

# Comparison of algorithms for solving the sign problem in the O(3) model in 1 + 1 dimensions at finite chemical potential

S. D. Katz,<sup>1,2</sup> F. Niedermayer,<sup>3</sup> D. Nógrádi,<sup>1,2</sup> and Cs. Török<sup>1,2</sup>

<sup>1</sup>*Institute for Theoretical Physics, Eötvös University,  
Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary*

<sup>2</sup>*MTA-ELTE “Lendület” Lattice Gauge Theory Research Group,  
Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary*

<sup>3</sup>*Institute for Theoretical Physics, Albert Einstein Center for Fundamental Physics Bern University,  
Sidlerstrasse 5, CH-3012 Bern, Switzerland*

(Received 17 November 2016; published 15 March 2017)

We study three possible ways to circumvent the sign problem in the O(3) nonlinear sigma model in 1 + 1 dimensions. We compare the results of the worm algorithm to complex Langevin and multiparameter reweighting. Using the worm algorithm, the thermodynamics of the model is investigated, and continuum results are shown for the pressure at different  $\mu/T$  values in the range 0–4. By performing  $T = 0$  simulations using the worm algorithm, the Silver Blaze phenomenon is reproduced. Regarding the complex Langevin, we test various implementations of discretizing the complex Langevin equation. We found that the exponentialized Euler discretization of the Langevin equation gives wrong results for the action and the density at low  $T/m$ . By performing a continuum extrapolation, we found that this discrepancy does not disappear and depends slightly on temperature. The discretization with spherical coordinates performs similarly at low  $\mu/T$  but breaks down also at some higher temperatures at high  $\mu/T$ . However, a third discretization that uses a constraining force to achieve the  $\phi^2 = 1$  condition gives correct results for the action but wrong results for the density at low  $\mu/T$ .

DOI: [10.1103/PhysRevD.95.054506](https://doi.org/10.1103/PhysRevD.95.054506)

## I. INTRODUCTION

Monte Carlo simulations of quantum field theories based on the path integral formalism play an important role in investigating the physics of various models nonperturbatively. However, standard numerical methods fail when the action becomes complex, and thus the probability interpretation of the weight  $e^{-S}$  and importance sampling cannot be applied. The problem is present in QCD at finite density or with a theta term and also arises in condensed matter physics, e.g. in the simulations of strongly correlated electronic systems [1]. To solve these complex action problems, several methods have been devised; for a review of different approaches and further references, see Refs. [2–5].

In the present paper, we compare and test three different methods, namely reweighting, the worm algorithm, and the complex Langevin in the case of the 1 + 1-dimensional O(3) model. Our focus is primarily on the applicability of the complex Langevin algorithm, since today it may seem that it is a promising approach to simulate even QCD [6,7], although several problems have not been solved yet. The idea behind the complex Langevin is stochastic quantization and originates from the work of Parisi, Wu, and Klauder from the 1980s [8–10]. But soon after its proposal, the first simulations revealed certain problems: the instability of the simulations with the absence of convergence (runaway trajectories) [11] and that even stable simulations may converge to a wrong limit [12]. These problems

hindered reliable calculations, but in the last decade, important improvements have been achieved. Runaway trajectories can now be eliminated e.g. using adaptive step size [13], and a formal justification of the algorithm as well as necessary and sufficient conditions for convergence to correct results have been established [14–16]. Roughly speaking, these suggest that if the probability distributions of the complexified variables fall sufficiently fast, then the results are correct. In order to reach this for gauge theories, the gauge cooling procedure was developed [17], which works perfectly in some models or at a certain parameter range but may fail in other models and parameter ranges [17–20]. In Ref. [20], where heavy dense QCD (HDQCD) was studied, it was argued that failure happens below a specific  $\beta$  value and, by increasing the temporal lattice size, one can get correct results at lower temperatures; in other words, continuum extrapolation may be feasible. The validity of this statement, however, is not entirely clear and may be model dependent. On the one hand, the above observation in HDQCD helped in exploring the phase diagram of the model [21], but on the other hand, in the case of full QCD, recent results [7] show that, using  $N_t = 4, 6, 8$  lattices, the breakdown of the complex Langevin prevents the exploration of the confined region. We note that e.g. for the 3D XY model, the breakdown of the complex Langevin also occurred around the phase boundary [22], but in that model, the question of continuum limit behavior cannot be addressed.

In the present paper, we investigate the 1 + 1-dimensional O(3) model for this purpose, which is not a gauge theory but asymptotically free; thus, the continuum behavior can be analyzed. We compare the results of different discretizations of the complex Langevin equation to the results of reweighting and the worm algorithm.

From the viewpoint of the sign problem, these two approaches are also interesting and can give insight into the properties of the O(3) model.

In this sense, reweighting is a well-defined approach, but with limited efficiency and reliability as the sign problem becomes more severe (and also an overlap problem appears). The worm algorithm, also called the dual variables approach [23], however, completely eliminates the sign problem of the model by introducing new, dual variables. The difficulty in this case is the rewriting of the model to these dual variables, but after it has been accomplished, effective simulations using the worm algorithm can be performed, and in fact many interesting models have been studied throughout the years [24–33]. Here, using the worm algorithm, we study the thermodynamic properties of the O(3) model.

Although the dual formalism of this model was introduced and studied [34,35] during the finalization of our work on the comparison of the different methods, we also introduce this formalism in this paper in order to give a consistent introduction to our notations.

In the following sections, after some introductory remarks about the O(3) model in Sec. II and the description of scale setting in Sec. III, we discuss these approaches in more detail: in Sec. IV the reweighting, in Sec. V the worm algorithm, and in Sec. VI the complex Langevin. In Sec. VII, we compare the results of the simulations. In the Appendix, we discuss in detail the updating steps of the worm algorithm.

## II. FORMULATION

The O(3) model in 1 + 1 dimensions has been widely studied in the past for several reasons, among others, because it has interesting features in common with four-dimensional non-Abelian gauge theories. Over the years, many important results have been achieved also numerically and—since the model is more or less tractable—analytically. Nonetheless, we do not give an overview here of the overall history of these results but mention only some facts that made this model attractive for us.

First of all, the coupling constant of the theory is dimensionless; thus, the theory is perturbatively renormalizable. It is asymptotically free in 1 + 1 dimensions [36,37], which enables us to study the continuum limit of the results obtained at finite lattice spacings. The O(3) model also has a nonperturbative mass gap generated dynamically. Moreover, similarly to QCD, the O(3) model also possesses instanton solutions [38,39].

The Lagrangian of the nonlinear O(3) model is

$$\mathcal{L} = \frac{1}{2g^2} (\partial_\mu \phi)^2, \quad (1)$$

where the fields obey the  $\sum_{i=1}^3 \phi_i^2 = 1$  condition in every space-time point.

The discretized action in 1 + 1 dimensions with periodic boundary condition is

$$\begin{aligned} S &= \frac{1}{g^2} \left( 2 \sum_x \phi_x^2 - \sum_{x,\mu=0,1} \phi_{x+\hat{\mu}} \phi_x \right) \\ &= 2\beta V - \beta \sum_{x,\mu=0,1} \phi_{x+\hat{\mu}} \phi_x, \end{aligned} \quad (2)$$

where we introduced  $\beta = 1/g^2$  and the lattice volume  $V = N_x \times N_t$ . After introducing the chemical potential to the rotations in the (12)-plane of O(3), the action becomes

$$S = 2\beta V - \beta \sum_x (\phi_{x+\hat{0}} e^{i\mu a t_{12}} \phi_x + \phi_{x+\hat{1}} \phi_x), \quad (3)$$

where  $t_{12}$  is the generator of the rotation in the (12)-plane of O(3).

## III. SCALE SETTING

Since in the later part of the paper we are interested in continuum extrapolations and physical quantities computed from the dimensionless quantities measured on the lattice, we need to determine the lattice spacing as a function of  $\beta$ , which we discuss in this section. In order to achieve this,  $\mu = 0$  simulations have been performed, for which we used the cluster algorithm [40–42] and measured the second moment correlation length  $\xi_2$  at zero temperature.  $\xi_2$  is defined through

$$\frac{1}{\xi_2} = \frac{\sin(\pi/N_t)}{\pi/N_t} \sqrt{\left( \frac{2M_0 a^2}{M_2} - \frac{4\pi^2}{N_t^2} \right)}, \quad (4)$$

where  $M_0$  denotes the zeroth moment and  $M_2$  denotes the second moment [43,44]:

$$M_{2n} = \left( \frac{N_t a}{2\pi} \right)^{2n} \sum_t \left( 2 \sin \left( \frac{\pi t}{N_t a} \right) \right)^{2n} C(t). \quad (5)$$

$C(t)$  is the two-point correlation function,  $C(t) = \sum_x \langle \sum_a \phi_a(x, t) \phi_a(0, 0) \rangle$ .  $\xi_2$  does not equal  $\xi = 1/ma$  but scales as  $\xi$  in the  $\beta \rightarrow \infty$  limit [43,45,46], and in infinite volume, the ratio  $\xi/\xi_2$  is very close to 1; it is 1.000826(1) [47]. The advantage of using  $\xi_2$  is that one does not need to fit any correlators this way. We can thus estimate the mass gap  $ma = 1/\xi$  with the help of  $\xi_2$  by running large volume, zero temperature simulations.

Actually, we used  $80 \times 80$  and  $100 \times 100$  lattices for  $0.9 \leq \beta \leq 1.57$ ,  $120 \times 120$  and  $140 \times 140$  lattices for  $1.58 \leq \beta \leq 1.62$ ,  $250 \times 250$  lattices for  $1.63 \leq \beta \leq 1.72$ , and  $400 \times 400$  lattices for  $1.73 \leq \beta \leq 1.85$ . The simulation points were chosen uniformly in the above  $\beta$  ranges with a  $\Delta\beta = 0.01$  distance from each other with  $10^6$  or  $2 \times 10^6$  cluster updates after thermalization, using every tenth for measurement. We studied the overlapping  $\beta$  regions as well and used larger lattices if deviations larger than errors between the smaller and larger volumes had occurred. The results are shown in Fig. 1.

#### IV. REWEIGHTING

Reweighting uses the idea of rewriting the partition function (and the expectation value of observables) in such a way that one needs to do simulations only at zero  $\mu$  and determine the configurations that can be relevant at finite  $\mu$ . This is done by measuring the weights of the configurations, which enables one to distinguish between them.

In the multiparameter reweighting approach [48], one reweights both in  $\beta$  and  $\mu$ , as we show it for the partition function of the O(3) model,

$$\begin{aligned} Z &= \int \prod_x d\phi_x \delta(\phi_x^2 - 1) e^{-S(\beta, \mu)} \\ &= \int \prod_x d\phi_x \delta(\phi_x^2 - 1) e^{-S(\beta_0, \mu_0=0)} w(\beta, \mu, \beta_0, \mu_0=0) \\ &= Z_0 \langle w \rangle_{\beta_0, \mu_0=0}, \end{aligned} \quad (6)$$

where  $w(\beta, \mu, \beta_0, \mu_0=0) = e^{S(\beta_0, \mu_0=0) - S(\beta, \mu)}$  is the weight and  $Z_0$  is the partition function for  $\beta_0$  and  $\mu_0=0$ . As one can see, this rewritten partition function can be simulated directly using standard methods since the action  $S(\beta_0, \mu_0=0)$  is real. Using reweighting, the expectation value of an  $O$  observable is the following:

$$\langle O(\beta, \mu) \rangle_{\beta, \mu} = \frac{\langle O(\beta, \mu) w(\beta, \mu, \beta_0, \mu_0=0) \rangle_{\beta_0, \mu_0=0}}{\langle w(\beta, \mu, \beta_0, \mu_0=0) \rangle_{\beta_0, \mu_0=0}}. \quad (7)$$

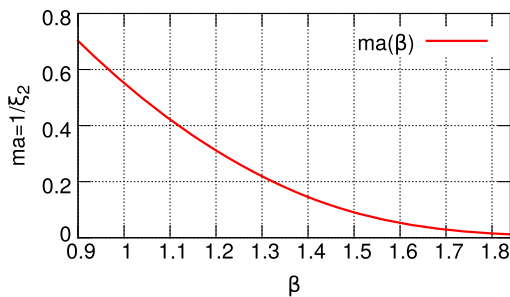


FIG. 1. The  $ma(\beta)$  scale for the  $1+1$ -dimensional O(3) model. (Errors are smaller than or compatible with the width of the line.) The parameters and details of the simulations are summarized in the text.

Although reweighting can reduce the sign problem, an overlap problem occurs in this case. That is, we have different important configurations at our “source” ( $\beta_0, \mu = 0$ ) ensemble and at the “target” ( $\beta, \mu$ ) ensemble. If the two sets just slightly overlap or do not overlap at all, then one rarely reaches the important configurations at  $\beta, \mu$  by simulating at  $\beta_0, \mu_0 = 0$ . In these cases, it can happen that one has many relatively small weights at the same order of magnitude, and only some with many magnitudes larger, and as a consequence collects only a tiny fraction of useful statistics during even long simulations. It was observed that multiparameter reweighting helps to reduce the overlap problem in the case of QCD and also can help to reduce the sign problem by doing reweighting on the so-called best reweighting lines [49,50]. These are defined as those curves that have the smallest standard deviations of  $\text{Re}(\ln w)$  on them. In the case of the O(3) model, we illustrate these lines in Fig. 2, and we show the overlap problem and the advantages of multiparameter reweighting in Fig. 3’s top and bottom panels, respectively. We also illustrate the severeness of the sign problem in Fig. 4, which is based on measurements on  $56 \times 14$  lattices. We used the cluster algorithm to simulate at  $\mu_0 = 0$ . Further results obtained by the multiparameter reweighting method can be found in Sec. VII, where we compare them to the worm and complex Langevin results.

#### V. WORM ALGORITHM

Another approach we employed in this study is the worm algorithm. In order to maintain generality, we will review the algorithm and the dual formulation in the O(N) case at  $d+1$  dimensions. In the O(N) case, the fields obey the  $\sum_{i=1}^N \phi_i^2 = 1$  condition in every space-time point. The discretized action in  $d+1$  dimensions is

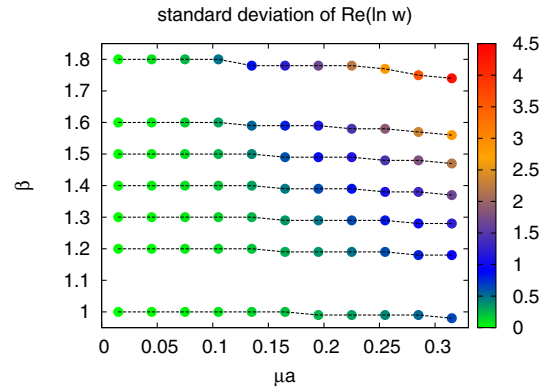


FIG. 2. The standard deviation of  $\text{Re}(\ln w)$  for reweighting from  $\beta_0 = 1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.8$  to different  $\beta, \mu$  values on  $56 \times 14$  lattices. We define the best reweighting lines (with dashed) as those that have the smallest standard deviations on them.

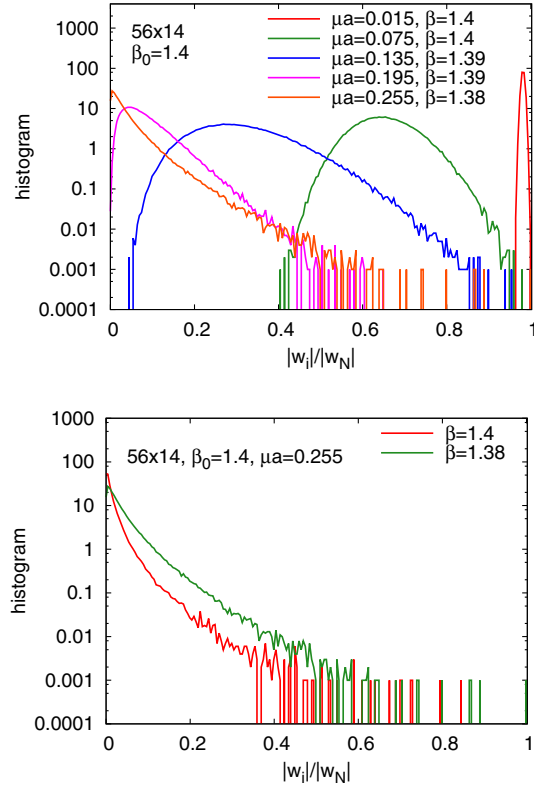


FIG. 3. Top: The histogram of normalized weights along the best reweighting line starting from  $\beta_0 = 1.4$ ,  $\mu_0 = 0$  shows the severeness of the overlap problem. At the horizontal axis, one finds  $|w_i|/|w_N|$  sorted as  $|w_1| < |w_2| \dots < |w_N|$ , while at the vertical axis, we show  $\#/N/\Delta w$ , where  $\#$  is the number of configurations that have normalized weights between  $|w_i|/|w_N|$  and  $|w_i|/|w_N| + \Delta w$ .  $N$  is the total number of configurations, and  $\Delta w = 0.005$ . Bottom: The figure illustrates how reweighting in  $\beta$  can help to enhance overlapping: the green curve shows that one has more configurations with greater weights when reweighting from  $\beta_0 = 1.4$ ,  $\mu_0 a = 0$  to  $\beta = 1.38$ ,  $\mu a = 0.255$  than in the case of e.g. standard reweighting.

$$S = \beta(d+1)V_{d+1} - \beta \sum_x \left( \phi_{x+\hat{0}} e^{i\mu a t_{12}} \phi_x + \sum_{i=1, \dots, d} \phi_{x+\hat{i}} \phi_x \right), \quad (8)$$

where now  $\beta = a^{d-1}/g^2$  and the  $d+1$ -dimensional lattice volume is  $V_{d+1} = N_s^d \times N_t$ .

Our goal in lattice simulations is to calculate expectation values. The worm algorithm performs this as counting different types of configurations, as we are going to explain later. The algorithm itself is based on the dual formulation of a model, which means in our case the characterization of configurations not with the  $\phi(x)$  continuous variables at every lattice point but with a set of discrete variables  $\{m_a\}_{a=(l;i=1, \dots, N)}$ , which “live” on links ( $l$ ). One  $\{m_a\}$  configuration matches to a certain, finite partial sum of an infinite sum (e.g. the expansion of the partition sum), and the algorithm jumps between these partial sums.

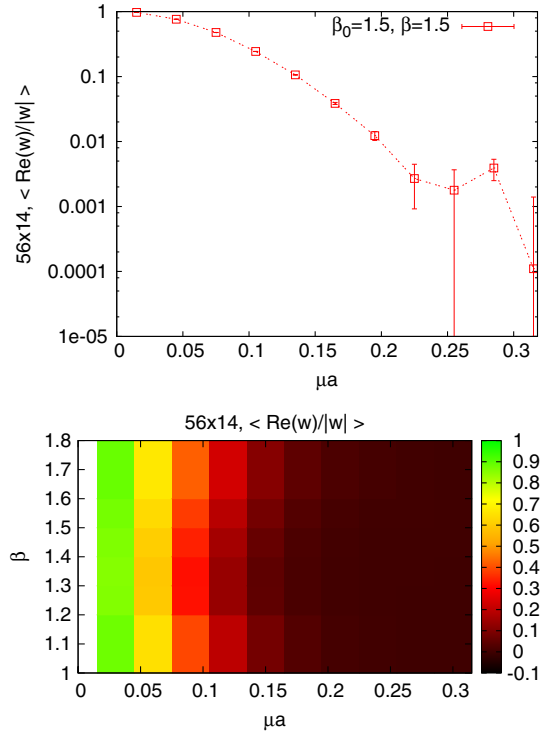


FIG. 4. Top: The figure shows the expectation value of the real part of the normalized weight at  $56 \times 14$  lattice when reweighting from  $\beta_0 = 1.5$ ,  $\mu_0 = 0$  to  $\beta = 1.5$  and  $\mu a$  values along the horizontal axis. Bottom:  $\langle \text{Re}(w)/|w| \rangle$  shows how hard the sign problem is in different  $\beta$ ,  $\mu$  regions.

In the following, we first review some conventional notations and integrals over the  $O(N)$  sphere, which is helpful for the derivation of weights. Then, at first, we restrict ourselves for the dual formulation of the case without the chemical potential and introduce the worm algorithm; then, we repeat the same procedure for the case with the chemical potential. Finally, we present continuum results in Secs. V D 2 and V D 4, and comparisons to reweighting and the complex Langevin can be found in Sec. VII.

### A. Integrals over $O(N)$ sphere

Consider an  $O(N)$  vector  $\phi$  of unit length,  $\phi^2 = 1$ . Averaging over the  $O(N)$  sphere, with the normalization condition  $\langle 1 \rangle = 1$ , one has  $\langle \phi_i \phi_j \rangle = \delta_{ij}/N$ ,  $\langle \phi_i \phi_j \phi_k \phi_l \rangle = (\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk})/(N(N+2))$ . For the general case, when one has  $k$  even number of vectors, one has

$$\langle \phi_{i_1} \phi_{i_2} \dots \phi_{i_k} \rangle = C_k^{(N)} (\delta_{i_1 i_2} \delta_{i_3 i_4} \dots \delta_{i_{k-1} i_k} + \text{perm}), \quad (9)$$

where

$$\begin{aligned} C_k^{(N)} &= \frac{1}{N(N+2)\dots(N+k-2)} \\ &= \frac{\Gamma(N/2)}{2^{k/2} \Gamma(N/2 + k/2)}. \end{aligned} \quad (10)$$



In the brackets in (9), there are altogether

$$\begin{aligned} M_k &= (k-1)!! \equiv 1 \cdot 3 \cdot \dots \cdot (k-3)(k-1) \\ &= 2^{-\frac{k}{2}+1} \frac{\Gamma(k)}{\Gamma(k/2)} \end{aligned} \quad (11)$$

terms for all possible pairings of the indices. This can be obtained e.g. from the recursion relation  $M_{k+2} = M_k + k(k-1)M_{k-2}$ , with  $M_2 = 1, M_4 = 3$ . To check this, one can contract in the last two indices  $i_{k-1}, i_k$ . The corresponding recursion relation reads

$$C_{k-2}^{(N)} = C_k^{(N)} \left( N + \frac{M_k}{M_{k-2}} - 1 \right) = C_k^{(N)} (N + k - 2). \quad (12)$$

It is interesting to note that for a free system, i.e. when the constraint  $\delta(\phi^2 - 1)$  is replaced by the Gaussian  $\exp(-\phi^2/2)$ , one gets  $C_k^{(N)} = 1$  for the weights in (9). This is in agreement with the relation

$$\lim_{N \rightarrow \infty} N^{k/2} C_k^{(N)} = \lim_{N \rightarrow \infty} \frac{N^{k/2}}{N(N+2)\dots(N+k-2)} = 1. \quad (13)$$

Collecting the powers of different components, one has

$$\begin{aligned} w(k_1, \dots, k_N) &\equiv \langle \phi_1^{k_1} \phi_2^{k_2} \dots \phi_N^{k_N} \rangle \\ &= \frac{\Gamma(N/2)}{\Gamma((k+N)/2)} \prod_{i=1}^N \frac{\Gamma((k_i+1)/2)}{\Gamma(1/2)} \\ &= \frac{1}{N(N+2)\dots(N+k-2)} \\ &\quad \times \prod_{i=1}^N (1 \cdot 3 \cdot \dots \cdot (k_i-1)), \end{aligned} \quad (14)$$

where all  $k_i$  are even and  $k = k_1 + \dots + k_N$ . In the last expression, the product is taken only for  $i$ 's for which  $k_i > 0$ . Also, obviously,  $w(0, 0, \dots, 0) = 1$ . Some useful relations obtaining the coefficients are (assuming that all powers  $k$  are even)

$$\int_{-\infty}^{\infty} dx e^{-x^2} x^k = \Gamma\left(\frac{k+1}{2}\right), \quad (15)$$

and

$$\begin{aligned} \int d^N x e^{-x^2} x_1^{k_1} \dots x_N^{k_N} &= \prod_{i=1}^N \Gamma\left(\frac{k_i+1}{2}\right) \\ &= \int_0^{\infty} dr e^{-r^2} r^{N+k-1} S_N \langle \phi_1^{k_1} \dots \phi_N^{k_N} \rangle \\ &= \frac{1}{2} \Gamma\left(\frac{N+k}{2}\right) S_N \langle \phi_1^{k_1} \dots \phi_N^{k_N} \rangle, \end{aligned} \quad (16)$$

where  $S_N$  is the surface of the  $N$ -dimensional sphere:  $S_N = 2\pi^{N/2}/\Gamma(N/2)$ .

## B. Strong coupling expansion without chemical potential

In the case without the chemical potential, for one link between neighbor lattice points  $x$  and  $y$ , one can write

$$\begin{aligned} e^{\beta\phi(x)\phi(y)} &= \sum_{m_1, \dots, m_N} \frac{\beta^m}{m_1! \dots m_N!} (\phi_1(x)\phi_1(y))^{m_1} \dots \\ &\quad \times (\phi_N(x)\phi_N(y))^{m_N}, \end{aligned} \quad (17)$$

where  $m = m_1 + \dots + m_N$ . Then, one can consider the sum

$$\begin{aligned} \sum_{i=1}^N \sum_{u, v \in \Lambda} \int_{\phi} \phi_i(u)\phi_i(v) e^{-S} \\ &= \sum_{\text{conf}} W(u, v, i; m) \\ &= e^{-\beta(d+1)V_{d+1}} \sum_{\text{conf}} \left( \prod_l \frac{\beta^{m^{(l)}}}{m_1^{(l)}! \dots m_N^{(l)}!} \right) \\ &\quad \times \left( \prod_x w(k_1(x), \dots, k_N(x)) \right), \end{aligned} \quad (18)$$

where  $\text{conf} = \{u, v, i; \{m_j^{(l)}\}_{j=1, \dots, N}^{\text{all } l \text{ links}}\}$  is the configuration: two distinguished points ( $u$  and  $v$ ), a component  $i$ , and the set  $\{m_j^{(l)}\}_{j=1, \dots, N}^{\text{all } l \text{ links}}$ . During the simulations, the configuration can change in different ways, which we discuss in the Appendix. One way is to change  $u$  to a neighboring site, meanwhile increasing/decreasing  $m_i^{(l)}$  along the  $l$  link that connects these two. We start from  $u = v$ , and the continuous path connecting  $u$  and  $v$  that appears this way is called the *worm*. The weights  $w(\dots)$  of Eq. (18) are given by Eq. (14). The value of  $k_j(x)$  depends on the position of  $x$ :  $k_j(x) = \hat{k}_j(x) + (\delta_{xu} + \delta_{xv})\delta_{ij}$ , where  $\hat{k}_j(x) = \sum_{x \in \partial l} m_j^{(l)}$ . Then, all  $k_j(x)$  values must be even. The ratios of the weights when one of the  $k_i$ 's is changed by  $\pm 2$  are

$$\frac{w(k_1+2, k_2, \dots, k_N)}{w(k_1, k_2, \dots, k_N)} = \frac{k_1+1}{k_1+N}, \quad (19)$$

$$\frac{w(k_1-2, k_2, \dots, k_N)}{w(k_1, k_2, \dots, k_N)} = \frac{k_1+N-2}{k_1-1}. \quad (20)$$

With the help of the sum (18) defined above, one can see that the partition sum is related to those configurations where the two ends of the worm coincide ( $u = v$ ); this gives explicitly  $V_{d+1}Z$ . The configurations with one distance between the two ends ( $c_1$ ) divided by the number of configurations where the two ends of the worm coincide

( $c_2$ ) give  $\frac{c_1}{c_2} = -\frac{2}{\beta} (\frac{\langle S \rangle}{V_{d+1}} - \beta(d+1))$ , from which one can determine  $\langle S \rangle$ , the expectation value of the action.

### C. Strong coupling expansion with chemical potential

The action with the chemical potential coupled to  $t_{12}$  is obtained from the standard one by replacing the interaction terms which couple the fields in the time direction according to  $\phi(x+\hat{0})\phi(x) \rightarrow \phi(x+\hat{0})e^{i\mu a t_{12}}\phi(x)$ . Therefore, the corresponding action is complex. For the O(2) non-linear sigma model, this problem was avoided in Ref. [27] using the worm algorithm: the terms in the strong-coupling expansion are real even in the presence of the chemical potential. We describe here the extension to the O( $N$ ) case for general  $N$ . Let us introduce  $\phi_{\pm} = \frac{1}{\sqrt{2}}(\phi_1 \pm i\phi_2)$ . Expressed through these variables, the scalar product is

$$\begin{aligned} \phi(x+\hat{0})\phi(x) &= \phi_-(x+\hat{0})\phi_+(x) + \phi_+(x+\hat{0})\phi_-(x) \\ &\quad + \phi_3(x+\hat{0})\phi_3(x) + \dots + \phi_N(x+\hat{0})\phi_N(x), \end{aligned} \quad (21)$$

and the action is

$$\begin{aligned} S &= \beta(d+1)V_{d+1} - \beta \sum_x \sum_{\nu=0}^d \left( e^{-\mu_\nu a} \phi_-(x+\hat{\nu})\phi_+(x) \right. \\ &\quad \left. + e^{\mu_\nu a} \phi_+(x+\hat{\nu})\phi_-(x) + \sum_{j=3}^N \phi_j(x+\hat{\nu})\phi_j(x) \right), \end{aligned} \quad (22)$$

where  $\mu_\nu = \mu\delta_{\nu,0}$ ,  $\nu = 0, 1, \dots, d$ . When one integrates over  $\phi$  at a given site, the nonvanishing contributions are all real and positive,

$$\begin{aligned} w(k_1, k_2, \dots, k_N) &= \langle (\phi_+ \phi_-)^{k_{12}} \phi_3^{k_3} \dots \phi_N^{k_N} \rangle \\ &= \frac{1}{2^{k_{12}}} \sum_{m=0}^{k_{12}} \binom{k_{12}}{m} \langle \phi_1^{2m} \phi_2^{2k_{12}-2m} \phi_3^{k_3} \dots \phi_N^{k_N} \rangle \\ &= \frac{\Gamma(N/2)}{\Gamma((k+N)/2)} 2^{-k_{12}} \Gamma(k_{12}+1) \prod_{i=3}^N \frac{\Gamma((k_i+1)/2)}{\Gamma(1/2)}, \end{aligned} \quad (23)$$

where  $k = 2k_{12} + k_3 + \dots + k_N$  and  $k_3, \dots, k_N$  are even.

The strong-coupling expansion for spatial neighbor sites is the same as in (17), and for temporal neighbor sites, it is

$$\begin{aligned} e^{\beta S^T e^{i\mu a t_{12}} S'} &= \sum_{m_+, m_-, m_3, \dots, m_N} \frac{\beta^m}{m_+! m_-! m_3! \dots m_N!} \\ &\quad \times (e^{\mu a} S_- S'_+)^{m_+} (e^{-\mu a} S_+ S'_-)^{m_-} \\ &\quad \times (S_3 S'_3)^{m_3} \dots (S_N S'_N)^{m_N}, \end{aligned} \quad (24)$$

where  $m = m_+ + m_- + m_3 + \dots + m_N$  and  $\mathbf{S} \equiv \phi(x)$ ,  $\mathbf{S}' \equiv \phi(x+\hat{0})$ . Consider then

$$\begin{aligned} &\sum_i \sum_{u,v \in \Lambda} \int_{\phi} \phi_i(u) \phi_i(v) e^{-S} \\ &= \sum_{\text{conf}} W(\{u, v, i; m\}; \mu) \\ &= e^{-\beta(d+1)V_{d+1}} \sum_{\text{conf}} \left( \prod_l \frac{\beta^{m^{(l)}} e^{(\mu a)^{(l)} m_+^{(l)}} e^{-(\mu a)^{(l)} m_-^{(l)}}}{m_+^{(l)}! m_-^{(l)}! \dots m_N^{(l)}!} \right) \\ &\quad \times \left( \prod_x w(k_{12}(x), k_3(x), \dots, k_N(x)) \right), \end{aligned} \quad (25)$$

where  $i = +, -, 3, \dots, N$  and  $(a\mu)^{(l)} = a\mu$  for timelike links and 0 for spatial links. In these expressions,

$$\begin{aligned} k_j(x) &= \hat{k}_j(x) + \delta_{xu} \delta_{j_i u} + \delta_{xv} \delta_{j_i v} \\ j &= +, -, 3, \dots, N, \end{aligned} \quad (26)$$

where  $(i_u, i_v) = (-, +), (3, 3), \dots, (N, N)$ , and different  $\hat{k}$ 's are defined as

$$\hat{k}_+(x) = \sum_{\nu=0}^d (m_+^{(x-\hat{\nu}, x)} + m_-^{(x, x+\hat{\nu})}) \quad (27)$$

$$\hat{k}_-(x) = \sum_{\nu=0}^d (m_-^{(x-\hat{\nu}, x)} + m_+^{(x, x+\hat{\nu})}) \quad (28)$$

$$\begin{aligned} \hat{k}_j(x) &= \sum_{x \in \partial l} m_j^{(l)} = \sum_{\nu=0}^d (m_j^{(x-\hat{\nu}, x)} + m_j^{(x, x+\hat{\nu})}) \\ j &= 3, \dots, N. \end{aligned} \quad (29)$$

The nonzero terms in (25) are those in which the same number of  $\phi_+(x)$  and  $\phi_-(x)$  factors are present; i.e.  $k_+(x) = k_-(x) = k_{12}$ , and the number of  $\phi_j(x)$  factors,  $k_j(x)$  (for  $j = 3, \dots, N$ ), is even at all sites  $x$ . The detailed steps of the worm algorithm based on these prescriptions are discussed in the Appendix. In the following subsections, we leave the general formalism and consider the O(3) model in 1 + 1 dimensions.

## D. Numerical results obtained with the worm algorithm

### 1. Check of the algorithm

In order to check the reliability of our algorithm, we have studied the spectrum of the O(3) model. The energy levels are characterized by the isospin quantum numbers  $I, I_3$  and the momentum  $p'$ . Let us consider the  $p' = 0$  case. Then, the smallest energy at zero  $\mu$  in a given sector is denoted by  $E(I)$ . The chemical potential splits the  $2I + 1$ -fold degeneracy, and the energy levels become  $E(I, I_3; \mu) = E(I) - \mu I_3$ , which has a minimum at  $I = I_3 = q(\mu)$ . By increasing  $\mu$ , larger  $q$  values are expected. Using the worm

algorithm and counting the number of + and - link variables connecting two time slices, one can determine  $q$  for that interval.

The two ends of the worm divide the periodic time direction into two parts: an interval with length  $\tau$  (where  $0 \leq \tau < N_t$ ) and charge  $I_3 = q$  and an interval with length  $N_t - \tau$  and charge  $I_3 = q - 1$ . Then, these give the leading contribution to the correlator  $C(\tau; q, q-1) \approx A_{q,q-1} \exp\{-E(q; \mu)\tau - E(q-1; \mu)(N_t - \tau)\} \propto \exp\{-(E(q) - E(q-1) - \mu)\tau\}$ , and thus by fitting the correlator, the energy differences can be determined. Choosing  $\mu \approx E(q) - E(q-1)$ , one obtains a long plateau in the effective mass plot. This way, one can follow the signal over a large interval in  $\tau$  to measure energies of higher excitations. These energy differences provide a strong consistency check, since we measure the same difference with different  $\mu$  values. In particular, we have measured the energy differences on  $16 \times 200$  lattices at  $\beta = 1.779$  using several values of  $\mu$  and obtained  $E(1) - E(0) = 0.0662(1)$ ,  $E(2) - E(1) = 0.1284(3)$ ,  $E(3) - E(2) = 0.1867(3)$ . Note that these agree roughly with the (approximate) rotator picture which is expected to hold for small spatial volumes [51–53]. The mass gap  $E(1) - E(0)$  agrees within the statistical error with the value cited in Ref. [54].

## 2. Pressure

Similar to what was mentioned at the end of Sec. VB, the Eq. (25) sum is related to the partition function if  $u = v$ , when it gives  $V_2 Z$ . For the action, one needs to calculate the ratio of two terms. In the denominator, there is  $Z$ , while in the numerator, there is  $-\frac{\beta}{2}(e^{\mu a} \times \#_1 + e^{-\mu a} \times \#_2 + \#_3 + \#_4)$ , where  $\#_1$  is the number of configurations with  $v = u + \hat{0}$  and  $i_u = -, i_v = +$ ;  $\#_2$  is the number of configurations with  $u - \hat{0} = v$  and  $i_u = -, i_v = +$ ;  $\#_3$  is the number of configurations with  $u \pm \hat{1} = v$  and  $i_u = -, i_v = +$ ; and  $\#_4$  is the number of configurations with  $u \pm \hat{v} = v$  with  $\nu = 0, 1$  and  $i_u, i_v = 3$ . The value of the numerator is constructed in the way that is suitable for the action (22). [The first term of (22), which is independent of the dual variables, was added to the averages at the end.]

For calculating thermodynamic quantities, we used lattices with  $N_x > N_t$  and measured the action after each worm movement. Since it is divergent as  $a \rightarrow 0$ , we renormalized it by subtracting  $\langle S(\beta, T = 0, \mu = 0) \rangle$ . For the latter, we used large symmetric lattices, in particular those that were used for determining the scale (see Sec. III). In order to eliminate the finite-size effects, we chose box sizes of  $ma(\beta_{pc})N_x \geq 5$ , where  $\beta_{pc}$  is the inflection point of the pressure. Since no phase transition is expected in this model, this  $\beta_{pc}$  is only a pseudocritical quantity. The chosen box sizes correspond to the aspect ratio  $N_x/N_t = 4$ , so we have used  $32 \times 8, 40 \times 10, 56 \times 14, 64 \times 16, 72 \times 18, 80 \times 20, 120 \times 30$  lattices for finite-temperature simulations. We used around  $4 \times 10^{10} \dots 1.2 \times 10^{11}$  local worm updates on

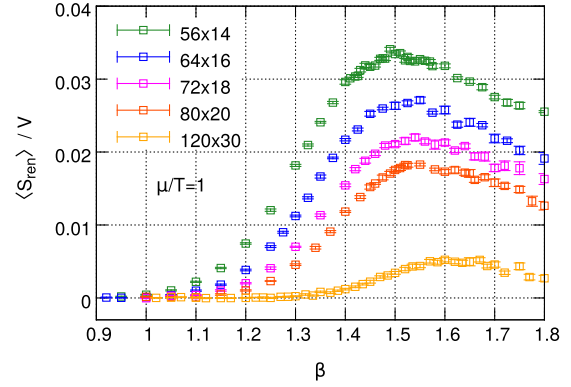


FIG. 5. The renormalized action divided by the lattice volume ( $V = N_x \times N_t$ ), measured by the worm algorithm.

these lattices, after  $(3 \dots 5) \times 10^5$  thermalization steps. Note that these numbers refer to the local change of the configuration. In order to compare the amount of updates to those of the Langevin simulations, one should divide them with the two-dimensional lattice volume. After calculating the action, we used the *integral method* [55] to obtain the pressure  $p(T)$ :

$$\begin{aligned} \frac{p}{T^2} &= \frac{N_t}{N_x} \log Z = \frac{N_t}{N_x} \int_{\beta_0}^{\beta} d\beta' \frac{\partial \log Z}{\partial \beta'} \\ &= \frac{N_t}{N_x} \int_{\beta_0}^{\beta} d\beta' \left\langle -\frac{\partial S}{\partial \beta'} \right\rangle. \end{aligned} \quad (30)$$

Since we defined  $S$  with  $\beta$  included,  $\partial S / \partial \beta$  is simply  $S / \beta$ . The pressure is also divergent, so we need to renormalize it using the expectation value of the renormalized action  $\langle S_{\text{ren}}(\beta, T, \mu) \rangle = \langle S(\beta, T, \mu) \rangle - \langle S(\beta, T = 0, \mu = 0) \rangle$  in the integrand of formula (30). In the following, we denote the renormalized pressure with  $p$ . Figure 5 shows the renormalized action density at  $\mu/T = 1$ , while Figs. 6 and 7 show the results for the renormalized pressure.

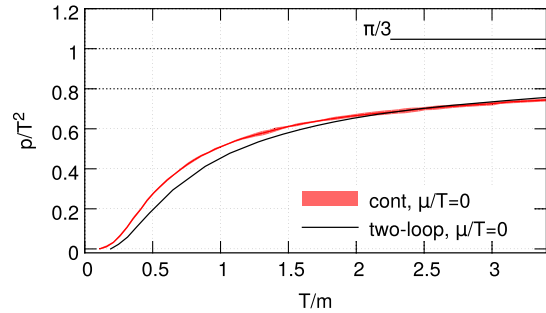


FIG. 6. Comparison of our lattice results and the two-loop calculation of Ref. [56] for  $p/T^2$  at  $\mu = 0$  in the continuum. The dashed line at  $\pi/3$  shows the asymptotic limit at high temperature.

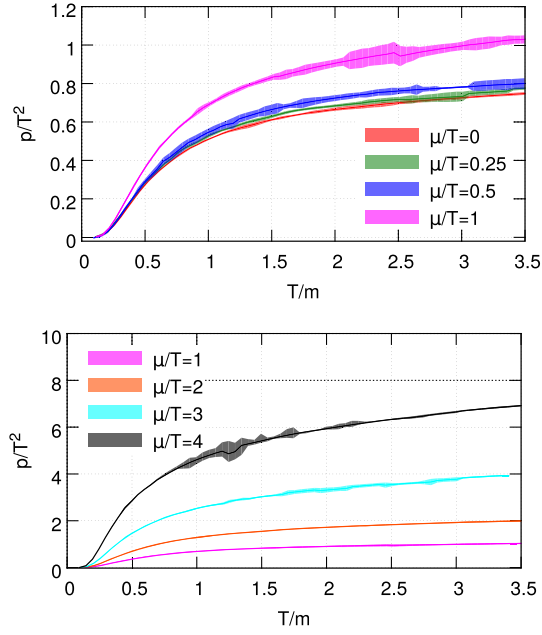


FIG. 7. Continuum results for  $p/T^2$  as a function of  $T/m$  for different  $\mu/T$  values.

### 3. Trace anomaly

Another quantity of interest is the trace anomaly (also called the interaction measure):

$$\frac{\theta}{T^2} = \frac{\epsilon - p}{T^2} = -\frac{N_t}{N_x} a \frac{\partial \log Z}{\partial a} = \frac{N_t}{N_x} \frac{\langle am \rangle}{\frac{\partial \langle am \rangle}{\partial \beta}} \left\langle \frac{S}{\beta} \right\rangle. \quad (31)$$

This quantity is also divergent; thus, renormalization is needed to obtain a finite value in the continuum, which is achieved simply by using the renormalized action  $S_{\text{ren}}$  instead of  $S$  in the above formula. Below, we show the continuum results for different  $\mu/T$  values as a function of temperature (Fig. 8).

We note that, similarly to the inflection point of the pressure, the peak position of the trace anomaly can also serve as a definition of a pseudocritical temperature ( $T_{\text{pc}}$ ) characterizing the transition in the O(3) model.<sup>1</sup>

### 4. Density

Another quantity we measured during the simulations was the isospin charge density, which is defined through

$$\begin{aligned} n &= \frac{T}{V_{sp}} \frac{\partial \log Z}{\partial \mu} = \frac{T}{V_{sp}} \frac{1}{Z} \frac{\partial Z}{\partial \mu} = \frac{T}{V_{sp}} \left\langle -\frac{\partial S}{\partial \mu} \right\rangle \\ &= m \frac{1}{N_t N_x} \frac{1}{am} \left\langle -\frac{\partial S}{\partial (\mu a)} \right\rangle, \end{aligned} \quad (32)$$

<sup>1</sup>Together with the  $T_{\text{pc}}$  calculated from the inflection point of the pressure, we show how the pseudocritical temperatures depend on  $\mu/T$  in Fig. 25.

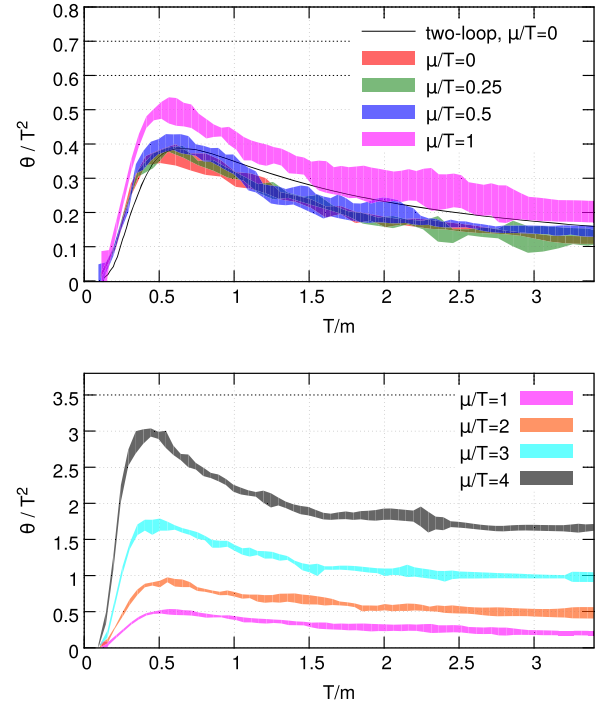


FIG. 8. Continuum results for  $\theta/T^2$  as a function of  $T/m$  for different  $\mu/T$  values. The two-loop result for  $\mu = 0$  is from Ref. [56].

where  $V_{sp}$  is the spatial volume which is simply  $N_x a$  in our case. The density does not need to be renormalized, because the divergent part of the action is independent of  $\mu$ . In the figures, we show the dimensionless ratio  $n/m$ . With the worm algorithm, the density can be calculated again as a ratio, which has  $Z$  in its denominator, and in the numerator, there is  $\frac{\beta}{2} (e^{\mu a} \times \#_1 - e^{-\mu a} \times \#_2)$ , where  $\#_1$  is the number of configurations with  $v = u + \hat{0}$  and  $i_u = -, i_v = +$  and  $\#_2$  is the number of configurations with  $u - \hat{0} = v$  and  $i_u = -, i_v = +$ . The value of the numerator is constructed in the way that is suitable for  $\partial S / \partial (\mu a)$  [see the derivative of (22) with respect to  $\mu a$ ].

As one can observe in Fig. 9,  $n/m$  depends almost linearly on  $T/m$ . Although we did not perform continuum extrapolation above  $T/m \approx 3.5$ , the numerical data from  $56 \times 14$  lattices show that this linear behavior also holds at higher temperature, at least up to  $T/m \approx 4.6$ . The configurations used for the finite density calculation were the same as for the pressure. We have also analyzed the low temperature behavior of the density as a function of  $\mu/m$ , where we observed the well-known Silver Blaze phenomenon (Fig. 10). We approached the  $T = 0$  continuum physics by running simulations at fixed  $\beta$  values, increasing the volume of the symmetric lattice; then, we extrapolated these  $T = 0$  results to the continuum (Fig. 11 and 12). Another way of obtaining the continuum results would be to run simulations at fixed low temperatures and take the continuum limit first, then extrapolate these finite



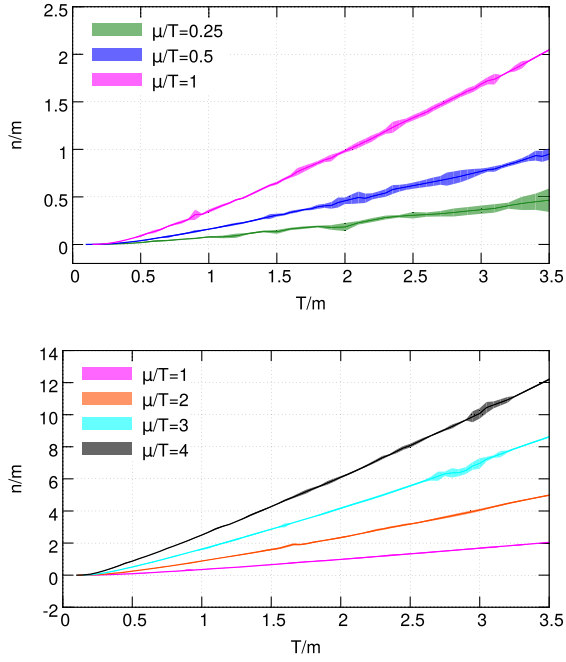


FIG. 9. The isospin charge density divided by  $m$  at finite temperature in the continuum limit.

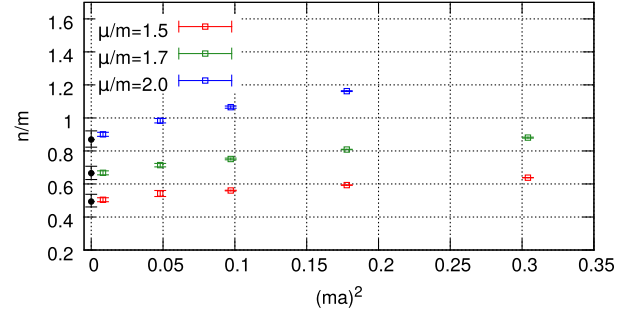


FIG. 11. Continuum extrapolation at some  $\mu/m$  at  $T/m = 0$ .

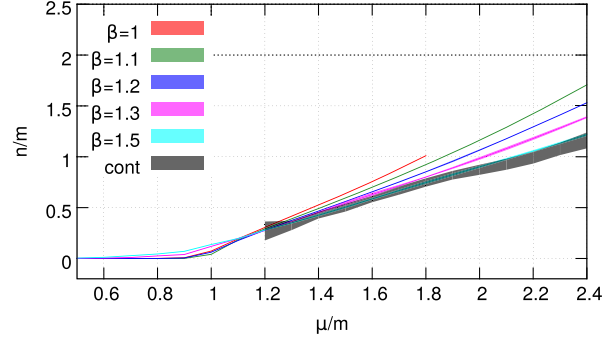


FIG. 12. Continuum limit for  $n/m$  at  $T/m = 0$  using lattices with fixed  $\beta$  values.

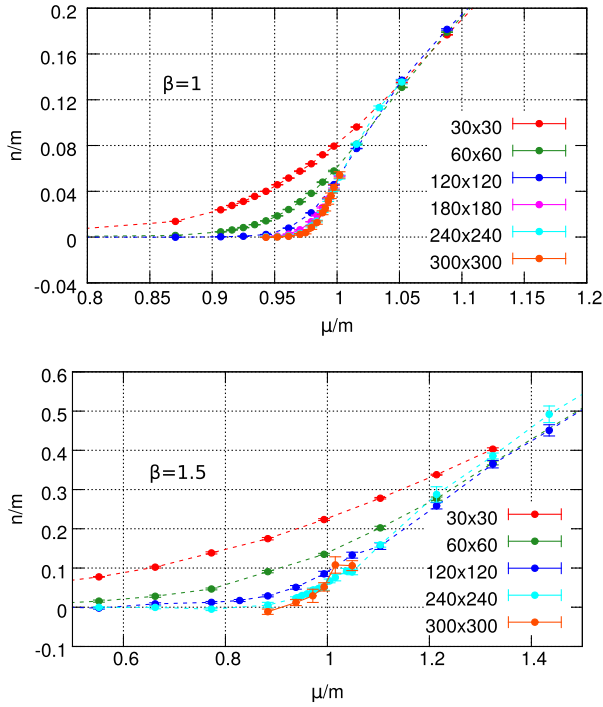


FIG. 10. The isospin charge density over  $m$  at low temperatures at  $\beta = 1$  and at  $\beta = 1.5$  for different lattice sizes. We used the  $T \rightarrow 0$  limit extrapolation based on these results to obtain the continuum limit. Although the analytical behavior of  $n/m$  is known at  $T = 0$  and infinite volume near  $\mu \sim m$  [57], our lattice results do not show this directly, because we are either far from the continuum (upper panel) or the temperature is not too small to reproduce this precisely (lower panel).

temperature continuum results to  $T = 0$ . We did not analyze in full detail this latter case but only performed simulations to obtain the density at two low temperatures:  $T/m = 0.01$  and  $T/m = 0.005$ . We compared these to the  $T = 0$  results in Fig. 13.

The parameters for these low temperature simulations can be found in Table I. We note that the thermalization took significantly more steps at low temperature as one increased the lattice size and  $\beta$ . For example, in the case of  $N_t = 500$ ,  $\beta = 1.3234$  for  $N_x ma = 10$ , thermalization took  $3 \times 10^9$  steps, but for  $N_x ma = 40$ , it was  $\sim 4.4$  times longer, and for  $N_x ma = 100$  (symmetric lattice), it was  $\sim 19$  times longer than for  $N_x ma = 10$ . Thermalization was analyzed using the values of density during the simulations.

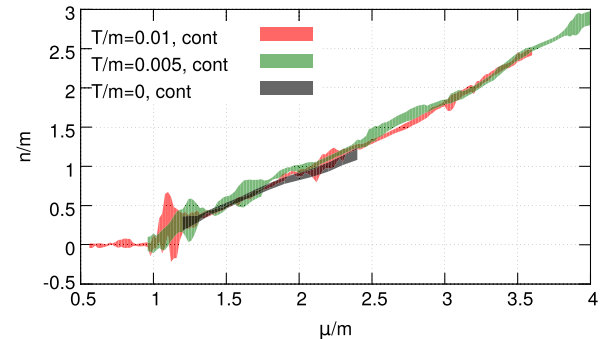


FIG. 13. Comparison of  $n/m$  continuum results at  $T/m = 0.01$ ,  $0.005$  and  $T/m = 0$ .

TABLE I. The set of parameters for low and zero temperature runs with the worm algorithm. We have run simulations at several  $\mu a$  using the above parameters. The number of used worm configurations for these runs was around  $(3\dots 9) \times 10^{10}$  after thermalization.

$T/m$	$\beta$	$ma$	$N_t = N_x$	
0	1	0.551	30, 60, 120, 180, 240, 300, 360	
	1.1	0.422	30, 60, 120, 180, 240, 300, 360	
	1.2	0.312	30, 60, 120, 180, 240, 300, 360, 500	
	1.3	0.219	30, 60, 120, 180, 240, 300, 360, 500	
	1.5	0.091	30, 60, 120, 180, 240, 300, 360	
$T/m$	$\beta$	$ma$	$N_t$	$N_x ma$
0.01	1.1789	0.333	300	approx. 10, 20, 40, 60,
	1.2644	0.25	400	100 (symmetric lattices)
	1.3234	0.2	500	
	1.3682	0.167	600	
0.005	1.2644	0.25	800	approx. 10, 20, 30, 40,
	1.2963	0.222	900	50, 60, 70
	1.3234	0.2	1000	
	1.3473	0.182	1100	

## VI. COMPLEX LANGEVIN ALGORITHM

We proceed with the realization of the complex Langevin algorithm for the O(3) model. The continuum complex Langevin equation for each component  $i = 1, 2, 3$  of a three-component scalar field variable is

$$\frac{\partial \phi_{x,i}(\tau)}{\partial \tau} = -\frac{\delta S[\phi; \tau]}{\delta \phi_{x,i}(\tau)} + \eta_{x,i}(\tau), \quad (33)$$

where  $\tau$  is the simulation time and  $\eta_{x,i}(\tau)$  is a Gaussian noise obeying the following relations:

$$\begin{aligned} \langle \eta_{x,i}(\tau) \eta_{x',j}(\tau') \rangle &= 2\delta_{ij} \delta_{xx'} \delta(\tau - \tau'), \\ \langle \eta_{x,i}(\tau) \rangle &= 0. \end{aligned} \quad (34)$$

The simplest discretization for Eq. (33) is the so-called Euler (or Euler-Maruyama) discretization,

$$\phi_{x,i}^{(n+1)} = \phi_{x,i}^{(n)} - \varepsilon \frac{\delta S}{\delta \phi_{x,i}}^{(n)} + \sqrt{\varepsilon} \eta_{x,i}^{(n)}, \quad (35)$$

where we denote the simulation steps with  $n$  and  $\varepsilon$  is a finite step size. However, Eq. (35) does not preserve the length of the  $\phi_x$  vector, so in order to simulate the O(3) model, we must somehow include the constraint  $\sum_i \phi_{x,i}^2 = 1$ , because the partition function for the O(3) model is

$$\begin{aligned} Z &= \int \prod_x d\phi_x \delta(\phi_x^2 - 1) e^{-S[\phi]} \\ &= \int \prod_x d\phi_x e^{-(S[\phi] - \sum_x \ln \delta(\phi_x^2 - 1))}. \end{aligned} \quad (36)$$

Usually the integration measure is not considered explicitly during the integration; one does not use the force arising from the constraint but uses other (general) coordinates or specific integration. For example, suitable general coordinates in our case are spherical coordinates, and an example for a specific integration scheme in Cartesian coordinates is the so-called Euler discretization in group space, which is used for example in complex Langevin (CL) simulations of SU(N) gauge groups [17]. In the following, we will study these approaches to integrating CL equations.

### A. Use of spherical coordinates

Using spherical coordinates  $\phi_x = (\sin \vartheta_x \cos \varphi_x, \sin \vartheta_x \sin \varphi_x, \cos \vartheta_x)$ ,  $Z$  becomes

$$\begin{aligned} Z &= \int \prod_{x_1} d\varphi_{x_1} \prod_{x_2} d\vartheta_{x_2} e^{-(S[\varphi, \vartheta] - \sum_x \ln \sin \vartheta_x)} \\ &= \int \prod_{x_1} d\varphi_{x_1} \prod_{x_2} d\vartheta_{x_2} e^{-S_{\text{eff}}[\varphi, \vartheta]}, \end{aligned} \quad (37)$$

where

$$\begin{aligned} S_{\text{eff}}[\varphi, \vartheta] &= 2\beta V - \beta \sum_{x,\nu} (\sin \vartheta_{x+\hat{\nu}} \sin \vartheta_x \cos(\varphi_{x+\hat{\nu}} - \varphi_x \\ &\quad - i\mu a \delta_{\nu,0}) + \cos \vartheta_{x+\hat{\nu}} \cos \vartheta_x) - \sum_x \ln \sin \vartheta_x. \end{aligned} \quad (38)$$

From this expression, one can deduce the drifts:

$$\begin{aligned} -\frac{\delta S_{\text{eff}}}{\delta \varphi_x} &= \beta \sum_{\nu} (\sin \vartheta_x [\sin \vartheta_{x+\hat{\nu}} \sin(\varphi_{x+\hat{\nu}} - \varphi_x \\ &\quad - i\mu a \delta_{\nu,0}) - \sin \vartheta_{x-\hat{\nu}} \sin(\varphi_x - \varphi_{x-\hat{\nu}} - i\mu a \delta_{\nu,0})]), \end{aligned} \quad (39)$$

and

$$\begin{aligned} -\frac{\delta S_{\text{eff}}}{\delta \vartheta_x} &= \beta \sum_{\nu} (\cos \vartheta_x [\sin \vartheta_{x+\hat{\nu}} \cos(\varphi_{x+\hat{\nu}} - \varphi_x \\ &\quad - i\mu a \delta_{\nu,0}) + \sin \vartheta_{x-\hat{\nu}} \cos(\varphi_x - \varphi_{x-\hat{\nu}} - i\mu a \delta_{\nu,0})] \\ &\quad - \sin \vartheta_x (\cos \vartheta_{x+\hat{\nu}} + \cos \vartheta_{x-\hat{\nu}})) - \frac{1}{\tan \vartheta_x}. \end{aligned} \quad (40)$$

Then, the discretized complex Langevin steps are

$$\varphi_x(n+1) = \varphi_x(n) + \varepsilon_n K_x^{(\varphi)}(n) + \sqrt{\varepsilon_n} \eta_x^{(\varphi)}(n), \quad (41)$$

$$\vartheta_x(n+1) = \vartheta_x(n) + \varepsilon_n K_x^{(\vartheta)}(n) + \sqrt{\varepsilon_n} \eta_x^{(\vartheta)}(n), \quad (42)$$

where  $K_x^{(\varphi)} = -\delta S_{\text{eff}}/\delta \varphi_x$  and  $K_x^{(\vartheta)} = -\delta S_{\text{eff}}/\delta \vartheta_x$ . In these equations,  $\eta_x^{(\varphi)}$  and  $\eta_x^{(\vartheta)}$  are real, Gaussian noises, and the finite step size  $\varepsilon_n$  is determined adaptively, so it also depends on  $n$ . As one can observe, the force

$-\delta S_{\text{eff}}[\varphi, \vartheta]/\delta \vartheta_x$  is singular because of the  $1/\tan \vartheta_x$  term. In order to avoid overflow during the simulations due to the singular forces, one can do a constrained simulation and truncate the configuration space to avoid the values of  $\vartheta_x$

$$\vartheta_x(n+1) := \begin{cases} 2\vartheta_{\text{LIM}} - \vartheta_x(n+1), & \text{if } \vartheta_x(n+1) < \vartheta_{\text{LIM}}, \\ \vartheta_x(n+1), & \text{if } \vartheta_{\text{LIM}} < \vartheta_x(n+1) < \pi - \vartheta_{\text{LIM}}, \\ 2(\pi - \vartheta_{\text{LIM}}) - \vartheta_x(n+1), & \text{if } \pi - \vartheta_{\text{LIM}} < \vartheta_x(n+1). \end{cases} \quad (43)$$

The threshold value for  $\vartheta$  was defined with a parameter  $\vartheta_{\text{LIM}}$ . Results shown later support the expectation that if  $\varepsilon$  and  $\vartheta_{\text{LIM}}$  are small enough, then the results are independent of their values.

### B. Integration in group space

Using ideas from the CL equation on SU(N) gauge groups [17], one can write a specific integration scheme that uses Cartesian coordinates but takes the constraint into account: the Euler discretization (which we call below the exponentialized Euler-Maruyama discretization). For the O(3) group elements  $O_x$ , this is the following:

$$O_x(n+1) = R_x(\varepsilon)O_x(n). \quad (44)$$

Since all  $\phi_x$  can be written with some  $\phi_0$  constant unit vector and with an  $O_x$  rotation matrix as  $\phi_x = O_x\phi_0$ , the above time evolution can turn into the time evolution of the original  $\phi_x$  variables.

The  $R_x(\varepsilon)$  in Eq. (44) can be written in different ways. It can be e.g.

$$R_x^{(1)}(\varepsilon) = \exp\left(\sum_a t_a(\varepsilon K_{ax} + \sqrt{\varepsilon}\eta_{ax})\right) \quad (45)$$

or

$$R_x^{(2)}(\varepsilon) = \prod_{a \in (1,2,3)} \exp(t_a(\varepsilon K_{ax} + \sqrt{\varepsilon}\eta_{ax})) \quad (46)$$

or

$$R_x^{(3)}(\varepsilon) = \prod_{a \in (1,2,3)} \exp(t_a \varepsilon K_{ax}) \exp(t_a \sqrt{\varepsilon} \eta_{ax}). \quad (47)$$

Since  $e^{A+B} \neq e^A e^B$ , when  $[A, B] \neq 0$ , these are not equivalent to each other at finite  $\varepsilon$ , but the difference in the simulation results is not detectable at the numerical precision and parameter set we used. Due to this, during our simulations, we used  $R_x^{(2)}$ , which is not the computationally cheapest version [it is  $R_x^{(1)}(\varepsilon)$ ] but the cheapest version that evolves the system in each direction in the tangent space individually one after another. In the

near zero and  $\pi$ . If the step size is not small enough, this is needed in order to get stable simulation runs.

We achieved this by reflecting the trajectories in the following way:

expressions above, the  $t_a$ s are the three generators of O(3) in the three-dimensional representation. The drift  $K_{ax}$  is

$$\begin{aligned} K_{ax} &= -D_{ax}S[O] = -\partial_a S[e^{at_a} O_x] \Big|_{a=0} \\ &= \beta(\phi_0^T O_{x+0}^T e^{i\mu a t_3} t_a O_x \phi_0 - \phi_0^T O_x^T t_a e^{i\mu a t_3} O_{x-\hat{0}} \phi_0 \\ &\quad + \phi_0^T O_{x+1}^T t_a O_x \phi_0 - \phi_0^T O_x^T t_a O_{x-\hat{1}} \phi_0) \\ &= \beta(\phi_{x+0}^T e^{i\mu a t_3} t_a \phi_x - \phi_x^T e^{i\mu a t_3} \phi_{x-\hat{0}} \\ &\quad + \phi_{x+1}^T t_a \phi_x - \phi_x^T t_a \phi_{x-\hat{1}}). \end{aligned} \quad (48)$$

Here,  $\phi^T$  and  $O^T$  denote the transpose of  $\phi$  and  $O$ , and  $\eta_{ax}$  is the usual Gaussian noise. The time evolution determined by Eq. (44) then can be written with  $\phi_x$  as

$$\phi_x^{(n+1)} = \prod_{a \in (1,2,3)} \exp[(\varepsilon_n K_{ax} + \sqrt{\varepsilon_n} \eta_{ax}) t_a] \phi_x^{(n)}. \quad (49)$$

In particular, we performed the updates by varying the order of the three matrix multiplications randomly. Higher order integrations, like Runge-Kutta [58], are also possible,

$$\begin{aligned} \phi'_x &= \exp[(\varepsilon_n K_{ax}[\phi^{(n)}] + \sqrt{\varepsilon_n} \eta_{ax}) t_a] \phi_x^{(n)} \\ \phi_x^{(n+1)} &= \exp\left[\frac{\varepsilon_n}{2} \left(1 + \frac{C_A \varepsilon_n}{6}\right) (K_{ax}[\phi^{(n)}] + K_{ax}[\phi']) t_a \right. \\ &\quad \left. + \sqrt{\varepsilon_n} \eta_{ax} t_a\right] \phi_x^{(n)}, \end{aligned} \quad (50)$$

where  $C_A = 2$  is the Casimir invariant for the three-dimensional representation of O(3).

### C. Direct method to include the constraint in Cartesian coordinates

Using Cartesian coordinates, the constraining force can be considered by using a term arising from  $-\sum_x \ln \delta(\phi_x^2 - 1)$ ; see Eq. (36). This is also singular, so one can attempt to approximate the Dirac  $\delta$  with a sharp Gaussian,  $(\sqrt{2\pi}b)^{-1} e^{-x^2/(2b^2)} \rightarrow \delta(x)$ , as  $b \rightarrow 0$ . The force is then

TABLE II. The set of parameters for the simulations with the complex Langevin algorithm using spherical coordinates. In the table,  $\epsilon$  refers to the largest step size in the runs, and  $\theta_{\text{LIM}}$  is the minimum distance between any  $\vartheta_x$  and 0 ( $\pi$ ). (See the text for the definition of  $\theta_{\text{LIM}}$ .) The  $\theta_{\text{LIM}}$  dependence of the results was analyzed on  $56 \times 14$  lattices at  $\mu/T = 1$  by using  $\theta_{\text{LIM}} = 10^{-3}, 10^{-5}, 10^{-7}, 10^{-8}, 10^{-11}$ . These are not listed in the table.

Method	$N_x \times N_t$	$\mu/T$	$\beta$	$\epsilon$	$\theta_{\text{LIM}}$	No. of trajectories
CL with spherical coordinates	$56 \times 14$	0, 0.25, 0.5, 1, 2, 3, 4	0.9...1.8	$5, 2, 1 \times 10^{-4}, 10^{-5}$	$10^{-5}$	1000...1600
	$72 \times 18$	1	0.9...1.8	$5, 1 \times 10^{-4}$	$10^{-5}$	

$$\begin{aligned}
K_x &= -\frac{\delta}{\delta\phi_x} \left( S[\phi] - \sum_y \ln \delta(\phi_y^2 - 1) \right) \\
&= \beta(\phi_{x+\hat{0}} e^{i\mu a t_3} + e^{i\mu a t_3} \phi_{x-\hat{0}} + \phi_{x+\hat{1}} + \phi_{x-\hat{1}}) \\
&\quad - \frac{2}{b^2} (\phi_x^2 - 1) \phi_x,
\end{aligned} \tag{51}$$

where the last term helps to keep the length of  $\phi_x$  near 1. Then, the fields evolve according to

$$\phi_{x,i}^{(n+1)} = \phi_{x,i}^{(n)} + \epsilon_n K_{x,i}^{(n)} + \sqrt{\epsilon_n} \eta_{x,i}^{(n)}, \tag{52}$$

where again  $\epsilon_n$  is the finite step size determined adaptively and the noise is Gaussian with  $\sqrt{2}$  width. Note that using this time evolution, the  $\sum_i \phi_{x,i} = 1$  condition is no longer true during the simulations, but the force can push the field into this direction. We refer to this time evolution as standard Euler-Maruyama discretization with Dirac  $\delta$  in the following.

In the next section, we discuss the results obtained using the various algorithms.

## VII. COMPARISON OF RESULTS

As was discussed in the Introduction, the complex Langevin algorithm may provide a feasible way to study sign problems in different models but may converge to wrong results, which would lower the reliability of the simulation when there are no alternative results in the problematic parameter range. The conventional reasoning of explaining the wrong results is that the justification of the complex Langevin [14–16] is not correct in that parameter range, because some observables develop long tailed distributions. (For details, we refer to Ref. [15].) However, in some models (e.g. in a random-matrix model [59,60]), it was observed that using different variables in describing the model can help to eliminate this problem. This can imply that the failure of the algorithm is not because physics has changed but because of some unknown technical details. One can think of systematic errors originating from e.g. the step size to zero limit, floating point roundoff errors, not taking the continuum limit, or wrong sampling of the configuration space, etc. In the following, we analyze some of these aspects in the case of the  $1+1$ -dimensional  $O(3)$  model. The results are compared to the worm results, which are referred to as the correct ones in the text.

Although we used the adaptive step size in all our simulations, this cannot replace the completion of the  $\epsilon \rightarrow 0$  extrapolation of the results. (Of course, its effect depends on the used numerical precision and the algorithm under study as well as other subtle circumstances. For example, as we will see, the simulations with spherical coordinates do not depend in a detectable way on  $\epsilon$  at the simulations with the used set of parameters for e.g. the action variable.) In the following, first, we discuss the results for the action  $S$  and the trace anomaly  $\theta$  and then for the density  $n/m$ .

### A. Action and trace anomaly

Using spherical coordinates to parametrize the model, we ran simulations at  $56 \times 14$  lattices, at several chemical potentials (see Table II). We analyzed the step size and  $\theta_{\text{LIM}}$

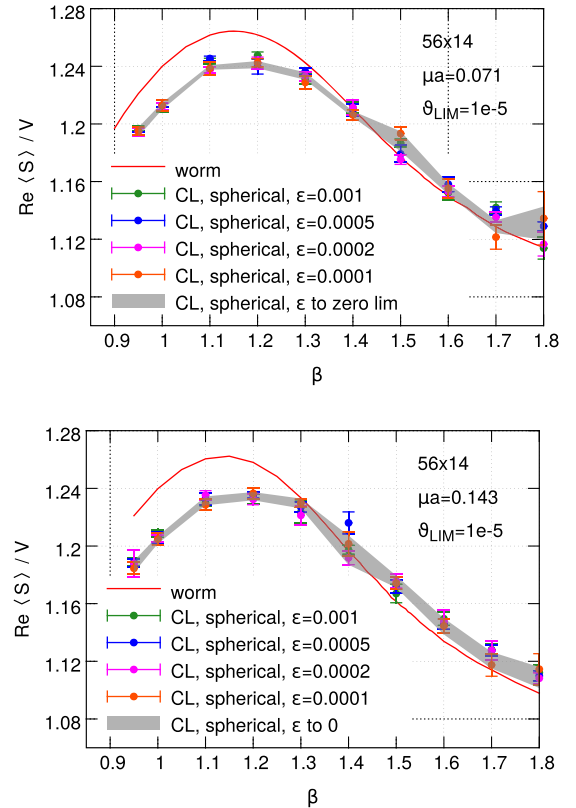


FIG. 14. CL results for  $\text{Re}\langle S \rangle / V$  with spherical coordinates. Top:  $\mu a = 0.071$  ( $\mu/T = 1$ ), bottom:  $\mu a = 0.143$  ( $\mu/T = 2$ ).



TABLE III. The set of parameters for the simulations with the complex Langevin algorithm using the group integration approach (exponentialized Euler-Maruyama method). In the table,  $\epsilon$  refers to the largest step size during the trajectories.

Method	$N_x \times N_t$	$\mu/T$	$\beta$	$\epsilon$	No. of trajectories
CL with group integration (exp. E-M method)	$32 \times 8$	0.25	0.9...1.8	$(5, 2, 1) \times 10^{-4}$	4500...5500
	$40 \times 10$	0.25, 0.5, 1	0.9...1.8	$(10, 5, 2, 1, 0.5) \times 10^{-4}$	3000...5000
	$56 \times 14$	0, 0.25, 0.5, 1, 2, 3, 4	0.9...1.8	$(10, 5, 2, 1) \times 10^{-4}$	3000...5500
	$64 \times 16$	0.25, 0.5, 1, 2, 3, 4	1.1...1.8	$(10, 5, 2, 1, 0.5) \times 10^{-4}$	2000...3500
	$72 \times 18$	0.25, 0.5, 1, 2, 3, 4	1.1...1.8	$(10, 5, 2, 1) \times 10^{-4}$	2500...5500
	$80 \times 20$	0.25, 0.5, 1, 2, 3, 4	1.1...1.8	$(10, 5, 2, 1) \times 10^{-4}$	3000...5500
	$120 \times 30$	0.5, 3	1.1...1.8	$(1, 0.5, 0.2, 0.1) \times 10^{-4}$	1000...2000
	$200 \times 50$	0.5, 3	1.2...1.8	$(1, 0.8, 0.5, 0.2, 0.1) \times 10^{-4}$	800...1200

dependence of the results. Although we found that there was no detectable step size dependence of the results, we extrapolated to zero step size at  $\mu a = 0.071$  and  $0.143$  using four step sizes. We also found that the results do not depend on  $\vartheta_{\text{LIM}}$  if it was chosen quite small. To test this at  $\mu a = 0.071$ , we used  $\vartheta_{\text{LIM}} = 10^{-3}, 10^{-5}, 10^{-7}, 10^{-8}, 10^{-11}$  and  $\epsilon = 0.0005$ . We found that at  $\mu a = 0$ , spherical CL results agree completely with the correct results. At  $56 \times 14, \mu a = 0.018$  they deviate below  $\beta \sim 1.3$ . This clearly shows that the wrong convergence property was

not the consequence of the severeness of the sign problem, because it was very mild at  $\mu a = 0.018$  (see Fig. 2, right panel). We note that CL results obtained with spherical coordinates seem to slightly deviate from the correct ones in the high  $\beta$  region as well, but these differences are not significant statistically (they are within 2 sigma). At high  $\mu/T$ , however, the deviations are significant, so at e.g.  $\mu a = 0.286, 56 \times 14$ , results for the action are wrong at all  $\beta$  values. Since this discretization was a bit problematic due to the singularity in the force and deviations at high  $\mu/T$  seemed discouraging, we did not test so carefully its continuum behavior or possible improvements. However, we mention that using  $72 \times 18$  lattices did not show any improvement at  $\mu/T = 1$ . We show some results in Fig. 14.

We have investigated the group integration approach (Sec. VIB) more carefully. First, we compared simulation results of using  $R^{(1)}, R^{(2)}$ , or  $R^{(3)}$  at  $56 \times 14$  lattices at  $\mu a = 0.071, \epsilon = 0.0005$  using 2000 Langevin trajectories at several  $\beta$  in the range  $0.9...1.8$ . We found complete agreement using these parameters. Then, we used  $R^{(2)}$  during our further simulations with the exponentialized Euler-Maruyama (E-M) discretization (abbreviated in the following and in the figures as *exp. E-M*). We carried out simulations on several lattice sizes and chemical potentials in the  $\beta$  range  $0.9...1.8$ . The parameters for

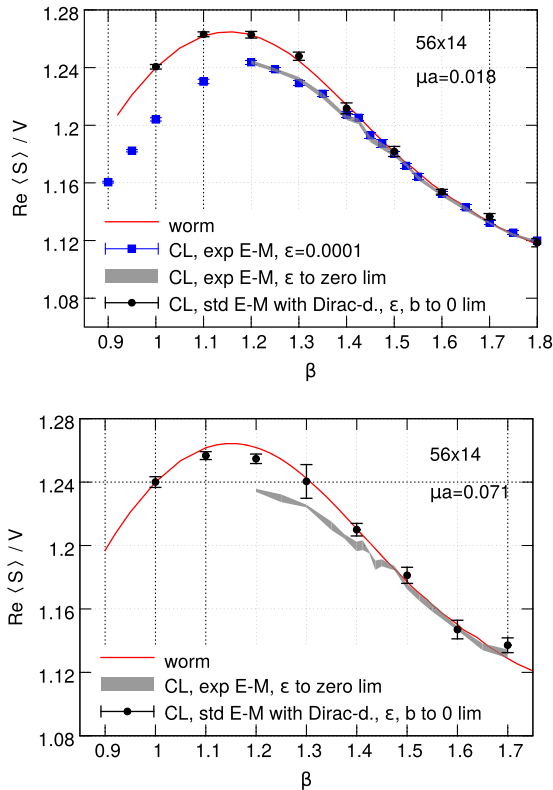


FIG. 15. The comparison of  $\text{Re}\langle S \rangle/V$  results for different algorithms: worm, CL with exponentialized E-M, and CL with standard E-M discretization with Dirac  $\delta$ . Top:  $\mu a = 0.018$  ( $\mu/T = 0.25$ ), bottom:  $\mu a = 0.071$  ( $\mu/T = 1$ ). Note that the exp. E-M method is wrong at low  $\beta$  even when the sign problem is still very mild.

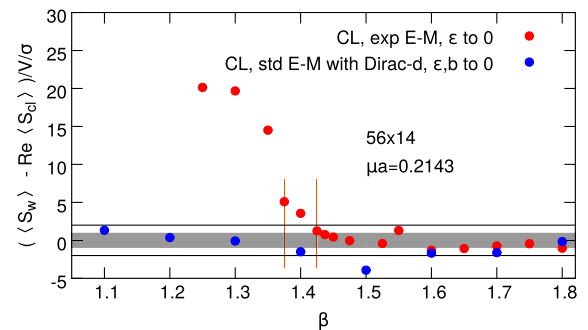


FIG. 16.  $(\langle S_w \rangle - \text{Re}\langle S_{\text{cl}} \rangle)/V/\sigma$  at  $\mu a = 0.143$  ( $\mu/T = 2$ ) (top) and at  $\mu a = 0.2143$  ( $\mu/T = 3$ ) (bottom). The gray band shows the 1 sigma interval, and the black lines show the 2 sigma interval. The orange lines show the beta threshold range, below which the exp. E-M complex Langevin fails.

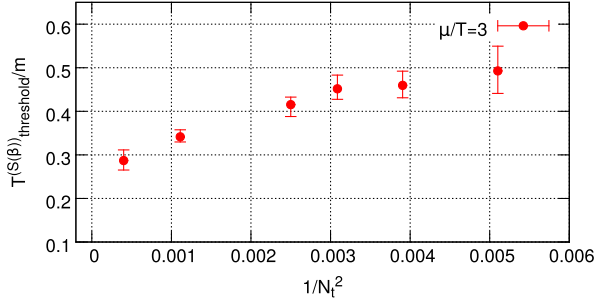


FIG. 17. The temperature  $T_{\text{threshold}}^{(S(\beta))}/m$  as a function of  $1/N_t^2$  at  $\mu/T = 3$ . Below  $T_{\text{threshold}}^{(S(\beta))}/m$ , the action densities obtained by the exp. E-M discretization start to become wrong. Using larger lattices, this temperature becomes lower, but the continuum extrapolation of this quantity cannot be made without assuming some functional dependence of it as a function of  $1/N_t$ , which would hinder drawing the conclusion.

these simulations can be found in Table III. Several initial step sizes were used during these simulations, and we extrapolated these to zero step size. (However, we did not find significant step size dependence of the results obtained with step sizes  $10^{-4}$  and smaller.)

We came to similar conclusions as with spherical coordinates: at  $\mu/T = 0$  complete agreement was found, then at nonzero  $\mu/T$ —even at  $\mu/T = 0.25$ —, a discrepancy in the low  $\beta$  region. We note, however, that the exp. E-M results do not deviate from the correct results at higher  $\beta$  values. In order to quantify the deviations and investigate their continuum behavior, we first calculated  $(\langle S_w \rangle - \text{Re}\langle S_{\text{cl}} \rangle)/V/\sigma$ , where  $\langle S_w \rangle$  is the worm result,  $\langle S_{\text{cl}} \rangle$  is the exp. E-M result in the  $\varepsilon \rightarrow 0$  limit,  $V = N_x \times N_t$  is the lattice volume, and  $\sigma = \sqrt{\Delta_w^2 + \Delta_{\text{cl}}^2}$ , where  $\Delta_w$ ,  $\Delta_{\text{cl}}$  are the full errors of the worm and CL simulations, respectively. ( $\Delta_w$  is just the statistical error, but  $\Delta_{\text{cl}}$  contains systematic errors because of the step size extrapolation.) After that, we determined a  $\beta$  region at each lattice size and  $\mu a$  parameter, when  $(\langle S_w \rangle - \text{Re}\langle S_{\text{cl}} \rangle)/V/\sigma$  started to be above 2. This

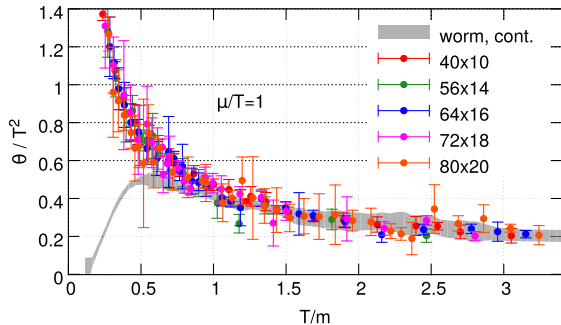


FIG. 18. The trace anomaly determined with the complex Langevin algorithm using the exp. E-M discretization and the worm algorithm. Although we do not plot the continuum limit of the complex Langevin results here, the results suggest that there is no improvement toward the continuum.

definition of the  $\beta$  range (the latest point under 2 sigma and two successive points above 2 sigma) is a bit ambiguous in the sense that it depends on the statistics, but we did not find significant differences between the results of some smaller statistics runs and the long runs. Figure 16 shows an example of the determination of this  $\beta$  range.

Then, the middle of this range with errors to cover the whole range was used to calculate  $T_{\text{threshold}}^{(S(\beta))}/m$ , the temperature below which CL converges to wrong results at the lattice under study. We show how these temperatures depend on the temporal lattice size at  $\mu/T = 3$  in Fig. 17. In that figure, one can see that these threshold temperatures become lower as  $N_t$  increases, but we do not know the scaling of this quantity as a function of the lattice spacing or  $N_t$ ; thus, we cannot extrapolate to the continuum without assumptions. In order to avoid these assumptions, we have calculated the continuum limit of  $(\langle S_w(T, \mu/T) \rangle - \text{Re}\langle S_{\text{cl}}(T, \mu/T) \rangle)/V$  and then determined the *continuum* threshold temperature  $T_{\text{threshold}}^{(S)}/m$  below which the continuum results deviate from zero. The results for these

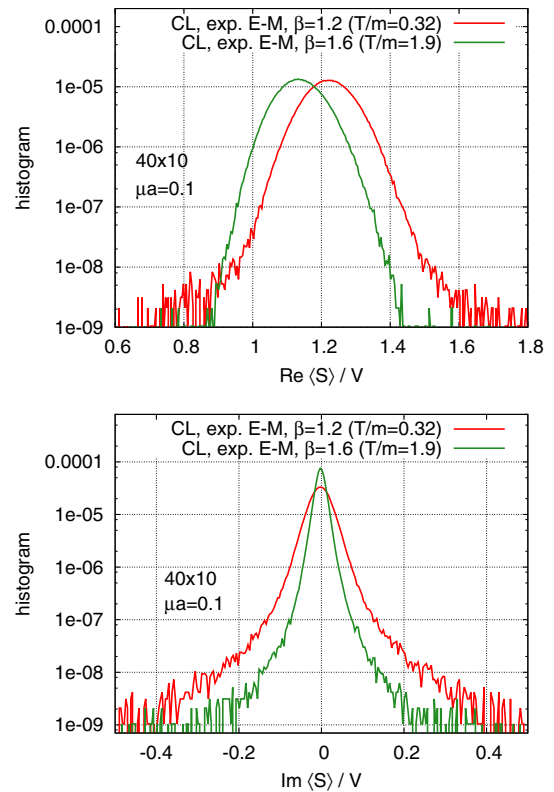


FIG. 19. The histogram of  $\text{Re}\langle S \rangle/V$  (top) and  $\text{Im}\langle S \rangle/V$  (bottom) at  $\beta = 1.2$  and  $\beta = 1.6$ . These  $\beta$  values correspond to  $T/m = 0.32$  and  $T/m = 1.9$ ; the former is such a temperature where the complex Langevin exp. E-M discretization develops wrong results for the action, and the latter is a temperature where it is correct. One can see that the histograms indeed show the usual concomitant sign of wrong results, that is, the longer tail of the distribution of the observables.

TABLE IV. The set of parameters for the simulations with the complex Langevin algorithm using the standard Euler-Maruyama discretization with directly including the constraint by approximating the Dirac  $\delta$ .  $\epsilon$  refers to the largest step size during the trajectories, and  $b$  is the width of the Gaussian approximating the Dirac  $\delta$  (see Sec. VIC).

Method	$N_x \times N_t$	$\mu/T$	$\beta$	$\epsilon$	$b$	No. of trajectories
CL with direct constraint	$40 \times 10$	0.25, 0.5, 1	0.9...1.8	$(5, 2, 1, 0.5, 0.2) \times 10^{-4}$	0.01...0.06	2000...3000
(standard E-M method	$56 \times 14$	0, 0.25, 0.5, 1, 2, 3, 4	0.9...1.8	$(5, 2, 1, 0.5) \times 10^{-4}$	0.015...0.05	1800...3000
with Dirac $\delta$ )	$64 \times 16$	0.25, 1, 2, 3	1.0...1.8	$(1, 0.8, 0.5) \times 10^{-4}$	0.02...0.05	2000...3500
	$72 \times 18$	1, 2, 3	1.0...1.8	$(1, 0.8, 0.5) \times 10^{-4}$	0.018...0.038	1600...2500
	$80 \times 20$	1, 2	1.1...1.8	$(2, 1, 0.8, 0.5) \times 10^{-4}$	0.02...0.04	1600...2500

temperatures (with further relevant temperature ranges discussed in the text) are shown in Fig. 25.

Of course, one can discuss the deviations of the correct and the CL action densities in terms of a more standard physical quantity which has a continuum limit: the trace anomaly. So, we have also used the trace anomaly and investigated what happens with the continuum limit of the “wrong” complex Langevin results (Fig. 18) and obtained similar threshold temperature values.

We also made some runs to test this approach against a change in computer numerical precision and the order of integration; that is, we used float (32-bit) and long double

(80-bit) precision and found that, although float and double differ from each other (float is wrong at all  $\beta$ ), double and long double are almost the same at the parameters used to clarify this ( $56 \times 14$ ,  $\mu/T = 0.071, 0.143$ ,  $\epsilon = 0.0005$ ,  $\sim 1800$  Langevin trajectories). We also tested exponentialized Runge-Kutta integration at double precision, but results did not improve.

We have also checked the shape of the distributions, and these are shown in Fig. 19. One can see that the distributions are narrower in the high temperature range.

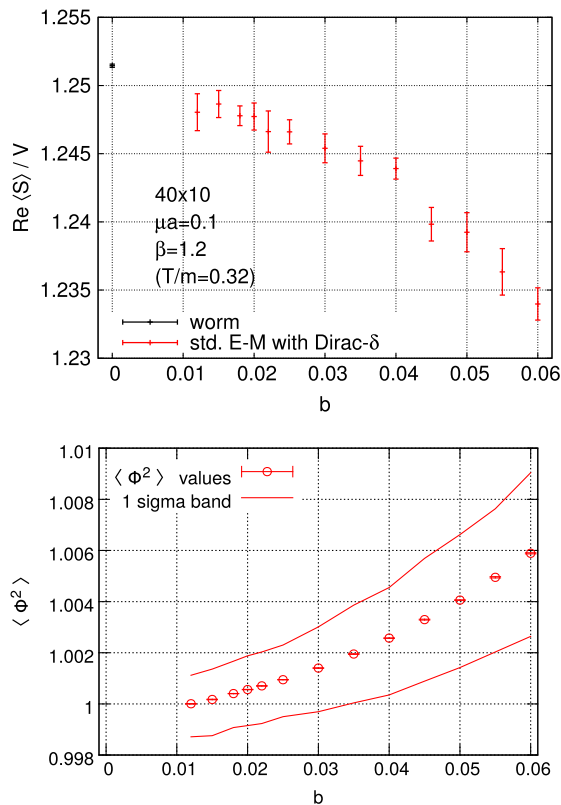


FIG. 20. Top: The figure shows how taking the  $b \rightarrow 0$  limit can help to obtain the correct results for the action with the standard E-M implementation (the starting  $\epsilon$  step size was  $5 \times 10^{-5}$  during the simulations to obtain the data points of the figure). Bottom: The effect of the reduction of the  $b$  parameter on the average length of  $\phi$  vectors. The lines show the one sigma interval.

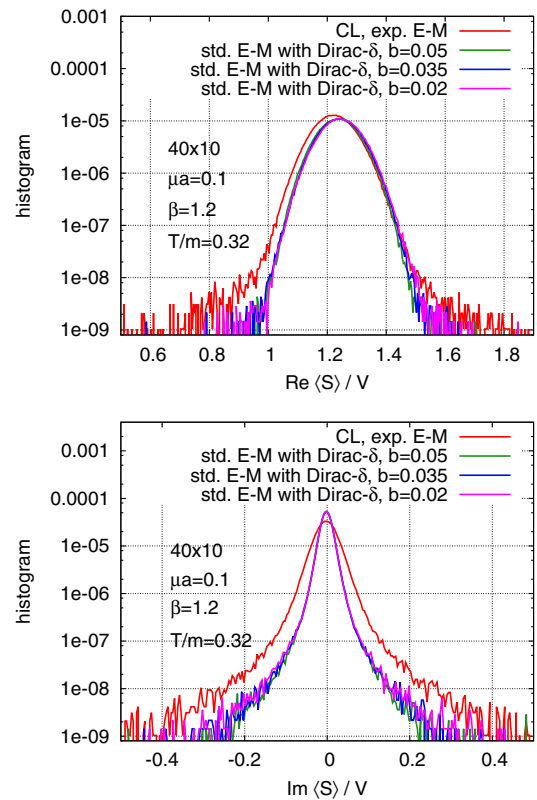


FIG. 21. Histograms of  $\text{Re}(S)/V$  (top) and  $\text{Im}(S)/V$  (bottom) at  $\beta = 1.2$  obtained with the exp. E-M implementation and with the standard E-M implementation with Dirac  $\delta$ . The std. E-M implementation develops quite similar distributions for the real part of the action, while narrower for the imaginary part. The shapes of the histograms do not change so much as one decreases  $b$  (the width of the Gaussian that approximates the Dirac  $\delta$ ), although results get closer to the correct ones.

Finally, let us discuss the results obtained using the standard Euler-Maruyama discretization with a Dirac  $\delta$  approximated with a Gaussian. We simulated with this algorithm at the parameters of Table IV. We used several initial step sizes and at each step size several  $b$  values ( $0.01 < b < 0.06$ ). Using double precision, we found that at a given step size, below a low  $b$  value, simulations became unstable, so there we used long double precision to set lower  $b$  values. At each step size, we extrapolated to  $b \rightarrow 0$ , then used these results to extrapolate in  $\varepsilon$ . As mentioned in Sec. VIC, this algorithm did not keep the  $\sum_i \phi_i^2 = 1$  constraint rigorously during the simulation. To characterize it quantitatively, we measured the length of the

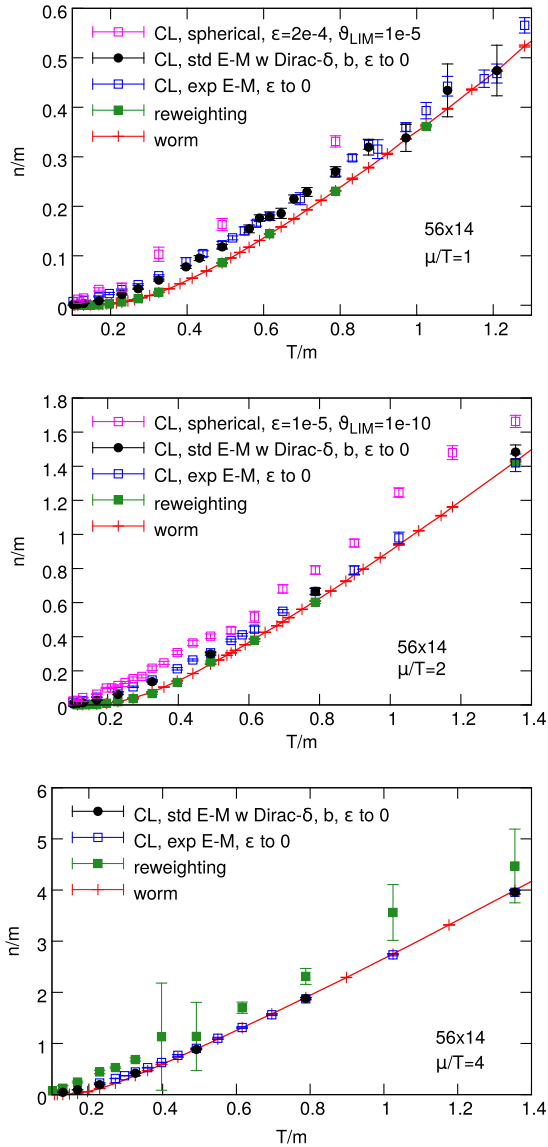


FIG. 22. The density at low temperature measured on  $56 \times 14$  lattices, at  $\mu/T = 1$  (top), at  $\mu/T = 2$  (middle), and at  $\mu/T = 4$  (bottom). At  $\mu/T = 4$ , the reweighting results become unreliable, while CL remains correct if the temperature is greater than  $0.37 \dots 0.5$ .

$\phi$  vectors over the lattice and found it is typically a bit larger than 1, but with smaller  $b$  and  $\varepsilon$  values, it can be reduced (see the lower panel of Fig. 20).

We found that after taking both the  $b$  to zero (Fig. 20, top) and  $\varepsilon$  to the zero limit, the results obtained by this method agree well with the correct results (Figs. 15 and 16). These results are interesting because when we accomplish the  $b$  to zero and  $\varepsilon$  to zero limits the used data may have distributions also with some nonexponential decay; see Fig. 21. The obtained histograms were compared to those of the exp. E-M discretization, and one can see that the decay of the standard E-M discretization with Dirac  $\delta$  seems to be sharper (Fig. 21). For completeness, we mention here that the errors coming from the two extrapolations became significantly larger, especially at small  $\beta\delta$  as the chemical potential increased.

## B. Density

In the present subsection, we review the results for the density [Eq. (32)] obtained by the different algorithms. For the worm results, we used the configurations, that were listed in Sec. V D2, and for the CL results we used the data of Tables II, III, and IV. For reweighting results, we used the cluster algorithm to simulate at  $\mu = 0$  and used  $3 \times 10^6$  updating steps.

We found that reweighting results agree well with the worm results below  $\mu a \sim 0.16$  on  $56 \times 14$  lattices, that is below  $\mu/T \sim 2.2$  (see Fig. 22). At higher  $\mu a$  values, the results have large error bars, and (apparent) deviations from correct results occur. This coincides with the fact that the sign problem became severe at these lattices around  $\mu a \sim 0.15$  (Fig. 4).

Regarding the different CL implementations, we found that both the exp. E-M integration and the standard E-M with Dirac  $\delta$  produced wrong results at low temperature (Fig. 22).

A threshold temperature ( $T_{\text{threshold}}^{(n/m)}/m$ ) can be defined similarly as we did in the case of the action: this is the

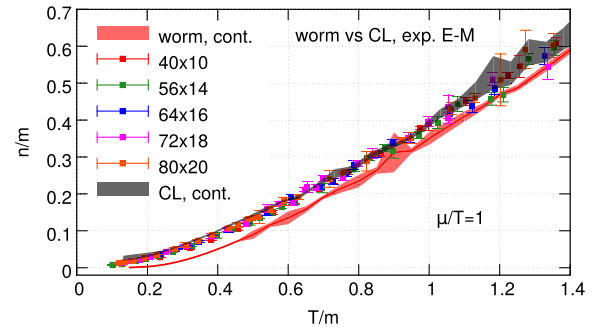


FIG. 23. Comparison of the continuum results for the density obtained by the worm and the CL, exp. E-M discretization at  $\mu/T = 1$ . The figure shows that the continuum extrapolation from the complex Langevin results differs from the worm continuum result.



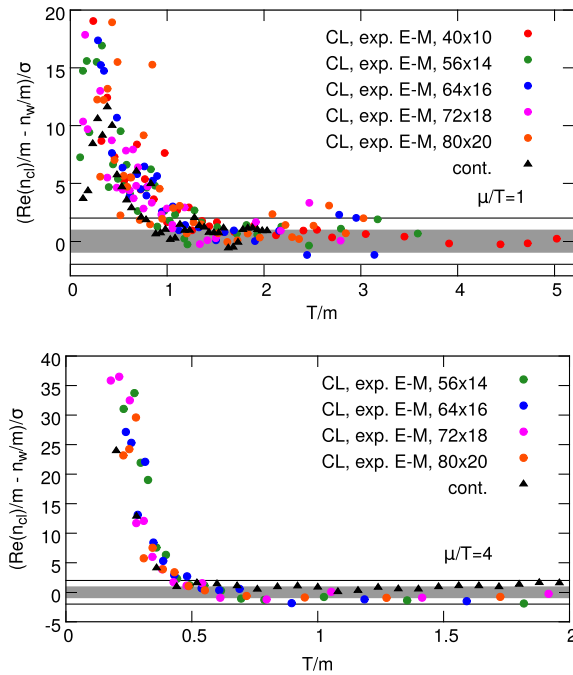


FIG. 24. The density differences of different complex Langevin results ( $n_{cl}/m$ ) and the worm ( $n_w/m$ ) divided by  $\sigma = \sqrt{\Delta_{cl}^2 + \Delta_w^2}$  at  $\mu/T = 1$  (top) and at  $\mu/T = 4$  (bottom). The figure shows that the threshold temperature does not reduce toward the continuum. The black triangles are calculated from the continuum CL density and the continuum worm density.

temperature below which continuum CL density results deviate from the continuum worm density results with 2 sigma significance. A comparison of the continuum results from the exp. E-M. CL and the worm algorithm can be seen

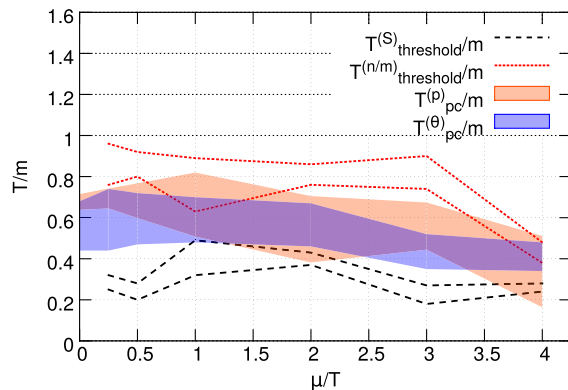


FIG. 25. The figure shows the temperatures  $T_{\text{threshold}}/m$  determined from the action differences and the density differences discussed in the text. Below  $T_{\text{threshold}}^{(S)}/m$ , the exp. E-M discretization develops wrong results for the action. The std. E-M discretization with Dirac  $\delta$  produces correct results for the action at all temperatures. However, both implementations give wrong results for the density below  $T_{\text{threshold}}^{(n/m)}/m$ . The blue and red solid bands show the pseudocritical temperatures determined from the inflection point of the pressure and from the maximum of the trace anomaly,  $T_{pc}^{(p)}/m$  and  $T_{pc}^{(\theta)}/m$ , respectively.

in Fig. 23. We note that in the low temperature region, the continuum extrapolation takes the CL results even further from the worm continuum results. By increasing the chemical potential, one can observe that  $T_{\text{threshold}}^{(n/m)}/m$  is approximately constant, then gets smaller (see Figs. 24 and 25). We note that the  $T_{\text{threshold}}^{(n/m)}/m$  values are approximately the same for the exp. E-M. discretization and for the standard E-M. discretization with Dirac  $\delta$ .

The spherical CL implementation also developed wrong results, but its threshold temperature seemed to be larger than that of the others. However, we note that in the case of the spherical formulation, we did not determine so carefully the threshold temperature.

The distributions of the real and imaginary parts of  $\partial S/\partial(\mu a)$  are again narrower, when different complex Langevin algorithms produce correct results. Comparing the distributions for this quantity of the exp. E-M and std. E-M with Dirac- $\delta$  implementations, we see that, although the latter is narrower, the results do not converge to the correct results, as they did in the case of the action.

## VIII. CONCLUSION

In the present paper, we studied the sign problem in the  $O(3)$  model. We used reweighting in order to investigate the severeness of the sign and overlap problems. We described a dual formalism and based on that a worm algorithm, which completely solves the sign problem. Using this, we have calculated the pressure, the trace anomaly, and the density at finite temperature. At zero and low temperatures, we reproduced the Silver Blaze phenomenon. We then analyzed the correctness of the complex Langevin as approaching the continuum limit. The failure of the complex Langevin in certain parameter ranges is argued to be the consequence of the developing long tailed distributions. However, it is an interesting question whether the wrong convergence property happens at a specific  $\beta$  value—as was found in HDQCD [20]—or at a given temperature. In the former case, continuum extrapolation could enable one to study the full phase diagram of the given model. According to recent results [7], the breakdown seems to prevent the exploration of the confined region in QCD. However, those simulations used only  $N_t = 4, 6, 8$  lattices. In the present paper, our main goal was to investigate whether taking the continuum limit in the  $1 + 1$ -dimensional  $O(3)$  model can help to overcome the wrong convergence property of the complex Langevin.

For this purpose, we have used three different approaches: describing the model with spherical coordinates, using the generators and integrating the CL equations in the group space (exp. E-M), and considering the  $\sum_i \phi_i^2 = 1$  constraint with a term containing the logarithm of a Dirac  $\delta$  in the action (standard Euler-Maruyama discretization with Dirac  $\delta$ ). We approximated the Dirac  $\delta$  by a Gaussian having a width  $\sim 1/b$ . We analyzed three observables: the action, the trace anomaly, and the density. Regarding the

action, we found that both the spherical coordinates and the exp. E-M discretization gave wrong results at the low  $\beta$  (low temperature) range. For the exp. E-M discretization, we analyzed whether taking the continuum limit can help to improve the results and found that, although the accessible temperature range became larger, one cannot explore the full low temperature region. However, we established that the so-called standard Euler-Maruyama discretization with the Dirac- $\delta$  method can give correct results for the action density after taking both the  $b$  to zero and  $\varepsilon$  to zero limits. We found that for the density observable, the situation is different: in this case, the latter way of integrating the CL equations gave wrong results also at low  $T/m$ , and the results did not seem to improve by taking larger and finer lattices.

### ACKNOWLEDGMENTS

The authors thank Dénes Sexty, Erhard Seiler, and Falk Bruckmann for useful comments on the manuscript. The work was supported by the Lendület program of the HAS (Program No. LP2012-44/2012) and by Grant No. OTKA-NF-104034 from OTKA.

### APPENDIX: WORM ALGORITHM UPDATING STEPS

In this Appendix, we describe our updating steps of the worm algorithm in detail. As in the beginning of Sec. V (before Sec. VD), we consider the general case  $O(N)$  in  $d + 1$  dimensions.

#### 1. Updating steps of the worm algorithm

When  $\mu$  is zero, we consider three simple updating steps:

- (1) Move the head of the worm from the position  $u$  along a link  $l$  to the new position  $u'$ . To keep the constraints, by this, we increase or decrease the corresponding link variable  $m = m_i^{(l)}$  by 1. (Here,  $i$  is the index of the worm.) The corresponding 2D possibilities are chosen with equal probabilities.

- (a) Propose increasing the link variable,  $m' = m + 1$ . The corresponding acceptance probability is  $p_{acc} = \min\{q, 1\}$ , where

$$q = \frac{\beta k_i(u') + 1}{m' k(u') + N}. \quad (\text{A1})$$

- (b) Propose decreasing the link variable,  $m' = m - 1$ . When  $m = 0$ , the move is rejected; otherwise, the corresponding acceptance probability is given by

$$q = \frac{m(k'(u) + N)}{\beta (k'_i(u) + 1)}. \quad (\text{A2})$$

- (2) When the two heads of the worm coincide, the worm can jump to a new position and change its index ( $u = v \rightarrow u' = v', i \rightarrow j$ ). In this case,

$$q = \frac{(k(u) + N - 2)(k_j(u') + 1)}{(k_i(u) - 1)(k(u') + N)} = \frac{(\hat{k}(u) + N)(\hat{k}_j(u') + 1)}{(\hat{k}_i(u) + 1)(\hat{k}(u') + N)}. \quad (\text{A3})$$

- (3) Propose increasing or decreasing a link variable  $m \equiv m_j^{(l)}$  by 2, without changing the worm variables  $u, v, i$ .

- (a) Propose increasing  $m \rightarrow m' = m + 2$ . The acceptance probability is given by

$$q = \frac{\beta^2}{m'(m' - 1)} \prod_{x \in \partial l} \frac{k_j(x) + 1}{k(x) + N}. \quad (\text{A4})$$

- (b) Propose decreasing  $m \rightarrow m' = m - 2$ . When  $m < 2$ , the proposal is rejected; otherwise, the acceptance probability is given by

$$q = \frac{m(m - 1)}{\beta^2} \prod_{x \in \partial l} \frac{k'(x) + N}{k'_j(x) + 1}. \quad (\text{A5})$$

#### 2. Worm update with chemical potential

According to the representation, Eq. (21) of the scalar product, the index of the head and tail of the worm could be  $(i_u, i_v) = (-, +), (3, 3), \dots, (N, N)$ ; hence, we have to distinguish two types of the worm. For the type  $(i_u, i_v) = (j, j)$ , the updating steps described in Appendix A.1 remain unchanged. The same is true for updating a link variable  $m_j^{(l)}$  for  $j = 3, \dots, N$ . Below, we consider the case  $(i_u, i_v) = (-, +)$ :

- (1) Moving the  $(-)$  end of the worm in direction  $+\hat{v}$  from  $u$  to  $u' = u + \hat{v}$ :

- (a) Propose increasing the variable  $m_+ \rightarrow m'_+ = m_+ + 1$  on the corresponding link:

$$q = e^{\mu_v} \frac{\beta k_{12}(u') + 1}{m'_+ k(u') + N}. \quad (\text{A6})$$

- (b) Propose decreasing the link variable  $m_- \rightarrow m'_- = m_- - 1$ . If  $m_- = 0$ , then the proposal is rejected; otherwise,

$$q = e^{\mu_v} \frac{m_- k'(u) + N}{\beta k'_{12}(u) + 1}. \quad (\text{A7})$$

- (2) Moving the  $(-)$  end of the worm in direction  $-\hat{v}$  from  $u$  to  $u' = u - \hat{v}$ :  
 (a) Propose increasing the link variable  $m_- \rightarrow m'_- = m_- + 1$ :

$$q = e^{-\mu_v} \frac{\beta k_1 2(u') + 1}{m'_- k(u') + N}. \quad (\text{A8})$$

- (b) Propose decreasing the link variable  $m_+ \rightarrow m'_+ = m_+ - 1$ . If  $m_+ = 0$ , then the proposal is rejected; otherwise,

$$q = e^{-\mu_v} \frac{m_+ k'(u) + N}{\beta k'_{12}(u) + 1}. \quad (\text{A9})$$

The acceptance probabilities for moving the  $(+)$  end of the worm are described by the same expressions; the case  $(+)$ ,  $\pm \hat{v}$  is equivalent to  $(-)$ ,  $\mp \hat{v}$  (both decrease/increase the  $Q_{12} = +1$  line by 1). However, because of the next updating step, one does not need to move the  $(+)$  head to satisfy ergodicity. Due to these facts, in our simulations, we did not update the  $(+)$  end but only the  $(-)$  end with  $2/3$  probability.

- (3) When the two ends of the worm coincide ( $u = v$ ), it can jump to a new position ( $u' = v'$ ) and change its index.  
 (a)  $(i, i) \rightarrow (j, j)$ :

$$q = \frac{(\hat{k}(u) + N)(\hat{k}_j(u') + 1)}{(\hat{k}_i(u) + 1)(\hat{k}(u') + N)}. \quad (\text{A10})$$

- (b)  $(-, +) \rightarrow (-, +)$ :

$$q = \frac{(\hat{k}(u) + N)(\hat{k}_{12}(u') + 1)}{(\hat{k}_{12}(u) + 1)(\hat{k}(u') + N)}. \quad (\text{A11})$$

- (c)  $(-, +) \rightarrow (j, j)$ :

$$q = \frac{(\hat{k}(u) + N)(\hat{k}_j(u') + 1)}{(\hat{k}_{12}(u) + 1)(\hat{k}(u') + N)}. \quad (\text{A12})$$

- (d)  $(i, i) \rightarrow (-, +)$ :

$$q = \frac{(\hat{k}(u) + N)(\hat{k}_{12}(u') + 1)}{(\hat{k}_i(u) + 1)(\hat{k}(u') + N)}. \quad (\text{A13})$$

- (4) Changing the link variables  $\pm$  on a given link simultaneously:

- (a)  $m_+ \rightarrow m'_+ = m_+ + 1, m_- \rightarrow m'_- = m_- + 1$ :

$$q = \frac{\beta^2}{m'_+ m'_-} \prod_{x \in \partial l} \frac{k_{12}(x) + 1}{k(x) + N}. \quad (\text{A14})$$

- (b)  $m_+ \rightarrow m'_+ = m_+ - 1, m_- \rightarrow m'_- = m_- - 1$  (if  $m_+ = 0$  or  $m_- = 0$ , then the proposal is rejected):

$$q = \frac{m_+ m_-}{\beta^2} \prod_{x \in \partial l} \frac{k'(x) + N}{k'_{12}(x) + 1}. \quad (\text{A15})$$

Note that these expressions with the chemical potential using the modified basis  $\phi_+, \phi_-, \phi_3, \dots, \phi_N$  look very similar to those discussed in the case without the chemical potential.

- 
- [1] E. Y. Loh, J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, *Phys. Rev. B* **41**, 9301 (1990).  
 [2] P. de Forcrand, *Proc. Sci.*, LAT2009 (2009) 010.  
 [3] G. Aarts, *Pramana* **84**, 787 (2015).  
 [4] G. Aarts, *Proc. Sci.*, CPOD2014 (2014) 012.  
 [5] D. Sexty, *Proc. Sci.*, LATTICE2014 (2014) 016.  
 [6] D. Sexty, *Phys. Lett. B* **729**, 108 (2014).  
 [7] Z. Fodor, S. D. Katz, D. Sexty, and C. Trk, *Phys. Rev. D* **92**, 094516 (2015).  
 [8] G. Parisi and Y. s. Wu, *Scientia Sinica* **24**, 483 (1981).  
 [9] G. Parisi, *Phys. Lett. B* **131**, 393 (1983).  
 [10] J. R. Klauder, *Acta Phys. Austriaca Suppl.* **25**, 251 (1983).  
 [11] J. Ambjorn and S. K. Yang, *Phys. Lett.* **165B**, 140 (1985).  
 [12] J. Ambjorn, M. Flensburg, and C. Peterson, *Nucl. Phys.* **B275**, 375 (1986).  
 [13] G. Aarts, F. A. James, E. Seiler, and I. O. Stamatescu, *Phys. Lett. B* **687**, 154 (2010).  
 [14] G. Aarts, E. Seiler, and I. O. Stamatescu, *Phys. Rev. D* **81**, 054508 (2010).  
 [15] G. Aarts, F. A. James, E. Seiler, and I. O. Stamatescu, *Eur. Phys. J. C* **71**, 1756 (2011).  
 [16] G. Aarts, P. Giudice, and E. Seiler, *Ann. Phys. (Amsterdam)* **337**, 238 (2013).  
 [17] E. Seiler, D. Sexty, and I. O. Stamatescu, *Phys. Lett. B* **723**, 213 (2013).  
 [18] H. Makino, H. Suzuki, and D. Takeda, *Phys. Rev. D* **92**, 085020 (2015).  
 [19] J. Bloch, J. Mahr, and S. Schmalzbauer, *Proc. Sci.*, LATTICE2013 (2014) 451.  
 [20] G. Aarts, L. Bongiovanni, E. Seiler, D. Sexty, and I. O. Stamatescu, *Proc. Sci.*, LATTICE2013 (2014) 451.  
 [21] G. Aarts, F. Attanasio, B. Jger, E. Seiler, D. Sexty, and I. O. Stamatescu, *Proc. Sci.*, LATTICE2014 (2014) 200.  
 [22] G. Aarts and F. A. James, *J. High Energy Phys.* **08** (2010) 020.

- [23] N. Prokof'ev and B. Svistunov, *Phys. Rev. Lett.* **87**, 160601 (2001).
- [24] M. G. Endres, *Phys. Rev. D* **75**, 065012 (2007).
- [25] M. G. Endres, *Proc. Sci.*, LAT2006 (2006) 133.
- [26] U. Wolff, *Nucl. Phys.* **B810**, 491 (2009).
- [27] D. Banerjee and S. Chandrasekharan, *Phys. Rev. D* **81**, 125007 (2010).
- [28] Y. D. Mercado, H. G. Evertz, and C. Gattringer, *Comput. Phys. Commun.* **183**, 1920 (2012).
- [29] K. Langfeld, *Phys. Rev. D* **87**, 114504 (2013).
- [30] Y. Delgado, C. Gattringer, and A. Schmidt, *Proc. Sci.*, LATTICE2013 (2014) 147.
- [31] E. V. Herland, T. A. Bojesen, E. Babaev, and A. Sudbo, *Phys. Rev. B* **87**, 134503 (2013).
- [32] C. Gattringer, T. Kloiber, and M. Müller-Preussker, *Phys. Rev. D* **92**, 114508 (2015).
- [33] L. P. Yang, Y. Liu, H. Zou, Z. Y. Xie, and Y. Meurice, *Phys. Rev. E* **93**, 012138 (2016).
- [34] F. Bruckmann, C. Gattringer, T. Kloiber, and T. Sulejmanpasic, *Phys. Lett. B* **749**, 495 (2015).
- [35] F. Bruckmann, C. Gattringer, T. Kloiber, and T. Sulejmanpasic, *Phys. Rev. Lett.* **115**, 231601 (2015).
- [36] A. M. Polyakov, *Phys. Lett. B* **59**, 79 (1975).
- [37] E. Brezin and J. Zinn-Justin, *Phys. Rev. B* **14**, 3110 (1976).
- [38] V. A. Novikov, M. A. Shifman, A. I. Vainshtein, and V. I. Zakharov, *Phys. Rep.* **116**, 103 (1984); *Fiz. Elem. Chastits At. Yadra* **17**, 472 (1986) [*Sov. J. Part. Nucl.* **17**, 204 (1986)].
- [39] F. Bruckmann, *Phys. Rev. Lett.* **100**, 051602 (2008).
- [40] R. H. Swendsen and J. S. Wang, *Phys. Rev. Lett.* **58**, 86 (1987).
- [41] U. Wolff, *Phys. Rev. Lett.* **62**, 361 (1989).
- [42] F. Niedermayer, *Phys. Rev. Lett.* **61**, 2026 (1988).
- [43] S. Caracciolo, R. G. Edwards, A. Pelissetto, and A. D. Sokal, *Nucl. Phys.* **B403**, 475 (1993).
- [44] D. Nogradi, *J. High Energy Phys.* **05** (2012) 089.
- [45] S. Caracciolo, R. G. Edwards, A. Pelissetto, and A. D. Sokal, *Nucl. Phys. B, Proc. Suppl.* **42**, 752 (1995).
- [46] S. Caracciolo, R. G. Edwards, A. Pelissetto, and A. D. Sokal, *Phys. Rev. Lett.* **75**, 1891 (1995).
- [47] J. Balog, M. Niedermaier, F. Niedermayer, A. Patrascioiu, E. Seiler, and P. Weisz, *Phys. Rev. D* **60**, 094508 (1999).
- [48] Z. Fodor and S. D. Katz, *Phys. Lett. B* **534**, 87 (2002).
- [49] Z. Fodor, S. D. Katz, and K. K. Szabo, *Phys. Lett. B* **568**, 73 (2003).
- [50] F. Csikor, G. I. Egri, Z. Fodor, S. D. Katz, K. K. Szabo, and A. I. Toth, *J. High Energy Phys.* **05** (2004) 046.
- [51] H. Leutwyler, *Phys. Lett. B* **189**, 197 (1987).
- [52] P. Hasenfratz, *Nucl. Phys.* **B828**, 201 (2010).
- [53] F. Niedermayer and C. Weiermann, *Nucl. Phys.* **B842**, 248 (2011).
- [54] J. Balog, F. Niedermayer, and P. Weisz, *Nucl. Phys.* **B824**, 563 (2010).
- [55] G. Boyd, J. Engels, F. Karsch, E. Laermann, C. Legeland, M. Lutgemeier, and B. Petersson, *Nucl. Phys.* **B469**, 419 (1996).
- [56] E. Seel, D. Smith, S. Lottini, and F. Giacosa, *J. High Energy Phys.* **07** (2013) 010.
- [57] F. Bruckmann, C. Gattringer, T. Kloiber, and T. Sulejmanpasic, *Phys. Rev. D* **94**, 114503 (2016).
- [58] G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, *Phys. Rev. D* **32**, 2736 (1985).
- [59] A. Mollgaard and K. Splittorff, *Phys. Rev. D* **88**, 116007 (2013).
- [60] A. Mollgaard and K. Splittorff, *Phys. Rev. D* **91**, 036007 (2015).