Finite-size scaling around the critical point in the heavy quark region of QCD

Atsushi Kiyohara,[1] Masakiyo Kitazawa[1,2] Shinji Ejiri[3] and Kazuyuki Kanaya[4]

[1]*Department of Physics, Osaka University, Toyonaka, Osaka 560-0043, Japan*
[2]*J-PARC Branch, KEK Theory Center, Institute of Particle and Nuclear Studies, KEK,
203-1, Shirakata, Tokai, Ibaraki 319-1106, Japan*
[3]*Department of Physics, Niigata University, Niigata 950-2181, Japan*
[4]*Tomonaga Center for the History of the Universe, University of Tsukuba,
Tsukuba, Ibaraki 305-8571, Japan*

Finite-size scaling is investigated in detail around the critical point in the heavy quark region of nonzero temperature QCD. Numerical simulations are performed with large spatial volumes up to the aspect ratio $N_s/N_t = 12$ at a fixed lattice spacing with $N_t = 4$. We show that the Binder cumulant and the distribution function of the Polyakov loop follow the finite-size scaling in the $Z(2)$ universality class for large spatial volumes with $N_s/N_t \geq 9$, while, for $N_s/N_t \leq 8$, the Binder cumulant becomes inconsistent with the $Z(2)$ scaling. To realize the large-volume simulations in the heavy quark region, we adopt the hopping parameter expansion for the quark determinant: We generate gauge configurations using the leading-order action including the Polyakov loop term for $N_t = 4$, and incorporate the next-to-leading order effects in the measurements by the multipoint-reweighting method. We find that the use of the leading-order configurations is crucially effective in suppressing the overlapping problem in the reweighting and thus in reducing the statistical errors.

I. INTRODUCTION

One of the interesting features of the medium described by quantum chromodynamics (QCD) is the existence of phase transitions of various orders. While the finite-temperature QCD transition is an analytic crossover at zero-quark chemical potential μ_q [1,2], this phase transition is expected to become first order in a dense medium with large μ_q [3]. The end point of the first-order transition is called the critical point (CP) at which the transition is of second order. The singularity in thermodynamic observables associated with the second order CP are believed to be useful in detecting the CP in heavy-ion collision experiments [4,5]. Accordingly, researches called the beam-energy scan are actively performed in experimental facilities all over the world to search for the critical fluctuations around the CP [5–7].

The order of the finite temperature QCD transition changes also with variation of quark masses [8]. For the $(2 + 1)$-flavor QCD, it is known that the crossover at the physical quark masses becomes of first order both in the light- and heavy-quark limits; the phase diagram representing this feature is known as the Columbia plot [9,10]. Revealing the nature of the phase transitions with the variation of quark masses is an important subject of QCD at nonzero temperature since it provides us with various insights into the transition at the physical quark masses.

Pinning down the boundaries of the first-order transitions in $(2 + 1)$-flavor QCD in the light [11–23] and heavy [24–31] quark regions is a longstanding subject in lattice QCD simulations. However, it has been found that the location of the boundaries are strongly dependent on the lattice cutoff of the simulations [22,23,31], and their quantitative determination in the continuum limit has not been established yet.

One of the difficulties in these analyses is that observables near the CP are strongly dependent on the spatial volume of the system. The spatial volume dependence is in part described by the finite-size scaling (FSS) [32]. However, the FSS is applicable only for describing the singular part of thermodynamic quantities that dominates over the nonsingular part only in the vicinity of the CP for sufficiently large spatial volumes. When the spatial volume is not large enough the FSS of observables is violated due to the contributions of the nonsingular part and this makes their analysis based on the FSS problematic. In fact,

although the CP of QCD is believed to belong to the three-dimensional $Z(2)$ universality class [8], a clear FSS in this universality class has not been observed in the latest numerical study around the CP in the light quark mass region [23]. In the heavy quark region, on lattices of $N_t = 6$ and 8 with the aspect ratio $N_s/N_t = 6$ and 8, the Binder cumulant B_4 is reported to be consistent with the $Z(2)$ FSS using data mostly in the crossover side, while data in the first order side as well as that on $N_t = 10$ lattice show a deviation from the $Z(2)$ FSS [31]. These results suggest the necessity of performing numerical analyses with even larger spatial volumes and with high statistics.

In the present study we focus on the CP in the heavy quark region and study the behavior of observables with large spatial volumes corresponding to the aspect ratios up to $N_s/N_t = 12$ using numeical simulations. To carry out analyses on the large spatial volumes with high precision, we fix the temporal lattice extent to be $N_t = 4$ in this study.

We also employ the hopping parameter expansion (HPE) to deal with the quark determinant. In this study, we generate gauge configurations using the leading order (LO) action of the HPE including the Polyakov-loop term for $N_t = 4$, and then incorporate the next-to-leading order (NLO) effects by a multipoint-reweighting method [33]. In our previous study we generated the configurations in quenched QCD (zeroth order of the HPE) and reweighted them to incorporate LO and NLO effects [27,29,30]. We show that the use of the LO action to generate configurations is quite effective in suppressing the overlapping problem of the reweighting method. This is essential for carrying out simulations with large system volumes as studied in the present paper. We also verify the convergence of the HPE by comparing the LO and NLO results.

We perform the Binder cumulant analysis of the Polyakov loop around the CP and find that the numerical results for $N_s/N_t \geq 9$ follow the FSS in the $Z(2)$ universality class well. On the other hand, inclusion of the data at $N_s/N_t \leq 8$ in the analysis gives rise to a statistically-significant deviation from the scaling behavior. We further investigate the scaling behavior of the distribution function of the Polyakov loop. We find that the distribution function follows the FSS in the $Z(2)$ universality class for $N_s/N_t \geq 8$. From the deviation pattern of the distribution function for $N_s/N_t = 6$ from the $Z(2)$ FSS, we discuss that the violation of the scaling behavior in the Binder cumulant is caused by the deviation in the tails of the distribution for small N_s/N_t. In this paper we consider the case of degenerated N_f flavors with $N_f = 1, 2, 3$. Generalization of the formalism to nondegenerate cases is straightforward.

This paper is organized as follows. In the next section we give a brief review on the FSS. We then explain the setup of our lattice simulation and analyses using the HPE in Sec. III. In Sec. IV, we determine the transition line and perform the Binder cumulant analysis to determine the location of the CP and the critical exponent. In Sec. V, we investigate the FSS of the distribution function of the Polyakov loop. The last section, Sec. VI, is devoted to a summary. In Appendix A, we give definition of cumulants. In Appendix B, we examine the effect of the smearing width used in the calculation of distribution functions. In Appendix C, the HPE of the quark determinant is calculated up to the NLO. In Appendix D, the convergence of the HPE is examined by comparing the Binder cumulants at the LO and the NLO.

II. FINITE-SIZE SCALING

Let us first give a brief review on the FSS and its application to the CP in the heavy quark region of QCD.

The heavy-quark limit of QCD corresponds to the $SU(3)$ Yang-Mills theory (quenched QCD). This theory has a first-order deconfinement phase transition at nonzero temperature T. When the quark mass m_q is finite (throughout this section we assume that the quark masses are degenerate) with increasing $1/m_q$, this first-order transition becomes weaker and eventually terminates at the CP, as schematically shown in the phase diagram on the $(T, 1/m_q)$ plane in Fig. 1(a). This CP, as well as that in the light quark region, is believed to belong to the $Z(2)$ universality class, i.e., the universality class of the three-dimensional Ising model [8].

Near the CP of the three-dimensional Ising model, the relevant scaling parameters are the reduced temperature t and external magnetic field h; extensive variables which conjugate to these parameters are the energy and the magnetization, respectively. As shown in Fig. 1(b), the CP is located at $(t, h) = (0, 0)$ and the first-order transition exists on the t axis for $t < 0$. The singular part of thermodynamic quantities near the CP is described by the scaling function of t and h. According to the universality, the singular part of thermodynamic quantities near the CP of heavy quark QCD is described by the same scaling function, where the scaling parameters t and h are encoded into the $(T, 1/m_q)$ plane as schematically shown in Fig. 1(a); the t axis is parallel to the first-order line at the CP while the direction of the h axis is not constrained from the universality.
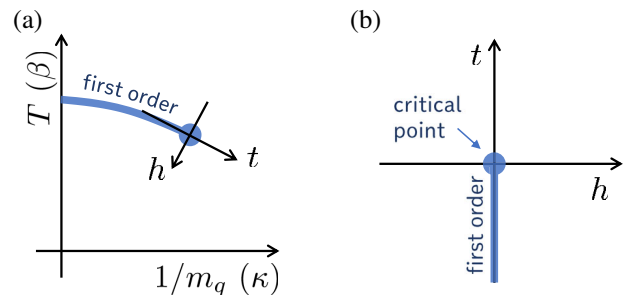

FIG. 1. Phase diagrams of (a) QCD in heavy-quark region, and (b) three-dimensional Ising model.

According to the FSS argument [32] the singular part of the dimensionless free energy $F(t, h, L^{-1})$ around the CP obtained at a finite volume $V = L^3$ has a scaling

$$F(t, h, L^{-1}) = F(tb^{y_t}, hb^{y_h}, L^{-1}b), \qquad (1)$$

for the arbitrary scale factor b. The values of the exponents y_t and y_h are specific for each universality class. In the $Z(2)$ universality class these parameters are numerically obtained as [32]

$$y_t = 1.588, \qquad y_h = -2.482. \qquad (2)$$

By setting $b = L$ one has

$$F(t, h, L^{-1}) = F(tL^{y_t}, hL^{y_h}, 1) \equiv \tilde{F}(tL^{y_t}, hL^{y_h}). \quad (3)$$

Derivatives of $F(t, h, L^{-1})$ with respect to t and h define the cumulants of the corresponding extensive variables. For example, cumulants of the magnetization M are given by

$$\langle M(t, h, L^{-1})^n \rangle_c = \partial_h^n F(t, h, L^{-1}), \qquad (4)$$

with $\partial_h = \partial/\partial h$. From Eq. (3), L dependence of the cumulants $\langle M^n \rangle_c$ near the CP are written as

$$\begin{aligned}
\langle M(t, h, L^{-1})^n \rangle_c &= \partial_h^n F(t, h, L^{-1}) \\
&= L^{ny_h} \partial_h^n \tilde{F}(tL^{y_t}, hL^{y_h}). \qquad (5)
\end{aligned}$$

In Ref. [34], it is suggested that the so-called (fourth-order) Binder cumulant

$$B_4(t, h, L^{-1}) = \frac{\langle M(t, h, L^{-1})^4 \rangle_c}{(\langle M(t, h, L^{-1})^2 \rangle_c)^2} + 3 \qquad (6)$$

plays a useful role to determine the location of the CP from numerical results obtained at finite L. When the distribution of M obeys the Gauss distribution or the distribution composed of two delta functions with an equal weight, we have

$$B_4 = \begin{cases} 3 & \text{Gauss distribution,} \\ 1 & \text{two delta functions,} \end{cases} \qquad (7)$$

respectively. Since the distribution of M approaches these functions in the $L \to \infty$ limit on the crossover and first-order lines at $h = 0$, respectively, B_4 should approach Eq. (7) on these lines in the $L \to \infty$ limit. Moreover, from Eq. (4), B_4 at $h = 0$ behaves as a function of t and L as

$$\begin{aligned}
B_4(t, 0, L^{-1}) &= \frac{\partial_h^4 F(t, 0, L^{-1})}{(\partial_h^2 F(t, 0, L^{-1}))^2} + 3 \\
&= \frac{\partial_h^4 \tilde{F}(tL^{y_t}, 0)}{(\partial_h^2 \tilde{F}(tL^{y_t}, 0))^2} + 3 \\
&= b_4 + ctL^{1/\nu} + \mathcal{O}(t^2), \qquad (8)
\end{aligned}$$

for small t, where $b_4 = \partial_h^4 \tilde{F}(0,0)/(\partial_h^2 \tilde{F}(0,0))^2 + 3$, $\nu = 1/y_t$, and c is a constant. Equation (8) shows that $B_4(t, 0, L^{-1})$ obtained for various L at $h = 0$ has a crossing at $t = 0$. The parameter b_4 is given only from $\tilde{F}(t, h)$ and thus are specific for each universality class. For the $Z(2)$ universality class, the value is known to be [32]

$$b_4 = 1.604. \qquad (9)$$

Equation (4) means that $F(t, h, L^{-1})$ is the cumulant generating function of M up to an additive constant. Then, as shown in Eqs. (A2) and (A4) in Appendix A, this function is related to the probability distribution function $p_M(M; t, h, L^{-1})$ of M as

$$e^{F(t, h' - h, L^{-1})} = c_F \int dM e^{h'M} p_M(M; t, h, L^{-1}), \quad (10)$$

where c_F is a constant determined from the normalization condition $\int dM p_M(M) = 1$. Here, let us define another probability distribution $\tilde{p}_M(M; t)$ as

$$e^{\tilde{F}(t, h')} = c_{\tilde{F}} \int dM e^{h'm} \tilde{p}_M(M; t). \qquad (11)$$

From Eq. (3), one finds

$$p_M(M; t, 0, L^{-1}) = L^{y_h} \tilde{p}_M(ML^{-y_h}; tL^{y_t}). \qquad (12)$$

When we consider magnetization per unit volume $m = M/V$, the probability distribution of m is given by

$$p_m(m; t, 0, L^{-1}) = L^{y_h - 3} \tilde{p}_M(mL^{3-y_h}; tL^{y_t}). \quad (13)$$

At the CP, $(t, h) = (0, 0)$, one finds from Eq. (13) that

$$p_m(m; 0, 0, L^{-1}) = L^{y_h - 3} \tilde{p}_M(mL^{3-y_h}; 0). \qquad (14)$$

Equation (13) also suggests that, when $\tilde{p}_M(M; t)$ has a local extremum at $M = \tilde{M}(t)$, $p_m(m; t, 0, L^{-1})$ has corresponding local extremum at

$$m = L^{y_h - 3} \tilde{M}(tL^{y_t}). \qquad (15)$$

This implies that the t and L dependences of the maximum of $p_m(m; t, 0, L^{-1})$ are described by a single function $\tilde{M}(t)$.

III. SETUP

A. Lattice action and basic observables

In this study we investigate the four-dimensional system described by the lattice action of QCD

$$S = S_g + S_q, \qquad (16)$$

with S_g and S_q being the gauge and quark actions. For S_g we employ the plaquette action

$$S_g = -6N_{\text{site}}\beta\hat{P}, \qquad (17)$$

with the gauge-coupling parameter $\beta = 2N_c/g^2$ and the space-time lattice volume $N_{\text{site}} = N_s^3 \times N_t$. The plaquette operator $\hat{P}$ is given by

$$\hat{P} = \frac{1}{6N_cN_{\text{site}}} \sum_{x,\mu<\nu} \text{Re}\,\text{tr}_C[U_{x,\mu}U_{x+\hat{\mu},\nu}U_{x+\hat{\nu},\mu}^\dagger U_{x,\nu}^\dagger], \qquad (18)$$

where $U_{x,\mu}$ is the link variable in the μ direction at site x, $x + \hat{\mu}$ is the next site in the μ direction from x, $N_c = 3$, and tr_C is the trace over color index.

For S_q, we adopt the Wilson quark action

$$S_q = \sum_{f=1}^{N_f} \sum_{x,y} \bar{\psi}_x^{(f)} M_{xy}(\kappa_f)\psi_y^{(f)}, \qquad (19)$$

with the Wilson quark kernel

$$M_{xy}(\kappa_f) = \delta_{xy} - \kappa_f B_{xy}, \qquad (20)$$

$$B_{xy} = \sum_{\mu=1}^{4} [(1 - \gamma_\mu)U_{x,\mu}\delta_{y,x+\hat{\mu}} + (1 + \gamma_\mu)U_{y,\mu}^\dagger\delta_{y,x-\hat{\mu}}], \qquad (21)$$

where x, y represent lattice sites. The color and Dirac-spinor indices are suppressed for simplicity. κ_f is the hopping parameter for the fth flavor. The bare quark mass m_f is related to κ_f as

$$\kappa_f = \frac{1}{2am_f + 8}, \qquad (22)$$

with the lattice spacing a. The matrix B_{xy} has nonzero values only when lattice sites x and y are located in adjacent sites. Therefore, this term represents the "hopping" of a quark between adjacent sites. The heavy quark limit $m_f \to \infty$ corresponds to $\kappa_f \to 0$.

In the following, we consider degenerated N_f flavors with a common hopping parameter $\kappa = \kappa_f$ corresponding to a common quark mass m_q—generalization to non-degenerate cases is straightforward. In this case, the expectation value of a gauge operator $\hat{O}(U)$ is calculated as

$$\langle\hat{O}(U)\rangle = \frac{1}{Z}\int \mathcal{D}U\mathcal{D}\psi\mathcal{D}\bar{\psi}\,\hat{O}(U)e^{-S_g - S_q}$$

$$= \frac{1}{Z}\int \mathcal{D}U\,\hat{O}(U)[\det M(\kappa)]^{N_f}e^{-S_g}$$

$$= \frac{1}{Z}\int \mathcal{D}U\,\hat{O}(U)e^{-S_g + N_f\ln\det M(\kappa)}, \qquad (23)$$

with the partition function $Z = \int \mathcal{D}Ue^{-S_g + N_f\ln\det M(\kappa)}$.

In the heavy quark limit $\kappa = 0$ ($m_q = \infty$), the deconfinement phase transition at nonzero temperature is characterized by the spontaneous symmetry breaking of the global $Z(3)$ center symmetry of the $SU(3)$ gauge symmetry. The most conventional choice for the order parameter of this phase transition is the Polyakov loop

$$\hat{\Omega} = \frac{1}{N_cN_s^3}\sum_{\vec{x}}\text{tr}_C[U_{\vec{x},4}U_{\vec{x}+\hat{4},4}U_{\vec{x}+2\cdot\hat{4},4}\cdots U_{\vec{x}+(N_t-1)\cdot\hat{4},4}], \qquad (24)$$

where the summation $\sum_{\vec{x}}$ is over the spatial lattice sites on one time slice. In the heavy quark limit, $\langle\hat{\Omega}\rangle = 0$ below the critical temperature T_c, while $\langle\hat{\Omega}\rangle$ takes a nonzero value at $T > T_c$. For finite m_q, the $Z(3)$ symmetry is explicitly broken by the quark term, and thus $\langle\hat{\Omega}\rangle$ becomes non-vanishing for all T. Even in this case, when m_q is sufficiently large, $\langle\hat{\Omega}\rangle$ jumps discontinuously at the first-order transition and thus can be used to detect the first-order transition line and its CP [27,29,30].

In Sec. V we study the scaling property of the distribution function of the real part of the Polyakov loop

$$\hat{\Omega}_R = \text{Re}\hat{\Omega}, \qquad (25)$$

defined by

$$p(\Omega_R) = \langle\delta(\Omega_R - \hat{\Omega}_R)\rangle. \qquad (26)$$

In Sec. III E, we also calculate the double-distribution function of $\hat{P}$ and $\hat{\Omega}_R$ defined by

$$p(P, \Omega_R) = \langle\delta(P - \hat{P})\delta(\Omega_R - \hat{\Omega}_R)\rangle. \qquad (27)$$

In numerical calculation of these distribution functions, because the statistics of the data is finite, we have to replace the delta functions by smeared ones with finite width. A conventional choice for this is the normalized Gauss function, $\delta(x) \simeq \exp[-(x/\Delta)^2]/(\Delta\sqrt{\pi})$ [27,29]. The width Δ should be large enough to have a statistically meaningful number of data at each point within the width, and simultaneously small enough to resolve the functional shape of the distribution function. Examining the resolution and the statistical error of distribution functions, we adopt $\Delta_{\Omega_R} = 0.002$ and $\Delta_P = 0.0001$ in this study. In Appendix B, we confirm that the resulting distribution

functions as well as other results of observables discussed in this study are stable under variations of the widths around these values.

B. Hopping parameter expansion

To calculate Eq. (23) around the heavy quark limit in the present study, we adopt the hopping parameter expansion (HPE) for $\ln \det M(\kappa)$,

$$\ln\left[\frac{\det M(\kappa)}{\det M(0)}\right] = -\sum_{n=1}^{\infty}\frac{1}{n}\mathrm{Tr}[B^n]\kappa^n, \qquad (28)$$

where the matrix B_{xy} is defined by Eq. (21) and Tr is the trace over all indices. In Eq. (28), the contribution at $\kappa = 0$ is subtracted for convenience. Since B_{xy} takes nonzero values only for adjacent lattice sites x and y, B^n is graphically represented by trajectories of n links [35]. Because of the trace in Eq. (28), nonvanishing contributions are given by closed trajectories. The lowest-order contributions of Eq. (28) start from $n = 4$, and all contributions for odd n vanish when N_t is even. By writing

$$-\ln\left[\frac{\det M(\kappa)}{\det M(0)}\right] = S_{\mathrm{LO}} + S_{\mathrm{NLO}} + \mathcal{O}(\kappa^8), \qquad (29)$$

one finds for $N_t = 4$ that the LO contribution is given by the plaquette and the Polyakov loop as

$$S_{\mathrm{LO}} = -2N_{\mathrm{c}}(48N_{\mathrm{site}}\hat{P} + 32N_s^3\hat{\Omega}_{\mathrm{R}})\kappa^4. \qquad (30)$$

The NLO term S_{NLO} consists of the six-step Wilson loops and bent Polyakov loops as

$$S_{\mathrm{NLO}} = -2N_{\mathrm{c}}(384\hat{W}_{\mathrm{rec}} + 768\hat{W}_{\mathrm{chair}} + 256\hat{W}_{\mathrm{crown}}$$
$$+ 192\mathrm{Re}\hat{\Omega}_1 + 96\mathrm{Re}\hat{\Omega}_2)N_{\mathrm{site}}\kappa^6. \qquad (31)$$

Here, $\hat{W}_{\mathrm{rec}}$, $\hat{W}_{\mathrm{chair}}$, and $\hat{W}_{\mathrm{crown}}$ represent the six-step Wilson loops of the rectangular, chair, and crown types, respectively, as illustrated in Fig. 2. $\hat{\Omega}_n$ are the bent Polyakov loops illustrated in Fig. 3, which run one step in a spatial direction, n steps in the temporal direction, and return to the original line. All the Wilson loops and Polyakov loop-type loops are normalized such that $\hat{W}_{\mathrm{rec}} = \hat{W}_{\mathrm{chair}} = \hat{W}_{\mathrm{crown}} = 1$
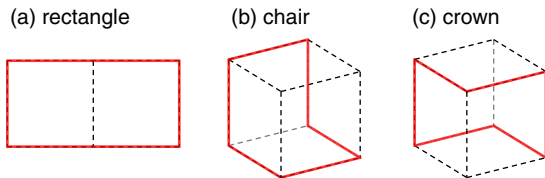

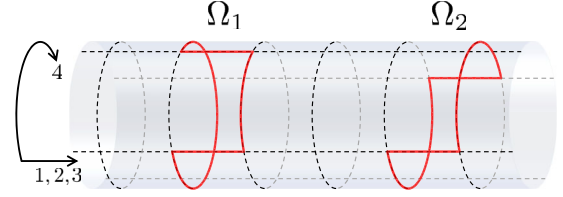FIG. 2. Six-step Wilson loops; rectangle (a), chair type (b), and crown type (c).

FIG. 3. Bent Polyakov loops, Ω_1 (left) and Ω_2 (right) for $N_t = 4$.

and $\hat{\Omega}_n = 1$ in the weak coupling limit, $U_{x,\mu} = 1$. Explicit definitions of these operators as well as the derivation of Eqs. (30) and (31) are given in Appendix C.

C. Numerical implementation with HPE

In this study, we generate the gauge configurations with respect to the action at the LO in the HPE, i.e.,

$$S_{\mathrm{g+ LO}} = S_g + N_{\mathrm{f}}S_{\mathrm{LO}}$$
$$= -6N_{\mathrm{site}}(\beta + 16N_{\mathrm{c}}N_{\mathrm{f}}\kappa^4)\hat{P} - 64N_{\mathrm{c}}N_{\mathrm{f}}N_s^3\kappa^4\hat{\Omega}_{\mathrm{R}}$$
$$= -6N_{\mathrm{site}}\beta^*\hat{P} - \lambda N_s^3\hat{\Omega}_{\mathrm{R}}, \qquad (32)$$

with

$$\beta^* = \beta + 16N_{\mathrm{c}}N_{\mathrm{f}}\kappa^4, \qquad \lambda = 64N_{\mathrm{c}}N_{\mathrm{f}}\kappa^4. \qquad (33)$$

We then perform the measurements at the NLO by incorporating the effect of S_{NLO} by the multipoint-reweighting method [29,33,36]. In this subsection we discuss the numerical implementation of these analyses.

In the Monte Carlo simulations of pure gauge theory, thanks to the locality of the action S_g, it is possible to adopt the pseudoheat-bath (PHB) and over-relaxation (OR) algorithms for updating gauge configurations. Focusing on a link variable $U_\mu(x)$, the dependence of S_g on $U_\mu(x)$ is given by

$$\Delta S_g(U_\mu(x)) = -\frac{\beta}{N_{\mathrm{c}}}\mathrm{Re}\,\mathrm{tr}_{\mathrm{C}}[U_\mu(x)X_\mu(x)], \qquad (34)$$

where the staple

$$X_\mu(x) = \sum_{\nu\neq\mu}\sum_{s=\pm 1}U_{x+\hat{\mu},s\nu}U_{x+s\hat{\nu},\mu}^\dagger U_{x,s\nu}^\dagger, \qquad (35)$$

with $U_{x,-\mu} = U_{x-\hat{\mu},\mu}^\dagger$ for $\mu > 0$, is graphically shown in Fig. 4(a). In the PHB and OR algorithms, the link variable $U_\mu(x)$ is updated according to the probability determined by Eq. (34). The fact that Eq. (34) is represented only by local variables near $U_\mu(x)$ makes this procedure efficient especially on the memory-distributed parallel computing.

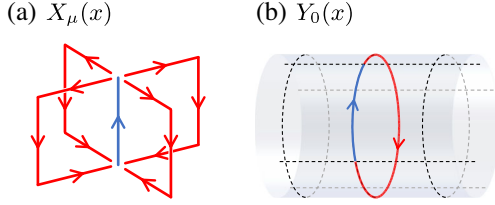(a) $X_\mu(x)$ (b) $Y_0(x)$

FIG. 4. Staple $X_\mu(x)$ and the operator $Y_0(x)$ in Eq. (37) corresponding to the link variable $U_\mu(x)$ shown by the blue line.

When the LO term, Eq. (30), is included into the action, the contribution of a temporal link $U_0(x)$ to S_{g+LO} is modified as

$$\Delta S_{g+LO}(U_\mu(x))$$
$$= -\frac{\beta^*}{N_c} \operatorname{Re} \operatorname{tr}_C \left[U_0(x) \left(X_0(x) + \frac{\lambda}{\beta^*} Y_0(x) \right) \right], \quad (36)$$

with

$$Y_0(x) = U_0(x+\hat{0})U_0(x+2\cdot\hat{0})U_0(x+3\cdot\hat{0}), \quad (37)$$

which is schematically shown in Fig. 4(b). For $N_t > 4$, $Y_0(x)$ is given by the product of $N_t - 1$ link variables along the temporal direction. The contribution of a spatial link to S_{g+LO} is unchanged from Eq. (34).

These results on $\Delta S_{g+LO}(U_\mu(x))$ suggest that the Monte Carlo updates of $U_\mu(x)$ can be performed by the PHB and OR efficiently even for S_{g+LO}, provided that the temporal direction is not separated into different parallel nodes and $Y_0(x)$ can be calculated efficiently. Satisfying this condition is not difficult to attain for large-volume simulations. Using this to our advantage, we perform updates of gauge fields at the LO by combining PHB and OR in this study.[1] Compared with the pure-gauge simulation, the increase of the numerical cost to deal with S_{g+LO} in this method is small since the additional multiplications of $SU(3)$ matrices required for an update are only $N_t - 2$ times for the temporal links and the cost to update the spatial links is unchanged.

In the measurement of observables, we include all the contribution of S_{NLO} using the multipoint-reweighting method.[2] The expectation value of a gauge observable $\hat{O}(U)$ at the NLO at the parameter set (β, κ) is given from the LO simulations at $(\tilde{\beta}, \tilde{\kappa})$ as

---

[1]After finishing our numerical analyses, we discovered that a similar idea is suggested in Ref. [37]. We thank F. Karsch for notifying us of this literature.
[2]In Ref. [30], the effect of S_{NLO} is included in part effectively by the effective NLO method. In this study, we deal with it exactly.

$$\langle \hat{O}(U) \rangle_{\beta,\kappa}^{NLO} = \frac{\int \mathcal{D}U \hat{O}(U) e^{-S_{g+LO}(\beta,\kappa)-S_{NLO}(\beta,\kappa)}}{\int \mathcal{D}U e^{-S_{g+LO}(\beta,\kappa)-S_{NLO}(\beta,\kappa)}}$$
$$= \frac{\int \mathcal{D}U \hat{O}(U) e^{-\delta S_{g+LO}-S_{NLO}(\beta,\kappa)} e^{-S_{g+LO}(\tilde{\beta},\tilde{\kappa})}}{\int \mathcal{D}U e^{-\delta S_{g+LO}-S_{NLO}(\beta,\kappa)} e^{-S_{g+LO}(\tilde{\beta},\tilde{\kappa})}}$$
$$= \frac{\langle \hat{O}(U) e^{-\delta S_{g+LO}-S_{NLO}(\beta,\kappa)} \rangle_{\tilde{\beta},\tilde{\kappa}}^{LO}}{\langle e^{-\delta S_{g+LO}-S_{NLO}(\beta,\kappa)} \rangle_{\tilde{\beta},\tilde{\kappa}}^{LO}}, \quad (38)$$

with

$$\delta S_{g+LO} = S_{g+LO}(\beta,\kappa) - S_{g+LO}(\tilde{\beta},\tilde{\kappa})$$
$$= -6N_{site}(\beta - \tilde{\beta} + 16N_c N_f(\kappa^4 - \tilde{\kappa}^4))\hat{P}$$
$$+ 64N_c N_f N_s^3(\kappa^4 - \tilde{\kappa}^4)\hat{\Omega}_R, \quad (39)$$

and $\langle \cdot \rangle_{\beta,\kappa}^{LO}$ is the expectation value taken with the action $S_{g+LO}(\beta,\kappa)$. We generate gauge configurations for the LO action S_{g+LO} at several values of $(\tilde{\beta}, \tilde{\kappa})$, and evaluate the expectation values to the NLO at various (β, κ) by the multipoint-reweighting method.

D. Simulation parameters

In this study, we perform Monte Carlo simulations with fixed temporal lattice size $N_t = 4$, while the spatial extent N_s is changed from 24 to 48. This allows us to perform simulations with large aspect ratio $N_s/N_t = LT$ up to 12, where L is the lattice size along the spatial direction in physical units. For each N_s, the gauge configurations are generated for 3 to 5 sets of (β^*, λ) shown in Table I, which are chosen so that β^* is close to the transition line at the LO. All numerical results shown in Secs. IV and V are generated by the multipoint-reweighting method from these configurations.

The gauge configurations are updated with the LO action S_{g+LO} using the PHB and OR algorithms as discussed in Sec. III C. Gauge configurations are updated by five OR steps after each PHB step. We measure observables every two sets of the PHB + OR updates, i.e., totally ten OR steps and two PHB steps. For all parameters we have performed 6×10^5 measurements in this way. In the following, we set $N_f = 2$ to show the numerical results unless otherwise stated.

In Monte Carlo simulations near a first-order transition, because the transition between the coexisting two phases becomes rare when the spatial volume of the system is large, observables averaged over the two phases tend to have quite long autocorrelations. In Fig. 5 we show the Monte Carlo time history of $\hat{\Omega}_R$ for $N_s/N_t = LT = 10$ and 12 at the smallest λ, at which the autocorrelation is the longest. The horizontal axis represents the Monte Carlo time in the unit of measurements. To estimate the autocorrelation time, in Fig. 6, we plot the integrated autocorrelation time of Ω_R defined as

TABLE I.   Simulation parameters; lattice size $N_s^3 \times N_t$, β^*, and λ. The value of κ corresponding to each λ is also listed for the case $N_f = 2$. The last column is for the integrated autocorrelation time estimated using Ω_R.

Lattice size	β^*	λ	$\kappa^{N_f=2}$	τ_{int}
$48^3 \times 4$	5.6869	0.004	0.0568	642(150)
	5.6861	0.005	0.0601	837(75)
	5.6849	0.006	0.0629	537(49)
$40^3 \times 4$	5.6885	0.003	0.0529	1448(160)
	5.6869	0.004	0.0568	685(113)
	5.6861	0.005	0.0601	630(60)
	5.6849	0.006	0.0629	416(33)
	5.6837	0.007	0.0653	310(24)
$36^3 \times 4$	5.6885	0.003	0.0529	936(113)
	5.6869	0.004	0.0568	459(49)
	5.6861	0.005	0.0601	511(46)
	5.6849	0.006	0.0629	364(23)
	5.6837	0.007	0.0653	278(16)
$32^3 \times 4$	5.6885	0.003	0.0529	646(62)
	5.6865	0.004	0.0568	307(34)
	5.6861	0.005	0.0601	401(32)
	5.6845	0.006	0.0629	270(18)
	5.6837	0.007	0.0653	225(15)
$24^3 \times 4$	5.6870	0.0038	0.0561	464(32)
	5.6820	0.0077	0.0669	250(18)
	5.6780	0.0115	0.0740	160(10)

$$\tau_{\text{int}} = \frac{1}{2} + \sum_{\tau=1}^{\tau_{\text{max}}} \frac{\langle \Omega_R(\tau)\Omega_R(0)\rangle}{\langle \Omega_R(0)\Omega_R(0)\rangle}, \qquad (40)$$

as a function of τ_{max} for the two Monte Carlo trajectories shown in Fig. 5, where τ represents the Monte Carlo time. We estimate the autocorrelation time from the value of τ_{int} at $\tau_{\text{max}} \simeq 4\tau_{\text{int}}$, i.e., the crossing point between τ_{int} and the dashed line in Fig. 6, at which the τ_{max} dependence is well saturated. The values of τ_{int} thus determined on each lattice are listed in Table I.

Throughout this study, we estimate the statistical errors of observables by the jackknife method unless otherwise stated, adopting the binsize of 10,000 measurements which is sufficiently larger than the estimated autocorrelation lengths. We checked that the statistical errors thus estimated are roughly stable within a variation of the binsize from 5,000 to 30,000 measurements.

E. Overlapping problem

Our main objective in using the LO action for configuration generation is to avoid the overlapping problem in the reweighting. In Fig. 7 we show the contour plot for the probability distribution function $p(P, \Omega_R)$ defined by Eq. (27), obtained at $\lambda = 0.003$ (red) and 0.007 (blue) on $40^3 \times 4$ lattices adopting $\Delta_{\Omega_R} = 0.002$ and $\Delta_P = 0.0001$ for the smearing widths. We checked that $p(P, \Omega_R)$ hardly changes under variation of Δ_{Ω_R} and Δ_P around this choice.
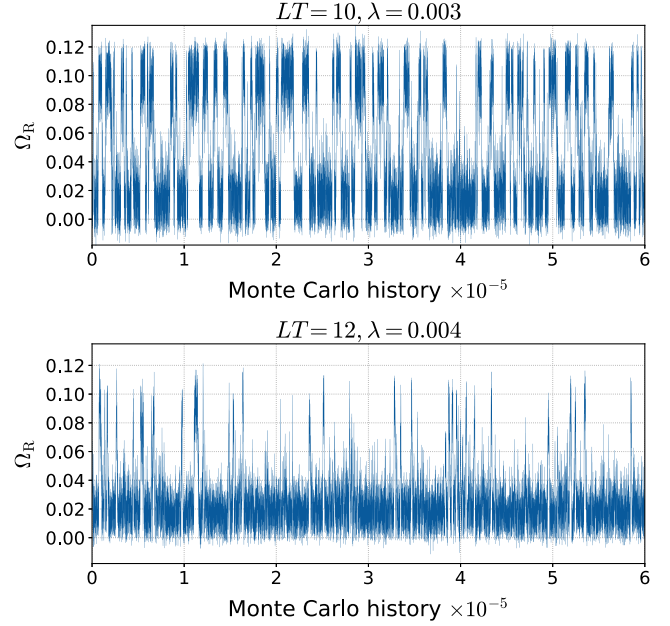

FIG. 5.   Monte Carlo time history of $\hat{\Omega}_R$ at the smallest λ for $LT = 12$ and $LT = 10$. Horizontal axis represents the Monte Carlo time in the unit of measurements which are made every two sets of PHB + OR updates.
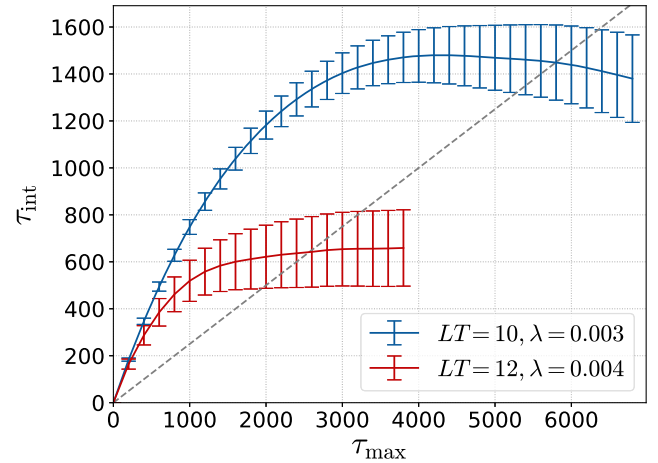

FIG. 6.   Integrated autocorrelation time τ_{int} of $\hat{\Omega}_R$ as a function of τ_{max} on the lattices shown in Fig. 5. The dashed line shows $\tau_{\text{max}} = 4\tau_{\text{int}}$.

The solid lines are the contours for the distribution measured on the LO configurations generated with $S_{g+\text{LO}}$. Each contour curve is drawn such that the probability inside the contour is 0.9, 0.7, $\cdots$, and 0.1.

In Fig. 7, we also show by the dotted lines the contours of $p(P, \Omega_R)$ at the NLO, which is obtained by reweighting the LO data at the same (β^*, λ). The meaning of the contours is the same as the solid lines. We find that the deviation of the NLO distribution from the original one at the LO is not significant, suggesting that the effects of the NLO contribution are not large. The large overlap of the LO
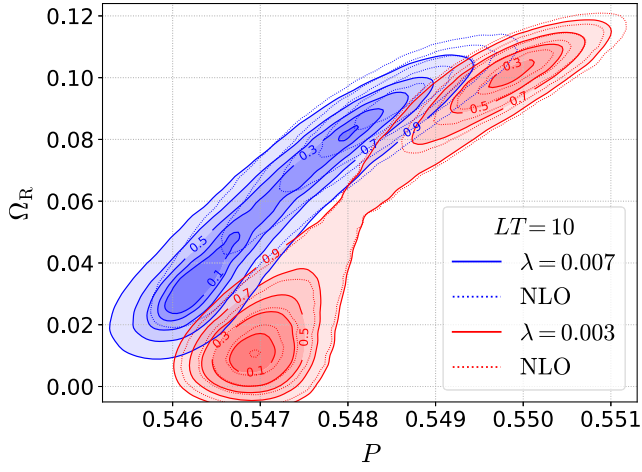
FIG. 7. Probability distribution $p(P, \Omega_R)$ of the gauge configurations on the $40^3 \times 4$ lattice for the simulation points at $\lambda = 0.003$ and 0.007. Solid lines represent the contours of $p(P, \Omega_R)$ obtained with the LO action. The label on each contour shows the probability inside the contour. Dotted lines represents the contour lines of $p(P, \Omega_R)$ with the NLO action at the same parameters obtained by reweighting the LO data.

and NLO distributions ensures that, for observables constructed from $\hat{P}$ and $\hat{\Omega}_R$, the NLO results obtained by reweighting the LO data at the same β^* and λ are statistically reliable. In the analysis of the CP, the overlapping of the distributions is even more improved after adjusting the parameters to the transition line.

On the other hand, from this figure, we find that the overlapping of the distributions at $\lambda = 0.003$ and 0.007 is quite poor—the regions with probability larger than 0.7 are not overlapping at all with each other. This means that, if we were to calculate observables at $\lambda = 0.007$ by reweighting data obtained at $\lambda = 0.003$, or vice versa, the statistical quality of the results would be quite low. In Refs. [27,29,30], the CP in heavy quark region was investigated on lattices with $N_s/N_t = 4$–6, by reweighting from pure gauge configurations, i.e., those obtained at $\lambda = 0$. Because the overlapping problem becomes quickly severe as the system volume becomes large, the same strategy is not applicable to

the present study in which much larger system volumes up to $N_s/N_t = LT = 12$ are simulated. Figure 7 shows that, to incorporate the NLO effects by reweighting, the use of the LO action for configuration generation is sufficiently effective in suppressing the overlapping problem. The smallness of the NLO effects in Fig. 7 further suggests that the effects of dynamical quarks are dominated by the LO term for these values of λ.

IV. BINDER CUMULANT ANALYSIS

A. Transition line

We first determine the location of the transition line that corresponds to $h = 0$ in terms of the Ising parameters; see Fig. 1. In the coupling parameter space (β^*, λ), we denote the transition line as $\beta^*_{\text{tr}}(\lambda)$. In this study, we determine β^*_{tr} at each λ adopting the following three conventional choices:

(i) Maximum of $\langle \Omega_R^2 \rangle_c$
(ii) Zero point of $\langle \Omega_R^3 \rangle_c$
(iii) Minimum of $B_4^\Omega = \langle \Omega_R^4 \rangle_c / \langle \Omega_R^2 \rangle_c^2 + 3$

In Fig. 8 we show the LT dependence of β^*_{tr} determined by these definitions for several values of λ. The figure shows that the maximum of $\langle \Omega_R^2 \rangle_c$ has a visible LT dependence. On the other hand, the zero point of $\langle \Omega_R^3 \rangle_c$ and the minimum of B_4^Ω do not have statistically-significant LT dependence for $LT \geq 8$. This result shows that the zero point of $\langle \Omega_R^3 \rangle_c$ and minimum of B_4^Ω are sufficiently close to the β^*_{tr} in the $L \to \infty$ limit in this range of LT. In the following, we employ the minimum of B_4^Ω for the definition of the transition line β^*_{tr} for each N_t. In Fig. 9, we show the transition line on the (β^*, λ) and (β, λ) planes obtained at $LT = 10$.

In Fig. 10, we show the distribution function of Ω_R, Eq. (26), on the transition line for several values of λ, where the delta function in Eq. (26) is smeared by the Gauss function as before with the width $\Delta_{\Omega_R} = 0.002$. As discussed in Appendix B, dependence of these results on Δ_{Ω_R} is well suppressed around this Δ_{Ω_R}. The shaded bands represent the statistical errors. At $\lambda = 0.003$, we see a clear two-peak structure in $p(\Omega_R)$ and find that the peaks become sharper as LT becomes larger. This behavior
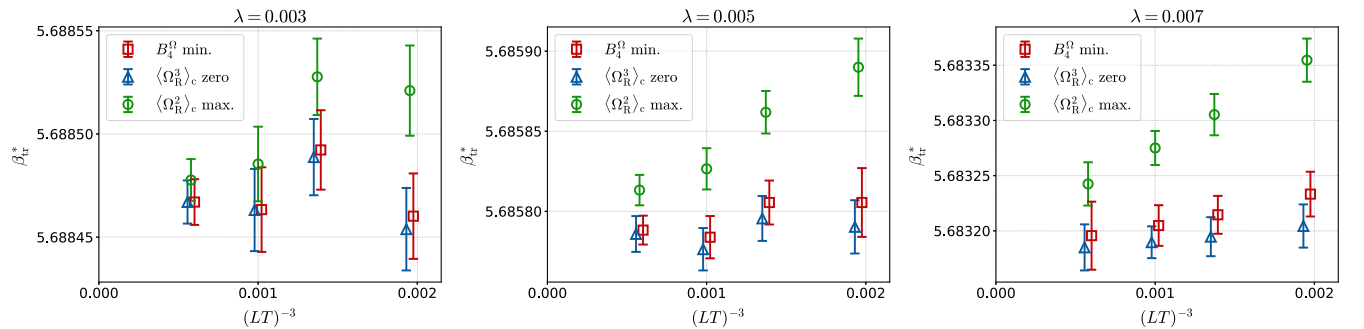

FIG. 8. Transition line $\beta^*_{\text{tr}}(\lambda)$ as function of the aspect ratio $LT = N_s/N_t$, determined through three definitions: (1) minimum of B_4^Ω, (2) zero of $\langle \Omega_R^3 \rangle_c$, and (3) maximum of $\langle \Omega_R^2 \rangle_c$. The results for $\lambda = 0.003, 0.005, 0.007$ are shown from left to right.
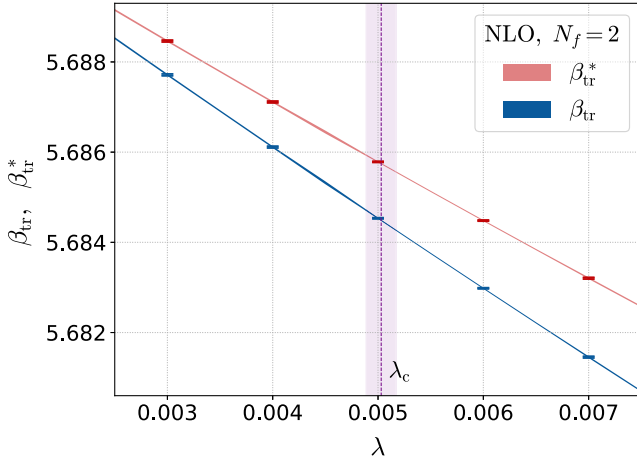
FIG. 9. Transition line β_{tr}^* on the (β, λ) and (β^*, λ) plane obtained at $LT = 10$. The points with the error bar are the results obtained on each measurement. The shaded areas show the result obtained by the multipoint-reweighting method with the error band representing the statistical error. The dotted vertical line shows the value of λ_c determined by the analysis in Sec. IV B.

suggests the first-order phase transition at $\lambda = 0.003$. At $\lambda = 0.007$, on the other hand, while two peaks are observed for $LT \leq 9$, the two peaks cease to exist as LT becomes large. This suggests the crossover transition in the $L \to \infty$ limit at this λ.

B. Binder cumulant

Next, let us determine the position of the CP on the (β, κ) plane. As discussed in Sec. II, it is convenient to employ the Binder cumulant B_4 of Ω_R

$$B_4^\Omega = \frac{\langle \Omega_R^4 \rangle_c}{\langle \Omega_R^2 \rangle_c^2} + 3. \tag{41}$$

This quantity approaches the known values given in Eq. (7) in the $L \to \infty$ limit depending on the order of the transition. Furthermore, provided that Ω_R corresponds to $m = M/V$ of the Ising model, B_4^Ω should obey Eq. (8) near the CP.

In the upper panel of Fig. 11, we show B_4^Ω along the transition line as a function of λ for five values of LT. λ is varied continuously by the multipoint-reweighting method. The lower panel is an enlargement of the upper panel around the crossing point. The figure shows that B_4^Ω has a crossing at $\lambda = \lambda_c \simeq 0.005$ and is an increasing (decreasing) function of LT for $\lambda > \lambda_c$ ($\lambda < \lambda_c$). The existence of the CP at $\lambda \simeq 0.005$ is suggested from this result.

To determine λ_c and the critical exponent ν quantitatively, we fit the numerical results of B_4^Ω using a fitting function motivated by Eq. (8),

$$B_4^\Omega(\lambda, LT) = b_4 + c(\lambda - \lambda_c)(LT)^{1/\nu}, \tag{42}$$

where b_4, λ_c, ν, and c are the fit parameters. In this study, we can vary λ continuously using the multipoint-reweighting method. However, because data at different λ on the same volume are correlated, it is not meaningful to use too many λ values. Using the data at three largest volumes, $LT = 12, 10, 9$, two or three λ values (6 or 9 data points, respectively) should be sufficient for the four parameter fit of Eq. (42). We thus repeat the fit for several choices of λ values, taking the covariance between data at different λ into account in the calculation of χ^2.

In Table II, we summarize the results of the fit using the data at three largest volumes, $LT = 12, 10, 9$, and at λ values listed in the left column of the table. The statistical error in the table is estimated by the standard chi-square analysis. The table shows that the value of χ^2/dof (dof = degrees of freedom) are smaller than unity in the fits with two λ values, but χ^2/dof is unacceptably large with three λ values, while all the results for the fitting parameters are consistent within errors. We choose the fit result for $\lambda = (0.0048, 0.0053)$ depicted by bold characters in Table II as the central value and include the uncertainty in the fits with two λ values as the systematic error.

We repeat the analysis also with other sets of system volumes. The results of the fits with four and five largest volumes, $LT = 12$–8 and 12–6, together with the fit with three largest volumes, are summarized in Table III. For the fit with four and five largest volumes, we now have 8 and
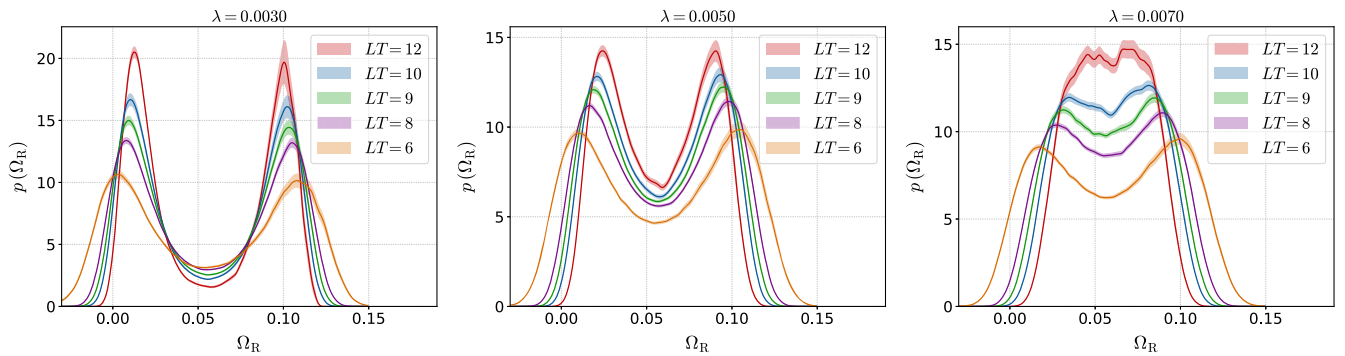

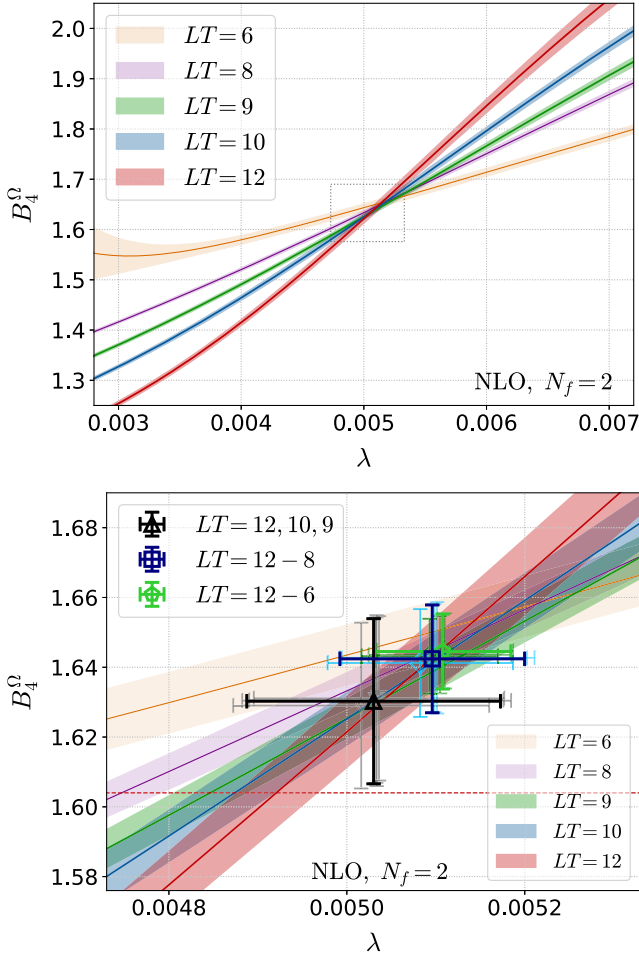FIG. 10. Distribution of Ω_R at $\lambda = 0.003$, 0.005 and 0.007.

FIG. 11. Binder cumulant B_4^Ω as a function of λ obtained at five $LT = N_s/N_t$. The statistical errors are shown by the shaded area. The lower panel is an enlargement of the upper panel around the crossing point, where the dotted rectangle in the upper panel represents the region of the lower panel. The points in the lower panel with error bars show the results of the four parameter fit with Eq. (42). See text for details.

10 data points for the four-parameter fit with two λ values. For these fits, we choose $\lambda = (0.0048, 0.0053)$ as the central value again.

TABLE II. Results of the four parameter fits of B_4^Ω with Eq. (42). Three largest volumes with $LT = N_s/N_t = 12, 10, 9$ are used for the fits. The left column shows the values of λ used for the fit.

Fit λ ($\times 10^4$)	b_4	λ_c	ν	χ^2/dof
49, 52	1.631(24)	0.00503(14)	0.614(49)	0.45
48, 53	**1.630(24)**	**0.00503(14)**	**0.614(48)**	**0.46**
45, 55	1.629(24)	0.00502(14)	0.622(48)	0.46
45, 50	1.630(24)	0.00503(15)	0.634(47)	0.48
50, 55	1.631(24)	0.00504(14)	0.610(51)	0.38
45, 50, 55	1.620(23)	0.00494(14)	0.626(41)	7.9

TABLE III. Results of the four-parameter fit of B_4^Ω using the data points at three and four largest and all volumes. The first parentheses are for statistical errors estimated by the jackknife method, and the second parentheses are for systematic errors due to the choice of λ values for the fit. See Sec. IV B for details.

$LT = N_s/N_t$	b_4	λ_c	ν	χ^2/dof
12, 10, 9	1.630(24)(2)	0.00503(14)(2)	0.614(48)(3)	0.46
12, 10, 9, 8	1.643(15)(2)	0.00510(10)(2)	0.614(29)(3)	0.37
12, 10, 9, 8, 6	1.645(11)(2)	0.00511(8)(2)	0.593(18)(3)	0.67

The results of b_4 and λ_c obtained by these fits are shown in the lower panel of Fig. 11. The thick symbols show the central value, and thin symbols with light colors show the results obtained by the variation of λ in Table II. The results with the three largest volumes are shown by black triangles, while those with four and five largest volumes are shown by blue squares and green pentagons, respectively. In the figure, b_4 expected from the $Z(2)$ universality, Eq. (9), is shown by the dashed horizontal line.

From Table III and Fig. 11, we find that when we adopt the fit with the three largest volumes, $LT \geq 9$, the fit result of b_4 is consistent with the $Z(2)$ value within about 1σ. On the other hand, when we include smaller volumes, $LT \geq 8$ or $LT \geq 6$, b_4 from the fits show a statistically significant deviation from the $Z(2)$ value. In Table III, we also summarize the results for the critical exponent ν. From the $Z(2)$ universality class, we expect $\nu = 1/y_t \approx 0.630$. We find that the result of the fit with $LT \geq 9$ is consistent with the $Z(2)$ value within the error, while the result of the fit with $LT \geq 6$ has a significant deviation from the $Z(2)$ value, though the values of χ^2/dof are all smaller than unity.

We thus conclude that the FSS in the $Z(2)$ universality class is confirmed when the system volume is large enough, $LT \geq 9$—lattices with $LT \leq 8$ are not large enough to apply an FSS analysis for B_4^Ω. The value of λ_c thus determined is also shown in Fig. 9.

In Appendix D, we perform the analysis of B_4^Ω at the LO of the HPE and compare the results with those at the NLO discussed in this section. We find that the LO result for λ_c is about 2.6% larger than the NLO value. This small difference suggests that the truncation error of the HPE is well under control at the NLO around λ_c.

C. Mixing with energylike observable

So far we have performed the analyses of B_4^Ω assuming that Ω_R corresponds to the magnetization $m = M/V$ of the Ising model. Although our numerical results thus far are in good agreement with this assumption, a possible mixing with the energylike observable [14,22] in Ω_R is not excluded in general. In this case, the behavior of B_4^Ω near the CP is modified from Eq. (8) as [22]

TABLE IV. Location of the critical point (β_c, κ_c) for various N_f. For λ_c, the first parentheses are for statistical errors and the second parentheses are for systematic errors from the fit as discussed in Sec. IV B. The errors for β_c and κ_c include the systematic errors.

N_f	β_c	κ_c	λ_c
1	5.68446(22)	0.0714(5)	0.00498(14)(2)
2	5.68453(22)	0.0602(4)	0.00503(14)(2)
3	5.68456(21)	0.0544(4)	0.00505(14)(2)

$$B_4^{\Omega}(\lambda, LT) = (b_4 + c(\lambda - \lambda_c)(LT)^{1/\nu})(1 + d(LT)^{y_t - y_h}). \tag{43}$$

To investigate the effect of this possible mixing, we try fits of B_4^{Ω} based on Eq. (43). We use the values of B_4^{Ω} at three λ for the fits to increase the number of data points. We find that the six-parameter fits with Eq. (43) with the fitting parameters b_4, λ_c, ν, c, d, $y_t - y_h$ are quite unstable, suggesting that χ^2 has many local minima. The model space of Eq. (43) would be too large against the data. As a next trial, we perform five-parameter fits with Eq. (43) by fixing $y_t - y_h = -0.894$. In this case, we find that χ^2 still has many local minima, and χ^2/dof becomes larger compared with the four-parameter fit. It is also found that the value of d is consistent with zero within the error for all trials with the variation of λ values. This suggests that the mixing of the energylike observable in Ω_R is negligible around the CP in the heavy quark region.

D. N_f dependence

In Table IV, we summarize our final results for the location of the CP, (β_c, κ_c). In the table, we also show the results for $N_f = 1$ and 3. We note that the N_f dependence of the HPE is trivial at the LO in the sense that N_f enters the action Eq. (32) at this order only through the combination $\lambda = 64 N_c N_f \kappa^4$ after the replacement $\beta \to \beta^*$. Therefore, λ_c does not depend on N_f. At the LO, this allows us to obtain the value of κ_c for various N_f from the value of κ_c at $N_f = 2$ [29]. Because such a simple scaling is no longer applicable at the NLO, we made individual numerical analyses at $N_f = 1$ and 3. From Table IV, we find that the results of λ_c are almost insensitive to N_f. This means that the NLO effects on λ_c are small.

V. DISTRIBUTION FUNCTION OF Ω_R

In this section, we study the scaling behavior of the distribution function $p(\Omega_R)$ to further investigate the consistency with the $Z(2)$ universality class around the CP.

A. Scaling of distribution function

Let us first focus on the LT dependence of $p(\Omega_R)$ at the CP. In the following, instead of $p(\Omega_R)$ itself, we study the effective potential defined from $p(\Omega_R)$,

$$V(\Omega_R; \lambda, LT) = -\ln p(\Omega_R)_{\lambda, LT}, \tag{44}$$

as this quantity is more convenient in comparing the results at different LT [27,29]. From Eq. (14), the LT dependence of $V(\Omega_R, \lambda, LT)$ at the CP will be described by a single function $\tilde{V}(x)$ as

$$V(\Omega_R; \lambda_c, LT) = \tilde{V}((\Omega_R - \langle\Omega_R\rangle)(LT)^{3-y_h}), \tag{45}$$

up to an additive constant, where $\langle\Omega_R\rangle$ is subtracted from Ω_R to adjust the center of the distribution.

To see if the scaling behavior of Eq. (45) is satisfied, we show in Fig. 12 the effective potential $V(\Omega_R; \lambda_c, LT)$ at the CP obtained at five values of LT, as a function of $(\Omega_R - \langle\Omega_R\rangle)(LT)^{3-y_h}$, where we set $3 - y_h = 0.518$. For the figure, we adjust the arbitrary constant
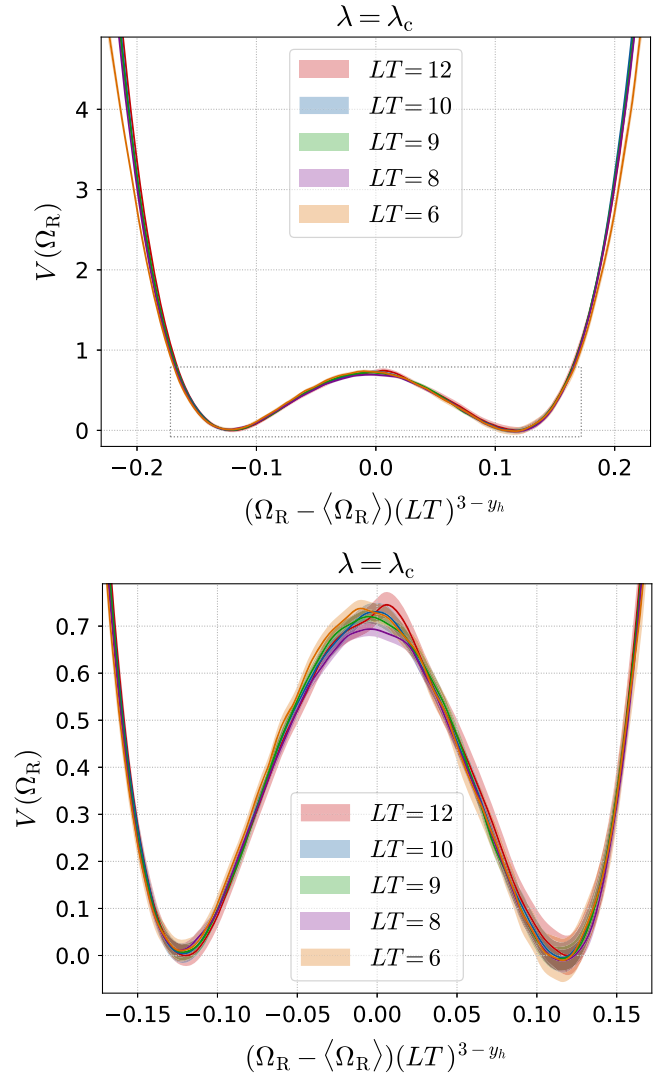

FIG. 12. Effective potential $V(\Omega_R) = -\ln p(\Omega_R)$. Bottom panel is an enlargement of the region enclosed by a dotted rectangle in the top panel.

term of $V(\Omega_R; \lambda_c, LT)$ such that $V(\Omega^{(1)}; \lambda_c, LT) + V(\Omega^{(2)}; \lambda_c, LT) = 0$, where $\Omega^{(1)}$ and $\Omega^{(2)}$ ($> \Omega^{(1)}$) are the values of Ω_R at the two local minima of $V(\Omega_R; \lambda_c, LT)$. No further adjustments are made in the figure. The error bands do not include the uncertainty of the additive constant. The lower panel is an enlargement of the region indicated by the dotted rectangle in the upper panel. From Fig. 12, we find that the numerical results for $LT = 8$–12 agree almost completely within the errors with the scaling relation Eq. (45). This result nicely supports the FSS in the $Z(2)$ universality class at the CP. From the upper panel of Fig. 12, we note that the effective potential for $LT = 6$ shows a clear deviation from the results for larger volumes at $\Omega_R \ll \Omega^{(1)}$ and $\Omega_R \gg \Omega^{(2)}$, while it agrees well with them in the range $\Omega^{(1)} \lesssim \Omega_R \lesssim \Omega^{(2)}$. This suggests that the deviation from the $Z(2)$ FSS by lattices with small LT, discussed in Sec. IV B, is due to that in the tails of the distribution $p(\Omega_R)$ for small LT.

B. Gap between the two minima

Using Eq. (13), the argument of Sec. V A on the effective potential can be extended away from the CP along the transition line. In this subsection, we study the gap between the two local minima of $V(\Omega_R; \lambda_c, LT)$,

$$\Delta\Omega = \Omega^{(2)} - \Omega^{(1)}. \qquad (46)$$

According to Eq. (15), this quantity should behave around the CP as

$$\Delta\Omega(\lambda, LT) = (LT)^{y_h - 3} \Delta\tilde{\Omega}((\lambda - \lambda_c)(LT)^{1/\nu}), \qquad (47)$$

provided that $p(\Omega_R)$ obeys the FSS.

In Fig. 13 we show the λ dependence of $\Omega^{(1)}$ and $\Omega^{(2)}$. As seen from Fig. 10, a clear two-peak structure of $p(\Omega_R)$ disappears when λ exceeds some value depending on LT. Even before the disappearance of the two peaks, identification of local maxima of $p(\Omega_R)$ becomes unstable because of statistical fluctuations. In Fig. 13 we thus truncate the plots for $\Omega^{(1)}$ and $\Omega^{(2)}$ at finite λ. The shaded areas in the figure represent statistical errors estimated by the jackknife method, for which we repeat the analysis of $\Omega^{(1,2)}$ for $p(\Omega_R)$ obtained in each jackknife sample with the smearing width of $\Delta_{\Omega_R} = 0.002$.[3] As shown in Appendix B, Δ_{Ω_R} dependence of these results is well suppressed at this Δ_{Ω_R}.

---

[3]We see that the errors in Fig. 13 become occasionally large. We find that this is due to statistical oscillations in the shape of $p(\Omega_R)$ around the peak: Though the oscillations are within the statistical errors, the peak position in each jackknife sample can jump discontinuously when oscillation appears just at the peak position as we vary λ. This makes the resulting jackknife error large there. From this observation, we think that these large errors are overestimated.
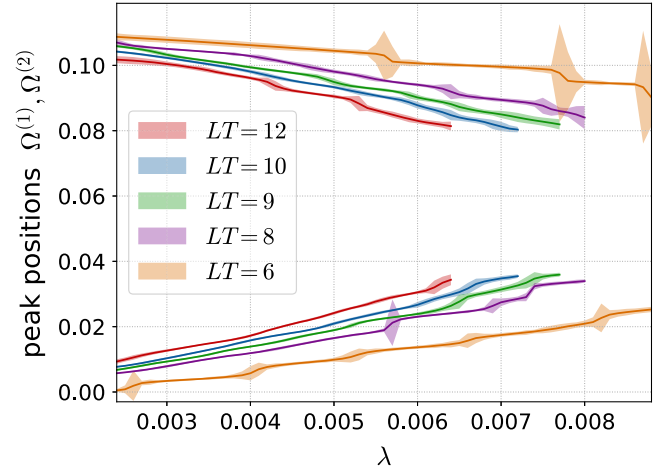

FIG. 13. Positions of peaks of the distribution function $p(\Omega_R)$ measured on the transition line.
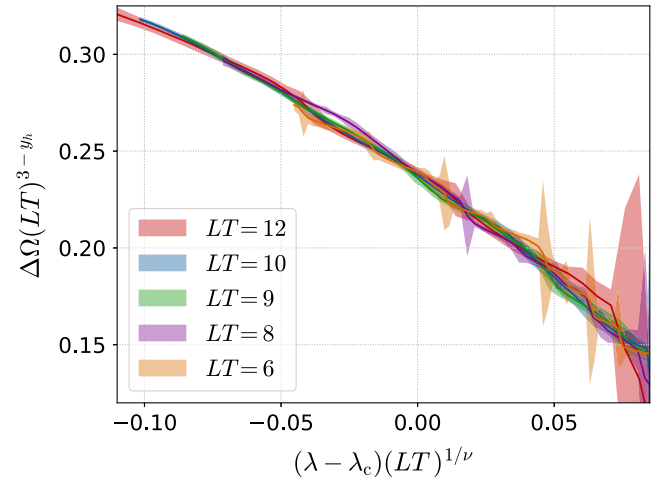

FIG. 14. Scaling of the gap $\Delta\Omega$ around T_c.

From Fig. 13 we extract $\Delta\Omega$ as a function of λ. In Fig. 14 we show $\Delta\Omega$ for five different volumes. To see the FSS, the vertical and horizontal axes are adjusted according to Eq. (47), where the $Z(2)$ values $3 - y_h = 0.518$ and $\nu = 0.630$, and $\lambda_c = 0.00503$ determined in the previous section are used. The figure shows that, for a wide range of $\lambda - \lambda_c$ and LT, the results of $\Delta\Omega$ obtained on different volumes are on top of each other within the errors. This supports the FSS of $p(\Omega_R)$ around the peak positions over a wide range of LT and λ.

It is interesting to note that the scaling behavior of $\Delta\Omega$ is observed even at $LT = 6$, although the FSS of B_4^Ω is violated already at $LT = 8$. As discussed in the previous subsection, we may understand this when the violation of the FSS for B_4^Ω is due to the violation in the tails of the distribution function $p(\Omega_R)$. As seen in Fig. 12, $V(\Omega_R)$ at various volumes agrees well for $\Omega^{(1)} \lesssim \Omega \lesssim \Omega^{(2)}$, even for small values of LT. As the higher-order cumulants are
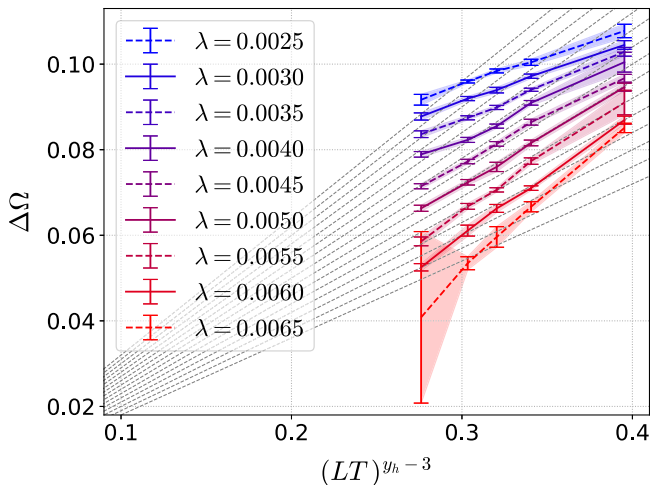
FIG. 15. Gap $\Delta\Omega$ around T_c as a function of $(LT)^{3-y_h}$. The $Z(2)$ value $3 - y_h = 0.518$ is assumed.

sensitive to the whole structure of the distribution, B_4^Ω will be more sensitive to the violation of FSS at the tails of the distribution. On the other hand, $\Delta\Omega$ is insensitive to them by definition. We also note that the statistical error of $\Delta\Omega$ is naturally small because it is defined by the peaks of the distribution. Therefore, $\Delta\Omega$ is useful in seeing the FSS around the CP.

From Eq. (47), $\Delta\Omega$ should behave linearly as a function of $(LT)^{y_h-3}$ at the CP, $\lambda = \lambda_c$. In Fig. 15, we show $\Delta\Omega$ on the transition line at various values of λ, as a function of $(LT)^{y_h-3}$. In the same figure, the dashed lines show linear functions $\Delta\Omega = k(LT)^{y_h-3}$ for various values of k. Figure 15 suggests that the linear behavior is realized at $\lambda \simeq 0.005$, which is consistent with our estimation $\lambda_c = 0.00503(14)(2)$ from the analysis of B_4^Ω.

C. Discussions

Let us comment on the relation of the present results with those given in Refs. [27,29,30]. In these studies, the CP is defined as the point at which the two-peak structure of $p(\Omega_R)$ disappears. On the lattice with finite LT, this leads to λ which is larger than our value of λ_c in the $L \to \infty$ limit. In fact, in Ref. [27], the location of the CP is estimated as $\kappa_c = 0.0658(3)\binom{+4}{-11}$ for $N_f = 2$ and $N_t = 4$, which is about 10% larger than that given in Table IV. Values of λ_c (κ_c) which are smaller than Ref. [30] for each N_t are also reported by a recent study of the CP in the heavy quark region on fine lattices ($N_t = 6$–10) using the Binder cumulant method [31]. Though the difference may be removed in the $L \to \infty$ limit, a careful extrapolation will be required. Because the FSS is clearly identified in this study, we think that the extrapolation to the large volume limit is stably performed with the present analysis.

We also note that the latent heat at the deconfinement transition in the $SU(3)$ Yang-Mills theory ($\kappa = 0$) has been

measured in Ref. [38] recently. It was found that the latent heat becomes larger with increasing the spatial volume. This may be attributed to a remnant of the FSS around the $Z(2)$ CP, like $\Delta\Omega$ studied in the present study.

VI. CONCLUSIONS

In this paper, we studied the distribution function of the Polyakov loop and its cumulants around the CP in the heavy quark region of QCD. Large volume simulations up to the aspect ratio $N_s/N_t = LT = 12$ have been carried out to see the finite-size scaling, while the lattice spacing is fixed to $N_t = 4$. We have performed the measurement of observables using the hopping parameter expansion for the quark determinant; the measurement has been performed at the next-to-leading order of the hopping parameter expansion by the multipoint-reweighting method evaluated on the gauge configurations generated for the action at the leading order. We found that this analysis is quite effective in reducing statistical errors by avoiding the overlapping problem of the reweighting method, while the numerical cost hardly changes from the pure Yang-Mills simulations. The convergence of the hopping parameter expansion at the next-to-leading order at the critical point is also verified by the comparison with the leading order result.

Using the data on $p(\Omega_R)$ thus obtained, we have performed the Binder cumulant analysis for determining the location of the critical point and evaluating the critical exponent. We found that the critical exponent ν and the value of the Polyakov-loop Binder cumulant B_4^Ω at the critical point is consistent with the $Z(2)$ universality class when $LT \geq 9$ data are used for the analysis. On the other hand, statistically-significant deviation from the $Z(2)$ scaling is observed when the data at $LT = 8$ is included, which suggests that this spatial volume is not large enough to apply the finite-size scaling.

The scaling behavior near the critical point is further studied using the structure of the distribution function of the real part of the Polyakov loop, $p(\Omega_R)$. We found that the structure of $p(\Omega_R)$ for various LT obeys the finite-size scaling well especially near the peaks of $p(\Omega_R)$. We have also proposed the use of the gap $\Delta\Omega$ between the peaks of $p(\Omega_R)$ for the finite-size scaling analysis. We showed that the λ and LT dependence of $\Delta\Omega$ is in good agreement with the $Z(2)$ scaling over a wide range of λ and LT. On the other hand, the deviation of $p(\Omega_R)$ around the tails of the distribution is observed on small lattices, which would give rise to the violation of the finite-size scaling of B_4^Ω in small volumes.

APPENDIX A: CUMULANTS

Let us consider a probability distribution function $p(x)$ of a stochastic variable x. Since $p(x)$ represents probability, it satisfies the normalization condition $\int dx\, p(x) = 1$.

The mth-order moment $\langle x^m \rangle$ of $p(x)$ is defined by

$$\langle x^m \rangle = \int dx\, x^m p(x). \tag{A1}$$

Using the moment-generating function

$$G(\theta) = \int dx\, e^{x\theta} p(x) = \langle e^{x\theta} \rangle, \tag{A2}$$

the moments are also given by

$$\langle x^m \rangle = \partial_\theta^m G(\theta)|_{\theta=0}, \tag{A3}$$

with $\partial_\theta = \partial/\partial\theta$.

The cumulants are defined from the cumulant-generating function

$$K(\theta) = \ln G(\theta) = \ln\langle e^{x\theta} \rangle, \tag{A4}$$

as

$$\langle x^m \rangle_c = \partial_\theta^m K(\theta)|_{\theta=0}. \tag{A5}$$

From Eqs. (A3) and (A5), one easily finds that the cumulants are related to the moments as

$$\langle x \rangle_c = \langle x \rangle, \tag{A6}$$

$$\langle x^2 \rangle_c = \langle x^2 \rangle - \langle x \rangle^2 = \langle \delta x^2 \rangle, \tag{A7}$$

$$\langle x^3 \rangle_c = \langle \delta x^3 \rangle, \tag{A8}$$

$$\langle x^4 \rangle_c = \langle \delta x^4 \rangle - 3\langle \delta x^2 \rangle^2, \tag{A9}$$

and etc. with $\delta x = x - \langle x \rangle$. The cumulants are useful in representing properties of $p(x)$ than the moments for various purposes. In particular, in statistical mechanics the cumulants of an extensive variable are extensive variables; see, for example, Ref. [5].

APPENDIX B: EFFECT OF SMEARING WIDTH ON THE DISTRIBUTION FUNCTION

In Secs. IV and V, we calculate the distribution function $p(\Omega_R)$ with smearing the delta function in Eq. (26) by the normalized Gauss function with the width Δ_{Ω_R}. With the statistics of this study, we choose $\Delta_{\Omega_R} = 0.002$ from an examination of the statistical error and resolution of $p(\Omega_R)$.

In this Appendix, we examine the Δ_{Ω_R} dependence of the numerical results, picking up the middle panel of Fig. 10, 13, and 14 as representative results. In Fig. 16, we compare these results with those obtained with $\Delta_{\Omega_R} = 0.001$ and 0.004. In Fig. 16, the results with $\Delta_{\Omega_R} = 0.001$ and 0.004 are shown by the dashed and dotted lines, while the solid lines show the results with $\Delta_{\Omega_R} = 0.002$. From this figure, we find that Δ_{Ω_R} dependence is very small around $\Delta_{\Omega_R} = 0.002$ and does not affect our conclusions.
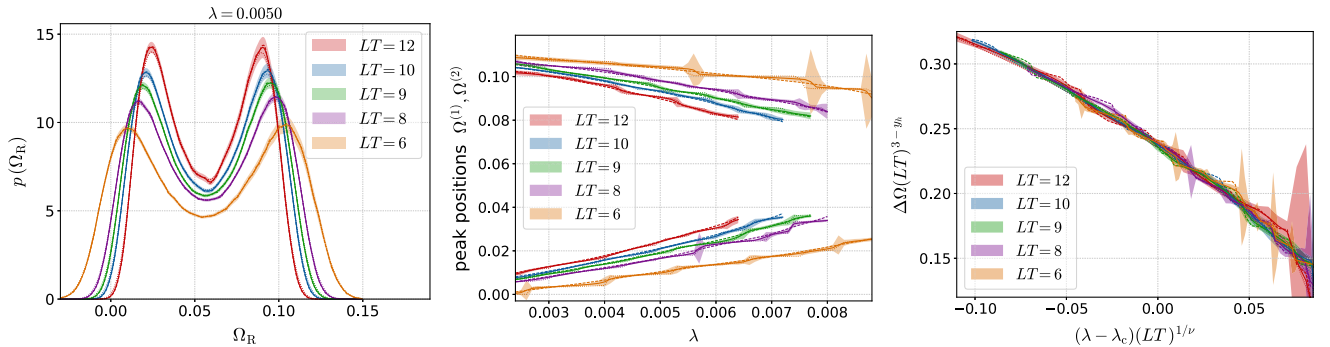

FIG. 16. Δ_{Ω_R} dependence of the results given in Secs. IV and V. The left, middle, and right panels correspond to the middle panel of Figs. 10, 13 and 14, respectively. The dashed and dotted lines in each panel shows the results with $\Delta_{\Omega_R} = 0.001$ and 0.004, respectively, while the solid lines with $\Delta_{\Omega_R} = 0.002$ are the same as the original figures.

APPENDIX C: LO AND NLO HOPPING PARAMETER EXPANSION

In this Appendix we derive Eqs. (30) and (31). Throughout this Appendix we assume a general value for N_t.

As in Eq. (28), the HPE of $\ln[\det M(\kappa)]$ is given by $\mathrm{Tr}[B^n]$. Since the matrix B only has nonzero contributions between neighboring lattice sites, $\mathrm{Tr}[B^n]$ are graphically represented by the closed trajectories with n links [35]. However, trajectories including "appendixes" shown in Fig. 17 do not contribute to the HPE because for such a path the product of the matrix in Dirac space vanishes at the tip of the appendix as $(1 - \gamma^\mu)(1 + \gamma^\mu) = 0$ [35]. With this exception, all possible closed trajectories composed of n links contribute to the HPE at the order κ^n. In the following, we calculate their contributions by classifying the trajectories by the shape.

Let us start from the plaquette, i.e., a 1×1 rectangle, which gives the lowest-order contribution to Eq. (28) at the order κ^4. The plaquette operator $\hat{P}$ in Eq. (18) is defined in such a way that $\langle \hat{P} \rangle = 1$ in the weak coupling limit with $U_\mu(x) = 1$. To satisfy this condition Eq. (18) has a coefficient $1/(N_c M_{\mathrm{plaq}}) = 1/18$, where M_{plaq} is the number of different plaquettes per lattice site; $M_{\mathrm{plaq}} = {}_4C_2 = 6$, which is the number of combinations of axes (μ, ν) at which the plaquettes are located.

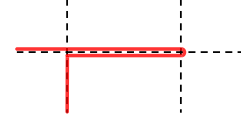The contribution from all plaquettes to Eq. (28) is calculated to be

FIG. 17. "Appendix" structure of trajectories.

$$-2N_c M_{\mathrm{plaq}} D_{\mathrm{plaq}} N_{\mathrm{site}} \kappa^4 \hat{P}. \tag{C1}$$

Here, D_{plaq} is the coefficient from the trace in the Dirac space

$$D_{\mathrm{plaq}} = \mathrm{tr}_{\mathrm{D}}[(1 - \gamma^\mu)(1 - \gamma^\nu)(1 + \gamma^\mu)(1 + \gamma^\mu)] = -8, \tag{C2}$$

where tr_{D} means the trace over the Dirac indices. The factor 2 in Eq. (C1) comes from two directions for each trajectory, which have to be distinguished in the HPE. The factor $1/n$ in Eq. (28) is canceled by the number of four starting points of a trajectory; this cancellation occurs for all trajectories.

Next, let us consider the Wilson loops of length 6 without windings along the temporal direction. At this order there are three types of trajectories; 1×2 rectangle, chair, crown, which are shown in Fig. 2. We define the operators, $\hat{W}_{\mathrm{rect}}$, $\hat{W}_{\mathrm{chair}}$, $\hat{W}_{\mathrm{crown}}$, corresponding to these trajectories as

$$\hat{W}_{\mathrm{rect}} = \frac{1}{N_c M_{\mathrm{rect}} N_{\mathrm{site}}} \sum_x \sum_{\mu \neq \nu} \mathrm{Re}\,\mathrm{tr}_{\mathrm{C}}[U_{x,\mu} U_{x+\hat{\mu},\mu} U_{x+2\cdot\hat{\mu},\nu} U^\dagger_{x+\hat{\mu}+\hat{\nu},\mu} U^\dagger_{x+\hat{\nu},\mu} U^\dagger_{x,\nu}], \tag{C3}$$

$$\hat{W}_{\mathrm{chair}} = \frac{1}{N_c M_{\mathrm{chair}} N_{\mathrm{site}}} \sum_x \sum_{\mu,\nu < \rho, \nu \neq \mu \neq \rho} \sum_{s,t=\pm 1} \mathrm{Re}\,\mathrm{tr}_{\mathrm{C}}[U_{x,s\nu} U_{x+s\hat{\nu},\mu} U^\dagger_{x+\hat{\mu},s\nu} U_{x+\hat{\mu},t\rho} U^\dagger_{x+t\hat{\rho},\mu} U^\dagger_{x,t\rho}], \tag{C4}$$

$$\hat{W}_{\mathrm{crown}} = \frac{1}{N_c M_{\mathrm{crown}} N_{\mathrm{site}}} \sum_x \sum_{\mu < \nu < \rho} \sum_{s,t=\pm 1} \mathrm{Re}\,\mathrm{tr}_{\mathrm{C}}[U_{x,\mu} U_{x+\hat{\mu},s\nu} U_{x+\hat{\mu}+s\hat{\nu},t\rho} U^\dagger_{x+s\hat{\nu}+t\hat{\rho},\mu} U^\dagger_{x+t\hat{\rho},s\nu} U^\dagger_{x,t\rho}], \tag{C5}$$

with $U_{x,-\mu} = U^\dagger_{x-\hat{\mu},\mu}$ and $U^\dagger_{x,-\mu} = U_{x-\hat{\mu},\mu}$ for $\mu > 0$. Equations (C3)–(C5) are defined so that $\langle \hat{W}_{\mathrm{rect}} \rangle = \langle \hat{W}_{\mathrm{chair}} \rangle = \langle \hat{W}_{\mathrm{crown}} \rangle = 1$ in the weak coupling limit again; to satisfy this condition, the operators are divided by $M_{\mathrm{rect}} = 12$, $M_{\mathrm{chair}} = 48$, and $M_{\mathrm{crown}} = 16$, respectively, corresponding to the number of trajectories per lattice site.

The contribution of these trajectories to the HPE of $\ln \det M(\kappa)$ is given by

$$-2N_c \sum_s M_s D_s N_{\mathrm{site}} \kappa^6 \hat{W}_s, \tag{C6}$$

with $s = \mathrm{rect}$, chair, crown. D_s is the coefficient from the trace in the Dirac space. For the 1×2 rectangle we have

$$D_{\mathrm{rect}} = \mathrm{tr}_{\mathrm{D}}[(1 - \gamma_\mu)(1 - \gamma_\mu)(1 - \gamma_\nu) \\ \times (1 + \gamma_\mu)(1 + \gamma_\mu)(1 + \gamma_\nu)] = -32, \tag{C7}$$

and similar manipulations lead to $D_{\mathrm{chair}} = D_{\mathrm{crown}} = -16$.

Next we consider the Polyakov-loop type operators having a winding along the temporal direction. The lowest-order contribution among them is the Polyakov loop, i.e., the straight lines of length N_t. To calculate its contribution, one needs to pay special attention to the fact

that there is only one independent Polyakov loop for each spatial coordinate on one time slice, not for each lattice site. Therefore, their contribution to the HPE is given by

$$2N_c D_{\text{pol}} N_s^3 \kappa^{N_t} \hat{\Omega}_R, \qquad (C8)$$

where the factor -1 is to be applied because of the antiperiodic boundary condition of the quark determinant. The real part of $\hat{\Omega}$ has to be taken after multiplying by the

factor two because two directions of a trajectory are independently taken into account. The factor from the Dirac trace for the Polyakov loop is calculated to be

$$D_{\text{pol}} = \text{tr}_D[(1 - \gamma_4)^{N_t}] = 2^{N_t + 1}. \qquad (C9)$$

Finally, we consider the contribution of the bent Polyakov loops shown in Fig. 3, whose explicit definitions are given by

$$\hat{\Omega}_1 = \frac{1}{N_c M_{\text{bent}} N_{\text{site}}} \sum_x \sum_{i=1}^3 \sum_{s=\pm 1} \text{tr}_C[U_{x,si} U_{x+s\hat{i},4} U^\dagger_{x+\hat{4},si} U_{x+\hat{4},4} U_{x+2\cdot\hat{4},4} \cdots U_{x+(N_t-1)\cdot\hat{4},4}], \qquad (C10)$$

$$\hat{\Omega}_2 = \frac{1}{N_c M_{\text{bent}} N_{\text{site}}} \sum_x \sum_{i=1}^3 \sum_{s=\pm 1} \text{tr}_C[U_{x,si} U_{x+s\hat{i},4} U_{x+s\hat{i}+\hat{4},4} U^\dagger_{x+2\cdot\hat{4},si} U_{x+2\cdot\hat{4},4} \cdots U_{x+(N_t-1)\cdot\hat{4},4}], \qquad (C11)$$

and so on. The factor $M_{\text{bent}} = 6$ is needed to make $\langle \hat{\Omega}_n \rangle = 1$ in the weak coupling limit. From the definition of $\hat{\Omega}_n$ we have $\hat{\Omega}_n = \hat{\Omega}_{N_t-n}$. Also, when N_t is even $\hat{\Omega}_{N_t/2}$ counts each trajectory twice, and thus its contribution to the HPE should be divided by 2. Bearing these facts in mind, the contribution from $\hat{\Omega}_n$ to the HPE of $\ln \det M(\kappa)$ is given by

$$2N_c M_{\text{bent}} D_{\text{bent}} N_{\text{site}} \kappa^{N_t+2} \left(\sum_{n=1}^{N_t/2-1} \text{Re}\hat{\Omega}_n + \frac{1}{2}\text{Re}\hat{\Omega}_{N_t/2} \right), \qquad (C12)$$

where the last term $\text{Re}\hat{\Omega}_{N_t/2}$ should be omitted for odd N_t. The contribution of the Dirac trace is calculated to be $D_{\text{bent}} = 2^{N_t+1}$.

Combining these results gives Eqs. (30) and (31).

APPENDIX D: COMPARISON OF LO AND NLO RESULTS

In this Appendix, to see the convergence of the HPE at the NLO, we repeat the analyses in Sec. IV B at the LO and compare its results with those at the NLO. In Fig. 18, we show the Binder cumulant B_4^Ω obtained at the LO along the transition line. In this figure, the NLO results of Fig. 11 are also shown by the thin dashed lines. We find that, though the difference between the LO and the NLO results grows as λ becomes larger, the deviation is within a few percent level around the crossing point λ_c.

In Fig. 18, we also show the result of the four-parameter fit with Eq. (42) using data at the LO on $LT = 12$, 10 and 9 lattices by the black square. The same result at the NLO is shown by gray circle for comparison. We find that the LO result of the fit is consistent with the NLO result

within statistical errors: The values of $b_4 = 1.630(24)$ and $\nu = 0.620(47)$ at the LO are hardly changed from the corresponding NLO values given in Table III. Though the central value of $\lambda_c = 0.00516(15)$ at the LO is about 2.6% larger than the NLO value, it is consistent with the NLO result within errors, suggesting that the truncation error of the HPE at the NLO is well suppressed in these quantities.

The success of the B_4^Ω fit together with the consistency of the fit results with the $Z(2)$ values suggests that the $Z(2)$ scaling is also realized with the LO action when the system volume is sufficiently large. This is reasonable since the scaling properties near the CP are insensitive to detailed structures of the theory.
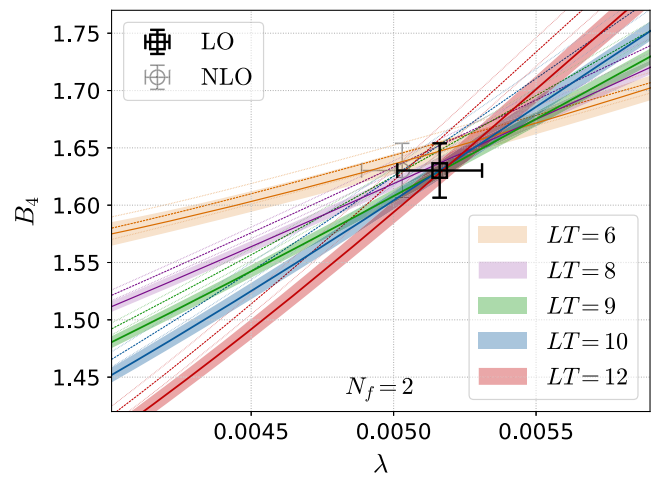

FIG. 18. Binder cumulant B_4 as a function of λ calculated at the LO. The results at the NLO are also plotted by the thin dashed lines. The square with errors shows the fit result with three largest volumes. The NLO result on the same fit result is also shown by the thin circle.

[1] Y. Aoki, G. Endrodi, Z. Fodor, S. D. Katz, and K. K. Szabo, Nature (London) **443,** 675 (2006).

[2] S. Ejiri, F. Karsch, E. Laermann, C. Miao, S. Mukherjee, P. Petreczky, C. Schmidt, W. Soeldner, and W. Unger, Phys. Rev. D **80,** 094505 (2009).

[3] M. Asakawa and K. Yazaki, Nucl. Phys. **A504,** 668 (1989).

[4] M. A. Stephanov, K. Rajagopal, and E. V. Shuryak, Phys. Rev. D **60,** 114028 (1999).

[5] M. Asakawa and M. Kitazawa, Prog. Part. Nucl. Phys. **90,** 299 (2016).

[6] A. Bzdak, S. Esumi, V. Koch, J. Liao, M. Stephanov, and N. Xu, Phys. Rep. **853,** 1 (2020).

[7] M. Bluhm *et al.*, Nucl. Phys. **A1003,** 122016 (2020).

[8] R. D. Pisarski and F. Wilczek, Phys. Rev. D **29,** 338 (1984).

[9] F. R. Brown, F. P. Butler, H. Chen, N. H. Christ, Z.-h. Dong, W. Schaffer, L. I. Unger, and A. Vaccarino, Phys. Rev. Lett. **65,** 2491 (1990).

[10] S. Gavin, A. Gocksch, and R. D. Pisarski, Phys. Rev. D **49,** R3079 (1994).

[11] Y. Iwasaki, K. Kanaya, S. Kaya, S. Sakai, and T. Yoshie, Z. Phys. C **71,** 343 (1996).

[12] Y. Iwasaki, K. Kanaya, S. Kaya, S. Sakai, and T. Yoshie, Phys. Rev. D **54,** 7010 (1996).

[13] S. Aoki *et al.* (JLQCD Collaboration), Nucl. Phys. B, Proc. Suppl. **73,** 459 (1999).

[14] F. Karsch, E. Laermann, and C. Schmidt, Phys. Lett. B **520,** 41 (2001).

[15] P. de Forcrand and O. Philipsen, Nucl. Phys. **B673,** 170 (2003).

[16] C. Bernard, T. Burch, E. B. Gregory, D. Toussaint, C. E. DeTar, J. Osborn, S. Gottlieb, U. M. Heller, and R. Sugar (MILC Collaboration), Phys. Rev. D **71,** 034504 (2005).

[17] P. de Forcrand and O. Philipsen, J. High Energy Phys. 01 (2007) 077.

[18] M. Cheng *et al.*, Phys. Rev. D **75,** 034506 (2007).

[19] X.-Y. Jin, Y. Kuramashi, Y. Nakamura, S. Takeda, and A. Ukawa, Phys. Rev. D **88,** 094508 (2013).

[20] S. Ejiri, R. Iwami, and N. Yamada, Phys. Rev. D **93,** 054506 (2016).

[21] A. Bazavov, H. T. Ding, P. Hegde, F. Karsch, E. Laermann, S. Mukherjee, P. Petreczky, and C. Schmidt, Phys. Rev. D **95,** 074505 (2017).

[22] X.-Y. Jin, Y. Kuramashi, Y. Nakamura, S. Takeda, and A. Ukawa, Phys. Rev. D **96,** 034523 (2017).

[23] Y. Kuramashi, Y. Nakamura, H. Ohno, and S. Takeda, Phys. Rev. D **101,** 054509 (2020).

[24] T. Banks and A. Ukawa, Nucl. Phys. **B225,** 145 (1983).

[25] T. A. DeGrand and C. E. DeTar, Nucl. Phys. **B225,** 590 (1983).

[26] C. Alexandrou, A. Borici, A. Feo, P. de Forcrand, A. Galli, F. Jegerlehner, and T. Takaishi, Phys. Rev. D **60,** 034504 (1999).

[27] H. Saito, S. Ejiri, S. Aoki, T. Hatsuda, K. Kanaya, Y. Maezawa, H. Ohno, and T. Umeda (WHOT-QCD Collaboration), Phys. Rev. D **84,** 054502 (2011); **85,** 079902(E) (2012).

[28] M. Fromm, J. Langelage, S. Lottini, and O. Philipsen, J. High Energy Phys. 01 (2012) 042.

[29] H. Saito, S. Ejiri, S. Aoki, K. Kanaya, Y. Nakagawa, H. Ohno, K. Okuno, and T. Umeda, Phys. Rev. D **89,** 034507 (2014).

[30] S. Ejiri, S. Itagaki, R. Iwami, K. Kanaya, M. Kitazawa, A. Kiyohara, M. Shirogane, and T. Umeda (WHOT-QCD Collaboration), Phys. Rev. D **101,** 054505 (2020).

[31] F. Cuteri, O. Philipsen, A. Schön, and A. Sciarra, Phys. Rev. D **103,** 014513 (2021).

[32] A. Pelissetto and E. Vicari, Phys. Rep. **368,** 549 (2002).

[33] A. M. Ferrenberg and R. H. Swendsen, Phys. Rev. Lett. **63,** 1195 (1989).

[34] K. Binder, Z. Phys. B **43,** 119 (1981).

[35] H. J. Rothe, *Lattice Gauge Theories: An Introduction* (World Scientific Publishing Company, Singapore, 1992), Vol. 43.

[36] R. Iwami, S. Ejiri, K. Kanaya, Y. Nakagawa, D. Yamamoto, and T. Umeda, Phys. Rev. D **92,** 094507 (2015).

[37] P. Hasenfratz, F. Karsch, and I. O. Stamatescu, Phys. Lett. **133B,** 221 (1983).

[38] M. Shirogane, S. Ejiri, R. Iwami, K. Kanaya, M. Kitazawa, H. Suzuki, Y. Taniguchi, and T. Umeda (WHOT-QCD Collaboration), Prog. Theor. Exp. Phys. **2021,** 013B08 (2021).