

Regularization of the complex Langevin method

Zhenning Cai^{1,*}, Yang Kuang^{2,†} and Hong Kiat Tan^{3,‡}

¹*Department of Mathematics, National University of Singapore, Level 4, Block S17,
10 Lower Kent Ridge Road, Singapore 119076*

²*School of Mathematics and Statistics, Guangdong University of Technology,
Guangzhou 510006, Guangdong, China*

³*Department of Mathematics, University of California, Los Angeles, 520 Portola Plaza,
Los Angeles, California 90095, USA*



(Received 1 November 2021; accepted 10 January 2022; published 28 January 2022)

The complex Langevin method, a numerical method used to compute the ensemble average with a complex partition function, often suffers from runaway instability. We study the regularization of the complex Langevin method via augmenting the action with a stabilization term. Since the regularization introduces biases to the numerical result, two approaches, named 2R and 3R methods, are introduced to recover the unbiased result. The 2R method supplements the regularization with regression to estimate the unregularized ensemble average, and the 3R method reduces the computational cost by coupling the regularization with a reweighting strategy before regression. Both methods can be generalized to the $SU(n)$ theory and are assessed from several perspectives. Several numerical experiments in the lattice field theory are carried out to show the effectiveness of our approaches.

DOI: [10.1103/PhysRevD.105.014508](https://doi.org/10.1103/PhysRevD.105.014508)

I. INTRODUCTION

The complex Langevin method is a numerical approach used to circumvent the numerical sign problem arising in the computation of ensemble averages with complex Boltzmann weights. Such issues may appear in the real-time quantum field theories [1,2], coupled quantum systems with chemical potentials such as the Hubbard model [3,4] and the quantum chromodynamics at finite density [5,6], and also the superstring theory [7,8]. In these applications, one usually encounters strong oscillations in high-dimensional functions, leading to significant cancellations when integrating the functions. As a result, the classical Monte Carlo method fails to work as the variance is large compared to the mean value, and such difficulty is known as the numerical sign problem [9].

The complex Langevin method, introduced in [10,11], tries to tame the numerical sign problem by using a straightforward extension of a classical sampling method called the Langevin method. The extension allows the samples to take complex values due to the complex Boltzmann weights. Unfortunately, the application of the complex Langevin method had been severely limited for a long time due to one of its major drawbacks: the method often diverges or converges to incorrect solutions [12]. The justification of the method and the understanding of its

failure were explored in several works [13–16], but the precise reason for the biased results remained unclear until recently [17,18]. In general, the failure of the complex Langevin method is due to the lack of control of excursions away from the real axis. Many efforts have been made in the past decade to restrict such excursions. For instance, the use of adaptive time steps is studied in [19] to avoid runaway trajectories; the method gauge cooling, which utilizes the gauge invariance to minimize the distances to the real axis, is proposed in [20] and has achieved many applications [21–23]; and the dynamical stabilization is introduced in [24] and tested in [25]. In the case where the method converges, the work [26] proposes an approximation technique to quantify the bias. Other attempts to improve the complex Langevin method include the coupling with Lefschetz thimbles [27], the deformation technique [28], etc. We invite the readers to refer to [29,30] for a comprehensive review of the recent advances.

In general, the complex Langevin method is still a numerical tool under construction. In this paper, we are going to carry out a deeper study of the aforementioned dynamical stabilization. The idea of dynamical stabilization is to add converging velocity fields to the complex Langevin equation to restrict the excursion of samples. Instead of working on the original approach introduced in [24], we will investigate a slightly improved version considered in [31], called the method of modified action. Here we will follow [17] and name this approach as the “regularization” of the complex Langevin method. Compared with the original approach, this method is easier

*matcz@nus.edu.sg

†ykuang@gdut.edu.cn

‡maxtanhk@math.ucla.edu

to be justified theoretically. Our focus will be on possible modifications upon regularization. These include coupling regularization with the reweighted complex Langevin method [32,33] and attempts to recover the unbiased result using regression. Our study is to be carried out via a deep look into a motivating example in the $U(1)$ one-link case, after which the method will be generalized to the $SU(n)$ theories and applied to several lattice field theories, including the 3D XY model [34], the Polyakov model [35], and the heavy dense QCD (quantum chromodynamics) [20].

The rest of the paper is organized as follows. In Sec. II, we briefly review the complex Langevin method and its general theory. In Sec. III, the $2R$ and $3R$ methods are presented for a motivating one-dimensional example in the $U(1)$ one-link case. Then these methods are generalized to multidimensional integrals in the $U(1)$ theory in Sec. IV and to the $SU(n)$ theories in Sec. V. In Sec. VI, the applications of these methods to several lattice field theories are discussed. Finally, the paper ends with some concluding remarks in Sec. VII.

II. A REVIEW OF THE COMPLEX LANGEVIN METHOD

Following [20], we use the notation $\{\cdot\}$ to denote the discrete field defined on a lattice. For instance, suppose $\{\phi\}$ is a three-dimensional real scalar lattice field. Then, $\{\phi\}$ contains a set of variables $\phi_x \in \mathbb{R}$ with x being a three-dimensional multi-index representing the lattice point. If the lattice has N points, then $\{\phi\}$ is essentially a vector in \mathbb{R}^N . For simplicity, we will also use ϕ_k , $k = 1, \dots, N$ to denote the components of $\{\phi\}$. In this section, we will provide a brief introduction to the complex Langevin method and its regularization. For introductory purposes, we will temporarily restrict ourselves to the scalar fields where $\phi_k \in \mathbb{R}$ or \mathbb{T} , where \mathbb{T} stands for the torus $\mathbb{T} = \mathbb{R} \bmod 2\pi$.

With the notations defined above, we are interested in computing the following ensemble average:

$$\langle O \rangle = \frac{1}{Z} \int_{\Omega} O(\{\phi\}) e^{-S(\{\phi\})} d\{\phi\}, \quad Z = \int_{\Omega} e^{-S(\{\phi\})} d\{\phi\}, \quad (1)$$

where $\Omega = \mathbb{R}^N$ or \mathbb{T}^N . When $S(\{\phi\})$ is real, we can regard Z as the partition function, so that the integral can be evaluated by the Langevin method [36]. Specifically, the Langevin equation associated with (1) is given by

$$d\phi_k = K_k(\{\phi\}) dt + dw_k, \quad K_k = -\frac{\partial S}{\partial \phi_k}, \quad k = 1, \dots, N. \quad (2)$$

Here w_k , $k = 1, \dots, N$ are independent Wiener processes satisfying $dw_k^2 = 2dt$ for each k . The Fokker-Planck equation of this stochastic process is

$$\frac{\partial P}{\partial t} + \sum_{k=1}^N \frac{\partial}{\partial \phi_k} (K_k P) = \sum_{k=1}^N \frac{\partial^2 P}{\partial \phi_k^2}, \quad (3)$$

where $P(\{\phi\}, t)$ represents the probability distribution of the field $\{\phi\}$ at time t . If $e^{-S(\phi)}$ is integrable and the stochastic process (2) is ergodic, then $P(\{\phi\}, t)$ will converge to the equilibrium distribution $\frac{1}{Z} e^{-S(\{\phi\})}$ as $t \rightarrow \infty$. As a result, we can approximate (1) by

$$\langle O \rangle \approx \frac{1}{N_{\text{sample}}} \sum_{m=1}^{N_{\text{sample}}} O(\{\Phi^{(m)}\}), \quad (4)$$

where $\{\Phi^{(m)}\}$, $m = 1, \dots, N_{\text{sample}}$ are the samples generated by simulating the Langevin equation (2) and choosing $\Phi_k^{(m)} = \phi_k(T + m\Delta T)$ for a sufficiently large T and sufficiently long time difference ΔT .

However, when $S(\{\phi\})$ is complex, the Langevin method is no longer valid. A possible indicator would be that the partition function Z is now complex valued. To handle such complex actions, the complex Langevin method [11,37] postulates stochastic equations of the same form as (2) with the trajectories of the process wandering in the complexified space $\Omega_{\mathbb{C}}$. Here

$$\Omega_{\mathbb{C}} = \begin{cases} \mathbb{C}^N, & \text{if } \Omega = \mathbb{R}^N \\ (\mathbb{T} + i\mathbb{R})^N, & \text{if } \Omega = \mathbb{T}^N \end{cases}. \quad (5)$$

Now we assume that both $O(\cdot)$ and $S(\cdot)$ can be extended to $\Omega_{\mathbb{C}}$ holomorphically. Thus, the stochastic process (2) is again well defined, and the complex Langevin method again approximates $\langle O \rangle$ using (4). Since $K_k(\cdot)$ can take complex values, the field $\{\phi\}$ becomes a complex field, which can also be represented by two real fields $\{\phi^R\}$ and $\{\phi^I\}$ with $\phi_k = \phi_k^R + i\phi_k^I$. The evolution of these two fields follows the complex Langevin equation:

$$\begin{cases} d\phi_k^R = K_k^R(\{\phi^R\}, \{\phi^I\}) dt + dw_k, & K_k = -\text{Re}K_k \\ d\phi_k^I = K_k^I(\{\phi^R\}, \{\phi^I\}) dt, & K_k^I = -\text{Im}K_k \end{cases}, \quad (6)$$

where K_k is again the partial derivative of S as defined in (2). The corresponding Fokker-Planck equation for the probability density function has the form $P(\{\phi^R\}, \{\phi^I\}, t)$, which evolves according to

$$\frac{\partial P}{\partial t} + \sum_{k=1}^N \left(\frac{\partial}{\partial \phi_k^R} (K_k^R P) + \frac{\partial}{\partial \phi_k^I} (K_k^I P) \right) = \sum_{k=1}^N \frac{\partial^2 P}{\partial (\phi_k^R)^2}. \quad (7)$$

The correctness of the complex Langevin method requires the following two conditions:

- (i) The stochastic process (6) is ergodic.

(ii) The following equation holds:

$$\begin{aligned} \lim_{t \rightarrow \infty} \int_{\Omega_c} O(\{\phi\}) P(\{\phi^R\}, \{\phi^I\}, t) d\{\phi\} \\ = \frac{1}{Z} \int_{\Omega} O(\{\phi\}) e^{-S(\{\phi\})} d\{\phi\}. \end{aligned} \quad (8)$$

As mentioned in the introduction, observations indicate that when the stochastic process is ergodic, the complex Langevin method may sometimes produce inaccurate results, which implies the possibility that (8) fails to hold. This occurs especially when $P(\cdot, \cdot, \infty)$ decays slowly, and the details have been studied in [17,38,39]. As a remedy, the method of dynamical stabilization proposed in [24] adds an artificial term to the imaginary part of the drift velocity K_k to suppress the tail of P . Specifically, in (6), K_k is chosen as

$$K_k(\{\phi\}) = -\partial_{\phi_k} S(\{\phi\}) - i\alpha_{DS} (\text{Im}\phi_k)^r, \quad \forall k = 1, \dots, N, \quad (9)$$

where r is an odd positive integer and α_{DS} is a positive parameter balancing the stabilizing effect and the bias introduced by this regularization. Since (9) is no longer the derivative of an analytic function, the justification of this approach remains open.

In this paper, we will focus on another type of regularization introduced in [17], in which a specific problem is studied. In the next section, we will conduct a deeper study of the regularization technique based on this motivating example and consider its possible extensions.

III. MOTIVATING EXAMPLE— REGULARIZATION OF COMPLEX LANGEVIN FOR THE $U(1)$ ONE-LINK MODEL

We will motivate the use of regularization and dynamic stabilization by applying our proposed method on the one-dimensional $U(1)$ one-link model studied in [17,40], where $\Omega = \mathbb{T}$. Following [17], we use x to denote the integral variable. The action is a 2π -periodic function:

$$S(x) = i\beta \cos x,$$

with $\beta \in \mathbb{R}^+$, and $x \in \mathbb{T}$.

A. Regularization of complex Langevin

Inspired from [17], the regularized action for this model is given by

$$S_s(x) = S(x) + \frac{sx^2}{2} = i\beta \cos x + \frac{sx^2}{2} \quad (10)$$

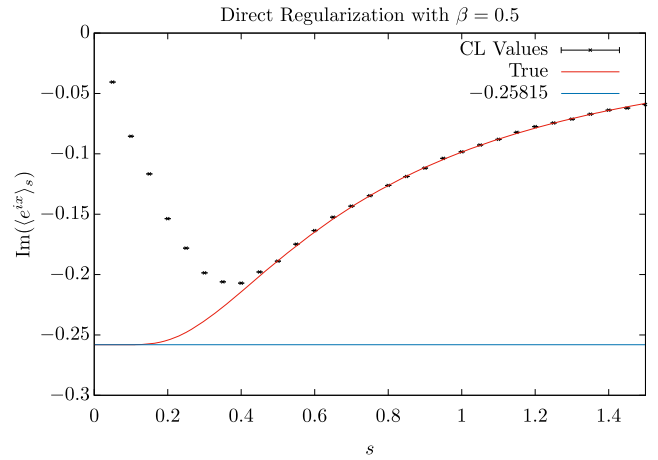


FIG. 1. The graph depicting the numerical results and true values of $\text{Im}(\langle e^{ix} \rangle_s)$ against s for $\beta = 0.5$.

for any given $s > 0$. As discussed in [17], for large values of s , the value of $\langle O \rangle$ agrees with its corresponding true value with a modified action. However, a divergence is observed for values of s close to 0.4 as we attempt to set s close to 0 to retrieve the true value of $\langle O \rangle$ with unregularized action. The results are also reproduced in Fig. 1. The authors have commented that the use of appropriate regression functions might have by extrapolating the results from $s > 0.4$ to obtain a decent estimate at $s = 0$. We will thus be following a similar argument while supplementing it with relevant regression functions with the appropriate mathematical justification.

Note that due to the presence of the regularizing term, the periodicity for $S(x)$ in x is destroyed. Therefore, in the definition of the observable, we will “unroll” the torus \mathbb{T} and change the integral domain to \mathbb{R} . Thus, under the modified action, we can rewrite equations (1) as

$$\langle O \rangle_s = \frac{1}{Z_s} \int_{\mathbb{R}} O(x) \exp(-S_s(x)) dx, \quad (11)$$

with

$$Z_s = \int_{\mathbb{R}} \exp(-S_s(x)) dx. \quad (12)$$

Upon complexification, we obtain the complex action

$$S_s(z) = S(z) + \frac{sz^2}{2}, \quad z \in \mathbb{C}. \quad (13)$$

The corresponding drift terms can be computed as follows¹:

¹Following the convention in (2) and in (90), we will represent $K_{x,s}$ as the drift term for the scalar field ϕ_x with regularized action. However, as in the one-dimensional case, as it is understood that we are only dealing with one field variable (which is written as x), we will drop the comma that separates x and s and simply write it as K_s .

$$\begin{aligned} K_s^R(x, y) &= -\operatorname{Re}(S'_s(x + iy)) = -\beta \cos x \sinh y - sx, \\ K_s^I(x, y) &= -\operatorname{Im}(S'_s(x + iy)) = \beta \sin x \cosh y - sy. \end{aligned} \quad (14)$$

As for the regularized action, two questions need to be answered:

- (i) What is the relation between the regularized observable $\langle O \rangle_s$ and the original observable $\langle O \rangle$?
- (ii) Can we apply the complex Langevin method to obtain the correct value of $\langle O \rangle_s$?

The following two sections will be devoted to the exploration of their answers.

1. Correct convergence under regularized action

Note that it is natural to expect that the regularized observable will converge to the original observable as the regularizing parameter vanishes:

$$\lim_{s \rightarrow 0^+} \langle O \rangle_s = \langle O \rangle. \quad (15)$$

$$\int_{\mathbb{R}} e^{imx} e^{-i\beta \cos x - s\frac{x^2}{2}} dx = \sum_{n \in \mathbb{Z}} i^n J_n(-\beta) \int_{\mathbb{R}} e^{i(n+m)x - s\frac{x^2}{2}} dx = \sum_{n \in \mathbb{Z}} i^n J_n(-\beta) \sqrt{\frac{2\pi}{s}} e^{-\frac{(m+n)^2}{2s}}. \quad (17)$$

Here, we have interchanged the infinite sum and the integral, which can be justified using dominated convergence theorem by computing

$$\left| \sum_{n=-N}^N i^n J_n(-\beta) e^{i(n+m)x - s\frac{x^2}{2}} \right| = e^{-s\frac{x^2}{2}} \left| \sum_{n=-N}^N i^n J_n(-\beta) e^{inx} \right| \leq 4e^{-s\frac{x^2}{2}} \quad (18)$$

for a given $N \in \mathbb{Z}^+$ that is large enough, as the finite sum inside the absolute sign tends to $e^{-i\beta \cos x}$ with modulus 1 if (16) is applied. The resulting upper bound in (18) is clearly integrable on \mathbb{R} for a fixed β and s .

Proposition 3.2. For the $U(1)$ one-link model, suppose the observable $O(x)$ is 2π periodic and absolutely continuous on $[0, 2\pi)$. In addition, if we demand that O is a $(1 + \alpha)$ -Hölder class function for some $\alpha > 0$, then, we have that (15) holds.

Proof.—First, we consider the Fourier series expansion of $O(x)$ given by

$$O(x) = \sum_{m \in \mathbb{Z}} \hat{O}_m e^{imx}. \quad (19)$$

Furthermore, from a standard result in harmonic analysis, we know that the convergence of the infinite series on the right-hand side of (14) is uniform, which thus implies that the Fourier series on the right can be used to represent O . From here, we apply Lemma 3.1:

$$\int_{\mathbb{R}} O(x) \exp(-S_s(x)) dx = \sum_{m \in \mathbb{Z}} \hat{O}_m \int_{\mathbb{R}} e^{imx} \exp(-S_s(x)) dx = \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} i^n \hat{O}_m J_n(-\beta) \cdot \sqrt{\frac{2\pi}{s}} e^{-\frac{(m+n)^2}{2s}}, \quad (20)$$

$$\begin{aligned} \int_{\mathbb{T}} O(x) \exp(-S(x)) dx &= \sum_{m \in \mathbb{Z}} \hat{O}_m \int_0^{2\pi} e^{imx} e^{-i\beta \cos x} dx = \sum_{m \in \mathbb{Z}} \hat{O}_m \int_0^{2\pi} e^{imx} \sum_{n \in \mathbb{Z}} i^n J_n(-\beta) e^{inx} dx \\ &= 2\pi \sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} i^n \hat{O}_m J_n(-\beta) \delta_{n,-m} = 2\pi \sum_{m \in \mathbb{Z}} i^m \hat{O}_{-m} J_m(-\beta), \end{aligned} \quad (21)$$

where the last equality of (20) is due to Lemma 3.1 and (21) utilizes (16) and the property that $J_n(x) = J_{-n}(x)$ for all integer n . Here, we note that in (20) and (21), we have swapped the relevant infinite series and integration. This can be justified using the dominated convergence theorem by considering the following partial sums for any $N \in \mathbb{N}$:

However, this is not immediately clear since we have changed the integration domain from $[0, 2\pi)$ to \mathbb{R} as we apply the regularization. Fortunately, the result above still holds under some mild conditions. To prove the limit (15), we need the following lemma:

Lemma 3.1. For any $m, \beta \in \mathbb{R}$,

$$\int_{\mathbb{R}} e^{imx} e^{-i\beta \cos x - s\frac{x^2}{2}} dx = \sum_{n \in \mathbb{Z}} i^n J_n(-\beta) \sqrt{\frac{2\pi}{s}} e^{-\frac{(m+n)^2}{2s}},$$

where J_n denotes the Bessel function of the first kind.

Proof.—Applying the Jacobi-Anger expansion

$$e^{-iz \cos \phi} = \sum_{n \in \mathbb{Z}} i^n J_n(-z) e^{in\phi} \quad \text{for all } \phi \in \mathbb{R}. \quad (16)$$

to (19), we have

$$\begin{aligned} \left| \sum_{m=-N}^N \hat{O}_m e^{imx} \exp(-S_s(x)) \right| &= \left| \sum_{m=-N}^N \hat{O}_m e^{imx - i\beta \cos(x) - s\frac{x^2}{2}} \right| \\ &\leq 2 \sum_{m=1}^N \frac{K}{|m|^{1+\alpha}} e^{-s\frac{x^2}{2}} + |\hat{O}_0| e^{-s\frac{x^2}{2}} \leq e^{-s\frac{x^2}{2}} (2K\zeta(1+\alpha) + |\hat{O}_0|), \end{aligned} \quad (22)$$

where $\zeta(\cdot)$ is the Riemann zeta function, and we have used the assumption that O is a $(1+\alpha)$ -Hölder class function for some $\alpha > 0$. This means that there exists a constant K such that

$$|\hat{O}_m| \leq \frac{K}{|m|^{1+\alpha}}. \quad (23)$$

Thus, from (22), we can see that the upper bound is clearly integrable on \mathbb{R} . Thus, by dominated convergence theorem, the aforementioned interchange is justified.

Equation (20) implies that

$$\lim_{s \rightarrow 0^+} \sqrt{\frac{s}{2\pi}} \int_{\mathbb{R}} O(x) \exp(-S_s(x)) dx = \sum_{n \in \mathbb{Z}} i^n \hat{O}_{-n} J_n(-\beta). \quad (24)$$

For Z_s and Z , one can use the same technique to deduce that

$$\lim_{s \rightarrow 0^+} \sqrt{\frac{s}{2\pi}} Z_s = J_0(-\beta), \quad Z = 2\pi J_0(-\beta). \quad (25)$$

It is now clear from (21), (24), and (25) that

$$\lim_{s \rightarrow 0^+} \langle O \rangle_s = \frac{\lim_{s \rightarrow 0^+} \sum_{m \in \mathbb{Z}} \hat{O}_m \sqrt{\frac{s}{2\pi}} \int_{\mathbb{R}} e^{imx} e^{-i\beta \cos x - s\frac{x^2}{2}} dx}{\lim_{s \rightarrow 0^+} \sqrt{\frac{s}{2\pi}} Z_s} = \frac{1}{J_0(-\beta)} \sum_{m \in \mathbb{Z}} \hat{O}_{-m} i^m J_m(-\beta) = \langle O \rangle, \quad (26)$$

which concludes the proof. ■

The result above justifies the regularization of the action—if s is chosen small and $\langle O \rangle_s$ can be correctly computed by the complex Langevin method, the value $\langle O \rangle_s$ can be regarded as an approximation of $\langle O \rangle$.

2. Correct numerical convergence for complex Langevin method

Despite the guarantee for correct convergence given in Proposition 3.2, numerical results from the complex Langevin method suggest otherwise. The numerical experiments on the regularized action have been carried out in [17], and we have repeated the same experiments for $\beta = 0.5$. The results are plotted in Fig. 1 for $O(x) = e^{ix}$, where we can observe a divergence between the true values represented by the red curve and the numerical results represented by the data points. The data points are obtained via numerical simulations with a fixed time step of $\Delta t = 3 \times 10^{-4}$ for values of s closer to 0 and $\Delta t = 1 \times 10^{-3}$ if otherwise, with each sample obtained after every 2000 steps for a total of 10^6 samples for each value of $s \in \{0.05k | 0 \leq k \leq 30, k \in \mathbb{Z}\}$.

In view of Proposition 3.2, we can deduce that such divergence between the true values and the numerical results must be due to the corresponding complexification of the Langevin dynamics. It is thus instructive to investigate the correct values of s in which the numerical results from our complex Langevin method agree with that from the original Langevin dynamics. The phenomenon that the correctness of complex Langevin changes with the parameter in the action has been observed and explained in a number of previous works [15,38,39]. In [39], it is demonstrated in another example that the correctness of the complex Langevin results can be guaranteed only when the probability density function is localized, meaning that for all $t > 0$, the solution of (6) always satisfies $y(t) \in [Y^-, Y^+]$ for some $Y^- < Y^+$. Our problem has a close similarity to the example in [39], and it can be expected that we also require the localization of the $y(t)$ to guarantee the correctness of complex Langevin. To confine the value of $y(t)$, we need the imaginary velocity K^I to satisfy $K_s^I(x, Y^-) > 0$ and $K_s^I(x, Y^+) < 0$ for all x . Note that the choice of 0 here is due to the fact that in all simulations of the complex Langevin dynamics, we will always set the

initial coordinates to be at the origin. This thus motivates the following proposition:

Proposition 3.3. For the $U(1)$ one-link model, given a fixed $s > 0$ and $\beta > 0$, if $s > 1.509\beta$, then there exist $Y^+ > 0$ and $Y^- < 0$ such that

$$K_s^I(x, Y^+) < 0 \quad \text{and} \quad K_s^I(x, Y^-) > 0. \quad (27)$$

Proof.—We first consider the case for $K_s^I(x, Y^+) < 0$. For any $(x, y) \in \mathbb{R}^2$, using the expression from (14), we are looking to solve the following inequality

$$K_s^I(x, y) = \beta \sin(x) \cosh(y) - sy < 0 \quad (28)$$

in the sense that there exists a $y = Y^+ > 0$ such that for all $x \in \mathbb{R}$, $K_s^I(x, Y^+) < 0$.

First, for this to hold for all x , it must thus hold at a point in which $\sin x$ is maximum, that is, it takes the value of 1, as $\cosh(y) > 0$ for all $y \in \mathbb{R}$. We define the new expression of K_s^I in which we replace $\sin(x)$ by 1 as \bar{K}^I . Thus, we are looking to solve for a region in the parameter space (β, s) such that such a Y^+ would be guaranteed. The strategy is as follows. First, we fix the parameters β and s and solve for the minimum value of this function \bar{K}^I at y_0 in terms of β and s . Since this minimum value is a function of β and s , we can in fact find such a region in the parameter space such that $\bar{K}^I(y_0) < 0$. Thus, since \bar{K}^I is minimized at y_0 and is negative, we then have for all $y \in [0, Y^+]$ with $Y^+ = y_0$ that $\bar{K}^I(y) < 0$ and thus $K_s^I(x, y) < 0$ for all $x \in \mathbb{R}$ and $y \in [0, Y^+]$. Therefore, we have y_0 as the required Y^+ that we are looking for. To apply this strategy, we first look at the corresponding function for \bar{K}^I :

$$\bar{K}^I(y) = \beta \cosh(y) - sy. \quad (29)$$

Using standard one-variable optimization techniques, we see that global minimum is attained at

$$y_0 = \sinh^{-1}\left(\frac{s}{\beta}\right). \quad (30)$$

Now we demand that the minimum value of \bar{K}^I be negative:

$$\beta \cosh(y_0) - sy_0 < 0. \quad (31)$$

Inserting (30) into the equation above and letting $\chi = s/\beta$, we can simplify the inequality (31) to

$$e^{\sqrt{1+(\frac{1}{\chi})^2}} - \chi - \sqrt{1+\chi^2} < 0, \quad (32)$$

which can be solved numerically to obtain

$$\chi > 1.509, \quad s > 1.509\beta, \quad (33)$$

and the proof is thus complete for this case.

For the other case in (27), we can use the same strategy to obtain the same sufficient condition $s > 1.509\beta$, which completes the proof of the proposition. ■

Indeed, as we can see from Fig. 1, for points after $s = 0.8 \approx 1.6\beta > 1.509\beta$, we can observe that the true values are coherent with the numerical values obtained. For $s \in (0.5, 0.7)$, although the numerical results from complex Langevin appear to be on the red curve, we believed that a small systematic bias has occurred.

In view of Propositions 3.2 and 3.3, it seems unlikely that we can obtain good numerical values of $\langle O \rangle$ solely with the use of a regularized action, as seen in Fig. 1. This thus motivates the following subsection, in which we will consider the fix of the regularized values.

B. Reweighted complex Langevin method with regularized action

In this subsection, we will introduce the reweighted complex Langevin method aimed at obtaining numerical results for $\langle O \rangle_s$.

In [32], the authors consider the action S_ξ with a parameter ξ . It then holds for any ξ and ξ_0 that

$$\frac{\int O(x) S_\xi(x) dx}{\int S_\xi(x) dx} = \frac{\int O(x) S_{\xi_0}(x) dx}{\int S_{\xi_0}(x) dx}. \quad (34)$$

Both the numerator and the denominator on the right-hand side can be approximated using the complex Langevin method with action $S_{\xi_0}(x)$. By choosing an appropriate ξ_0 , one may get a better approximation of $\langle O \rangle$ as compared to applying the complex Langevin method directly to the left-hand side of (34). In our case, the regularized action includes a regularizing parameter s , in which we know that the true value could be generated at $s = 0$. This inspires us to develop our algorithm according to the following proposition:

Proposition 3.4. The following equality holds for all $s, s_0 \geq 0$:

$$\langle O(x) \rangle_s = \frac{\langle O(x) \exp(\frac{(s_0-s)x^2}{2}) \rangle_{s_0}}{\langle \exp(\frac{(s_0-s)x^2}{2}) \rangle_{s_0}}. \quad (35)$$

This proposition is a direct result of (34) by setting ξ to be s . Thus, from Proposition 3.3, as long as we pick $s_0 > 1.509\beta$ and $s > 0$, we are guaranteed that the numerical values of two integrals in the ratio obtained using the complex Langevin method for the right-hand side of 3.4 have no biases. By equality (35), we can thus obtain an accurate numerical value of $\langle O \rangle_s$ even for $s < s_0$. Setting $s \rightarrow 0$ in (35), we thus have an accurate numerical value of $\langle O \rangle$.

Following Proposition 3.4, we carry out numerical experiments by fixing $s_0 = 0.8$ and compute $\langle O \rangle_s$ for

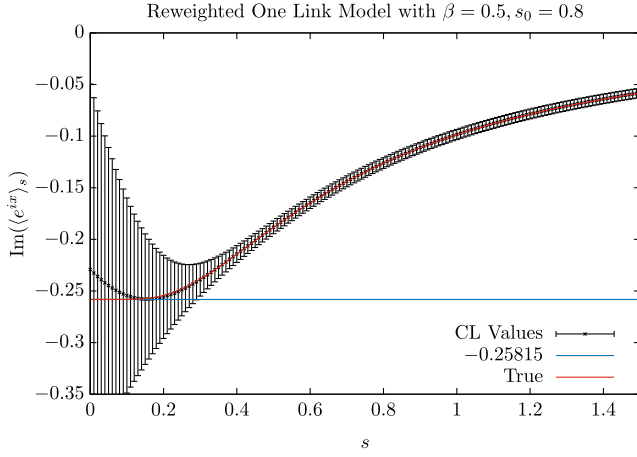


FIG. 2. The graph depicting the divergence of numerical results obtained from reweighted complex Langevin method and true values of $\text{Im}(\langle e^{ix} \rangle_s)$ against s for $\beta = 0.5$ due to large standard errors; -0.25815 represents the true value at $s = 0$.

$s \in (0, 1.5)$. The numerical values were generated using a fixed time step of $\Delta t = 10^{-3}$, with each sample obtained after every 2000 steps for a total of 10^7 samples at $s_0 = 0.8$. Values of $\langle O \rangle_s$ for $s \neq 0.8$ were obtained from this set of points generated via the equation (35) above. Furthermore, the corresponding error bars were generated using a M out of a N naïve bootstrap method at each s , with $M = 20000$ and $N = 10^7$, repeated for $n = 10000$ times. The results and the estimated error bars are plotted in Fig. 2. Indeed, we observe that for s around 0.5, we have obtained a better approximation of $\langle O \rangle$. However, two worrying phenomena have also surfaced from this experiment. Namely,

- (i) The numerical value of $\langle O \rangle_s$ deviates from the true value when s gets smaller than 0.5.
- (ii) As s reduces, the estimated standard error start to grow dramatically from s around 0.4.

Nonetheless, it can be shown that the divergence for $\langle O \rangle$ is due to a large standard error, in which the standard error for $\langle O \rangle_s$ grows as the value of s decreases from s_0 to 0. This thus provides motivation for the following section, in which the introduction of a mathematically motivated regression model aims to obtain an improved numerical estimate for $\langle O \rangle$.

C. Coupling reweighted complex Langevin method with regularized action, with regression

As mentioned at the start of this section, an important question to address would be the choice of the regressors that we should use to perform regression. Will a simple polynomial regression work? What would be considered as appropriate regressors? To answer these questions, we refer back to Fig. 1. The graph above shows the graph of the true curve of $\langle O \rangle_s$ for $O = e^{ix}$ with $\beta = 0.5$ in red. As observed, the curve becomes very flat when s is close to zero, which implies that the higher-order derivatives of $\text{Im}(\langle e^{ix} \rangle_s)$ might be 0 at $s = 0$. This is not a fact captured by arbitrary

polynomial regressors. Thus, if such an observation is true, we would have to turn to other regressors. This motivates the proposition below.

Proposition 3.5. For the $U(1)$ one-link model with regularized action, for any observable O satisfying the conditions in Proposition 3.2 and for any given $k \in \mathbb{Z}^+$, we have

$$\frac{d^k}{ds^k} \langle O \rangle_s |_{s=0} = 0. \quad (36)$$

Proof.—The proof of this proposition continues from the proof of Proposition 3.2. From (17), the numerator for $\langle O \rangle_s$ constitutes a sum over m of the expression in (17). The denominator however, consists of the $m = 0$ term in (17). Multiplying both the numerator and the denominator by a factor of $\sqrt{\frac{s}{2\pi}}$, we have that

$$\langle O \rangle_s = \frac{\sum_{m \in \mathbb{Z}} \sum_{n \in \mathbb{Z}} i^n J_n(-\beta) \hat{O}_m e^{-\frac{(m+n)^2}{2s}}}{\sum_{n \in \mathbb{Z}} i^n J_n(-\beta) e^{-\frac{n^2}{2s}}} := \frac{f_1(s)}{f_2(s)}. \quad (37)$$

As the given function above is clearly infinitely differentiable, we take the derivative with respect s on both sides to obtain

$$\frac{d}{ds} \langle O \rangle_s = \frac{f_1'(s)f_2(s) - f_2'(s)f_1(s)}{(f_2(s))^2}, \quad (38)$$

where

$$f_1'(s) = \sum_{(n+m) \in \mathbb{Z} \setminus \{0\}} i^n J_n(-\beta) \hat{O}_m \left(\frac{(m+n)^2}{s^2} \right) e^{-\frac{(m+n)^2}{2s}},$$

$$f_2'(s) = \sum_{n \in \mathbb{Z} \setminus \{0\}} i^n J_n(-\beta) \left(\frac{n^2}{s^2} \right) e^{-\frac{n^2}{2s}}. \quad (39)$$

Here, we used $\mathbb{Z} \setminus \{0\}$ since if n or $n + m$ is equals to 0 before differentiating, the corresponding term in the infinite series is a constant due to the absence of the exponential factor and disappears upon differentiation. By writing down (39), we have explicitly swapped the derivative and the infinite sum. This can be justified using a standard result in analysis (see Theorem 7.17 in [41]) as follows. First, we restrict our attention to $[0, s_r]$ for s_r large enough.² Then, we will proceed to show that the derivative of the sequence of partial sums converges uniformly on $[0, s_r]$. Below, we will verify the conditions for $f_1'(s)$, in which a simpler case will thus hold for $f_2'(s)$. The uniform convergence can be verified using Weierstrass M test by first computing

²“Large enough” can be understood in the sense that it is sufficient for our numerical simulations and that $s_r > s_0$.

$$\begin{aligned}
 & \sum_{(n+m) \in \mathbb{Z} \setminus \{0\}} \left| i^n J_n(-\beta) \hat{O}_m \left(\frac{(m+n)^2}{s^2} \right) e^{-\frac{(m+n)^2}{2s}} \right| \\
 & \leq \sum_{(m+n) \in \mathbb{Z} \setminus \{0\}} |J_n(-\beta)| \left((1 - \delta_{0m}) \frac{K}{|m|^{1+\alpha}} + \delta_{0m} |\hat{O}_0| \right) \frac{4}{(m+n)^2} \left(\frac{(m+n)^2}{2s} \right)^2 e^{-\frac{(m+n)^2}{2s}} \\
 & \leq 4 \left(\sum_{n \in \mathbb{Z}} |J_n(-\beta)| \right) \sum_{m \in \mathbb{Z}} \left((1 - \delta_{0m}) \frac{K}{|m|^{1+\alpha}} + \delta_{0m} |\hat{O}_0| \right) = M(\beta) [2K\zeta(1+\alpha) + |\hat{O}_0|] < +\infty, \quad (40)
 \end{aligned}$$

where we have used the following facts:

- (i) O is a $(1+\alpha)$ -Hölder class function for some $\alpha > 0$, for Fourier coefficients \hat{O}_n , where $n \neq 0$, and that $|\hat{O}_0|$ is bounded. These two cases are separated by using the Kronecker delta symbol δ_{0m} .
 - (ii) Since $m+n \neq 0$, then we have $\frac{4}{(m+n)^2} \leq 1$ for all $m, n \in \mathbb{Z}$.
 - (iii) $x^2 e^{-x} \leq \frac{4}{e^2} \leq 1$ for all $x \geq 0$.
 - (iv) $M(\beta) := \sum_{n=-\infty}^{\infty} |J_n(-\beta)| < +\infty$ for any $\beta > 0$.³
- With (40), the aforementioned interchange in (39) is justified. Furthermore, since $\lim_{s \rightarrow 0^+} f_1'(s) = 0$ and $\lim_{s \rightarrow 0^+} f_2'(s) = 0$, and $\lim_{s \rightarrow 0^+} f_1(s) < +\infty$ and $\lim_{s \rightarrow 0^+} f_2(s) < +\infty$, then we have $\frac{d}{ds} \langle O \rangle_s |_{s=0} = 0$.

Now, assume that (36) has been proven for all $k = 1, \dots, K$ for some $K > 0$. By the general Leibniz rule,

$$f_1^{(K+1)}(s) = \sum_{k=0}^{K+1} f_2^{(K+1-k)}(s) \frac{d^k}{d^k s} \langle O \rangle_s. \quad (41)$$

Using a similar logic as in (39), we can write down the higher order derivatives of f_1 and f_2 below:

$$\begin{aligned}
 f_1^{(q)}(s) &= \sum_{(n+m) \in \mathbb{Z} \setminus \{0\}} i^n J_n(-\beta) \hat{O}(m) h_{1,q}(s) e^{-\frac{(m+n)^2}{2s}} \quad \text{and} \\
 f_2^{(r)}(s) &= \sum_{n \in \mathbb{Z} \setminus \{0\}} i^n J_n(-\beta) h_{2,r}(s) e^{-\frac{n^2}{2s}}, \quad (42)
 \end{aligned}$$

where both $h_{1,k}(s)$ and $h_{2,k}(s)$ refer to polynomials in $\frac{1}{s}$ of degree $2k$. Note that the interchanges between the infinite sums and the q and r th order derivatives are still justified.

This is because, similar to (40), the presence of $e^{-\frac{(m+n)^2}{2s}}$ will always be able to overwhelm any polynomials in $\frac{1}{s}$ and create $\frac{1}{(m+n)^\gamma}$ for a sufficiently large γ . Taking the limit $s \rightarrow 0^+$ on both sides of (41), we obtain

$$0 = f_2(0) \lim_{s \rightarrow 0^+} \frac{d^{K+1}}{d^{K+1} s} \langle O \rangle_s. \quad (43)$$

³This can be observed from its asymptotic behavior for a large $|n|$, such that for a fixed β , $|J_n(-\beta)| \sim \frac{1}{\Gamma(|n|+1)} \left(\frac{|\beta|}{2}\right)^{|n|}$.

Thus (36) holds for $k = K + 1$ since $f_2(0) \neq 0$. By the principle of mathematical induction, (36) holds for all positive integer k . ■

Remark.—The proposition above indicates that the function $\langle O \rangle_s$ is not analytic at $s = 0$, which is essentially due to the fact that both integrals of $O(x) \exp(-S_s(x))$ and $\exp(-S_s(x))$ are not analytic at $s = 0$. To view the reason in a more straightforward way, we can write both integrands as $Q(x) \exp(-sx^2/2)$, where $Q(x)$ is a 2π -periodic function, whose Fourier transform is the sum of Dirac delta functions on integers. When s is small, the Fourier transform of $\exp(-sx^2/2)$, which has the form $s^{-1/2} \exp(-\xi^2/(2s))$, is a narrow Gaussian peaks at the origin. Thus, by taking the convolution, the Fourier transform of $Q(x) \exp(-sx^2/2)$ can be depicted as a number of Gaussians centered at all integer frequencies. Since the parameter s appears as the denominator inside the exponential function, the Fourier transform (and thus the integral) is obviously nonanalytic with respect to s .

Upon acknowledging the information presented in Proposition 3.5, we can investigate the structure of $\langle O \rangle_s$ presented in (37). From here, we can consider regressors in the form of a ratio of sum of exponential functions as summarized below:

Proposition 3.6. For any observable O satisfying the conditions as stated in Proposition 3.2, an appropriate rational representation would be

$$\langle O \rangle_s = \frac{\sum_{k=0}^{\infty} a_k e^{-\frac{k^2}{2s}}}{\sum_{k=0}^{\infty} b_k e^{-\frac{k^2}{2s}}}. \quad (44)$$

Proof.—This follows directly by considering all possible integer combinations of both the numerator and the denominator in (37). ■

Henceforth, we can use the rational representation in (44) and consider the following regression model:

$$\langle O \rangle_s = \frac{\sum_{k=0}^M a_k e^{-\frac{k^2}{2s}}}{1 + \sum_{k=1}^M b_k e^{-\frac{k^2}{2s}}} := \frac{A_M(s)}{B_M(s)}. \quad (45)$$

In the regression model above, the coefficients a_k and b_k are obtained by minimizing the objective function:

$$\operatorname{argmin}_{a_k, b_k} \int_{s_{\min}}^{s_{\max}} |A_M(s) - f(s)B_M(s)|^2 ds, \quad (46)$$

where s_{\min} and s_{\max} are the lower and upper bounds of s for which we can obtain the value of $\langle O \rangle_s$ via simulation, and $f(s)$ refers to a certain approximation of $\langle O \rangle_s$ for $s \in [s_{\min}, s_{\max}]$. In our experiments below, $f(s)$ is chosen as a polynomial of s and is obtained via least squares approximation.

D. Numerical results

In this subsection, we will attempt to include simulations and regressions conducted for $O(x) = e^{ix}$ with $\beta = 0.5$. Here, we note that, from [17], the exact value of $\langle e^{ix} \rangle$ at $\beta = 0.5$ is given by

$$\langle e^{ix} \rangle = \frac{I_1(-0.5i)}{I_0(-0.5i)} = 0 - 0.25815i.$$

First, we will attempt to obtain an estimate for $\langle e^{ix} \rangle$ through the use of regularization and regression. Note that we have shown that for $s \geq 0.8$, the numerical values

obtained from the complex Langevin method are accurate. In addition, as inspired from Fig. 1, we start to see a divergence between the true curve and the complex Langevin values at $s \approx 0.4$. Thus, we will attempt to include the complex Langevin values for two different cases, $s \geq 0.4$ and $s \geq 0.8$, and apply the regression model in Proposition 3.6 with different values of M for each case. The numerical values were generated using a fixed time step of $\Delta t = 10^{-4}$, with each sample obtained after every 2000 steps for a total of 10^6 samples for each value of $s \in \{0.01k | 40 \leq k \leq 150, k \in \mathbb{Z}\}$. Due to fluctuations present in the raw dataset, we have employed an interpolation using a quartic polynomial in s to average out the fluctuations prior to solving the optimization problem (46). Here, we note that this is consistent with the original regression model in Proposition 3.6, in which for s not close to 0, we do not have an issue with a flat curve as s approaches 0 and can therefore approximate such as expression with an appropriate polynomial. The relevant data points and regression curves are summarized in Fig. 3.

Next, we will illustrate the possible advantage obtained by reweighting our observables in accordance

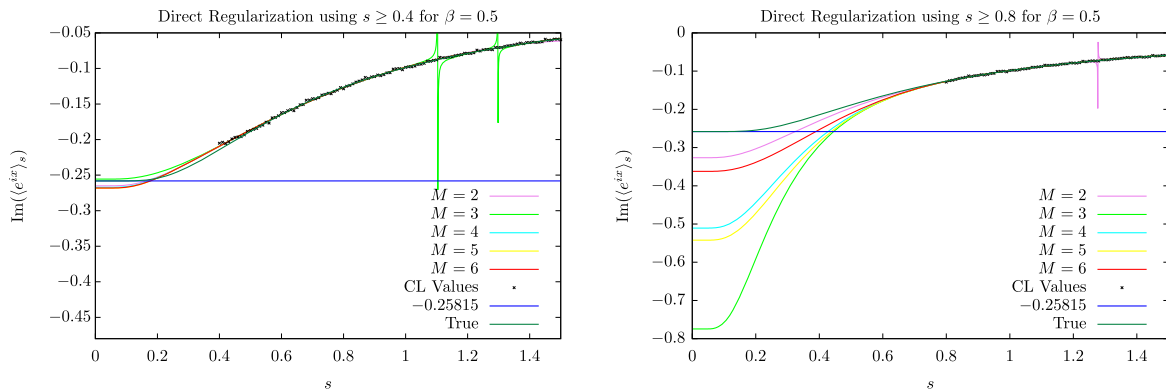


FIG. 3. Numerical results obtained by performing the relevant regression on observables via direct regularization. The left plot utilizes data points for $s \geq 0.4$ and the right plot utilizes data points for $s \geq 0.8$.

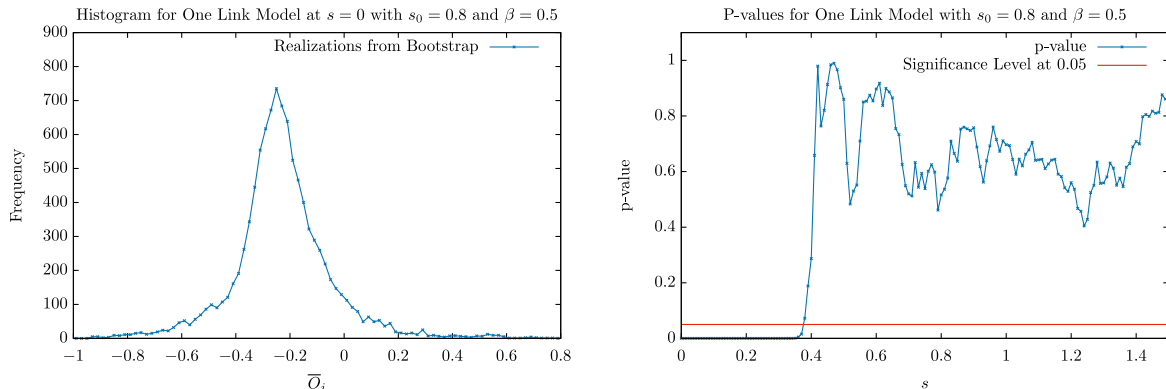


FIG. 4. The figure on the left depicts the histogram of the obtained distribution for \bar{O}_i at $s = 0$. The figure on the right corresponds to the corresponding p values obtained for $0 \leq s \leq 1.5$.

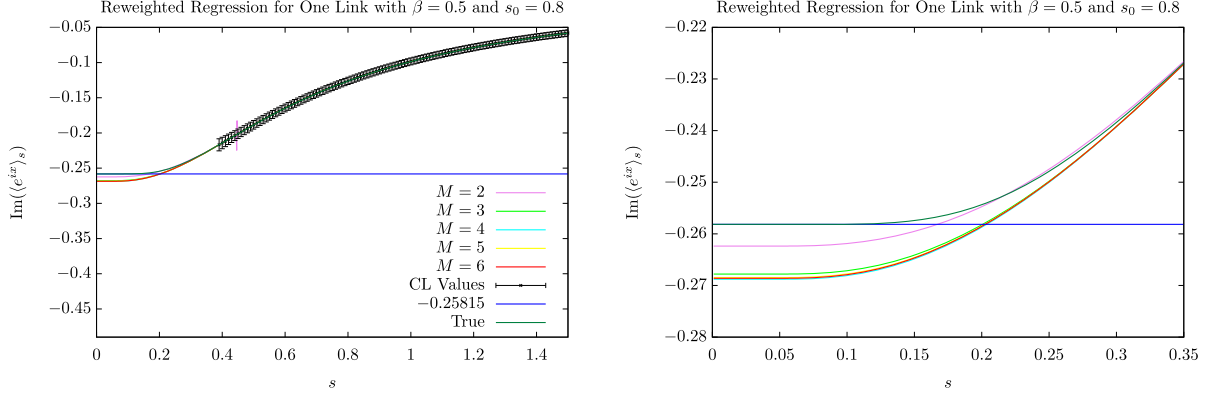


FIG. 5. Numerical results obtained by performing the relevant regression on reweighted observables. The plot on the right enlarges the behavior of curves of different M for small values of s and values of relevant observable close to -0.25815 .

to Proposition 3.4. As observed in Fig. 2, starting from $s = 0.4$, the corresponding standard error grows rapidly as s goes to 0. Thus, this portion of data may not be suitable to be used in the regression. We would therefore like to remove part of the information from our dataset. The criterion for this is based on the p value and will be described in the following paragraph.

The growth of the standard error as s approaches 0 can be explained as follows. First, we label each realization of the mean of the observable e^{ix} as \bar{O}_i for each iteration of the bootstrap method. Next, we obtain a histogram for the \bar{O}_i , as shown in the left diagram of Fig. 4. From there, we can observe that although the distribution looks somewhat symmetric and normal, the distribution of the \bar{O}_i seems to be concentrated more at its mean. We can support this with the use of a Kolmogorov-Smirnov test, conducted against a normal distribution at each value of s . If the p value at a given s happens to be below 0.05, then we will reject the null hypothesis that the underlying distribution for $\langle O \rangle_s = \langle e^{ix} \rangle_s$ is normal and concluding that the underlying distribution at that value of s is non-normal. Under the generalized central limit theorem, an instance in which a mean distribution converges to a non-normal distribution must correspond to the fact that the underlying population has an infinite variance. Thus, as seen from the right diagram of Fig. 4, we will perform regression using points generated for $s \geq 0.39$, which corresponds to values of s with p value greater than or equals to 0.05. The results are summarized below.

The numerical results for both methods, direct regression with regularized action (2R method) and regressing reweighted observables (3R method),⁴ are summarized in Table I. Note that we have excluded the estimates obtained by the 2R method using data points with $s \geq 0.8$ as we can

⁴The three Rs mentioned here correspond to regularization, regression, and reweighting. The missing R in the first method corresponds to regression done without any sort of reweighting.

see from Fig. 3 that the values of $\text{Im}\langle e^{ix} \rangle$ predicted for the given values of M are both inaccurate and imprecise.

We summarize some of the relevant key observations from Figs. 3–5 and Table I below.

- (i) From Table I, for a fixed $M \geq 3$, we can see that both methods are on par in terms of their accuracy in estimating the value of $\text{Im}\langle e^{ix} \rangle$.
- (ii) However, the matched performance of the 2R Method for $s \geq 0.4$ as mentioned is on top of the fact that we have used *a priori* information on the divergence of the numerical values generated using our complex Langevin algorithm as in Fig. 1. This piece of information might not be available for general $U(1)$ models such as the 3D XY Model. Alternatively, one might be able to obtain such information via other methods, such as the analysis of boundary terms as in [17,26].
- (iii) In the absence of *a priori* information, as mentioned, the results obtained via the 2R method for $s \geq 0.8$ are both inaccurate and imprecise. This can be already be seen from the right plot in Fig. 3. Therefore, to get better results using the 2R method, one may consider including some biased results with acceptable errors [e.g., $\langle O \rangle_s$ with $s \in (0.4, 0.8)$ in this example].
- (iv) In addition, we can see from Fig. 1 that the true curve (for $\text{Im}\langle e^{ix} \rangle_s$ as a function of s) diverges from

TABLE I. Estimates of $\text{Im}\langle e^{ix} \rangle$ for various methods at $\beta = 0.5$. Note that the % discrepancy is calculated with respect to the true value of -0.25815 .

M	Direct Reg $s \geq 0.4$ (Fig. 3; 2R Method)	% Disp	Reweighted (Fig. 5; 3R Method)	% Disp
2	-0.265004	2.7	-0.262370	1.6
3	-0.255572	1.0	-0.267792	3.7
4	-0.267196	3.5	-0.268786	4.1
5	-0.268420	4.0	-0.268468	4.0
6	-0.268089	3.9	-0.268611	4.1

the numerical value at $s \approx 0.4$, which is consistent with our non-normality test as explained for Fig. 4. As of now, we are not sure if this is a coincidence or if there are sufficient mathematical grounds to justify such a phenomenon.

Remark.—In theory, only samples drawn from populations of infinite variance will give rise to stable distributions (instead of normal distributions) under the limit of a large sample size. However, depending on the type of bootstrapping method used, there are other types of distribution in which a simple naïve bootstrapping method, like the one we have employed, might fail. Furthermore, as shown in [42], even for population distributions with infinite population variance that are suitably well behaved, we note that the resulting distribution might not be that of a stable distribution, but rather, a random distribution. Nonetheless, the success of the naïve bootstrapping method for large values of s indicates that we might not face issues that we will with standard counterexamples to naïve bootstrapping methods, such as the extreme order statistics but, rather, the issue can be attributed to infinite population variance.

IV. GENERALIZATION TO MULTIDIMENSIONAL INTEGRALS IN THE $U(1)$ THEORY

In this section, we will attempt to extend the method used for the $U(1)$ one-link model to integrals on \mathbb{T}^N , $N \geq 1$ with more general actions. This includes the $U(1)$ lattice field theory, where N equals the number of links. As in Sec. II, we will use $\{\phi\}$ to denote the collection of all the N variables ϕ_1, \dots, ϕ_N , and each ϕ_k is a variable in \mathbb{T} . The action and the observable will be denoted by $S(\{\phi\})$ and $O(\{\phi\})$, respectively. To extend the regularization to multidimensional models, we adopt the uniform regularization in all directions, that is, the regularized observable $\langle O \rangle_s$ is defined by

$$\langle O \rangle_s = \frac{1}{Z_s} \int_{\mathbb{R}^N} O(\{\phi\}) \times \exp \left(-S(\{\phi\}) - \frac{s}{2} \sum_{j=1}^N \phi_j^2 \right) d\phi_1 \cdots d\phi_N, \quad (47)$$

with Z_s defined by

$$Z_s = \int_{\mathbb{R}^N} \exp \left(-S(\{\phi\}) - \frac{s}{2} \sum_{j=1}^N \phi_j^2 \right) d\phi_1 \cdots d\phi_N. \quad (48)$$

The extension to multidimensional models is done by generalizing Propositions 3.2–3.6 whenever possible, giving the necessary proofs unless a given proposition generalizes clearly.

We start off by generalizing Proposition 3.2 as follows:

Proposition 4.1. Suppose that both $e^{-S(\{\phi\})}$ and $O(\{\phi\})$ are C^∞ functions on \mathbb{T}^N . Then it holds that

$$\lim_{s \rightarrow 0^+} \langle O \rangle_s = \langle O \rangle.$$

Remark.—The smoothness of $O(\{\phi\})$ and $e^{-S(\{\phi\})}$ guarantees the interchangeability of the infinite sum and the integrals. Similar to the one-dimensional case (Proposition 3.2), the C^∞ requirement can be weakened to a certain Hölder class. However, this C^∞ condition is satisfied in most applications of the complex Langevin method, due to the analytic extensibility of both functions to the complexified space.

Proof.—The proof will largely mimic that of Proposition 3.2. We can write down the multidimensional Fourier series of $O(\{\phi\})$ and $e^{-S(\{\phi\})}$ in the following form:

$$\begin{aligned} O(\{\phi\}) &= \sum_{\vec{m}=(m_1, \dots, m_N) \in \mathbb{Z}^N} \hat{O}_{\vec{m}} \exp \left(i \sum_{j=1}^N m_j \phi_j \right), \\ e^{-S(\{\phi\})} &= \sum_{\vec{k}=(k_1, \dots, k_N) \in \mathbb{Z}^N} \hat{\beta}_{\vec{k}} \exp \left(i \sum_{j=1}^N m_j \phi_j \right). \end{aligned} \quad (49)$$

In particular, analogous to Proposition 3.2, the C^∞ regularity allows for the interchange of the relevant infinite sum and the N -dimensional integral. Thus, it is sufficient to only consider integrals in the following form:

$$\begin{aligned} & \int_{\mathbb{R}^N} \exp \left(i \sum_{j=1}^N m_j \phi_j \right) \exp \left(-S(\{\phi\}) - \frac{s}{2} \sum_{j=1}^N \phi_j^2 \right) d\phi_1 \cdots d\phi_N \\ &= \int_{\mathbb{R}^N} \left[\sum_{\vec{k}=(k_1, \dots, k_N) \in \mathbb{Z}^N} \hat{\beta}_{\vec{k}} \exp \left(i \sum_{j=1}^N k_j \phi_j \right) \right] \left(\prod_{j=1}^N e^{i m_j \phi_j - \frac{s}{2} \phi_j^2} \right) d\phi_1 \cdots d\phi_N, \\ &= \sum_{\vec{k}=(k_1, \dots, k_N) \in \mathbb{Z}^N} \hat{\beta}_{\vec{k}} \int_{\mathbb{R}^N} \prod_{j=1}^N e^{i(m_j+k_j)\phi_j - \frac{s}{2} \phi_j^2} d\phi_1 \cdots d\phi_N, \\ &= \left(\frac{2\pi}{s} \right)^{N/2} \sum_{\vec{k}=(k_1, \dots, k_N) \in \mathbb{Z}^N} \hat{\beta}_{\vec{k}} \exp \left(\frac{1}{2s} \sum_{j=1}^N (m_j + k_j)^2 \right). \end{aligned} \quad (50)$$

Here we have again interchanged the summation and the integral, which can be justified by an argument similar to (18) using the smoothness of $e^{-S(\{\phi\})}$. Thus, we have

$$\langle O \rangle_s = \frac{1}{Z_s} \left(\frac{2\pi}{s} \right)^{N/2} \sum_{\vec{m} \in \mathbb{Z}^N} \hat{O}_{\vec{m}} \sum_{\vec{k} \in \mathbb{Z}^N} \hat{\beta}_{\vec{k}} \exp\left(-\frac{1}{2s} \sum_{j=1}^N (m_j + k_j)^2\right), \quad Z_s = \left(\frac{2\pi}{s} \right)^{N/2} \sum_{\vec{k} \in \mathbb{Z}^N} \hat{\beta}_{\vec{k}} \exp\left(-\frac{1}{2s} \sum_{j=1}^N k_j^2\right), \quad (51)$$

from which we can obtain the limit

$$\lim_{s \rightarrow 0^+} \langle O \rangle_s = \frac{1}{\hat{\beta}_0} \sum_{\vec{m} \in \mathbb{Z}^N} \hat{O}_{\vec{m}} \hat{\beta}_{-\vec{m}}. \quad (52)$$

It remains to show that the right-hand side of (52) equals $\langle O \rangle$. To this end, we write $\langle O \rangle$ as

$$\langle O \rangle = \frac{(2\pi)^{-N} \int_{\mathbb{T}^N} O(\{\phi\}) e^{-S(\{\phi\})} d\{\phi\}}{(2\pi)^{-N} \int_{\mathbb{T}^N} e^{-S(\{\phi\})} d\{\phi\}}. \quad (53)$$

By definition, it is clear that β_0 is the (unweighted)⁵ mean value of $e^{-S(\{\phi\})}$, which equals the denominator in (53). Since

$$O(\{\phi\}) e^{-S(\{\phi\})} = \left(\sum_{\vec{k} \in \mathbb{Z}^N} \hat{O}_{\vec{k}} \prod_{j=1}^N e^{ik_j \phi_j} \right) \left(\sum_{\vec{k} \in \mathbb{Z}^N} \hat{\beta}_{\vec{k}} \prod_{j=1}^N e^{ik_j \phi_j} \right) = \sum_{\vec{k} \in \mathbb{Z}^N} \left(\sum_{\vec{m} \in \mathbb{Z}^N} \hat{O}_{\vec{m}} \hat{\beta}_{\vec{k}-\vec{m}} \right) \prod_{j=1}^N e^{ik_j \phi_j}, \quad (54)$$

we see that the right-hand side of (52) is the Fourier coefficient of the zero-frequency mode in the expansion of $O(\{\phi\}) e^{-S(\{\phi\})}$, and thus corresponds to the numerator in (53). We have thus completed the proof. ■

Next, we note the range of $s(\beta)$ depends largely on the functional form of $S(\{\phi\})$. Thus, Proposition 3.3 does not generalize easily. Instead, we will proceed to generalize Proposition 3.4. However, this is straightforward, as we only have to replace $|x|^2$ with the sum of $|\phi_j|^2$, as elaborated below:

Proposition 4.2. The following equality holds for all $s, s_0 \geq 0$:

$$\langle O(\{\phi\}) \rangle_s = \frac{\langle O(\{\phi\}) \exp\left(\frac{s_0-s}{2} \sum_{j=1}^N |\phi_j|^2\right) \rangle_{s_0}}{\langle \exp\left(\frac{s_0-s}{2} \sum_{j=1}^N |\phi_j|^2\right) \rangle_{s_0}}. \quad (55)$$

As the proof is completely analogous to that in Proposition 3.4, we shall skip the proof.

We can also observe a similar phenomenon as Proposition 3.5 for all possible $S(\{\phi\})$ satisfying the relevant conditions for uniform convergence of their corresponding Fourier series. In particular, the following proposition generalizes this.

⁵Here, “unweighted” mean of an expression f refers to $\langle f \rangle_{\text{unweighted}} = \frac{\int_D f(x) dx}{\int_D 1 dx}$, where D is the appropriate domain for consideration. One can compare this expression to the expression for weighted means as in (47) and (48).

Proposition 4.3. Let the action $S(\{\phi\})$ and any observable O satisfying the conditions in Proposition 4.1 be given. For any given $k \in \mathbb{Z}^+$, we have

$$\frac{d^k}{ds^k} \langle O \rangle_s \Big|_{s=0} = 0. \quad (56)$$

Proof.—As the proof for this proposition is largely similar to that in Proposition 3.5, we will sketch the key differences here. From (51), we have $\langle O \rangle_s = f_1(s)/f_2(s)$, where

$$f_1(s) = \sum_{\vec{m} \in \mathbb{Z}^N} \hat{O}_{\vec{m}} \sum_{\vec{k} \in \mathbb{Z}^N} \hat{\beta}_{\vec{k}} \exp\left(-\frac{1}{2s} \sum_{j=1}^N (m_j + k_j)^2\right),$$

$$f_2(s) = \sum_{\vec{k} \in \mathbb{Z}^N} \hat{\beta}_{\vec{k}} \exp\left(-\frac{1}{2s} \sum_{j=1}^N k_j^2\right). \quad (57)$$

The structure of this function is similar to (37), and we can use exactly the same approach to prove that all the derivative at $s = 0$ vanishes. Here the interchangeability of the derivative and the series follows the fact that both O and S are infinitely, so that the coefficients $\hat{O}_{\vec{k}}$ and $\hat{\beta}_{\vec{k}}$ decay faster than $1/|\vec{k}|^\alpha$ for any $\alpha > 0$. ■

Based on the expression (57), it is straightforward to derive the following proposition, which is similar to Proposition 3.6.

Proposition 4.4. For any observable O satisfying the conditions as stated in Proposition 4.1, there exist coefficients a_k and b_k for all $k \in \mathbb{N}$ such that

$$\langle O \rangle_s = \frac{\sum_{k=0}^{\infty} a_k e^{-\frac{k}{2s}}}{\sum_{k=0}^{\infty} b_k e^{-\frac{k}{2s}}}. \quad (58)$$

Note that a key difference between (57) and (37) is that the sum over the components $\sum_{j=1}^N$ does not exist in the one-link model. In fact, the integer $\sum_{j=1}^N k_j^2$ can take any value from 0 to N , so that when N is large, the zero terms in both summations in (58) will appear only for a large k . To accommodate for such cases, we have included all possible integer coefficients of $-\frac{1}{2s}$ instead as shown in (58).

In view of the proposition above, an appropriate regression model is given by

$$\langle O \rangle_s = \frac{\sum_{k=0}^M a_k e^{-\frac{k}{2s}}}{1 + \sum_{k=1}^M b_k e^{-\frac{k}{2s}}}. \quad (59)$$

Note that for any model under the $U(1)$ lattice field theory framework, similar to the regression model in (45), we can minimize a similar objective function as described in (46) to obtain the corresponding regression coefficients.

V. EXTENSION TO THE $SU(n)$ THEORY

This section introduces how we can extend results from the previous sections to the $SU(n)$ theory, and presents results for the one-dimensional problems. Before we introduce the regularization method in $SU(n)$ theory, we will briefly review the complex Langevin method under the $SU(n)$ gauge theory.

Consider the set $\{U_k : k = 1, \dots, N\} \in [SU(n)]^N$, where N is the total number of lattice points. Given the action $S(\{U\})$, the expectation value for the observable $O(\{U\})$ is given by

$$\langle O \rangle = \frac{\int_{[SU(n)]^N} O(\{U\}) \exp(-S(\{U\})) d\{U\}}{\int_{[SU(n)]^N} \exp(-S(\{U\})) d\{U\}}. \quad (60)$$

Let $w_{a,k}$ for $a = 1, \dots, n^2 - 1$ be independent Brownian motions. Upon complexifying the configuration space $[SU(n)]^N$ to $[SL(n, \mathbb{C})]^N$, the complex Langevin method for $SU(n)$ theory can be described by the complex stochastic process:

$$dU_k = - \sum_{a=1}^{n^2-1} i\lambda_a [U_k D_{a,k} S(\{U\}) dt + U_k \circ dw_{a,k}],$$

$$U_k \in SL(n, \mathbb{C}), \quad k = 1, \dots, N, \quad (61)$$

where \circ stands for the Stratonovich interpretation of the stochastic integral and λ_a , $a = 1, \dots, n^2 - 1$ are the infinitesimal generators of the $SU(n)$ group satisfying the orthogonality

$$\text{tr}(\lambda_a \lambda_b) = 2\delta_{ab}, \quad \forall a, b = 1, \dots, n^2 - 1,$$

and $D_{a,k}$ denotes the left Lie derivative operator defined as

$$D_{a,k} S(\{U\}) = \lim_{\epsilon \rightarrow 0} \frac{S(\{\tilde{U}^\epsilon\}) - S(\{U\})}{\epsilon}, \quad (62)$$

where $\{\tilde{U}^\epsilon\}$ denotes the field with links defined by

$$\tilde{U}_l^\epsilon = \exp(i\epsilon \delta_{kl} \lambda_a) U_l, \quad l = 1, \dots, N.$$

To solve (61), we mimic standard methods and apply the following scheme to update the links as follows:

$$U_k^{(j+1)} = \exp\left(- \sum_{a=1}^{n^2-1} i\lambda_a (D_{a,k} S^{(j)} \Delta t + \eta_{a,k} \sqrt{\Delta t})\right) U_k^{(j)},$$

$$k = 1, \dots, N, \quad (63)$$

where $U_k^{(j)}$ is the link at time instance t_j , $\Delta t = t_{j+1} - t_j$ is time step and each $\eta_{a,k}$ is normally distributed with mean 0 and variance 2. The scheme is similar to the Euler-Maruyama method, while the exponential map is applied to keep the solution from leaving the Lie group.

Below, we will generalize the techniques of regularization, reweighting, and regression to the $SU(n)$ group theory. They will be introduced in the following three subsections.

A. Regularization

To demonstrate how the regularization can be generalized to the $SU(n)$ theory, we would like to restate the method for the $U(1)$ theory in the language of group theories. Here we regard $U(1)$ as the unit circle on the complex plane, so that we can establish the one-to-one map between $U \in U(1)$ and $x \in \mathbb{T}$ by $U = \exp(ix)$. Thus, we can write down the integrals over \mathbb{T} as integrals over $U(1)$. For instance,

$$Z = \int_{\mathbb{T}} \exp(-S(x)) dx$$

$$= \int_{U(1)} \exp(-S(-i \log U)) dU, \quad (64)$$

where we have assumed that $S(\cdot)$ is periodic with period 2π , so that the value of $S(-i \log U)$ is unique. In Eq. (64), we can also consider $\log U$ as an element in the Lie algebra of $U(1)$, i.e., $\mathfrak{g} = i\mathbb{R}$. When we apply the regularization, an extra term $sx^2/2$ is added to the action, whose counterpart should be $-s(\log U)^2/2$ if we represent the regularization term using the variable in $U(1)$. However, since $sx^2/2$ is no longer periodic with respect to x , the expression $-s(\log U)^2/2$ becomes multivalued, so simply adding this term to the action will cause ambiguities.

To address this problem, we can rewrite (11) and (12) in the following form:

$$\begin{aligned} \langle O \rangle_s &= \frac{1}{Z_s} \int_{U(1)} O(-i \log U) \sum_{\substack{g \in \mathfrak{g} \\ \text{s.t. } \exp(g) = U}} \exp\left(-S(-i \log U) + \frac{sg^2}{2}\right) dU, \\ Z_s &= \int_{U(1)} \sum_{\substack{g \in \mathfrak{g} \\ \text{s.t. } \exp(g) = U}} \exp\left(-S(-i \log U) + \frac{sg^2}{2}\right) dU. \end{aligned} \tag{65}$$

Here the summation is taken over all logarithms of U , corresponding to unrolling the torus \mathbb{T} to the real axis \mathbb{R} . Our generalization of the regularization to the $SU(n)$ theory will be based on the form of (65).

Formally, for the $SU(n)$ theory, the regularized action can be written as

$$S_s(\{U\}) = S(\{U\}) - \frac{s}{2} \sum_{k=1}^N \text{tr}[(\log U_k)^2], \tag{66}$$

where $\log U_k$ is an element in the Lie algebra $\mathfrak{su}(n)$. Here the regularization term is chosen as the trace of the matrix square since its square root defines a norm on $\mathfrak{su}(n)$. However, the ambiguity again comes from the nonuniqueness of the matrix logarithm. Therefore, we mimic (65) to write down the regularized observable as

$$\begin{aligned} \langle O \rangle_s &= \frac{1}{Z_s} \int_{[SU(n)]^N} O(\{U\}) \sum_{\substack{g_1 \in \mathfrak{su}(n) \\ \text{s.t. } \exp(g_1) = U_1}} \cdots \sum_{\substack{g_N \in \mathfrak{su}(n) \\ \text{s.t. } \exp(g_N) = U_N}} \exp\left(-S(\{U\}) + \sum_{k=1}^N \frac{s}{2} \text{tr}(g_k^2)\right) dU, \\ Z_s &= \int_{[SU(n)]^N} \sum_{\substack{g_1 \in \mathfrak{su}(n) \\ \text{s.t. } \exp(g_1) = U_1}} \cdots \sum_{\substack{g_N \in \mathfrak{su}(n) \\ \text{s.t. } \exp(g_N) = U_N}} \exp\left(-S(\{U\}) + \sum_{k=1}^N \frac{s}{2} \text{tr}(g_k^2)\right) dU. \end{aligned} \tag{67}$$

Following the idea in Proposition 3.2, it is expected that when $s \rightarrow 0^+$, the regularized observable $\langle O \rangle_s$ will converge to $\langle O \rangle$ in (60). We leave the rigorous proof as our further work and we focus on the implementation of the regularized method for $SU(n)$ theory in this section.

To formulate the complex Langevin method for the complexified action (66), we need to first formulate the Langevin method for real actions. The general idea of the numerical scheme follows (63), while the derivative of the action $D_{a,k} S$ should be replaced with $D_{a,k} S_s$ defined by

$$\begin{aligned} D_{a,k} S_s(\{U\}) &= D_{a,k} S(\{U\}) \\ &\quad - \frac{s}{2\epsilon} \lim_{\epsilon \rightarrow 0} \frac{\text{tr}[(\log(e^{i\lambda_a \epsilon} U_k))^2 - (\log U_k)^2]}{\epsilon}. \end{aligned} \tag{68}$$

The numerical scheme is fully determined once the Lie derivative is determined. Meanwhile, the generalization to the complex action follows naturally as the formula of the numerical scheme is unchanged, and each link U_k automatically falls into the complexification of $SU(n)$, i.e., the special linear group $SL(n, \mathbb{C})$. Below, we will focus on the computation of (68) with $U_k \in SL(n, \mathbb{C})$ and resolve the complication caused by the multivalued logarithmic function.

In (68), $g_k := \log U_k$ can take any matrix logarithm of U_k , while $\log(e^{i\lambda_a \epsilon} U_k)$ must be the matrix logarithm that is

closest to g_k such that the limit exists. The value of the limit is given in the following proposition:

Proposition 5.1. For any $U_k \in SL(n, \mathbb{C})$, it holds that

$$\lim_{\epsilon \rightarrow 0} \frac{\text{tr}[(\log(e^{i\lambda_a \epsilon} U_k))^2 - (\log U_k)^2]}{\epsilon} = 2i \text{tr}(\lambda_a \log U_k).$$

Proof.—Since $\text{tr}(A^2 - B^2) = \text{tr}(A + B)(A - B)$ for any matrices A and B , we can simplify the limit as follows:

$$\begin{aligned} &\lim_{\epsilon \rightarrow 0} \frac{\text{tr}[(\log(e^{i\lambda_a \epsilon} U_k))^2 - (\log U_k)^2]}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0} \text{tr} \left([\log(e^{i\lambda_a \epsilon} U_k) + \log U_k] \frac{\log(e^{i\lambda_a \epsilon} U_k) - \log U_k}{\epsilon} \right), \\ &= 2 \lim_{\epsilon \rightarrow 0} \text{tr} \left(\log U_k \frac{\log(e^{i\lambda_a \epsilon} U_k) - \log U_k}{\epsilon} \right). \end{aligned} \tag{69}$$

Let $\Delta_1^\epsilon = \log(e^{i\lambda_a \epsilon} U_k) - \log U_k$ and

$$\begin{aligned} \Delta_{i+1}^\epsilon &= [\log U_k, \Delta_i^\epsilon] = (\log U_k) \Delta_i^\epsilon - \Delta_i^\epsilon (\log U_k), \\ & \quad i = 1, 2, \dots \end{aligned} \tag{70}$$

It can be derived from the cyclic property of the matrix trace that

$$\begin{aligned} \text{tr}[(\log U_k) \Delta_{i+1}^\epsilon] &= \text{tr}[(\log U_k)^2 \Delta_i^\epsilon - (\log U_k) \Delta_i^\epsilon (\log U_k)] = 0, \\ i &= 1, 2, \dots \end{aligned} \quad (71)$$

According to the differential formula of the exponential map [43, Theorem 5], we have

$$\begin{aligned} \sum_{i=1}^{+\infty} \frac{(-1)^{i+1}}{i!} \Delta_i^\epsilon &= U_k^{-1} e^{i\lambda_a \epsilon} U_k - I + \mathcal{O}(\epsilon^2) \\ &= i\epsilon U_k^{-1} \lambda_a U_k + \mathcal{O}(\epsilon^2). \end{aligned}$$

Using (71), we can left multiply the above equation by $\log U_k$ and then take the trace to obtain

$$\begin{aligned} \text{tr}[(\log U_k) \Delta_1^\epsilon] &= i\epsilon \text{tr}[(\log U_k) U_k^{-1} \lambda_a U_k] + \mathcal{O}(\epsilon^2) \\ &= i\epsilon \text{tr}(\lambda_a \log U_k) + \mathcal{O}(\epsilon^2). \end{aligned} \quad (72)$$

Inserting this equation into (69) concludes the proof. ■
By this proposition, the drift term (68) becomes

$$D_{a,k} S_s(\{U\}) = D_{a,k} S(\{U\}) - i s \text{tr}(\lambda_a \log U_k), \quad (73)$$

based on which the update of the links (63) becomes

$$U_k^{(j+1)} = \exp\left(-\sum_{a=1}^{n^2-1} i\lambda_a (D_{a,k} S^{(j)} \Delta t - \underline{i s \text{tr}(\lambda_a \log U_k^{(j)}) \Delta t} + \eta_{a,k} \sqrt{\Delta t})\right) U_k^{(j)}, \quad k = 1, \dots, N, \quad (74)$$

where the underlined term produces a pullback velocity so that the excursion away from $SU(n)$ can be restricted.

We now consider the determination of $g_k = \log U_k$. We first assume that U_k is diagonalizable, i.e., $U_k = R\Theta R^{-1}$ for some $R \in SL(n, \mathbb{C})$ and $\Theta = \text{diag}(\theta_1, \dots, \theta_n)$. Then

$$g_k = \log U_k = \log(R\Theta R^{-1}) = R(\log \Theta) R^{-1}. \quad (75)$$

The nonuniqueness of g_k comes from the nonuniqueness of $\log \Theta$. If $\Xi = \text{diag}(\xi_1, \dots, \xi_n)$ is a diagonal matrix satisfying

$$\exp(\xi_i) = \theta_i, \quad i = 1, \dots, n \quad \text{and} \quad \sum_{i=1}^n \xi_i = 0, \quad (76)$$

then for any $K_1, \dots, K_n \in \mathbb{Z}$ satisfying $K_1 + \dots + K_n = 0$, we have

$$\Xi_k^{(j+1)} = \text{argmin}_{\Xi \in \mathbb{C}^{n \times n}} \|\Xi - \Xi_k^{(j)}\|_F,$$

$$\text{s.t. } \Xi \text{ is a Jordan normal form and } \exists R \in \mathbb{C}^{n \times n} \text{ such that } \exp(R\Xi R^{-1}) = U_k^{(j+1)}. \quad (78)$$

When $j=0$, we simply choose $\Xi_k^{(0)}$ such that the imaginary parts of all its diagonal elements locate in $[-\pi, \pi)$. Here, we comment that due to the existence of the stochastic term, $\Xi_k^{(j+1)}$ may be distant from $\Xi_k^{(j)}$. Thus, our strategy does not produce the ‘‘correct’’ choice of $\Xi_k^{(j+1)}$. However, such a probability decreases exponentially as Δt decreases, which will not affect the order of accuracy for the numerical method. Note that in this approach, we need to keep track of the evolution of $\Xi_k^{(j)}$,

$$\exp(\Xi + 2\pi i \Gamma) = \Theta, \quad (77)$$

where $\Gamma = \text{diag}(K_1, \dots, K_n)$. Hence, the matrix $R(\Xi + 2\pi i \Gamma) R^{-1}$ is a candidate of g_k . Note that here we require $\text{tr} \Xi = \text{tr} \Gamma = 0$ since g_k is required to be an element in $\mathfrak{sl}(n, \mathbb{C})$, which contains all traceless $n \times n$ matrices. If U_k is nondiagonalizable, then Θ will contain Jordan blocks and Ξ becomes an upper-triangular matrix. Nonetheless, the relation (77) still holds and still plays the role that leads to multiple values of matrix logarithms.

To resolve this issue, we would like to determine a specific $\Xi_k^{(j)}$ for each j and k such that the diagonal of $\exp(\Xi_k^{(j)})$ consists of all the eigenvalues of U_k , which will further determine the matrix logarithm. In order to maintain the continuity of the dynamics, we choose $\Xi_k^{(j+1)}$ by minimizing its difference from $\Xi_k^{(j)}$. In other words, we determine $\Xi_k^{(j+1)}$ via

which guarantees that the summations in (67) are taken into account.

To summarize, we describe the algorithm of the complex Langevin method for the $SU(n)$ theory below.

Like in the regularized $U(1)$ theory, when the regularization parameter s is small, the regularized method may still converge to biased results. Our improvements, including reweighting and regularization, will be studied in the following two subsections.

B. Reweighting

The idea of reweighting follows from our motivating example, as in (35). We represent the observable $\langle O \rangle_s$ as

$$\langle O \rangle_s = \frac{\langle O \exp(S_{s_0} - S_s) \rangle_{s_0}}{\langle \exp(S_{s_0} - S_s) \rangle_{s_0}}, \quad (79)$$

where

$$S_{s_0} - S_s = -\frac{s_0 - s}{2} \sum_{k=1}^N \text{tr}[(\log U_k)^2] \quad (80)$$

according to the definition (66). To compute the numerator and the denominator of (79), Algorithm 1 can still be applied, and the logarithms in (80) can still be found by using $\Xi_k^{(j)}$ at each time step [see (78)]. However, when N is large, the value of $S_{s_0} - S_s$ in (80) might be a large positive number if $s_0 > s$ [note that $\text{tr}[(\log U_k)^2] < 0$]. Consequently, its exponent $\exp(S_{s_0} - S_s)$ will be so huge that it will be difficult to handle using double-precision floating-point numbers. Therefore, we shift $S_{s_0} - S_s$ by its average, so that (79) can be computed as

$$\langle O \rangle_s = \frac{\langle O \exp(S_{s_0} - S_s - \langle S_{s_0} - S_s \rangle_{s_0}) \rangle_{s_0}}{\langle \exp(S_{s_0} - S_s - \langle S_{s_0} - S_s \rangle_{s_0}) \rangle_{s_0}}. \quad (81)$$

This requires us to record the values of O and $S_{s_0} - S_s$ for each sample, so that we can first compute $\langle S_{s_0} - S_s \rangle$, and then use the result to evaluate (81). In our tests, the magnitude of the shifted exponents turns out to be acceptable after applying such a trick.

C. Regression

For the $U(1)$ theory, the expression we used in the regression has a fractional form (58), which is derived based on the Fourier expansion of functions on $U(1)$. Similarly, for the $SU(n)$ theory, suppose that ψ_k , $k = 1, 2, \dots$ form an orthonormal set of basis:

Algorithm 1. Complex Langevin method for the regularized action.

Input: Initial field $\{U^{(0)}\}$, time step Δt , total number of time steps J , number of time steps to reach the invariant measure J_0 , number of time steps between two samples ΔJ

- 1 Set $j \leftarrow 0$, $N_{\text{sample}} \leftarrow 0$ and $\langle O \rangle_s \leftarrow 0$;
 - 2 For each k , let Ξ_k be the diagonal matrix with diagonal elements being the logarithms of the eigenvalues of $U_k^{(0)}$, and compute its logarithm $\log U_k^{(0)} = R \Xi_k^{(0)} R^{-1}$;
 - 3 **for** $j \leftarrow 0$ **to** J **do**
 - 4 | Use the result of $\log U_k^{(0)}$ to evolve the solution by (74);
 - 5 | **if** $j > J_0$ **then**
 - 6 | | $\langle O \rangle_s \leftarrow \langle O \rangle_s + O(\{U^{(j+1)}\})$;
 - 7 | | $J_0 \leftarrow J_0 + \Delta J$, $N_{\text{sample}} \leftarrow N_{\text{sample}} + 1$;
 - 8 | -- Compute $\Xi_k^{(j+1)}$ for each k according to (78), and find $\log U_k^{(j+1)}$ based on the result of $\Xi_k^{(j+1)}$;
 - 9 $\langle O \rangle_s \leftarrow \langle O \rangle_s / N_{\text{sample}}$;
- Output:** $\langle O \rangle_s$

$$\int_{SU(n)} \psi_k(U) \psi_l(U) dU = \delta_{kl}. \quad (82)$$

Then the expression used in the regularization can be determined by computing

$$\int_{SU(n)} \psi_k(U) \sum_{\substack{g \in \mathfrak{su}(n) \\ \text{s.t. } \exp(g) = U}} \exp\left(\frac{s}{2} \text{tr} g^2\right) dU, \quad k = 1, 2, \dots \quad (83)$$

For the $SU(2)$ theory, the integral (83) can be calculated using the isomorphism between $SU(2)$ and the three-sphere \mathbb{S}^3 . The basis functions $\psi_k(U)$ can be chosen as the generalized spherical harmonics defined on \mathbb{S}^3 , which are denoted by $Y_{l_1 l_2 l_3}(\psi, \theta, \varphi)$ with (ψ, θ, φ) being the hyperspherical coordinates. The precise form of $Y_{l_1 l_2 l_3}$ can be found in [44]. To compute the integral, we write (83) as an integral on the three-dimensional linear space $\mathfrak{su}(2)$. With proper changes of variables, the integral (83) can be transformed into

$$I_{l_1 l_2 l_3} := \sqrt{\frac{2l_2 + 1}{4\pi} \frac{(l_2 + l_1)!}{(l_2 - l_1)!}} \int_{\mathbb{R}} \int_{\mathbb{R}} \int_{\mathbb{R}} \left(\frac{g_1 + ig_2}{\sqrt{g_1^2 + g_2^2}} \right)^{l_1} P_{l_2}^{-l_1} \left(\frac{g_3}{|g|} \right) \frac{Q_{l_3}^{l_2}(\sin |g|, \cos |g|)}{|g| \sqrt{g_1^2 + g_2^2}} \exp(-s|g|^2) dg_1 dg_2 dg_3, \quad (84)$$

where $|g| = \sqrt{g_1^2 + g_2^2 + g_3^2}$, and $P_{l_2}^{-l_1}(\cdot)$ is the associated Legendre function. Here g_1, g_2, g_3 can be considered as the coefficients of the Pauli matrices $\sigma_1, \sigma_2, \sigma_3$ when representing the elements in $\mathfrak{su}(2)$. Since $\text{tr}(\sigma_i \sigma_j) = 2\delta_{ij}$, we get s instead of $s/2$ as the parameter in the exponent of (84), which differs slightly from the $U(1)$ theory. The polynomial $Q_{l_3}^{l_2}(\cdot, \cdot)$, $l_3 \geq l_2 \geq 0$ has degree l_3 , and is defined by the recurrence relations

$$Q_{l_3+1}^{l_2}(x, y) = 2\sqrt{\frac{(l_3+2)(l_3+1)}{(l_3-l_2+1)(l_3+l_2+2)}} y Q_{l_3}^{l_2}(x, y) - \sqrt{\frac{(l_3+2)(l_2+l_3+1)(l_3-l_2)}{l_3(l_3-l_2+1)(l_3+l_2+2)}} Q_{l_3-1}^{l_2}(x, y),$$

with the initial condition

$$Q_{l_3}^{l_3} = \sqrt{\frac{(2l_3 + 2)!!}{(2l_3 + 1)!!}} \frac{x^{l_3}}{\sqrt{\pi}}.$$

The recurrence relation shows that $Q_{l_3}^{l_2}(x, y)$ has the form $x^{l_2} R_{l_3}^{l_2}(y)$, where $R_{l_3}^{l_2}(\cdot)$ is a polynomial of degree $l_3 - l_2$.

For the three-dimensional integral (84), one can use spherical coordinates to further simplify it. Upon integrating out the spherical angles, we obtain

$$I_{l_1 l_2 l_3} = \begin{cases} \frac{\pi^2}{2^{2l_2}} \sqrt{\frac{2l_2+1}{4\pi}} \binom{l_2}{l_2/2}^2 \int_0^{+\infty} (\sin g)^{l_2} R_{l_3}^{l_2}(\cos g) \exp(-sg^2) dg, & \text{if } l_1 = 0 \text{ and } l_2 \text{ is even,} \\ 0, & \text{otherwise.} \end{cases}$$

When $l_1 = 0$ and l_2 is even, we define the polynomial $\bar{R}_{l_3}^{l_2}(x) = (1 - x^2)^{l_2/2} R_{l_3}^{l_2}(x)$. Then, we have

$$I_{l_1 l_2 l_3} = \frac{\pi^2}{2^{2l_2}} \sqrt{\frac{2l_2+1}{4\pi}} \binom{l_2}{l_2/2}^2 \int_0^{+\infty} \bar{R}_{l_3}^{l_2}(\cos g) \exp(-sg^2) dg, \quad l_1 = 0 \quad \text{and} \quad l_2 \text{ is even.}$$

The integral above is the linear combination of $s^{-1/2} e^{-\frac{k^2}{4s}}$, $k = 0, 1, \dots, l_3$. Thus, similar to Propositions 4.1 and 4.2, we conclude that in the $SU(2)$ theory, the regularized observable $\langle O \rangle_s$ has the form

$$\langle O \rangle_s = \frac{\sum_{k=0}^{+\infty} a_k e^{-\frac{k^2}{4s}}}{\sum_{k=0}^{+\infty} b_k e^{-\frac{k^2}{4s}}}. \quad (85)$$

This expression can then be used in the regression model via a truncation of both infinite series.

For the $SU(n)$ theory with $n > 2$, we have not found a straightforward way to evaluate the integral (83). Instead, we conjecture that the form (85) holds for all the $SU(n)$ theories, and we will use the approximation

$$\langle O \rangle_s = \frac{\sum_{k=0}^M a_k e^{-\frac{k^2}{4s}}}{1 + \sum_{k=1}^M b_k e^{-\frac{k^2}{4s}}} \quad (86)$$

in our regression model when carrying out numerical tests.

Remark.—Although the precise form of $\langle O \rangle_s$ remains unclear, it is reasonable to believe that $\langle O \rangle_s$ is also non-analytic at $s = 0$ in the $SU(n)$ theory. Similar to the $U(1)$ theory, both integrals in the definition of $\langle O \rangle_s$ [see (67)] can be written as the integral of $Q(\{U\}) \exp(\sum_{k=1}^N \frac{s}{2} g_k^2)$, where $Q(\{U\})$ is defined on the compact group $[SU(n)]^N$, which can be decomposed into the linear combination of countable discrete ‘‘Fourier modes.’’ Here each ‘‘Fourier mode’’ can be understood as an eigenfunction of the Laplace-Beltrami operator on $[SU(n)]^N$. Meanwhile, when s approaches zero, the exponential part approaches a constant, meaning that in the frequency space, this function approaches a Dirac delta function at the zero point. This implies that s should appear in the denominator inside the

exponential function. In addition, the factor $Q(\{U\})$ spreads the Dirac-like function of frequency to all the discrete modes. With this understanding, we think the postulate (85) for all $SU(n)$ theories is sensible, although it is possibly up to a scalar multiple in the exponent.

VI. APPLICATIONS IN THE LATTICE FIELD THEORY

We are now ready to carry out numerical simulations for the lattice field theories. For the D -dimensional lattice, the each lattice point is denoted by a periodic multi-index

$$x \in X := (\mathbb{Z}/l_0\mathbb{Z}) \times \cdots \times (\mathbb{Z}/l_{D-1}\mathbb{Z}), \quad (87)$$

where l_i refers to the length of the lattice in the $(i + 1)^{\text{th}}$ component of x . For a scalar field $\{\phi\}$, the variables will be denoted as ϕ_x ; for a vector field $\{U\}$, we denote the variables using $U_{x,\mu}$, where $\mu \in \{0, 1, \dots, D - 1\}$. For simplicity, we use $x \pm \hat{\mu}$ to denote the multi-index that adds/subtracts the μ th component of x by 1. For instance,

$$\begin{aligned} x + \hat{0} &= (x_0 + 1, x_1, \dots, x_{D-1}), \\ x - \hat{1} &= (x_0, x_1 - 1, \dots, x_{D-1}). \end{aligned} \quad (88)$$

In what follows, three models in the lattice field theory will be studied. In order to achieve better results from regression, we will derive more suitable regression models for specific problems whenever possible.

A. 3D XY model

For the 3D XY model, we have $D = 3$ and its action reads

$$S(\{\phi\}) = -\beta \sum_{x \in X} \sum_{\nu=0}^2 \cos(\phi_x - \phi_{x+\hat{\nu}} - i\mu\delta_{\nu,0}), \quad (89)$$

where the field variables $\phi_x \in \mathbb{T}$, and μ is the chemical potential. For simplicity, we shall work with the assumption that the length of the lattice l in each dimension to be equal, that is $l = l_i$ for $i \in \{0, 1, 2\}$. Note that one can attempt to generalize some of the arguments to lattices with differing lattice sizes in each dimension with a somewhat similar argument. For a scalar field, the size of the lattice, denoted by $N = l^3$, corresponds to the number of field variables we will be considering in our integration. When $\text{Re } \mu \neq 0$, the action becomes complex, and the complex Langevin method is applied to solve this model. This method fails even for small μ in this model [17,34], and its failure was carefully analyzed in [34], with the effect of the boundary terms being discussed in [26]. Furthermore, according to our numerical experiments, the complex Langevin dynamics diverges even for a simple Euler-Maruyama method without any modifications such as using adaptive time-stepping algorithms. In addition, to impose the presented regularization method on this model, we simply add the regularization term $-\frac{s}{2} \sum_x \phi_x^2$ with $s > 0$ as discussed in Sec. IV.

For this model, the complex drift force corresponding to the variable ϕ_x upon regularization is given by

$$\begin{aligned} K_{x,s} &= -\frac{\partial S_s}{\partial \phi_x} \\ &= -\beta \sum_{\nu=0}^2 [\sin(\phi_x - \phi_{x+\hat{\nu}} - i\mu\delta_{\nu,0}) \\ &\quad + \sin(\phi_x - \phi_{x-\hat{\nu}} + i\mu\delta_{\nu,0})] - s\phi_x. \end{aligned} \quad (90)$$

We will mainly focus on two observables: the action density

$$\begin{aligned} \langle S \rangle &= -\beta \frac{\partial \ln Z}{\partial \beta} \\ &= -\beta \left\langle \sum_x \sum_{\nu=0}^2 \cos(\phi_x - \phi_{x+\hat{\nu}} - i\mu\delta_{\nu,0}) \right\rangle \end{aligned} \quad (91)$$

and the number density

$$\langle n \rangle = \frac{\partial \ln Z}{\partial \mu} = \left\langle i\beta \sum_x \sin(\phi_x - \phi_{x+\hat{0}} - i\mu) \right\rangle. \quad (92)$$

As part of the general framework for $U(1)$ lattice field theory, we can expect Propositions 4.1–4.4 to hold. What remains is to determine an equivalent interval of s depending on a given β and μ such that we are guaranteed correct convergence. In addition, for this specific model, we are able to derive a better regression model as compared to the

general model in Proposition 4.4. The former is summarized in a proposition that follows.

Proposition 6.1. Let β and μ be given. Let η_0 be the smallest real number such that the following inequality holds

$$\begin{aligned} (e^{|\mu|} + 2)Y(\eta_0; \mu) - (e^{-|\mu|} + 2)Y(\eta_0; \mu)^{-1} \\ - \eta_0 \log(Y(\eta_0; \mu)) \leq 0, \end{aligned} \quad (93)$$

whereby

$$Y(\eta; \mu) = \frac{\eta + \sqrt{\eta^2 - 4(e^{|\mu|} + 2)(e^{-|\mu|} + 2)}}{2(e^{|\mu|} + 2)}.$$

Then, if $s > 2\eta_0\beta$, the imaginary part of the field $\{\phi^I\}$ is bounded for any realization of the complex Langevin dynamics.

Here, by following the notation in (2) and (14), the first index in the subscript of $K_{x,s}$ represents the drift term for the corresponding scalar field ϕ_x , while s in the second index indicates that this drift term is obtained from a regularized action.

Proof.—We begin the proof by decomposing the drift term in (90) as follows:

$$\begin{aligned} K_{x,s}^R &= -\beta \sum_{\nu=0}^2 [\sin(\phi_x^R - \phi_{x+\hat{\nu}}^R) \cosh(\phi_x^I - \phi_{x+\hat{\nu}}^I - \mu\delta_{\nu,0}) \\ &\quad + \sin(\phi_x^R - \phi_{x-\hat{\nu}}^R) \cosh(\phi_x^I - \phi_{x-\hat{\nu}}^I + \mu\delta_{\nu,0})] - s\phi_x^R, \\ K_{x,s}^I &= -\beta \sum_{\nu=0}^2 [\cos(\phi_x^R - \phi_{x+\hat{\nu}}^R) \sinh(\phi_x^I - \phi_{x+\hat{\nu}}^I - \mu\delta_{\nu,0}) \\ &\quad + \cos(\phi_x^R - \phi_{x-\hat{\nu}}^R) \sinh(\phi_x^I - \phi_{x-\hat{\nu}}^I + \mu\delta_{\nu,0})] - s\phi_x^I. \end{aligned} \quad (94)$$

Following a similar strategy as to Proposition 3.3 but in a generalized case, we need to show that for s large enough, the support of the probability density in the imaginary variables ϕ^I is compact. Thus, we do so by first removing any dependence in $\{\phi^R\}$ by finding an upper bound for $K_{x,s}^I$ for a given $x \in X$ that holds for all ϕ_x^R . An instructive upper bound for $K_{x,s}^I$ would be

$$\begin{aligned} K_{x,s}^I &\leq \beta \sum_{\nu=0}^2 (\sinh(|\phi_x^I| + |\phi_{x+\hat{\nu}}^I| + |\mu\delta_{\nu,0}|) \\ &\quad + \sinh(|\phi_x^I| + |\phi_{x-\hat{\nu}}^I| + |\mu\delta_{\nu,0}|)) - s\phi_x^I. \end{aligned} \quad (95)$$

Let $\overline{K_{x,s}^I}$ be the right-hand side of the inequality above, which only depends on the imaginary part of the field variables $\{\phi^I\}$. Note that $\overline{K_{x,s}^I}$ can be viewed as a function on \mathbb{R}^N . As a generalization to Proposition 3.3, we will need

to find an N -dimensional object H that contains the origin⁶ such that, along the boundary of the object ∂H , we have⁷

$$\sum_{x \in X} K_{x,s}^I \cdot \hat{n}_x|_{\partial H} < 0, \quad (96)$$

where $\{\hat{n}\}$ represents the outward-oriented unit vector normal of the surface ∂H . For this proof, our choice of H would be an N -dimensional hypercube centered at the origin with length $2C$ in each dimension. Here, we reserve the freedom of choice on $C > 0$, which would be chosen to close the proof for this proposition. We note that the surfaces of this hypercube are $(N-1)$ -dimensional finite planes given by

$$\Pi_x^\pm = \{\{\phi^I\} \in \mathbb{R}^N | \phi_x^I = \pm C \text{ and } \forall y \neq x, |\phi_y^I| < C\} \quad (97)$$

with a total of $2N$ of such planes, indexed by a sign on its superscript and $x \in X$ corresponding to the imaginary part of the field variable ϕ_x in which the value of C is achieved. Note that since the outward oriented unit vector normal to Π_x^\pm only has a component in the x direction and that $K_{x,s}^I \leq \overline{K_{x,s}^I}$, it is sufficient to show that $\overline{K_{x,s}^I}(\{\phi|_{\Pi_x^\pm}\}) < 0$ if $\phi_x^I = +C$ and $\overline{K_{x,s}^I}(\{\phi|_{\Pi_x^\pm}\}) > 0$ if $\phi_x^I = -C$ for all x . Thus, applying the relevant inequalities from (97) for Π_x^\pm , we have

$$\begin{aligned} \overline{K_{x,s}^I}|_{\Pi_x^\pm} &< \beta \sum_{\nu=0}^2 (\sinh(|\phi_x^I| + |\phi_{x+\hat{\nu}}^I| + |\mu\delta_{\nu,0}|)) \\ &\quad + \sinh(|\phi_x^I| - |\phi_{x-\hat{\nu}}^I| + |\mu\delta_{\nu,0}|) - s\phi_x^I \\ &< \beta(2 \sinh(2C + |\mu|) + 4 \sinh(2C)) - sC \\ &:= K_{\text{upper}}^I. \end{aligned} \quad (98)$$

Thus, $\overline{K_{x,s}^I}|_{\Pi_x^\pm} < 0$ if $K_{\text{upper}}^I \leq 0$. Thus, for a given β , s , and μ , we can conduct a one variable optimization on C and deduce that K_{upper}^I is minimized at $C = C_1$ with

$$C_1 = \frac{1}{2} \log \left(\frac{\frac{s}{2\beta} + \sqrt{\left(\frac{s}{2\beta}\right)^2 - 4(e^{|\mu|} + 2)(e^{-|\mu|} + 2)}}{2(e^{|\mu|} + 2)} \right). \quad (99)$$

A remark here is that minimization is consistent with the fact that we can utilize the freedom of choice of C for any given β , μ , and s . Thus, we want to lower the upper bound of K_{upper}^I as much as possible so that $K_{\text{upper}}^I < 0$ can be achieved with a smaller value of s . This is analogous to picking the choice of y_0 in the $U(1)$ one-link model in

⁶This condition is necessary since all our complex Langevin simulations always starts from the origin.

⁷In the one-link model, the one-dimensional object is the line segment $[Y^-, Y^+]$ that contains the origin 0, with the boundary being the points $y = Y^+$ and $y = Y^-$.

Proposition 3.3. Substituting C_1 to the expression of $\overline{K_{\text{upper}}^I}|_{\Pi_x^\pm}$ in (98) yields

$$\begin{aligned} K_{\text{upper}}^I|_{\Pi_x^\pm} &= \beta((e^{|\mu|} + 2)Y(\eta; \mu) - (e^{-|\mu|} + 2)Y(\eta; \mu)^{-1} \\ &\quad - \eta \log(Y(\eta; \mu))), \end{aligned} \quad (100)$$

where $Y(\eta; \mu)$ is as defined in (93) and $\eta := \frac{s}{2\beta}$. Thus, demanding $K_{\text{upper}}^I \leq 0$ is equivalent to solving the inequality as described in (93). In other words, we describe the minimum value in which (93) holds as η_0 . Thus, this is equivalent to

$$\eta = \frac{s}{2\beta} \geq \eta_0,$$

and thus $s \geq 2\eta_0\beta$. Alternatively, we can just choose $s > 2\eta_0\beta$. A symmetric and analogous argument holds for the case of Π_x^- and for all x . This concludes the proof. ■

Furthermore, as promised, we will attempt to obtain a better regression model as compared to that in (58), summarized in the following proposition:

Proposition 6.2. We consider the 3D XY model for the action density and the number density observables as defined in (91) and (92). For both observables, we can improve the representation of $\langle O \rangle_s$ to

$$\langle O \rangle_s = \frac{\sum_{k=0}^{\infty} a_k e^{-\frac{k}{s}}}{\sum_{k=0}^{\infty} b_k e^{-\frac{k}{s}}}. \quad (101)$$

Proof.—It is obvious that both observables and the action are C^∞ functions. Thus the derivations in the proof of Proposition 4.1 work in the 3D XY model. By observing that the action S and the number density n on a discrete lattice takes the form of a difference in two neighboring field variables $\phi_x - \phi_{x+\hat{\nu}}$ as compared to individual field variables, we thus rework some of the steps in (50). For the 3D XY model, with the action given in (89), we can rewrite the exponentiation of the negative of it as

$$\begin{aligned} e^{-S(\{\phi\})} &= \prod_{x \in X} \prod_{\nu=0}^2 e^{\beta \cos(\phi_x - \phi_{x+\hat{\nu}} - i\mu\delta_{\nu,0})} \\ &= \prod_{x \in X} \prod_{\nu=0}^2 \sum_{m \in \mathbb{Z}} [i^m e^{m\mu\delta_{\nu,0}} J_m(-i\beta)] e^{im(\phi_x - \phi_{x+\hat{\nu}})} \\ &= \sum_{\vec{k} \in \mathcal{K}} \beta_{\vec{k}} e^{i \sum_{x \in X} k_x \phi_x}, \end{aligned} \quad (102)$$

where we have again used the Jacobi-Anger expansion (16) and the resultant Fourier series with its coefficients represented by $\beta_{\vec{k}}$ is analogous to that in the second line of Eq. (50). Here the range of the summation \mathcal{K} is given as a proper subset of \mathbb{Z}^N defined by

$$\mathcal{K} = \left\{ \vec{k} \in \mathbb{Z}^N : \sum_{x \in X} k_x = 0 \right\}.$$

The appearance of \mathcal{K} can be explained as follows. Since the inner sum represents different Fourier frequency modes for $\phi_x - \phi_{x+\hat{\nu}}$, with each satisfying the fact that $\sum_{x \in X} k_x = 0$, and the fact that the product of exponentials is given by the exponential of the sum of the individual arguments, such a property is preserved in the resulting Fourier series, and thus have the given property for \mathcal{K} . Here, we note that the observables n or S follow a similar structure. This thus implies that, analogous to (51) and (57), we have

$$\begin{aligned} \langle O \rangle_s &= \frac{1}{Z_s} \sum_{\vec{m} \in \mathcal{K}} \sum_{\vec{k} \in \mathcal{K}} \hat{O}_{\vec{m}} \beta_{\vec{k}} \exp \left(-\frac{1}{2s} \sum_{x \in X} (m_x + k_x)^2 \right), \\ Z_s &= \sum_{\vec{k} \in \mathcal{K}} \beta_{\vec{k}} \exp \left(-\frac{1}{2s} \sum_{x \in X} k_x^2 \right). \end{aligned} \quad (103)$$

For every $\vec{k} \in \mathbb{Z}^N$, the sum $\sum_{x \in X} k_x^2$ is odd/even if and only if $\sum_{x \in X} k_x$ is odd/even since k_x^2 and k_x always share the same parity. Then for $\vec{k} \in \mathcal{K}$, we know that $\sum_{x \in X} k_x^2$ is even so that in (103), each exponential term in the series expansion of Z_s has the form $\exp(-l/s)$ for some $l \in \mathbb{Z}$, where the factor 2 drops off by reduction of the fraction.

The same can be done to the expansion of $\langle O \rangle_s$ since $\vec{m} + \vec{k} \in \mathcal{K}$, which leaves with us the expression in (101). ■

Remark.—We note that the concept of counting the parity of k_x and m_x is inspired by the representation of the partition function into bonds as proposed by the Worm algorithm in [45,46].

With reference to the above proposition, the appropriate regression model is thus given by

$$\langle O \rangle_s = \frac{\sum_{k=0}^M a_k e^{-\frac{k}{s}}}{1 + \sum_{k=1}^M b_k e^{-\frac{k}{s}}}, \quad (104)$$

for a given integer M .

We apply the above regression model to two examples on a $8 \times 8 \times 8$ lattice with parameters $\beta = 0.2, \mu = \sqrt{0.1}$ and $\beta = 0.7, \mu = \sqrt{0.1}$, respectively, for the two observables of interest, the number density scaled with the chemical potential μn , and the action density S . Similar to the $U(1)$ one-link model, we will compare the numerical results obtained with that from standard methods from current literature. The results for the $2R$ method are summarized in Fig. 6 and Table II.

Here, we summarize some of the key points from Table II as follows. First, we note that the range of s used might not be consistent with what is obtained in Proposition 4.4, $s_0 > 2\eta_0\beta = 2(7.8)(0.2) = 3.12$ for $\beta = 0.2$ and $\mu = \sqrt{0.1}$, while we have included points as close as $s = 0.6$ in our

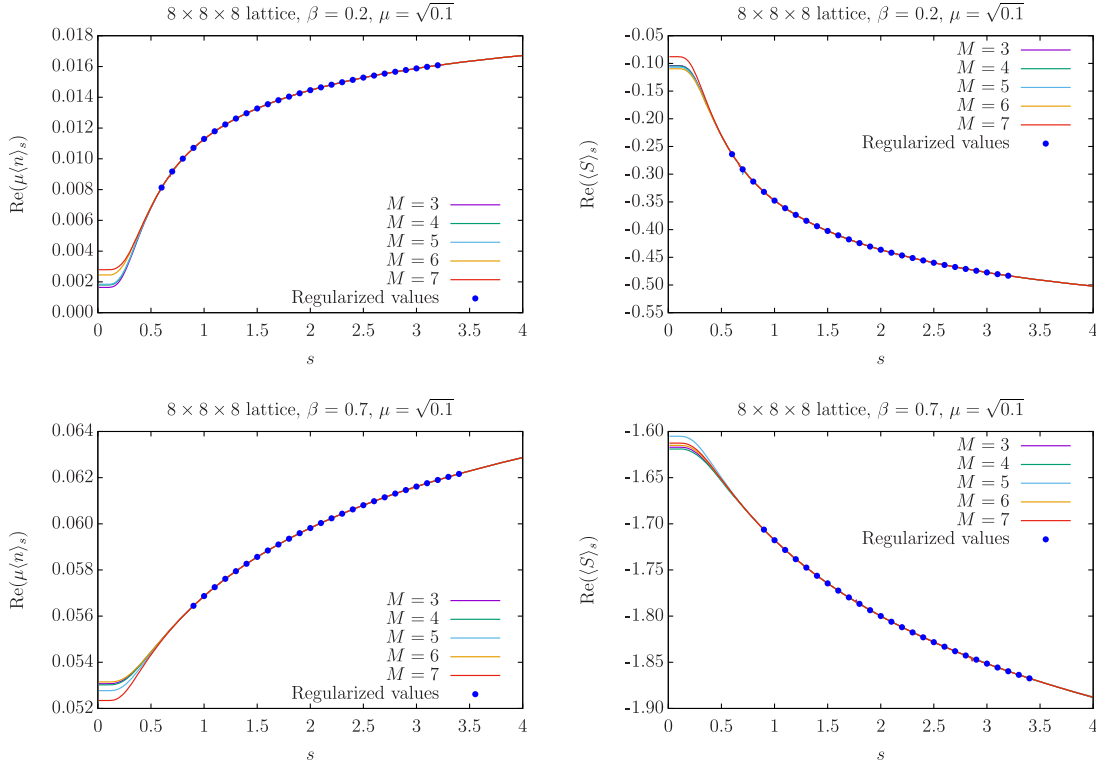


FIG. 6. Number density (left) and action density (right) for $\beta = 0.2$ (top) and 0.7 (bottom) of 3D XY model. M represents the number of expansion terms used for extrapolation (101).

TABLE II. Estimates for $\text{Re}\mu\langle n \rangle$ and $\text{Re}\langle S \rangle$ using various methods from the current literature and from our $2R$ method as in Fig. 6.

	$\beta = 0.2, \mu = \sqrt{0.1}$		$\beta = 0.7, \mu = \sqrt{0.1}$	
	$\text{Re}\mu\langle n \rangle$	$\text{Re}\langle S \rangle$	$\text{Re}\mu\langle n \rangle$	$\text{Re}\langle S \rangle$
Original complex Langevin method from [26]	0.001840	-0.07929	0.04926	-1.5268
Corrected complex Langevin method from [26]	0.0003858	-0.06716	0.04905	-1.5240
Worldline method from [26]	1.5495×10^{-7}	-0.06230	0.04898	-1.5240
Best for $2R$ method	0.001641	-0.08791	0.05234	-1.6052
	$2R$ method with different values of M			
$M = 3$	0.001641	-0.10397	0.05307	-1.6170
$M = 4$	0.001797	-0.1046	0.05303	-1.6191
$M = 5$	0.001849	-0.1069	0.05277	-1.6052
$M = 6$	0.002455	-0.1097	0.05315	-1.6151
$M = 7$	0.002793	-0.08791	0.05234	-1.6126

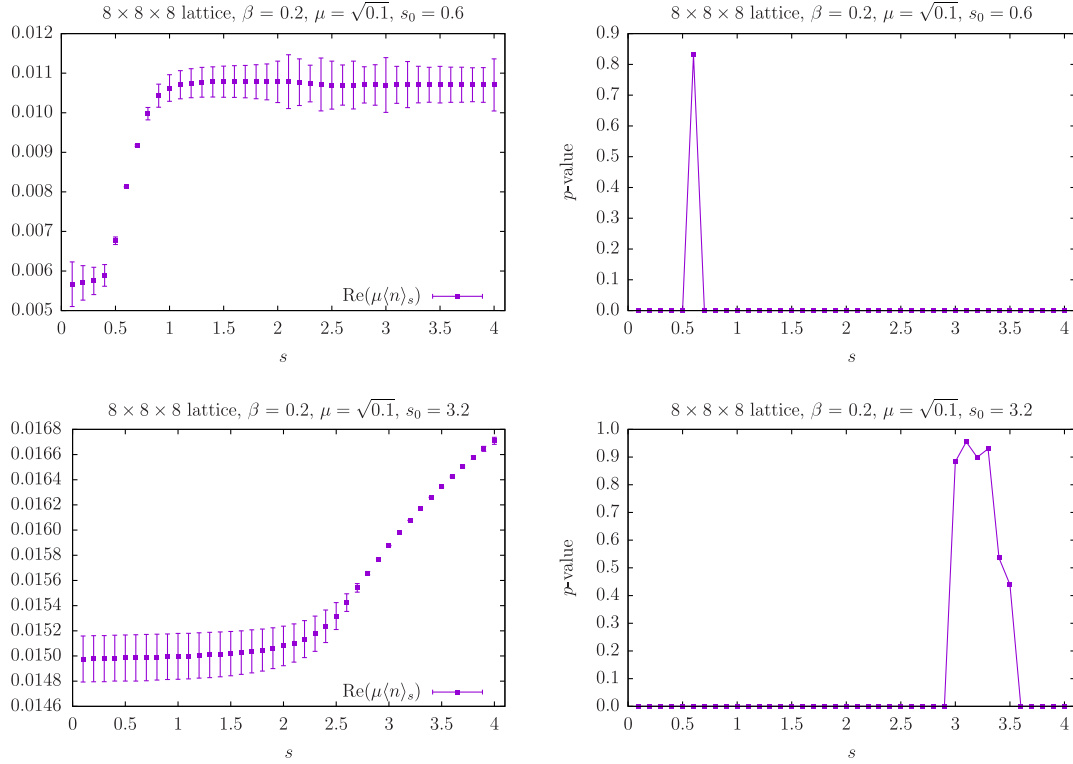
regression. Next, we note that the results obtained from our $2R$ method are similar to those obtained from the original complex Langevin method. Although they might not do as well as compared to the corrected complex Langevin method from [26] and the worldline method, the results obtained are still not too far off from these methods. Furthermore, we note that the results obtained using different values of M for the $2R$ method are generally stable, while according to our simulations, the original complex Langevin method for $s = 0$ suffers from instabilities, and the result of the original complex Langevin method from [26] may require adaptive time stepping. Nonetheless, one should note that by picking $s = 0.6$, a value which is way off from our guaranteed region of $s > s_0$, there might be an unquantifiable bias that possibly grows as we pick values of s closer to 0. Such a possibility is inferred for the case of the $U(1)$ one-link model, in which the difference between the true value and the complex Langevin method grows significantly as s falls below a certain threshold s_1 ($\approx 0.4 < s_0$) and gets larger as it approaches 0. Despite the inability to obtain a result of better accuracy as compared to the worldline and the corrected complex Langevin method, the relatively simple structure and the improved generalizability of the $2R$ method might still serve as a method that we can use to corroborate with alternative methods in the current literature.

Next, we shall explain the rationale of choosing the $2R$ method over the $3R$ method despite the latter having better success with the $U(1)$ one-link model. Recall that for our $3R$ method, we will have to choose a reference s_0 . From Fig. 7, we can observe that the estimates for $\text{Re}\mu\langle n \rangle$ without regression is better for $s_0 = 0.6$ at 0.005666 as compared to that in $s_0 = 3.2$ at 0.01498. Here, better is defined as how close our results are to the results generated from the worldline method, at ~ 0 for $\text{Re}\mu\langle n \rangle$ at $\beta = 0.2$, $\mu = \sqrt{0.1}$ and an estimate without regression is obtained by quoting the value at $s = 0.1$ directly for an estimate as regression including this point will likely not predict the

value at $s = 0$ to be too far off from it. The reason for this difference might be as follows. We note that there is a trade-off between errors arising from regression and errors arising from having s_0 that is not sufficiently large enough to guarantee possible correct convergence. In this case here, $s_0 = 3.2$ is too far off from our point of interest at $s = 0$, and thus may result in a large error resulting from regression. This error might be larger than the error arising from inaccurate simulations with $s_0 = 0.6$ and thus accounts for such a phenomenon. Nonetheless, the values obtained for both choices of s_0 fail to surpass that obtained from the $2R$ method with even the worst value at $M = 7$ at 0.002793.

On top of that, we note that the benefits of regression might be limited for the 3D XY model. As seen from the p values on the right diagrams in Fig. 7, the phenomenon of infinite variance appears relatively quick for both cases, with p value dropping to a value close to 0 as soon as s hits 0.5 for $s_0 = 0.6$ and a similar phenomenon at $s = 2.9$ for $s_0 = 3.2$. Thus, restricting our regression points for which the p value is at least 0.05 would heavily restrict the number of points that can be used, and therefore reduces the prediction ability of the model at $s = 0$. This is also why we did not perform regression for this model, as the number of points might not be sufficient to determine the regression coefficients.

In both the $2R$ and the $3R$ method, a key ingredient would be the regression method in the final step. Thus, in view of improving results obtained from regression, we list down a few plausible explorations. One includes improving the regression method itself as it was done using ordinary regression techniques on highly nonlinear models as such that in (104), where the dependent variable only appears after a ratio of sums of the exponential of the inverse of its independent variable. Another possible exploration would be to reduce the distance of s_0 from 0 by evaluating the expectation of a modified observable. Yet another possible exploration would be to combine the use of the corrected complex Langevin method with our $2R$ or $3R$ method. The


 FIG. 7. Reweighted results for $\beta = 0.2$ with parameter $s_0 = 0.6$ (top) and 3.2 (bottom).

interested reader is invited to try out some of these explorations in view of improving the accuracy of the proposed method.

B. Polyakov chain model

In this subsection, we discuss the results for the one-dimensional Polyakov chain model [47], whose action is given by

$$S(\{U\}) = -\text{tr}(\beta_1 U_1 \cdots U_N + \beta_2 U_N^{-1} \cdots U_1^{-1}), \quad \beta_1, \beta_2 \in \mathbb{R}. \quad (105)$$

The observable of interest is $O_l(\{U\}) = \text{tr}([U_1 \cdots U_N]^l)$ for $l \in \mathbb{Z}_+$. Due to the gauge invariance, this model can actually be reduced to the one-link model ($N = 1$) by gauge fixing [20]. Thus, to test the performance of the 2R method, we choose to simulate the original problem without fixing the gauge.

Note that this model works for both $U(1)$ and $SU(n)$ theories. For the $U(1)$ theory, the trace operator reduces to the identity operator. Let $U_k = \exp(i\theta_k)$, $k = 1, \dots, N$. Then, the expectation of the regularized observable can be represented by

$$\langle O_l \rangle_s = \frac{1}{Z_s} \int_{\mathbb{R}^N} e^{il(\theta_1 + \cdots + \theta_N)} \exp\left(\beta_1 e^{i(\theta_1 + \cdots + \theta_N)} + \beta_2 e^{-i(\theta_1 + \cdots + \theta_N)} - \frac{s}{2}(\theta_1^2 + \cdots + \theta_N^2)\right) d\theta_1 \cdots d\theta_N, \quad (106)$$

where

$$Z_s = \int_{\mathbb{R}^N} \exp\left(\beta_1 e^{i(\theta_1 + \cdots + \theta_N)} + \beta_2 e^{-i(\theta_1 + \cdots + \theta_N)} - \frac{s}{2}(\theta_1^2 + \cdots + \theta_N^2)\right) d\theta_1 \cdots d\theta_N.$$

By the series expansion of the exponential function, we get that

$$Z_s = \int_{\mathbb{R}^N} \sum_{j=-\infty}^{+\infty} \alpha_j e^{ij(\theta_1 + \cdots + \theta_N)} \exp\left(-\frac{s}{2}(\theta_1^2 + \cdots + \theta_N^2)\right) d\theta_1 \cdots d\theta_N = \left(\frac{2\pi}{s}\right)^{N/2} \sum_{j=-\infty}^{+\infty} \alpha_j \exp\left(-\frac{j^2 N}{2s}\right),$$

where

$$\alpha_j = \sum_{k=\max(0,j)}^{+\infty} \frac{\beta_1^k \beta_2^{k-j}}{k!(k-j)!}, \quad j \in \mathbb{Z}.$$

Similarly, the regularized observable $\langle O_l \rangle_s$ given in (106) can be expanded as

$$\langle O_l \rangle_s = \frac{1}{Z_s} \left(\frac{2\pi}{s} \right)^{N/2} \sum_{j=-\infty}^{+\infty} \alpha_j \exp\left(-\frac{(j+l)^2 N}{2s}\right). \quad (107)$$

Inspired by the analysis above, one can choose the expression

$$\langle O_l \rangle_s = \frac{\sum_{k=0}^M a_k e^{-\frac{Nk^2}{2s}}}{\sum_{k=0}^M b_k e^{-\frac{Nk^2}{2s}}} \quad (108)$$

in the regression. For $SU(n)$ theories, a similar expression will be used, while s will be replaced with $2s$ due to the reason stated in Sec. V C.

For the general $SU(n)$ theory, the Lie derivative of (105) can be derived as

$$\begin{aligned} D_{a,k} S(\{U\}) &= -i\beta_1 \text{tr}(U_1 \cdots U_{k-1} \lambda_a U_k \cdots U_N) \\ &\quad + i\beta_2 \text{tr}(U_N^{-1} \cdots U_k^{-1} \lambda_a U_{k-1}^{-1} \cdots U_1^{-1}). \end{aligned}$$

In our simulation, we would like to focus on the $SU(3)$ model, which has also been studied in [20,47]. Following [20], we choose β_1 and β_2 to be 2.27 and 2.04, respectively. In this case, the exact values of $\langle O_l \rangle$ for $l = 1, 2, 3$ obtained were

$$\langle O_1 \rangle = 2.0957, \quad \langle O_2 \rangle = 0.3761, \quad \langle O_3 \rangle = -0.5269,$$

which have been calculated in [47]. For all the numerical tests, we chose the fixed time step $\Delta t = 5 \times 10^{-4}$ and simulate the complex Langevin dynamics up to $T = 2$. Then, we drew one sample for every 10 time steps until 20 million samples were collected. These samples were used to estimate the expectations of the observables.

Following [48], we use the quantity

$$\Delta F = \frac{1}{N} \sum_{i=1}^N \text{tr}(U_i U_i^\dagger - I) \quad (109)$$

to measure the extent of excursion away from $[SU(3)]^N$. Here I stands for the 3×3 identity matrix. Results for some values of s and N are given in Fig. 8. For instance, when $N = 16$, the complex Langevin dynamics quickly diverges if no regularization is applied. Even when $s = 4$, we can still observe a few spikes of the curve at the magnitude of 10^{-4} , which may indicate convergent but biased results. For $s = 8$ and 16, the deviation from $SU(3)$ is well suppressed, so that the regularized observables computed from the samples are likely to be reliable. However, as N increases, the regularization for $s = 8$ may no longer be sufficient. The middle diagram of Fig. 8 shows that for $s = 8$, the complex Langevin dynamics fails to converge when $N = 64$ and 128. Even with $N = 32$, the few spikes on the curve of ΔF may imply possibly biased results. A reasonable modification appears to be setting s to be proportional to N , as displayed in the right diagram of Fig. 8. This agrees with the analysis for (108), in which s also scales with N .

We now focus on the case $N = 16$. For the $2R$ method, we display the results for s ranging from 4 to 34 in Fig. 9. The dashed horizontal line denotes the reference solution. As previously observed, smaller values of s will lead to unstable complex Langevin dynamics. In these examples, the estimated values of $\langle O_l \rangle_s$ for $s = 4$ and $l = 1, 2, 3$ are very close to the exact solution. These indicate the existence of an example where the regularization can work well without further corrections. However, the reliability of this method is hard to judge if the exact solution is unknown.

We then consider the $3R$ method and display the results in Fig. 10. Here we select $s_0 = 8$ and $s_0 = 16$ for which the results are likely to be accurate according to the evolution of ΔF . In all the cases, we observe that when s decreases from s_0 to 0, the numerical results first move toward and then deviate from the exact solutions, which behave similarly to the $U(1)$ one-link model shown in Fig. 2.

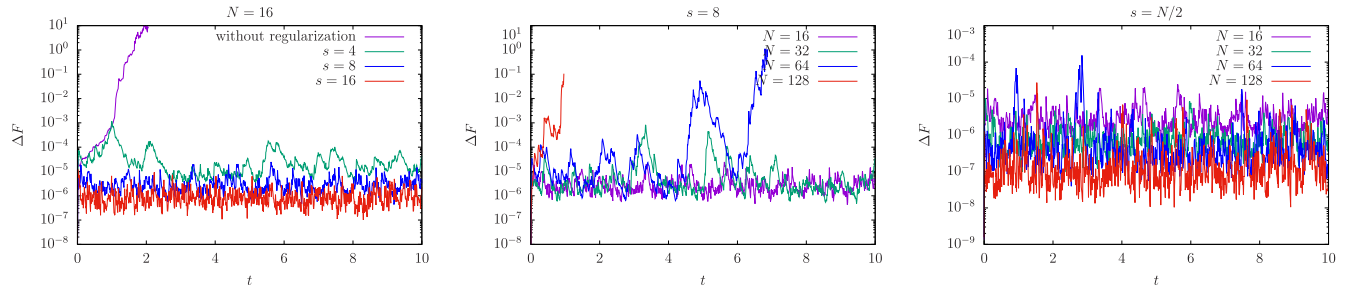


FIG. 8. Restriction of regularization method ΔF . Left: different s for $N = 16$. Middle: fixed $s = 8$ for $N = 16, 32, 64, 128$. Right: $s = N/2$.

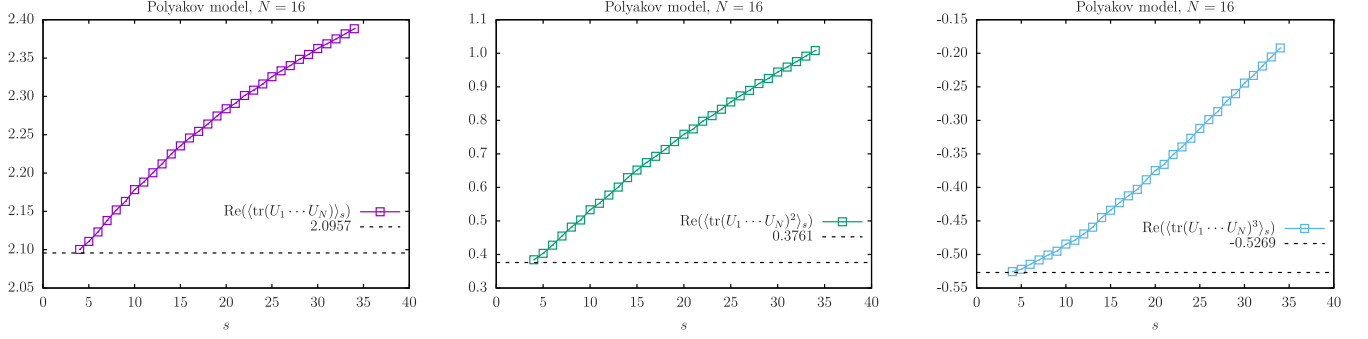
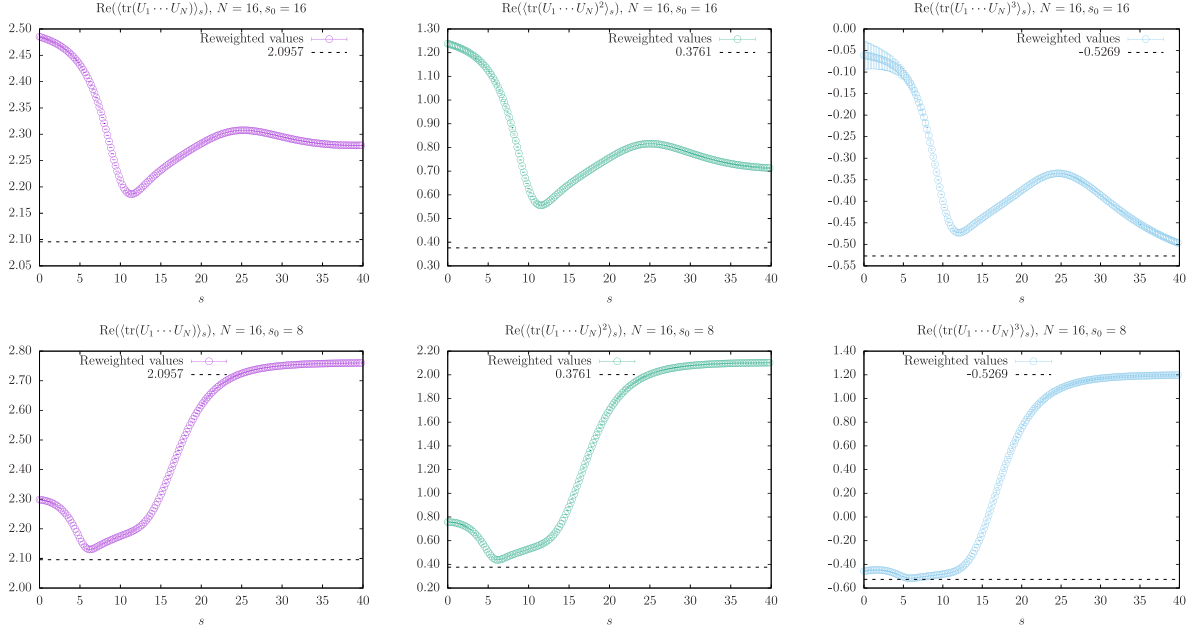


FIG. 9. Regularized results for the problem (105).


 FIG. 10. Reweighted regularization of Polyakov model. $s_0 = 16$ (top) and $s_0 = 8$ (bottom).

This again confirms that, for the high-dimensional integrals, reweighting might fail to provide desired solutions.

In light of the observations above, we thus consider the use of the $2R$ method to extrapolate the observables. As discussed previously, the proposed regression model is given by

$$\langle O_I \rangle_s = \frac{\sum_{k=0}^M a_k e^{-\frac{Nk^2}{4s}}}{1 + \sum_{k=1}^M b_k e^{-\frac{Nk^2}{4s}}}. \quad (110)$$

We will use data points with $s \in [10, 34]$ in the extrapolation. Note that some of the points closer to $s = 0$ in Fig. 10 were discarded to avoid including points with significant biases. The extrapolations obtained using $M = 3, 4, 5, 6, 7$ were plotted in Fig. 11. For all three

observables, the results obtained were generally acceptable, though the values at $s = 0$ were slightly underestimated.

C. Heavy dense QCD

In this section, we consider a more realistic numerical example originating from the heavy dense QCD model at finite potential, studied in several works [6,48,49]. The field is discretized on a four-dimensional lattice indexed by $x := (x_0, \mathbf{x}) \in X$, where $\mathbf{x} = (x_1, x_2, x_3)$ [see also (87)]. We consider a vector field $\{U\}$ defined on the lattice, and each $U_{x,\mu}$ is a member of the $SU(3)$ group. The action of the heavy dense QCD model is given by

$$S(\{U\}) = -\ln \det M_\mu(\{U\}) + S_B(\{U\}),$$

where $S_B(\{U\})$ is defined by

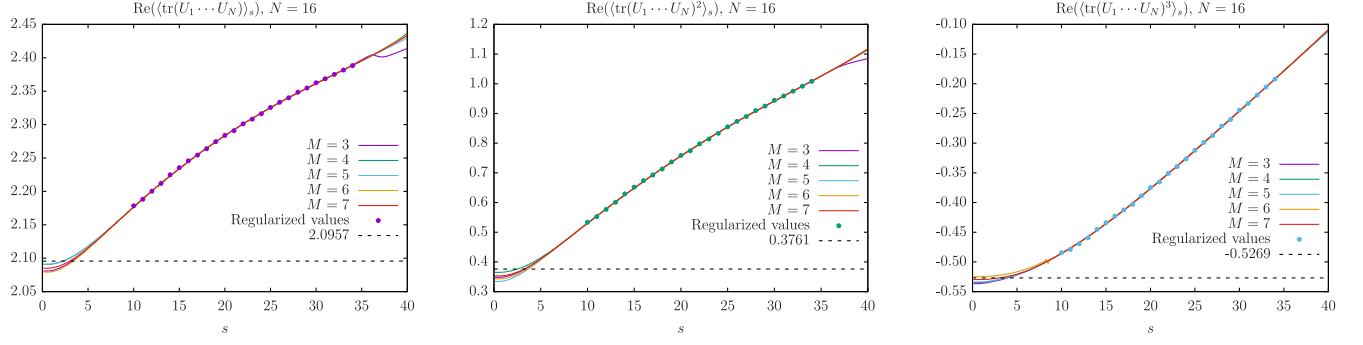
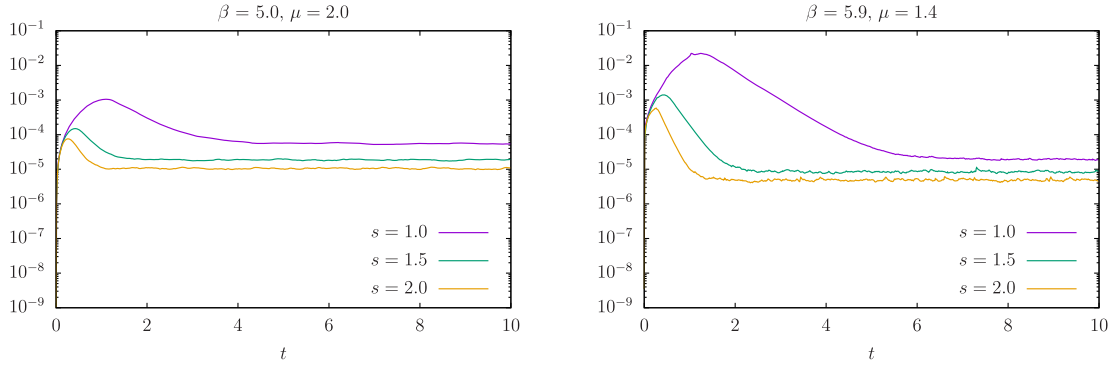
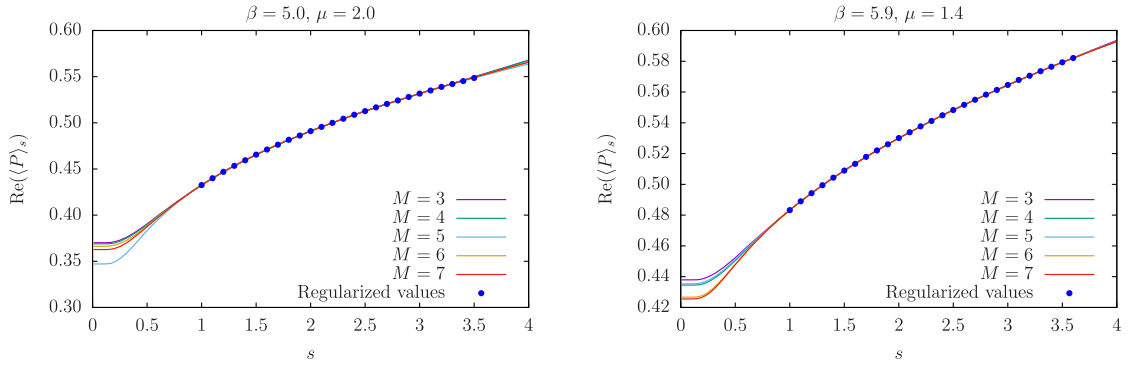


FIG. 11. Results of Polyakov model using the 2R method.


 FIG. 12. The evolution of ΔF . Left: $\beta = 5.0$, $\mu = 2.0$; right: $\beta = 5.9$, $\mu = 1.4$.

 FIG. 13. Numerical results for the heavy dense QCD model. Left: $\beta = 5.0$, $\mu = 2.0$; right: $\beta = 5.9$, $\mu = 1.4$.

$$S_B(\{U\}) = -\beta \sum_{x \in X} \sum_{\nu_1 < \nu_2} \left(\frac{1}{6} [\text{tr}(U_{x,\nu_1} U_{x+\hat{\nu}_1,\nu_2} U_{x+\hat{\nu}_2,\nu_1}^{-1} U_{x,\nu_2}^{-1}) + \text{tr}(U_{x,\nu_2} U_{x+\hat{\nu}_2,\nu_1} U_{x+\hat{\nu}_1,\nu_2}^{-1} U_{x,\nu_1}^{-1})] - 1 \right),$$

and $\det M_\mu$ is the fermionic determinant with μ being the chemical potential, whose definition is

$$\det M_\mu(\{U\}) = \prod_{\mathbf{x}} \det(I + C \mathcal{P}_{\mathbf{x}}(\{U\}))^2 \times \det(I + C' [\mathcal{P}_{\mathbf{x}}(\{U\})]^{-1})^2,$$

where $C = [2\kappa \exp(\mu)]^{l_0}$, $C' = [2\kappa \exp(-\mu)]^{l_0}$ with κ being the hopping parameter, and

$$\mathcal{P}_{\mathbf{x}}(\{U\}) = \prod_{t=1}^{l_0} U_{(t,\mathbf{x}),0}.$$

We refer the readers to [6] for the Lie derivatives of this action.

In our simulation, we worked with two sets of parameters, namely $\mu = 2.0$, $\beta = 5.0$ and $\mu = 1.4$, $\beta = 5.9$. The value of κ was set to be 0.12. The observable of interest is given by

$$P(\{U\}) = \frac{1}{3l_1 l_2 l_3} \sum_{\mathbf{x}} \text{tr} \mathcal{P}_{\mathbf{x}}.$$

The parameters of the lattice used in our numerical tests is given by $l_0 = 6$, $l_1 = l_2 = l_3 = 8$, and the time step is fixed to be $\Delta t = 2 \times 10^{-5}$.

For the heavy dense QCD model, we only considered the 2R method. To determine the range of s to be adopted in the regression model, we again study the evolution of the deviation from $SU(n)$. This is defined in a similar way as (109), given by

$$\Delta F = \frac{1}{4l_0 l_1 l_2 l_3} \sum_{\mu=0}^3 \sum_{x \in X} \text{tr}(U_{x,\mu} U_{x,\mu}^\dagger - I).$$

Once again, we observe from Fig. 12 that larger values of s result in smaller deviations. In both cases, using $s = 1$ reduces the deviation to the magnitude of 10^{-5} , for which we expect that the results may contain sufficiently small biases and can be used in the regression model.

To estimate the expectation, we used 6.4 million samples in our tests. By computing the regularized observable $\langle P \rangle_s$ for s ranging from 1 to 3.5, we perform extrapolation based on the expression (86) with $b_0 = 1$. The results are provided in Fig. 13. In both cases, the regression results using $M = 6$ and $M = 7$ give similar estimates, whose values at $s = 0$ can be considered as approximations of $\langle P \rangle$. Note that for $\beta = 5.0$, the complex Langevin method is generally considered to be not applicable in the standard

literature. Thus, the result of our approximation obtained by the 2R method remains to be validated. For the case with $\beta = 5.9$, our estimate agrees with that in [20], in which gauge cooling is applied to stabilize the dynamics.

VII. CONCLUSION

We have performed an in-depth study of the regularization of the complex Langevin method. It is demonstrated that the regularization can produce significant bias in some cases, and we have proposed a few extensions to the regularized complex Langevin method:

- (1) The 3R method, which performs regression based on the results of the reweighting method proposed in [32].
- (2) The 2R method, which performs regression based on the regularized results with a number of different parameters.

The computational cost of the 2R method is higher than the other two approaches, since multiple complex Langevin dynamics have to be simulated for different regularization constants. The reweighting method and its 3R extension works well in the one-link toy model. However, it is observed that in the high-dimensional case, the results of the reweighting were reliable only for a very small range of parameters. The best results are obtained from the 2R method, which has successfully simulated one example in lattice QCD for which the original complex Langevin method was known to be inapplicable. We expect that this approach can also be applied to the actions with poles, which will be studied in our future works.

ACKNOWLEDGMENTS

Z. C. was supported by the Academic Research Fund of the Ministry of Education of Singapore under Grant No. R-146-000-291-114.

-
- [1] J. Berges, Sz. Borsanyi, D. Sexty, and I.-O. Stamatescu, Lattice simulations of real-time quantum fields, *Phys. Rev. D* **75**, 045007 (2007).
 - [2] A. Alexandru, G. Başar, P. F. Bedaque, S. Vartak, and N. C. Warrington, Monte Carlo Study of Real Time Dynamics on the Lattice, *Phys. Rev. Lett.* **117**, 081602 (2016).
 - [3] S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar, Numerical study of the two-dimensional Hubbard model, *Phys. Rev. B* **40**, 506 (1989).
 - [4] J.-L. Wynen, E. Berkowitz, S. Krieg, T. Luu, and J. Ostmeier, Machine learning to alleviate Hubbard-model sign problems, *Phys. Rev. B* **103**, 125153 (2021).
 - [5] S. Muroya, A. Nakamura, C. Nonaka, and T. Takaishi, Lattice QCD at finite density: An introductory review, *Prog. Theor. Phys.* **110**, 615 (2003).
 - [6] G. Aarts and I.-O. Stamatescu, Stochastic quantization at finite chemical potential, *J. High Energy Phys.* **09** (2008) 018.
 - [7] K. N. Anagnostopoulos, T. Azuma, Y. Ito, J. Nishimura, and S. K. Papadoudis, Complex Langevin analysis of the spontaneous symmetry breaking in dimensionally reduced super Yang-Mills models, *J. High Energy Phys.* **02** (2018) 151.
 - [8] J. Nishimura and A. Tsuchiya, Complex Langevin analysis of the space-time structure in the Lorentzian type IIB matrix model, *J. High Energy Phys.* **06** (2019) 077.

- [9] E. Y. Loh, Jr., J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar, Sign problem in the numerical simulation of many-electron systems, *Phys. Rev. B* **41**, 9301 (1990).
- [10] J. Klauder, A Langevin approach to fermion and quantum spin correlation functions, *J. Phys. A* **16**, L317 (1983).
- [11] G. Parisi, On complex probabilities, *Phys. Lett.* **131B**, 393 (1983).
- [12] J. Ambjørn, M. Flensburg, and C. Peterson, The complex Langevin equation and Monte Carlo simulations of actions with static charges, *Nucl. Phys.* **B275**, 375 (1986).
- [13] H. Gausterer, Complex Langevin: A numerical method?, *Nucl. Phys.* **A642**, c239 (1998).
- [14] G. Aarts, E. Seiler, and I.-O. Stamatescu, Complex Langevin method: When can it be trusted?, *Phys. Rev. D* **81**, 054508 (2010).
- [15] K. Nagata, J. Nishimura, and S. Shimasaki, Argument for justification of the complex Langevin method and the condition for correct convergence, *Phys. Rev. D* **94**, 114515 (2016).
- [16] L. L. Salcedo, Does the complex Langevin method give unbiased results?, *Phys. Rev. D* **94**, 114505 (2016).
- [17] M. Scherzer, E. Seiler, D. Sexty, and I.-O. Stamatescu, Complex Langevin and boundary terms, *Phys. Rev. D* **99**, 014512 (2019).
- [18] E. Seiler, Complex langevin: Boundary terms at poles, *Phys. Rev. D* **102**, 094507 (2020).
- [19] G. Aarts, F. A. James, E. Seiler, and I.-O. Stamatescu, Adaptive stepsize and instabilities in complex Langevin dynamics, *Phys. Lett. B* **687**, 154 (2010).
- [20] E. Seiler, D. Sexty, and I.-O. Stamatescu, Gauge cooling in complex Langevin for lattice QCD with heavy quarks, *Phys. Lett. B* **723**, 213 (2013).
- [21] G. Aarts, E. Seiler, D. Sexty, and I.-O. Stamatescu, Simulating QCD at nonzero baryon density to all orders in the hopping parameter expansion, *Phys. Rev. D* **90**, 114505 (2014).
- [22] J. B. Kogut and D. K. Sinclair, Applying complex Langevin simulations to lattice QCD at finite density, *Phys. Rev. D* **100**, 054512 (2019).
- [23] M. Hirasawa, A. Matsumoto, J. Nishimura, and A. Yosprakob, Complex Langevin analysis of 2D $U(1)$ gauge theory on a torus with a θ term, *J. High Energy Phys.* **09** (2020) 023.
- [24] F. Attanasio and B. Jäger, Dynamical stabilisation of complex Langevin simulations of QCD, *Eur. Phys. J. C* **79**, 16 (2019).
- [25] F. Attanasio and B. Jaeger, Testing dynamic stabilization in complex Langevin simulations, *Proc. Sci., LATTICE2016* (2017) 053.
- [26] M. Scherzer, E. Seiler, D. Sexty, and I.-O. Stamatescu, Controlling complex Langevin simulations of lattice models by boundary term analysis, *Phys. Rev. D* **101**, 014501 (2020).
- [27] J. Nishimura and S. Shimasaki, Combining the complex Langevin method and the generalized Lefschetz-thimble method, *J. High Energy Phys.* **06** (2017) 023.
- [28] K. Nagata, J. Nishimura, and S. Shimasaki, Complex Langevin calculations in finite density QCD at large μ/T with the deformation technique, *Phys. Rev. D* **98**, 114513 (2018).
- [29] C. E. Berger, L. Rammelmüller, A. C. Loheac, F. Ehmman, J. Braun, and J. E. Drut, Complex Langevin and other approaches to the sign problem in quantum many-body physics, *Phys. Rep.* **892**, 1 (2020).
- [30] F. Attanasio, B. Jäger, and F. P. G. Ziegler, Complex Langevin simulations and the QCD phase diagram: Recent developments, *Eur. Phys. J. A* **56**, 251 (2020).
- [31] A. C. Loheac and J. E. Drut, Third-order perturbative lattice and complex Langevin analyses of the finite-temperature equation of state of nonrelativistic fermions in one dimension, *Phys. Rev. D* **95**, 094502 (2017).
- [32] J. Bloch, Reweighting complex langevin trajectories, *Phys. Rev. D* **95**, 054509 (2017).
- [33] J. Bloch, J. Meisinger, and S. Schmalzbauer, Reweighted complex Langevin and its application to two-dimensional QCD, *Proc. Sci., LATTICE2016* (2017) 046.
- [34] G. Aarts and F. A. James, On the convergence of complex Langevin dynamics: The three-dimensional XY model at finite chemical potential, *J. High Energy Phys.* **08** (2010) 020.
- [35] A. M. Polyakov, Thermal properties of gauge fields and quark liberation, *Phys. Lett.* **72B**, 477 (1978).
- [36] G. Parisi, Y. S. Wu *et al.*, Perturbation theory without gauge fixing, *Sci. Sin* **24**, 483 (1981).
- [37] J. R. Klauder, Stochastic quantization, in *Recent Developments in High-Energy Physics* (Springer, New York, 1983), pp. 251–281.
- [38] G. Aarts, P. Giudice, and E. Seiler, Localised distributions and criteria for correctness in complex Langevin dynamics, *Ann. Phys. (Amsterdam)* **337**, 238 (2013).
- [39] Z. Cai, X. Dong, and Y. Kuang, On the validity of complex langevin method for path integral computations, *SIAM J. Sci. Comput.* **43**, A685 (2021).
- [40] J. Berges and D. Sexty, Real-time gauge theory simulations from stochastic quantization with optimized updating, *Nucl. Phys.* **B799**, 306 (2008).
- [41] W. Rudin, *Principles of Mathematical Analysis*, 3rd ed. (McGraw-Hill, New York, 1976).
- [42] K. B. Athreya, Bootstrap of the mean in the infinite variance case, *Ann. Stat.* **15**, 724–731 (1987).
- [43] W. Rossmann, *Lie Groups: An Introduction Through Linear Groups* (Oxford University Press on Demand, New York, 2006), Vol. 5.
- [44] A. Higuchi, Symmetric tensor spherical harmonics on the N-sphere and their application to the de Sitter group $SO(N, 1)$, *J. Math. Phys. (N.Y.)* **28**, 1553 (1987).
- [45] N. Prokof'ev and B. Svistunov, Worm algorithm for problems of quantum and classical statistics, [arXiv:0910.1393](https://arxiv.org/abs/0910.1393).
- [46] D. Banerjee and S. Chandrasekharan, Finite size effects in the presence of a chemical potential: A study in the classical nonlinear $o(2)$ sigma model, *Phys. Rev. D* **81**, 125007 (2010).
- [47] Z. Cai, Y. Di, and X. Dong, How does gauge cooling stabilize complex Langevin?, *Commun. Comput. Phys.* **27**, 1344 (2020).
- [48] X. Dong, Z. Cai, and Y. Di, Alternating descent method for gauge cooling of complex Langevin simulations, *Phys. Rev. D* **102**, 054518 (2020).
- [49] G. Aarts, F. Attanasio, B. Jäger, and D. Sexty, The QCD phase diagram in the limit of heavy quarks using complex Langevin dynamics, *J. High Energy Phys.* **09** (2016) 087.