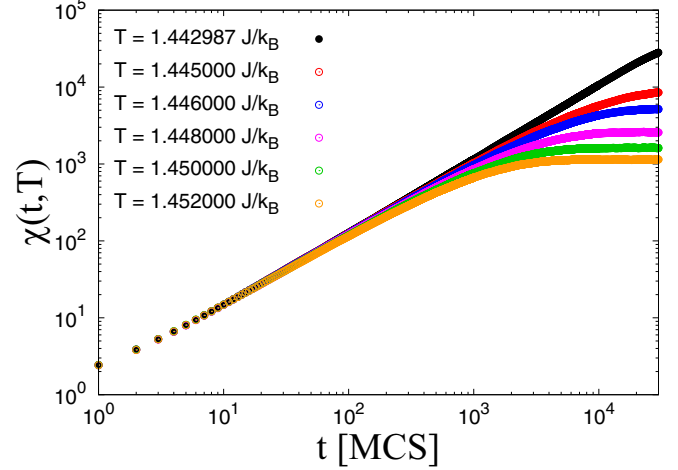

FIG. 3. Simulation-time dependence of the magnetic susceptibility for $L = 200$ at T_c [8] and various temperatures used for the temperature scaling in the Metropolis algorithm.

These data are scaled with Eq. (5) in Fig. 2, namely, the scaling plot of $\chi(t, T)(T - T_c)^\gamma$ versus $ct^\sigma + \ln(T - T_c)^\gamma$ in a semilog scale using $T_c = 1.442987(2) \, J/k_B$ and $\sigma = 0.47(1)$ evaluated in Ref. [8]. Since we take only the data rather far away from T_c, precise evaluation of T_c is difficult within the present scheme. It is also the case in the relaxation exponent σ. This exponent is characteristic to the critical relaxation in cluster algorithms, and appearance of it in Eq. (5) is just a trace of behaviors at T_c. Then it should be determined from the critical-relaxation data, not from the off-critical ones. The fitting parameters γ and c are estimated by minimizing the mutual residuals of these data. Although every two sets of the data can be scaled with each other, they are not independent and error bars cannot be evaluated in a simple way. Then we average the mutual residuals between the nearest-neighbor temperatures, determine the range of fitting by minimizing the averaged residual as shown by arrows in Fig. 2, and obtain

$$\gamma = 1.3945 \pm 0.0019, \quad c = 1.2595 \pm 0.0043. \quad (7)$$

Combining this estimate with $\gamma/\nu = 1.972 \pm 0.007$ obtained from the NE-ES at T_c [8], we have

$$\nu = 0.707 \pm 0.003. \quad (8)$$

B. Metropolis algorithm

Next, we test the temperature scaling in local-update algorithms (6) based on the Metropolis algorithm, and compare it with the standard off-critical scaling (3) for the same data. Here we also consider the magnetic susceptibility and take $Q(t, T) \to \chi(t, T)$ and $x_Q \to \gamma$ in these formulas. We take $L = 200$ and 3×10^4 MCS, and average 2×10^4 RNS. The raw data at T_c [8] and for various temperatures (from $T = 1.445 \, J/k_B$ to $1.452 \, J/k_B$) are plotted in a log-log scale in Fig. 3. Since the power-law relaxation at T_c is much slower than the stretched-exponential critical relaxation in the SW algorithm, much longer MCS are required and therefore the system size is reduced. The data at T_c still show a power-law behavior at $t = 3 \times 10^4$. When we attempt to evaluate T_c
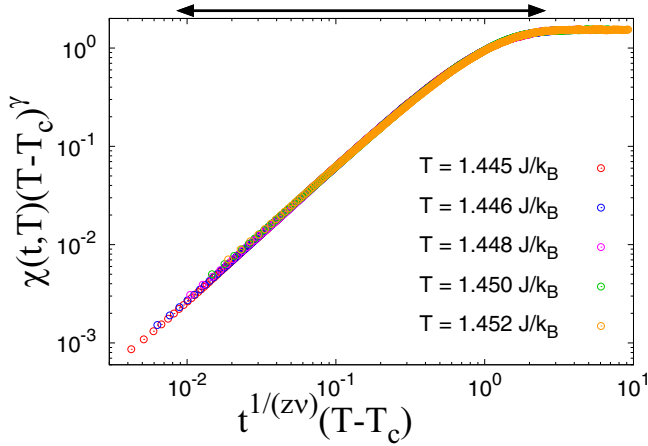
FIG. 4. Temperature scaling plot of the data in Fig. 3 using $T_c = 1.442987 \, J/k_B$ [8] with $\gamma = 1.4039(32)$ and $z\nu = 1.4866(64)$ in a log-log scale. The arrow specifies the range of data used for the fitting. Here the data for $t = 1, 2$ MCS are not included.

with the conventional NER, relaxation data at $T = 1.443 \, J/k_B$ cannot be distinguished from the present data at T_c, and the resolution of T_c becomes of one order lower than the one in Ref. [8]. In comparison with the previous subsection, the lowest temperature for scaling is increased in response to reduction of the system size, and the highest one is the same.

These data are scaled with Eq. (6) in Fig. 4, namely, the scaling plot of $\chi(t, T)(T - T_c)^\gamma$ versus $t^{1/(z\nu)}(T - T_c)$ in a log-log scale using $T_c = 1.442987(2) \, J/k_B$ [8]. The fitting parameters γ and $z\nu$ are estimated by minimizing the mutual residuals of these data. Since the relaxation process is much slower than that in the previous subsection, the number of data is further increased. When all the data are scaled with an equal weight, the contribution in the vicinity of equilibrium becomes dominant and the functional form in the whole simulation-time regions cannot be reproduced anymore. Then we reduce the density of data as sparse as that for 51–100 MCS in a log scale by averaging the sequential data points. That is, we take 100 points for 1–100 MCS, 50 points for 101–200 MCS, 60 points for 201–500 MCS, 50 points for 501–1000 MCS, 50 points for 1001–2000 MCS, 60 points for 2001–5000 MCS, 50 points for 5001–10 000 MCS, 50 points for 10 001–20 000 MCS, and 20 points for 20 001–30 000 MCS; totally we take 490 points for 1–30 000 MCS for the fitting. Based on this set of data and the fitting scheme similarly to that in the previous subsection, we have

$$\gamma = 1.4039 \pm 0.0032, \quad z\nu = 1.4866 \pm 0.0064. \quad (9)$$

Combining this estimate with ν in Eq. (8), we arrive at

$$z = 2.10 \pm 0.01. \quad (10)$$

Finally, we analyze the same data (those in Fig. 3 after the above thinning-out process) with the standard off-critical scaling (3), namely, the scaling plot of $t^{-\gamma/(z\nu)}\chi(t, T)$ versus $t^{1/(z\nu)}(T - T_c)$ as shown in Fig. 5. Using $T_c = 1.442987(2) \, J/k_B$ [8] and the above fitting scheme, we have

$$\gamma = 1.4024 \pm 0.0057, \quad z\nu = 1.4773 \pm 0.0067. \quad (11)$$

Combining this estimate with ν in Eq. (8), we obtain

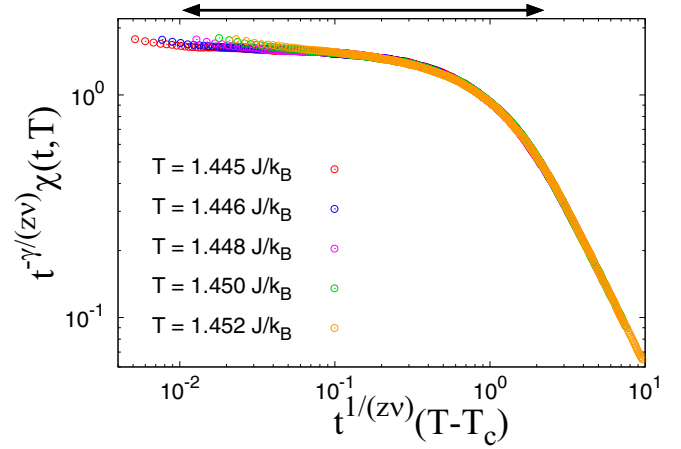$$z = 2.09 \pm 0.01. \quad (12)$$

FIG. 5. Off-critical scaling plot of the data in Fig. 3 based on the DFSS using $T_c = 1.442987 \, J/k_B$ [8] with $\gamma = 1.4024(57)$ and $z\nu = 1.4773(67)$ in a log-log scale. The arrow specifies the range of data used for the fitting. Here the data for $t = 1$–3 MCS are not included.

V. DISCUSSION

According to the most precise evaluation of the critical exponents of the 3D classical Heisenberg model to date [13], the exponents treated in the present article were given by $\gamma = 1.3957(22)$ and $\nu = 0.7113(11)$ by equilibrium Monte Carlo simulations. Our estimate of γ based on the SW algorithm (7) is comparable with this one. Although ours of ν (8) is rather underestimated, it is still within the error bar. Note that this tendency is not due to the present analysis, but the one based on the NE-ES at T_c, $\gamma/\nu = 1.972(7)$ [8]. From the estimates in Ref. [13], it is given by $\gamma/\nu = 1.962(4)$, and the underestimation of ν simply originates from the overestimation of γ/ν. Actually, in Ref. [13] the above MC analysis was coupled with the high-temperature expansion analysis, and they obtained more precise estimates $\gamma = 1.3960(9)$ and $\nu = 0.7112(5)$. Our estimate of γ is still consistent with it, even though it is rather underestimated.

The tendency of underestimation can be understood from the finite-size behavior of physical quantities in the vicinity of equilibrium. As explained in the previous section, the temperature scaling is based on the diverging behavior of physical quantities, e.g., $\chi(T) \sim (T - T_c)^{-\gamma}$ for $T \to T_c$. However, such a behavior is observed only in the thermodynamic limit, and in finite systems it saturates as $\chi(L, T_c) \sim L^{\gamma/\nu}$ even at $T = T_c$. Then, when the data too close to T_c in comparison with L are taken for the fitting, those become smaller than the ones expected from Eq. (5), which results in the underestimation of γ. On the other hand, the data far from T_c do not converge as sharply as a power with respect to $T - T_c$. When the data too far away from T_c are used for the fitting, those become larger than the ones expected from Eq. (5), which also causes the underestimation of γ.

Our estimate of γ based on the temperature scaling in the Metropolis algorithm (9) is overestimated (it is consistent with the previous estimate within 2σ). Although that based on the conventional off-critical scaling in the Metropolis algorithm (11) is consistent with the previous one, it is due to large error

bars and the most probable value itself is comparable with the one in Eq. (9) and is also overestimated. This tendency of overestimation originates from the early-time nonequilibrium behavior, not from the difference of system sizes ($L = 560$ and 200 in the SW and Metropolis algorithms, respectively). When we make a similar analysis with the $L = 200$ data in the SW algorithm, the tendency of underestimation is rather enhanced. As shown in Fig. 4, the data in the vicinity of equilibrium are actually not used in the fitting in the Metropolis algorithm. The dynamical critical exponent z is specific to the power-law critical relaxation in local-update algorithms, and the present estimate (10) may be comparable with that in the 3D Ising model, $z = 2.055(10)$ [14]. There were no previous studies on the dynamical critical behaviors in the 3D classical Heisenberg model, and we cannot argue this slight discrepancy in z too seriously at present.

Although the temperature scaling holds in both the SW and Metropolis algorithms, combination with the SW algorithm seems much better in the present analysis. Much larger systems can be treated owing to faster relaxation, and therefore critical phenomena can be evaluated more precisely. Moreover, the origin of the discrepancy from the previous estimate can be understood naturally. In addition, the temperature scaling can be compared with the conventional off-critical scaling in the Metropolis algorithm. While the two fitting parameters are separated in the temperature scaling, they are coupled in the conventional off-critical scaling. Then the error bar becomes twice larger in the latter, even though the most probable value of the estimate is comparable.

In the present article, we proposed the following procedure to determine critical phenomena with the cluster NER scheme:

(1) Determine T_c by the NE-ES on the magnetization and/or magnetic susceptibility.

(2) Determine β/ν and γ/ν by the NE-ES together with the above T_c.

(3) Determine γ by the temperature scaling using the above T_c.

(4) Evaluate other critical exponents through the scaling relations.

This is a minimum procedure, and precise evaluation of β within the present scheme seems difficult at present, as explained in the Appendix. However, from the scaling relation $\alpha + 2\beta + \gamma = 2$ and the hyperscaling relation $2 - \alpha = d\nu$, we have $2\beta/\nu + \gamma/\nu = d$. That is, evaluation of β is actually not necessary for the study on critical phenomena. If the critical exponent ν can be estimated from the temperature scaling of the correlation length ξ, the universality class can be identified only with the present scheme. Nevertheless, precise evaluation of T_c is not possible within this scheme, and the NE-ES of the critical relaxation is indispensable for the cluster NER.

VI. SUMMARY

In the present article, we proposed an extended scaling theory in the nonequilibrium relaxation process called temperature scaling, and we confirmed this theory on the magnetic susceptibility in the 3D classical Heisenberg model. When the temperature scaling was combined with the Swendsen-

Wang (SW) algorithm, it worked very well and our estimate of the critical exponent $\gamma = 1.3945(19)$ is comparable with the previous best estimate. When it was combined with the Metropolis algorithm, it worked as well as the conventional off-critical scaling, but not as well as the case with the SW algorithm, because of limitation of system sizes owing to slow relaxation.

APPENDIX: MAGNETIZATION IN THE SW ALGORITHM

Even if Monte Carlo simulations are started from the perfectly ordered state, the sign of magnetization may change in each step by a global flip of large clusters in the relaxation process in cluster algorithms. When the data of different random-number sequences are averaged, cancellation of signs takes place and the averaged results become meaningless. Then, in the cluster NER, we take the absolute value of magnetization. Here we start from the perfectly ordered state, simulate the $L = 560$ system during 225 MCS with the SW algorithm, and average 4×10^4 RNS. The relaxation data for various temperatures (from $T = 1.360 \, J/k_B$ to $1.435 \, J/k_B$ and at T_c) are displayed in Fig. 6.

Although the data at T_c decay on a stretched-exponential curve and do not arrive at equilibrium at $t = 225$ MCS, other data for $T < T_c$ seem to be already in equilibrium at that simulation time. Such relaxation behaviors are described by the following formula:

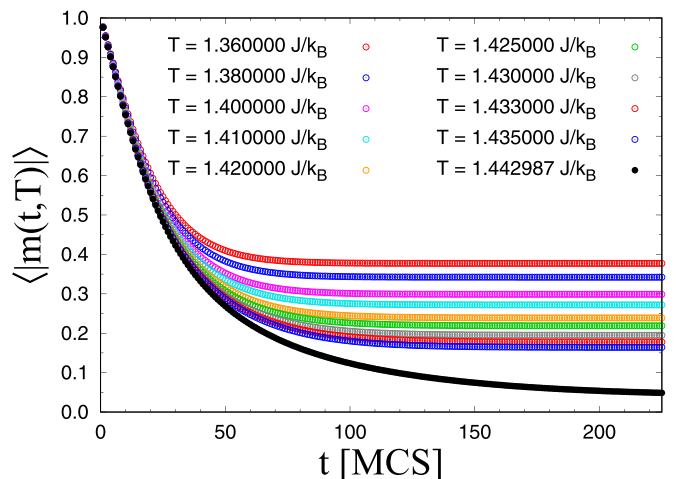$$\langle |m(t, T)| \rangle = m_s(T) + A(T) \exp[-C(T)t], \tag{A1}$$

FIG. 6. Simulation-time dependence of the absolute value of magnetization for $L = 560$ at T_c and various temperatures in the SW algorithm. The magnetization decreases monotonically as the temperature increases.
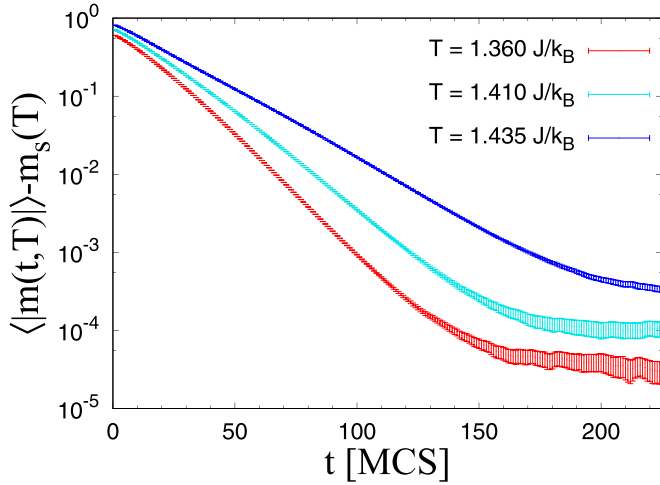
FIG. 7. Simulation-time dependence of the decaying part of the absolute value of magnetization based on Eq. (A1) in a semilog scale at $T = 1.360 \, J/k_B$, $1.410 \, J/k_B$, and $1.435 \, J/k_B$ (from bottom to top).
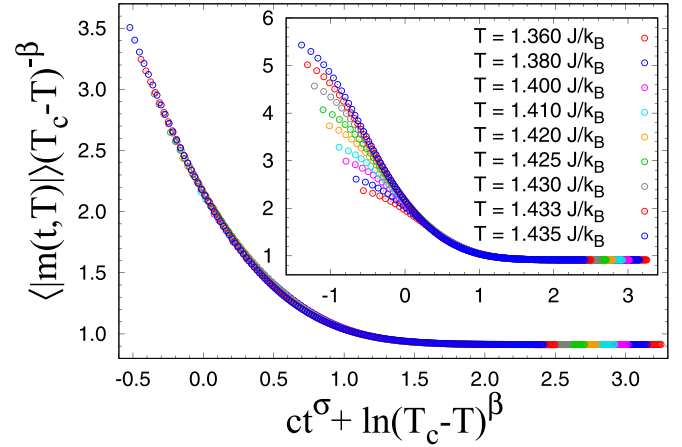

FIG. 8. Temperature scaling plot of the data in Fig. 6 after 16 MCS using $T_c = 1.442987 \, J/k_B$ and $\sigma = 0.47$ [8] with $\beta = 0.3553(10)$ and $c = 0.3244(3)$. The plot including the relaxation data for the initial 15 MCS is shown in the inset.

with the spontaneous magnetization $m_s(T)$ and fitting parameters $A(T)$ and $C(T)$. This formula was confirmed in the 2D Ising model in the Wolff algorithm [6], while the stretched-exponential relaxation was reported in the local-update algorithms [16,17]. This relaxation formula is verified in Fig. 7 by fitting the data with Eq. (A1) and plotting $\langle |m(t, T)| \rangle - m_s(T)$ versus t in a semilog scale at $T = 1.360 \, J/k_B$, $1.410 \, J/k_B$, and $1.435 \, J/k_B$ (from bottom to top). Linearity of the data reveals validity of Eq. (A1), and variance of the initial value and slope of the data represents explicit temperature dependence of the parameters $A(T)$ and $C(T)$ in Eq. (A1), respectively. Such nontrivial T dependence other than that of $m_s(T)$ makes a scaling analysis based on Eq. (A1) difficult. Nevertheless, the temperature scaling still holds on this quantity. From the stretched-exponential critical relaxation from the perfectly ordered state, $\langle |m(t, T)| \rangle \sim \exp(-ct^\sigma)$, and the temperature dependence in equilibrium, $\langle |m(t = \infty, T)| \rangle \sim (T_c - T)^\beta$, we have

$$\langle |m(t, T)| \rangle \sim (T_c - T)^\beta f_{tsc}[ct^\sigma + \ln(T_c - T)^\beta]. \quad \text{(A2)}$$

The data in Fig. 6 are scaled with Eq. (A2) in Fig. 8. Although the initial-time data are not scaled well owing to the discrepancy with the exponential decay (A1) as shown in the inset of Fig. 8, the scaling formula (A2) actually holds very well for the data from 16 MCS (in the main panel of Fig. 8). Similarly to the temperature scaling of the magnetic susceptibility, we minimize the mutual residuals of these data using $T_c = 1.442987(2) \, J/k_B$ and $\sigma = 0.47(1)$ [8]. We find that the averaged residuals are minimized when all the data in

the main panel of Fig. 8 are used for the fitting, and we have

$$\beta = 0.3553 \pm 0.0010, \quad c = 0.3244 \pm 0.0003. \quad \text{(A3)}$$

Although the error bars seem small enough, this estimate is not consistent with the most precise estimate to date, $\beta = 0.3689(3)$ [13].

The background of this discrepancy can be explained by the evaluation of β from the temperature dependence of $m_s(T)$ in Eq. (A1). Up to the leading term, it is given by $m_s(T) = B_1(T_c - T)^\beta$, and using all the data for $T < T_c$ in Fig. 6, we have $\beta = 0.3574(2)$. This estimate is not so different from that in Eq. (A3), and not consistent with the one in Ref. [13], either. On the other hand, when we take the next-order term into account as

$$m_s(T) = B_1(T_c - T)^\beta + B_2(T_c - T)^{2\beta}, \quad \text{(A4)}$$

we obtain

$$\beta = 0.3691 \pm 0.0010, \quad \text{(A5)}$$

$$B_1 = 0.988 \pm 0.005, \quad B_2 = -0.107 \pm 0.007. \quad \text{(A6)}$$

This estimate is consistent with the one in Ref. [13], and the coefficient of the next-order term is about 10% of that of the leading term. These results tell that the next-order term is crucial for the description of the critical phenomena in the 3D classical Heisenberg model based on the temperature dependence of the magnetization, and that the temperature-scaling formalism based only on the leading term of the temperature dependence of physical quantities is not suitable for the magnetization, at least in the present model. This mechanism is independent of the update algorithms, and therefore we do not consider the Metropolis algorithm here.

[1] As a review article, Y. Ozeki and N. Ito, J. Phys. A: Math. Theor. **40**, R149 (2007).

[2] M. Suzuki, Phys. Lett. A **58**, 435 (1976); Prog. Theor. Phys. **58**, 1142 (1977).

[3] P. C. Hohenberg and B. I. Halperin, Rev. Mod. Phys. **49**, 435 (1977).

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

[5] U. Wolff, Phys. Rev. Lett. **62**, 361 (1989); Nucl. Phys. B **322**, 759 (1989).

[6] Y. Nonomura, J. Phys. Soc. Jpn. **83**, 113001 (2014).

[7] Y. Nonomura and Y. Tomita, Phys. Rev. E **92**, 062121 (2015).

[8] Y. Nonomura and Y. Tomita, Phys. Rev. E **93**, 012101 (2016).

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

[10] Y. Tomita and Y. Nonomura, Phys. Rev. E **98**, 052110 (2018).

[11] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953); see also

J. E. Gubernatis, Phys. Plasmas **12**, 057303 (2005) (historical review).

[12] F. Y. Wu, Rev. Mod. Phys. **54**, 235 (1982).

[13] M. Campostrini, M. Hasenbusch, A. Pelissetto, P. Rossi, and E. Vicari, Phys. Rev. B **65**, 144520 (2002).

[14] N. Ito, K. Hukushima, K. Ogawa, and Y. Ozeki, J. Phys. Soc. Jpn. **69**, 1931 (2000).

[15] M. Matsumoto and T. Nishimura, ACM Trans. Model. Comput. Simul. **8**, 3 (1998); the Mersenne Twister Home Page maintained by M. Matsumoto, http://www.math.sci.hiroshima-u.ac.jp/∼m-mat/MT/emt.html.

[16] N. Ito, Physica A **192**, 604 (1993); **196**, 591 (1993).

[17] P. Grassberger and D. Stauffer, Physica A **232**, 171 (1996); D. Stauffer, *ibid.* **244**, 344 (1997).