Exponential reduction of the sign problem at finite density in the 2+1D XY model via contour deformations

Matteo Giordano,[1] Kornél Kapás,[1] Sándor D. Katz,[1,2] Attila Pásztor,[1] and Zoltán Tulipánt[1]

[1] *ELTE Eötvös Loránd University, Institute for Theoretical Physics, Pázmány Péter sétány 1/A, H-1117, Budapest, Hungary*

[2] *MTA-ELTE Theoretical Physics Research Group, Pázmány Péter sétány 1/A, H-1117 Budapest, Hungary.*

We study the 2+1 dimensional XY model at nonzero chemical potential μ on deformed integration manifolds, with the aim of alleviating its sign problem. We investigate several proposals for the deformations, and considerably improve on the severity of the sign problem with respect to standard reweighting approaches. We present numerical evidence that the reduction of the sign problem is exponential both in μ^2 and in the spatial volume. We also present a new approach to the optimization procedure based on reweighting, that sensibly reduces its computational cost.

I. INTRODUCTION

Euclidean quantum field theories with a finite chemical potential generally suffer from a complex action problem: the path integral weights are complex, and therefore cannot be interpreted as the Boltzmann weights of a classical statistical mechanical system. In QCD, this complex action problem is a severe roadblock for first principles understanding of the physics of neutron stars, supernovae, as well as heavy ion collisions at lower collision energies. In some theories the sign problem can be solved by a reformulation of the theory in different variables, such that in the new variables the weights are manifestly real and positive [1–5]. This has not been achieved in QCD so far.

The existence of a sign problem does not make simulations completely impossible. In the presence of a complex action problem, simulations can still be carried out by standard Monte Carlo methods in the phase-quenched (PQ) theory, with Boltzmann weights proportional to $\left|e^{-S}\right|$, or - assuming that the grandcanonical partition function is real - the sign-quenched (SQ) theory, with weights proportional to $\left|\mathrm{Re}\, e^{-S}\right|$ [1]. After such simulations have been carried out, the ratio of the simulated and target partition functions, as well as the expectation values of different operators O, can in principle be reconstructed via the formulas:

$$\frac{Z}{Z_{\mathrm{PQ}}} = \langle \cos\theta \rangle_{\mathrm{PQ}}\,, \qquad \frac{Z}{Z_{\mathrm{SQ}}} = \left\langle \frac{\cos\theta}{|\cos\theta|} \right\rangle_{\mathrm{SQ}}\,,$$
$$\langle O \rangle = \frac{\left\langle O e^{i\theta} \right\rangle_{\mathrm{PQ}}}{\left\langle e^{i\theta} \right\rangle_{\mathrm{PQ}}} = \frac{\left\langle O \frac{e^{i\theta}}{|\cos\theta|} \right\rangle_{\mathrm{SQ}}}{\left\langle \frac{\cos\theta}{|\cos\theta|} \right\rangle_{\mathrm{SQ}}}, \tag{1}$$

where we introduced the phase of the complex action θ as $e^{-S} = \left|e^{-S}\right| e^{i\theta}$ and we assumed that the partition function Z is real. This leads to large cancellations when the phase $e^{i\theta}$ has large fluctuations, a problem that is generally referred to as a sign problem. The severity of the sign problem can be measured by the partition function ratios Z/Z_{PQ} and Z/Z_{SQ}, and as long as these quantities are under numerical control one can reconstruct expectation values in the desired theory reliably. In fact, reweighting from the phase and sign quenched theories is starting to become feasible in QCD [6–9]. However, the range of applicability - both in the chemical potential and in the physical volume - of such an approach is severely constrained by the sign problem, which requires an exponential increase of the statistics both as a function of the volume V and of the squared chemical potential μ^2 (for small μ). It is therefore desirable to develop methods that either solve or at least alleviate the sign problem. Even if only the second goal is achieved, this could still drastically increase the range of parameters that reweighting methods can reach.

A set of methods that try to deal with the sign problem is based on complexification of the fields. There are two broadly defined approaches of this type. In the first one - complex Langevin [10–14] - an N-dimensional integral over real fields is enlarged to a $2N$-dimensional integral over the real components of the complexified fields. In the second one - contour deformations - the integral still remains N-dimensional, but the integration manifold is deformed to a different manifold of the same dimension. In this paper, we pursue this second approach.

In most cases of interest, the path integral weights are holomophic functions of the fields.[2] In such a case, any integration manifold that is in the same homology class as the undeformed manifold leads to the same partition function [17]. However, the phase and sign quenched integrands are not holomorphic, and therefore the phase and sign quenched partition functions are not invariant under such contour deformations. Thus, it could be possible to bring closer to unity the ratios Z/Z_{PQ} and Z/Z_{SQ}, measuring the severity of the sign problem, by such con-

[1] This is true provided that the phase- or sign-quenched partition function is given by a convergent integral, or can be made convergent with an infinitesimal regulator, like for the case of real-time dynamics.

[2] A notable exception is lattice QCD with rooted staggered fermions [15, 16].

tour deformations. One could then perform simulations in the contour-deformed phase or sign quenched theory, and perform a reweighting via Eq. (1) to get results in the target theory.

There are different ways to deform integration contours to make sign problems milder. First, there are methods based on Lefschetz thimbles [17–23]. However, Lefschetz thimbles are in general not the numerically optimal integration contours [24], i.e. they are in general not the contours with the largest $Z/Z_{\rm PQ}$, and therefore there is no need to solely concentrate on them. Thus, it is a worthwhile effort to explore different contours, even in a more ad hoc manner, leading to a second class of methods: path optimization methods. The main idea of these methods is to parametrize the deformed integration manifold by some finite number of parameters, which are then adjusted to make the sign problem as mild as possible. In the context of the sign problem at nonzero chemical potential, the path optimization method was applied to a one-dimensional oscillating integral [25], the 0+1D ϕ^4 theory [26], the 0+1D PNJL model [27], 0+1D QCD [28], the 1+1D ϕ^4 model [29], the 1+1D Thirring model [30], the 2+1D Thirring model [31], and Bose gases of several dimensions [32]. Other applications of the sign optimization method include improving the signal-to-noise ratio of noisy observables at zero chemical potential in 0+1D scalar field theory and 1+1D U(1) gauge theory [33] and in 1+1D SU(2) and SU(3) gauge theory [34], and reduction of the sign problem in 1+1D U(1) gauge theory with a complex coupling constant [35].

In this paper we apply the path optimization method to the 2+1 dimensional XY model. The choice of the model is mainly motivated by the fact that it shares several technical features with QCD, so that insight obtained here can hopefully be applied there as well. First, the integration variables take values in a compact space, which requires a slightly different treatment of contour deformations than non-compact integration domains do. Second, this theory has Roberge-Weiss periodicity at imaginary chemical potential [36, 37]. Third, the model also shares the property of QCD that the complex Langevin approach fails for small coupling β [38, 39]. Fourth, like in lattice QCD, the dependence of the action on the chemical potential is non-linear.

Finally, like in QCD, at finite lattice spacing the effects of a chemical potential should saturate at large μ, where the temporal dependence of the relevant field configurations tends to become trivial and μ essentially drops out of expectation values; in this limit the sign problem becomes mild.

Differently from QCD, however, the XY model can be rewritten in the worldline formulation to be free of a sign problem [3, 40, 41], allowing for direct simulations using a worm algorithm, and making explicit comparisons possible with the path optimization method.

Our goals in this work are the following. First, we investigate how much improvement can be made to the severity of the sign problem in this model by very simple ansätze for the contour deformations. Second, we study the chemical potential and volume dependence of the improvement on the sign problem, and so on the statistics required for reliable reweighting, achieved with the optimized contours. Third, we wish to see whether the optimization procedure can be performed more efficiently, avoiding a full Monte Carlo simulation at each step of the iteration searching for the optimum, by reweighting to nearby contours and only occasionally generating new ensembles.

The results for all three of these inquiries turn out to be quite encouraging. First - as we will see - a rather drastic improvement can be achieved in the severity of the sign problem, even with relatively simple ansätze. Second, we present numerical evidence that the improvement is exponential, i.e., it reduces the exponent of the severity of the sign problem. And third, a drastic improvement can be achieved in the numerical cost of the optimization by the proposed reweighting procedure.

The plan of the paper is the following. In Section II we briefly discuss the contour deformation approach to the 2+1 dimensional XY model at finite chemical potential. In Section III we provide details on the optimization procedure and on the different parametrizations used to ameliorate the severity of the sign problem. In Section IV we illustrate the chemical potential and volume dependence of the achieved improvement. In Section V we present a modified optimization procedure based on reweighting to different contours from a fixed ensemble, that allows us to reduce computational costs. We summarize our conclusions in Section VI.

II. CONTOUR DEFORMATIONS FOR 2+1D XY MODEL AT NONZERO CHEMICAL POTENTIAL

The action of the 2+1D XY model with nonzero chemical potential [38, 40–42] is

$$S = -\beta \sum_x \sum_{n=0}^{2} \cos(\varphi_x - \varphi_{x+\hat{n}} + \imath\mu\delta_{n0}), \tag{2}$$

where the sum runs over all lattice sites and directions, with 0 identified as the temporal direction. Periodic boundary conditions are imposed in every direction. The partition function,

$$\begin{aligned} Z(\mu) &= \int_{-\pi}^{\pi} \mathrm{d}\varphi_{(000)} \cdots \int_{-\pi}^{\pi} \mathrm{d}\varphi_{(N_0N_1N_2)}\, e^{-S} \\ &\equiv \int_{\mathcal{M}_0} \mathcal{D}\varphi\, e^{-S} \end{aligned} \tag{3}$$

can be interpreted as a complex contour integral in each φ_x with endpoints at $-\pi$ and π, and so as an integral over an $N_0N_1N_2$-dimensional manifold $\mathcal{M}_0 = [-\pi,\pi]^{N_0N_1N_2}$ embedded in a $2N_0N_1N_2$ dimensional space.

For each φ_x, we will consider contours in the strip of the complex plane satisfying $-\pi \leq \mathrm{Re}\,\varphi_x \leq \pi$, with the

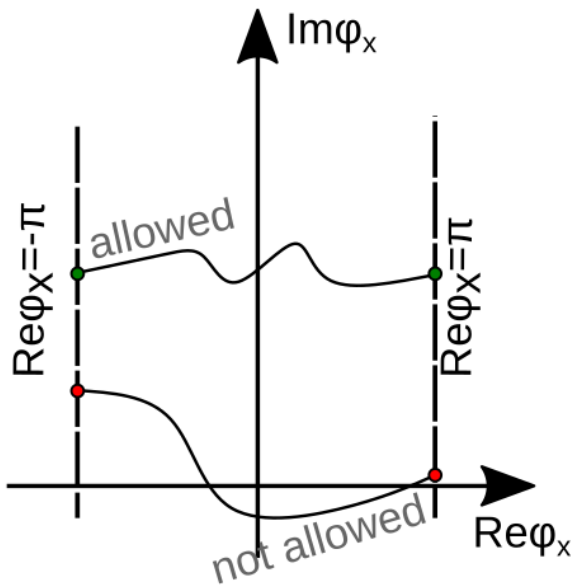

FIG. 1. Illustration of an allowed and not allowed contour deformation for a single field variable.

points $-\pi + i\,\mathrm{Im}\,\varphi_x$ and $\pi + i\,\mathrm{Im}\,\varphi_x$ identified, i.e., a cylinder. As long as the contours remain smooth on this space, i.e., if they are smooth curves from $-\pi + ix$ to $\pi + ix$ for some $x \in \mathbb{R}$, the partition function remains unchanged. This can be shown by connecting the original and the shifted contours using line segments perpendicular to the real axis, whose contributions cancel out thanks to periodicity of the integrand in $\mathrm{Re}\,\varphi_x$. Allowed contour deformations for a single field variable are illustrated in Fig. 1.

Denoting the deformed manifold $\mathcal{M}$, and parametrizing it with real parameters t_x, we have for any $\mathcal{M}$ satisfying the requirements above

$$\begin{aligned} Z(\mu) &= \int_{\mathcal{M}_0} \mathcal{D}\varphi e^{-S} = \int_{\mathcal{M}} \mathcal{D}\varphi e^{-S} \\ &= \int \mathcal{D}t \det J\, e^{-S} \equiv \int \mathcal{D}t\, e^{-S_{\mathrm{eff}}}, \end{aligned} \tag{4}$$

where J is the Jacobian matrix, with elements

$$J_{xy} = \frac{\partial \varphi_x}{\partial t_y}, \tag{5}$$

and the effective action is defined as $S_{\mathrm{eff}} = S - \ln\det J$. Exploiting the reality of the partition function we can write it in a manifestly real form,

$$Z(\mu) = \int \mathcal{D}t\, \cos S^I_{\mathrm{eff}}\, e^{-S^R_{\mathrm{eff}}}, \tag{6}$$

where S^R_{eff} and S^I_{eff} are the real and imaginary parts of the effective action. This enables us to make use of the sign reweighting approach [8, 9, 43, 44] by using the absolute value of the integrand, $|\cos S^I_{\mathrm{eff}}|\, e^{-S^R_{\mathrm{eff}}} \geq 0$, as a weight in importance sampling. This method has a slightly milder sign problem than the phase reweighting method [9, 43]. The corresponding expectation values will be denoted as $\langle \ldots \rangle_{\mathrm{SQ},\mu}$. The severity of the sign problem is measured by the average sign,

$$\langle \varepsilon \rangle_{\mathrm{SQ},\mu} = \frac{\int \mathcal{D}t\, \cos S^I_{\mathrm{eff}}\, e^{-S^R_{\mathrm{eff}}}}{\int \mathcal{D}t\, |\cos S^I_{\mathrm{eff}}|\, e^{-S^R_{\mathrm{eff}}}}. \tag{7}$$

While the numerator of this expression is invariant under changing integration contours, the denominator is instead altered. Hence, deforming contours leaves the target partition function unchanged but it has an effect on the severity of the sign problem.

III. OPTIMIZATION OF THE INTEGRATION MANIFOLD

In this paper we work at $\beta = 0.4$, that for $\mu = 0$ is in the disordered phase of the model [38, 40, 41]. As a preliminary check, we performed a scan in μ using the sign problem-free worldline formulation and the worm algorithm of Ref. [40]. Using a standard finite size scaling analysis, we observe a phase transition at $\mu_c^2 \approx 0.54$.

Our goal is to find integration contours that give a larger $\langle \varepsilon \rangle_{\mathrm{SQ},\mu}$ than without contour deformation. As such, we need to maximize the expectation value of the sign, or alternatively, minimize a cost function, with respect to some coefficients p_i that parametrize the contours. As our cost function, we choose the ratio of the number of configurations with negative and positive $\cos S^I_{\mathrm{eff}}$,

$$\frac{N_-}{N_+} = \frac{\int \mathcal{D}t\, \Theta(-\cos S^I_{\mathrm{eff}})\, |\cos S^I_{\mathrm{eff}}|\, e^{-S^R_{\mathrm{eff}}}}{\int \mathcal{D}t\, \Theta(\cos S^I_{\mathrm{eff}})\, |\cos S^I_{\mathrm{eff}}|\, e^{-S^R_{\mathrm{eff}}}}. \tag{8}$$

Its gradient can be computed easily,

$$\mathcal{F}_i \equiv \frac{\partial}{\partial p_i}\left(\frac{N_-}{N_+}\right) = \frac{N_-}{N_+}\Big(\langle F_i \rangle_- - \langle F_i \rangle_+\Big), \tag{9}$$

where $\langle \ldots \rangle_\pm$ means averaging only over configurations with positive/negative $\cos S^I_{\mathrm{eff}}$, and

$$F_i \equiv -\frac{\partial S^R_{\mathrm{eff}}}{\partial p_i} - \tan S^I_{\mathrm{eff}} \frac{\partial S^I_{\mathrm{eff}}}{\partial p_i}. \tag{10}$$

We perform optimization using a simple gradient descent algorithm. At every optimization step we update coefficients by subtracting the gradient of the cost function with a multiplication factor,

$$\boldsymbol{p}^{(j+1)} = \boldsymbol{p}^{(j)} - \alpha_j \boldsymbol{\mathcal{F}}^{(j)}, \tag{11}$$

where $\boldsymbol{p}^{(j)}$ is the vector of coefficients obtained at the jth step, $\boldsymbol{\mathcal{F}}^{(j)}$ is the value of the gradient obtained using $\boldsymbol{p}^{(j)}$, and α_j is

$$\alpha_j = \frac{\left|\left(\boldsymbol{p}^{(j)} - \boldsymbol{p}^{(j-1)}\right)^T \cdot \left(\boldsymbol{\mathcal{F}}^{(j)} - \boldsymbol{\mathcal{F}}^{(j-1)}\right)\right|}{||\boldsymbol{\mathcal{F}}^{(j)} - \boldsymbol{\mathcal{F}}^{(j-1)}||^2}. \tag{12}$$

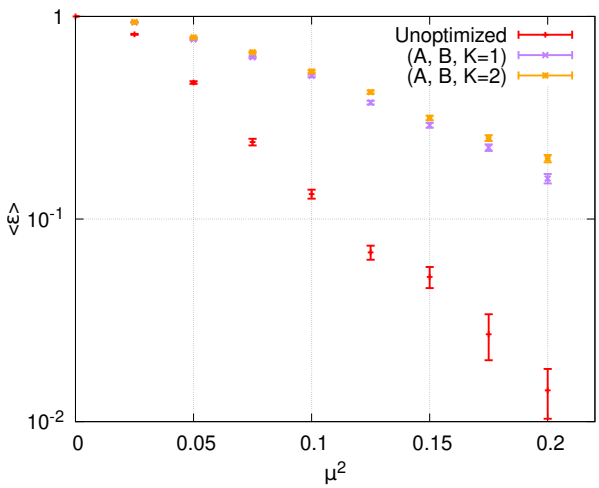

FIG. 2. Severity of the sign problem on an $\Omega = 8^3$ lattice with $\beta = 0.4$. Unimproved results are shown in red and the results optimized obtained using Eq. (13) for $K = 1$ and 2 are shown in purple and orange, respectively.

The algorithm terminates when the relative change of the coefficients decreases below a prescribed tolerance value.

We note that since we always start the optimization near the undeformed integration manifold, our procedure explores only the vicinity of the original contour, and might not detect an optimum lying beyond a "potential barrier".

For simplicity, we only consider contours where the imaginary part is a single-valued function of the real part. We thus parametrize the integration manifold by the real parts of the field variables on the different lattice sites, denoted by t_x. Periodicity of the action in the real part of φ_x restricts the imaginary part of φ_x on the deformed manifold to be a Fourier series in t_x, with coefficients that in general depend on x and on the other field variables. As our first attempt, we parametrized each complex φ_x using only the corresponding t_x. However, the sign problem could not be improved this way, as using this parametrization the original contour turned out to be the local minimum of the cost function. We then experimented with different ansätze that couple neighbouring sites in simple ways, all of which could improve the sign problem, but by quite different amounts. As a further simplification, we considered contours parametrized as Fourier series in sines and cosines of $t_x - t_y$.'

A. Ansätze without temporal translational invariance

Since the chemical potential affects directly only the interaction between temporal nearest neighbors, it is reasonable to consider deformed manifolds that depend only on their difference. Given the translation invariance of the system, it is also natural to expect that the optimal deformation is also translationally invariant. For more generality, we allow for a non-translationally invariant optimum in the temporal direction, using the same coefficients for every φ_x with the same temporal coordinate x_0. We later check our assumptions on spatial translation invariance and on neglecting spatial neighbors. We then considered parametrizations of the following general form,

$$\begin{aligned}\varphi_x(t_x, t_{x+\hat{0}}) = t_x + \imath\Big\{ & A_{0,x_0} \\ & + \sum_{k=1}^{K} \Big[A_{k,x_0} \cos\big(k(t_x - t_{x+\hat{0}})\big) \\ & + B_{k,x_0} \sin\big(k(t_x - t_{x+\hat{0}})\big)\Big]\Big\},\end{aligned} \tag{13}$$

leading to the following effective action:

$$\begin{aligned}S_{\text{eff}} = S &- \sum_x \ln \frac{\partial \varphi_x}{\partial t_x} \\ &- \sum_{x_1, x_2} \ln \left(1 - (-1)^{N_0} \prod_{x_0} \frac{1 - \partial\varphi_x/\partial t_x}{\partial\varphi_x/\partial t_x} \right).\end{aligned} \tag{14}$$

We used the choices $K = 1$ and 2 for the cut-off on the Fourier series, which we will denote with $(A, B, K = 1)$ and $(A, B, K = 2)$, respectively. Performing optimization on a lattice of size $\Omega \equiv N_0 N_1 N_2 = 8^3$ for $\beta = 0.4$ we find that the sign problem is substantially improved. In each iteration 10^6 configurations were generated to compute the gradient. The initial values for the Fourier coefficients were chosen to be zero for the smallest simulated μ^2 value, and for each subsequent μ^2 to be equal to the final values obtained in the previous optimization round.

Figure 2 shows the average sign achieved with optimization for $K = 1$ and 2, along with the unoptimized results for comparison. There is significant improvement in the sign problem even when using only a first-order Fourier series. When going up to second order there is a marginal increase in $\langle\varepsilon\rangle_{\text{SQ},\mu}$. The optimal values of the Fourier coefficients for $\mu^2 = 0.15$ and $K = 2$ are listed in Tabs. I and II of Appendix A.

The coefficients of the constant and of the sine terms are two-three orders of magnitude smaller than the coefficients of the cosine terms; they also fluctuate around zero as a function of x_0, with a standard deviation larger than their average, while the coefficients of the cosine terms have roughly the same value on every time slice, with $A_{1,x_0} < 0$ and $A_{2,x_0} > 0$. This remains true at every simulated value of μ^2. This suggests that there is little gain in allowing for x_0-dependent Fourier coefficients. Furthermore, carrying out the optimization on $\Omega = 8 \times 4^2, 8 \times 6^2$ and 8×10^2 lattices we have found that the values of the optimal coefficients are close to those obtained on the $\Omega = 8^3$ lattice.

B. Ansätze with a triangular Jacobian matrix

With the appearance of additional parameters the calculation of the Jacobian generally becomes more involved. Especially interesting from the computational point of view are the parametrizations which keep the Jacobian simple. For this reason, we considered ansätze with a triangular Jacobian matrix (see also Ref. [34]), that have the benefit of having a very simple form of the effective action :

$$S_{\text{eff}} = S - \sum_x \ln \frac{\partial \varphi_x}{\partial t_x}. \tag{15}$$

As we show below, this can be achieved at the price of losing translational invariance. This is expected to reduce the amount of improvement that can be gained by path deformation, especially at larger μ [32]. Nonetheless, there may be a trade off with computational costs when the Jacobian gets very complicated, in particular in systems where the dependence on μ becomes milder at large μ.

We considered 3 different ansätze of this type. First, we set

$$\begin{aligned}\varphi_x(t_x, t_{x+\hat{0}}, t_{x+\hat{1}}, t_{x+\hat{2}}) = \\ t_x + \imath \sum_{k=1}^{K} \sum_{n=0}^{2} \theta_n^{(1)} \Big[\bar{A}_{k,n} \cos(k(t_x - t_{x+\hat{n}})) \\ + \bar{B}_{k,n} \sin(k(t_x - t_{x+\hat{n}})) \Big],\end{aligned} \tag{16}$$

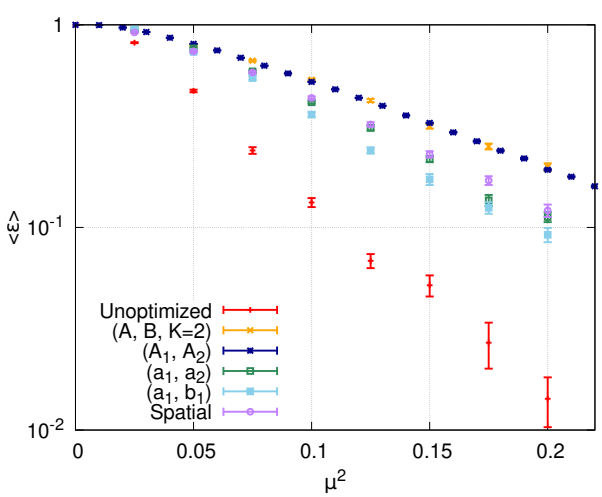

FIG. 3. Severity of the sign problem on an $\Omega = 8^3$ lattice with $\beta = 0.4$. Unimproved results are shown in red. The optimized result achieved with the parametrizations given in Eq. (13) with $K = 2$ are shown in orange. The optimized results achieved with Eq. (16) for $K = 2$ are shown in purple. The result from Eq. (18) with a_1 and a_2 non-zero are shown in green, while with a_1 and b_1 non-zero are shown in light blue. Results with the fully translationally invariant (A_1, A_2) parametrization are shown in blue.

where

$$\theta_i^{(j)} = \begin{cases} 1 & \text{if } x_i < N_i - j, \\ 0 & \text{otherwise.} \end{cases} \tag{17}$$

Here, we constrained the φ_x on the slices $x_i = N_i - 1$ for $i = 0, 1, 2$ to stay independent of the real parts of their neighbors in the ith direction, so as to make the Jacobian matrix triangular. Except for these points we used global Fourier coefficients, i.e., the same coefficients across all the other lattice sites. This leads to the cancellation of the constant term since the action depends only on the difference of the field variables. The parametrization Eq. (16) allows us to check the consequences of explicitly breaking temporal translation invariance, and to assess whether the optimal choice of contours is affected by the inclusion of spatial neighbors in the parametrization.

We carried out optimization with this ansatz on lattices of sizes $\Omega = 8 \times 4^2, 8 \times 6^2, 8^3$ and 8×10^2 with $\beta = 0.4$. The optimal coefficients are shown in Tab. III of Appendix A for $\mu^2 = 0.15$ and $K = 2$. We find that the optimal values of $\bar{A}_{k,i}$ and $\bar{B}_{k,i}$ for $i = 1, 2$ and $\bar{B}_{k,0}$ are much smaller than those of $\bar{A}_{k,0}$, which in turn are in agreement with those obtained for A_{k,x_0} using the ansatz of Eq. (13).

This suggests that the contribution of spatial neighbors can be neglected in the parametrization. Once again, we observed that the optimal coefficients are to a good approximation independent of the spatial size of the system.

We have also investigated the effect of including the second nearest neighbor in the temporal direction in the parametrization of φ_x. Once again, constraints were imposed similarly to Eq. (16) to make the Jacobian simple, and global coefficients were used,

$$\begin{aligned}\varphi_x(t_x, t_{x+\hat{0}}, t_{x+2\hat{0}}) = \\ t_x + \imath \sum_{k=1}^{K} \Big[\theta_0^{(1)} a_k \cos\big(k(t_x - t_{x+\hat{0}})\big) \\ + \theta_0^{(2)} b_k \cos\big(k(t_x - t_{x+2\hat{0}})\big) \Big].\end{aligned} \tag{18}$$

We omitted the sine terms since their coefficients remained close to zero with the previously tested parametrizations. Optimization was performed using Eq. (18) with two setups: using $K = 2$ and adjusting a_1 and a_2 with b_ks set to zero, later referred to as (a_1, a_2) optimization; or using $K = 1$ and adjusting both a_1 and b_1, that we call (a_1, b_1) optimization.

In Fig. 3 we compare the optimized results obtained with the parametrizations of Eqs. (16) and (18) with the unoptimized results, and with the optimized results previously obtained using Eq. (13). Comparing results obtained with Eq. (16) for $K = 2$ and the (a_1, a_2) parametrization, it is clear that including spatial neighbors leads only to a marginal improvement, at the cost of a much more complicated Jacobian. It is also apparent that including the second-order term a_2 gives a larger increase in the optimized $\langle \varepsilon \rangle_{\text{SQ},\mu}$ in the range of the

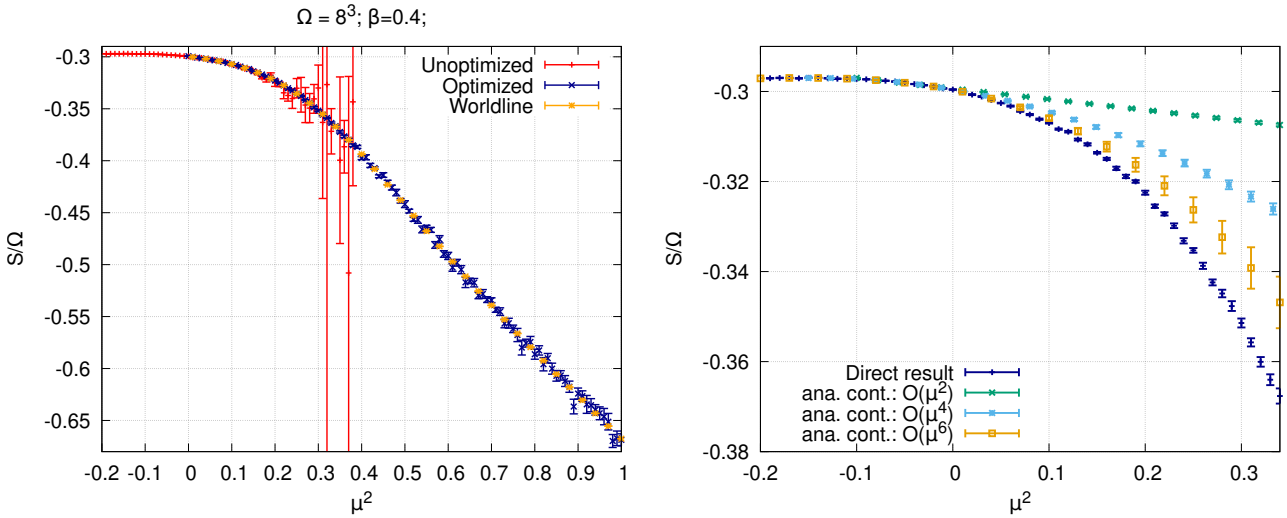

FIG. 4. Left: Average action density on an $\Omega = 8^3$ lattice with $\beta = 0.4$. The unoptimized results, including those obtained with a local heat bath at $\mu^2 \leq 0$, are shown in red, the (A_1, A_2) -optimized results in blue, and the worldline results in orange. Right: Average action density with the (A_1, A_2)-optimized contours, compared with analytic continuation with polynomial ansätze up to orders μ^2, μ^4 and μ^6 for small chemical potentials.

simulated μ values, than including the first-order term b_1. Hence, going to second order in the Fourier series is more important than including second neighbors in the parametrization.

It is also clear that our parametrizations with a triangular Jacobian matrix lead to a significant but not overwhelming loss in the improvement of the sign problem compared to the ansatz of Eq. (13). This discrepancy can probably be attributed to the constraints introduced in Eqs. (16) and (18). This is similar to the conclusions of Ref. [32] for Bose gases. In our model, parametrizations with the triangular Jacobian are not significantly cheaper to simulate the fully translationally invariant ansatz, so there is no gain in breaking translational invariance at the boundaries of the lattice, due to the sizeable difference in the improvement achieved. In other models, there might be a less obvious trade-off.

With all of our two-parameter ansätze, scans of $\langle\varepsilon\rangle_{\mathrm{SQ},\mu}$ in the space of Fourier coefficients reveal a rather simple landscape. There is a single optimum located on a small plateau where the average sign changes slowly. As an example, scans of the average sign in the space of coefficients a_1 and a_2 are shown in Appendix B.

C. Fully translationally invariant ansatz

Among our ansätze, the one that achieved the greatest improvement on the sign problem with the least amount of parameters was given by a translationally invariant version of Eq. (13), with $K = 2$, the constant and sine coefficients set to zero, and $A_{k,x_0} = A_k$ for all x_0. This will be denoted as the (A_1, A_2) parametrization.

IV. VOLUME AND CHEMICAL POTENTIAL DEPENDENCE OF THE SIGN PROBLEM ON OPTIMIZED MANIFOLDS

As a sanity check, we calculated the average action density as

$$\frac{\langle S\rangle}{\Omega} = -\frac{\beta}{\Omega}\frac{\partial}{\partial\beta}\ln Z = \frac{\langle\varepsilon(S^R + S^I \tan S^I_{\mathrm{eff}})\rangle_{\mathrm{SQ},\mu}}{\Omega\,\langle\varepsilon\rangle_{\mathrm{SQ},\mu}}. \quad (19)$$

We first present the unoptimized and the (A_1, A_2)-optimized average action density in Fig. 4. In both cases 10^8 configurations were used. In the left panel of Fig. 4 we also compare with results from the sign problem-free worldline formalism. We see good agreement between the different predictions. Predictions with the (a_1, a_2)-optimized scheme also agree in the range where the sign problem of the scheme is manageable. In the right panel of Fig. 4 we also compare with analytic continuation from $\mu^2 < 0$ with polynomial ansätze of increasing order. The expansion converges rather slowly, even at small chemical potentials, way before the phase transition to the ordered phase.

The significant improvement of the sign problem achieved by path optimization is clearly visible in the left panel of Fig. 5, where we compare the average sign in the unoptimized case and in the optimized case with parametrizations (a_1, a_2) and (A_1, A_2). The ratio of this quantity between our best parametrization and the unoptimized case is already of order 10^2 at $\mu^2 = 0.3$, beyond which the unoptimized approach fails. For the (A_1, A_2) parametrization, the path optimization method instead works well also deep in the ordered phase at $\mu^2 > \mu_c^2 \approx 0.54$. An exponential fit $\langle\varepsilon\rangle_{\mathrm{SQ},\mu} \sim e^{-C^{(\mu)}\mu^2}$

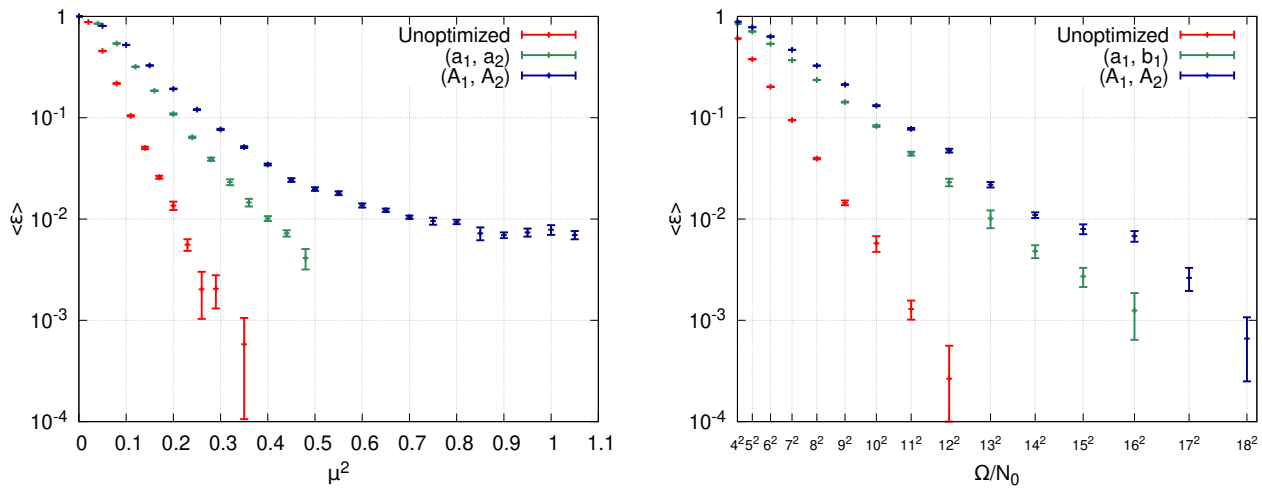

FIG. 5. Left: Dependence on μ^2 of the average sign for $\beta = 0.4$ and $\Omega = 8^3$ for the unoptimized calculations (red), and for the optimized calculations with (a_1, a_2) (green) and (A_1, A_2) (blue) parametrizations. Right: Volume dependence of the unoptimized (red) and (a_1, a_2)-optimized (green) and (A_1, A_2)-optimized (blue) average signs. Results were obtained at $\beta = 0.4$, $\mu^2 = 0.15$ with the temporal size fixed at $N_0 = 8$.

in the range $[0.1, 0.25]$ yields $C^{(\mu)}_{\rm unopt} \approx 24$, $C^{(\mu)}_{(a_1,a_2)} \approx 13$, and $C^{(\mu)}_{(A_1,A_2)} \approx 10$. Notice that since the sign problem should become mild at large μ, we expect the average sign to reach a minimum as a function of μ^2 (at fixed volume), and then increase. This probably explains the flattening of the (A_1, A_2) curve, starting from around $\mu^2 \approx 0.4$, which is likely related to the transition to the ordered phase. An eyeball estimate leads to expect several orders of magnitude of improvement in the central region where the sign problem is at its strongest.

In the right panel of Fig. 5 we show the dependence of the average sign on the volume for the unoptimized and for the (a_1, a_2) and (A_1, A_2)-optimized cases. A clear exponential decrease $\langle\varepsilon\rangle_{{\rm SQ},\mu} \sim e^{-C^{(V)}V}$ is visible, with $C^{(V)}_{\rm unopt} \approx 0.0073$, $C^{(V)}_{(a_1,a_2)} \approx 0.0032$, and $C^{(V)}_{(A_1,A_2)} \approx 0.0031$. The improvement of the sign problem by path optimization is exponential in the volume, reducing the "badness" $C^{(V)}$ of the volume scaling by 50%.

V. OPTIMIZATION WITH REWEIGHTING

Generating new configurations at every optimization step can be computationally too expensive in more complicated models. A possible way to decrease the computational cost is to generate a set of configurations before starting the optimization procedure, and use the same set at every step to compute the gradient of the cost function through reweighting. This would allow us to save time on the generation of configurations. In particular, using a fixed set of sufficiently decorrelated configurations instead of generating a set of similar quality at every step of the procedure saves us time required for the costly evaluation of the effective action on discarded configurations.

A convenient choice for the initial set of configurations, $\{\mathcal{C}_0\}$, is to generate them at the currently investigated value of μ^2 using the optimal parameters obtained through optimization at the previous value of μ^2; for the smallest nonzero μ^2, one can simply use those obtained with standard simulations with real phase variables at $\mu = 0$. We denote these parameters by $\boldsymbol{p}^{(0)}$, and the corresponding weights by $r[\mathcal{C}_0] = |\cos S^I_{\rm eff}[\mathcal{C}_0; \boldsymbol{p}^{(0)}]|\, e^{-S^R_{\rm eff}[\mathcal{C}_0;\boldsymbol{p}^{(0)}]}$. When computing the expectation values required by the optimization procedure, we need to reweight each configuration $\mathcal{C}_0$ in our fixed set from its original weight, $r[\mathcal{C}_0]$, to a new weight $w_n[\mathcal{C}_0] = |\cos S^I_{\rm eff}[\mathcal{C}_0; \boldsymbol{p}^{(n)}]|\, e^{-S^R_{\rm eff}[\mathcal{C}_0;\boldsymbol{p}^{(n)}]}$, corresponding to the updated values of the contour coefficients $\boldsymbol{p}^{(n)}$ obtained at the nth optimisation step. Setting for simplicity $S_{{\rm eff},n}[\mathcal{C}_0] \equiv S_{\rm eff}[\mathcal{C}_0; \boldsymbol{p}^{(n)}]$ (and omitting $\mathcal{C}_0$ in the following), the cost function at the nth step is

$$\left.\frac{N_-}{N_+}\right|_n \equiv \left.\frac{N_-}{N_+}\right|_{\boldsymbol{p}^{(n)}} = \frac{\langle \Theta(-\cos S^I_{{\rm eff},n})\, \frac{w_n}{r}\rangle_r}{\langle \Theta(\cos S^I_{{\rm eff},n})\, \frac{w_n}{r}\rangle_r}, \tag{20}$$

and its gradient with respect to the contour coefficients, needed for the $(n+1)$th step, is

$$\frac{\partial}{\partial p_i}\left(\frac{N_-}{N_+}\right)\bigg|_n = \left.\frac{N_-}{N_+}\right|_n \left[\frac{\langle \Theta(-\cos S^I_{{\rm eff},n})\, F_{i,n}\, \frac{w_n}{r}\rangle_r}{\langle \Theta(-\cos S^I_{{\rm eff},n})\, \frac{w_n}{r}\rangle_r} - \frac{\langle \Theta(\cos S^I_{{\rm eff},n})\, F_{i,n}\, \frac{w_n}{r}\rangle_r}{\langle \Theta(\cos S^I_{{\rm eff},n})\, \frac{w_n}{r}\rangle_r}\right], \tag{21}$$

where $\langle \ldots \rangle_r$ denotes averaging over $\{\mathcal{C}_0\}$ with the original weights r, and $F_{i,n}$ are the F_i of Eq. (10) evaluated with coefficients $\boldsymbol{p}^{(n)}$.

One downside of this method is that it might introduce an overlap problem, which we need to monitor throughout the optimization procedure. This can be done by looking at the numerical results for the denominators in Eq. (21). When these fall below a prescribed tolerance level and the overlap problem becomes too severe, say, after the n_*th step, we generate a new set of configurations $\{\mathcal{C}_*\}$ with the current values of the parameters, $\boldsymbol{p}^{(n_*)}$, replace $\{\mathcal{C}_0\}$ with $\{\mathcal{C}_*\}$ and $\boldsymbol{p}^{(0)}$ with $\boldsymbol{p}^{(n_*)}$, and proceed with the updates as before, essentially restarting the process but form a better position. Thus, by construction, the optimization procedure never lets the overlap problem between the different contours become severe. Note that this overlap problem in the optimization by no means can bias the final results, since even on suboptimal contours the integral is guaranteed to be the same by the multi-dimensional Cauchy theorem. On the other hand, it could lead to a loss in the improvement on the sign problem: that is why the second step of the procedure - generating new configurations when the overlap problem becomes too severe - is useful.

As one can see in Fig. 6, this modified procedure yielded similarly good results as the procedure used previously, which required the generation of new configurations at every step of the iteration. Here we used the parametrization of Eq. (13). We observed a good agreement between the values of the coefficients obtained with the two methods, and an almost equally good reduction of the severity of the sign problem. The optimal values for $\beta = 0.4$, $\Omega = 8^3$ and $\mu^2 = 0.15$ are shown in Tab. IV of Appendix A. With the new method, using a tolerance of 0.7 we had to regenerate configurations only four times during the contour optimization for the eight chemical potentials in Fig. 6 (three of which at the first, larger step in μ), thus considerably speeding up the procedure. Moreover, the fact that we achieve almost the same gain in the severity of the sign problem as with the full-fledged procedure dicussed previously provides an indirect indication that the overlap problem is not severe in this case: had it been, a much smaller gain would have been obtained.

For the volume used here, we barely detected the overlap problem between different contours in the numerical data. For larger volumes one expects this overlap problem to become severe more often, increasing the frequency at which the procedure needs to be restarted. However, the numerical evidence suggests that the sign problem is much more severe, even on the optimized contours. The overlap problem might then not even appear in practice, as by the time enough statistics is gathered to fight the sign problem, one also has the statistics to reweight to nearby contours. Moreover, as noted above, the optimal parameters for this model show little dependence on the volume, and one may achieve a sizeable reduction of the sign problem by carrying out the optimization procedure on small volumes, where the use of the reweighting technique discussed here can make it even more efficient.

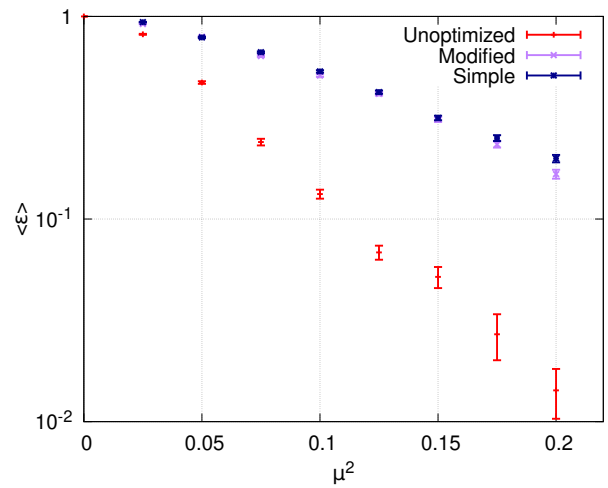

FIG. 6. Comparison of the average sign achieved with the simple and the modified optimization method at $\beta = 0.4$ and $\Omega = 8^3$ with the ansatz of Eq. (13).

VI. DISCUSSION

In this paper we studied the path optimization method for reducing the severity of the sign problem in the $2+1$ dimensional XY model with nonzero chemical potential. We used simple parametrizations for the complexified field variables. We have shown that the optimized manifold exhibits an explicit temporal translational invariance. Exploiting this property allows us to use significantly fewer optimization parameters. Furthermore, we have found that the optimal choice of contours appears to be independent of the spatial size of the lattice. Such a feature can be utilized to make the optimization procedure computationally less expensive as it would be sufficient to find the optimal contours for a small lattice. Then simulations can be carried out for larger lattices using the same contours.

We have shown numerical evidence that the reduction of the sign problem is exponential both in the chemical potential and the volume - i.e. it considerably reduces the exponents characterizing its severity. This was achieved without changing the number of parameters with the volume, keeping the number of optimizable parameters at a small fixed value. We have empirically found that the optimal ansätze respect the translational invariance of the original theory. Furthermore, on the optimized integration manifolds, it was possible to simulate also on the other side of the transition to the ordered phase.

We have also demonstrated that it is sufficient to generate configurations only at the start of the optimization procedure. Then, the same set of configurations is used to compute the gradient of the cost function at each step. As the contours are updated, it is necessary to reweight from the generated distribution to the one corresponding to the new contours. When the overlap between the

two distributions significantly decreases, it is preferable to generate new configurations, in order not to lose optimizing power. With this approach the computation time of the optimization can be significantly reduced as compared to a method where new configurations are generated at every iteration, with no significant loss in the reduction of the severity of the sign problem.

The success of our approach in this simple, yet nontrivial, purely bosonic model is a strong motivation to apply it to more complicated settings, especially in the presence of dynamical fermions. In this case there are additional difficulties, as the fermions effectively induce non-ultralocal interactions among bosonic fields, which may require some physical insight on the shape of the optimized contours.

ACKNOWLEDGEMENTS

This work was partly supported by the NKFIH grant KKP-126769. K.K. was supported by the ÚNKP-21-3-II-ELTE-625 New National Excellence Program of the Ministry of Human Capacities of Hungary. A.P. is supported by the J. Bolyai Research Scholarship of the Hungarian Academy of Sciences and by the ÚNKP-21-5 New National Excellence Program of the Ministry for Innovation and Technology from the source of the National Research, Development and Innovation Fund.

Appendix A: Optimal coefficient values

Here we present the optimal Fourier coefficients obtained at $\beta = 0.4$ and $\mu^2 = 0.15$ on $\Omega = 8 \times 4^2, 8 \times 6^2, 8^3$ and 8×10^2 lattices for $K = 2$. The optimal values for the parametrization described by Eq. (13) are given in Tabs. I and II while Tab. III shows the ones for Eq. (16). We compare the coefficients obtained with the reweighting technique introduced in Sec. V to those that were found using simple optimization in Tab. IV.

term	value, 8×4^2	value, 8×6^2	term	value, 8×4^2	value, 8×6^2
$A_{0,0}$	3.68×10^{-3}	-2.96×10^{-3}	$A_{0,4}$	6.15×10^{-4}	-1.18×10^{-3}
$\boldsymbol{A_{1,0}}$	$\mathbf{-1.08 \times 10^{-1}}$	$\mathbf{-1.19 \times 10^{-1}}$	$\boldsymbol{A_{1,4}}$	$\mathbf{-1.12 \times 10^{-1}}$	$\mathbf{-1.15 \times 10^{-1}}$
$B_{1,0}$	-6.57×10^{-4}	-9.40×10^{-4}	$B_{1,4}$	-9.48×10^{-4}	-1.90×10^{-3}
$\boldsymbol{A_{2,0}}$	$\mathbf{1.34 \times 10^{-2}}$	$\mathbf{1.29 \times 10^{-2}}$	$\boldsymbol{A_{2,4}}$	$\mathbf{1.29 \times 10^{-2}}$	$\mathbf{1.13 \times 10^{-2}}$
$B_{2,0}$	1.77×10^{-3}	-1.79×10^{-3}	$B_{2,4}$	6.92×10^{-4}	-2.93×10^{-4}
$A_{0,1}$	4.35×10^{-3}	-4.15×10^{-3}	$A_{0,5}$	-5.45×10^{-3}	1.69×10^{-3}
$\boldsymbol{A_{1,1}}$	$\mathbf{-1.15 \times 10^{-1}}$	$\mathbf{-1.13 \times 10^{-1}}$	$\boldsymbol{A_{1,5}}$	$\mathbf{-1.06 \times 10^{-1}}$	$\mathbf{-1.14 \times 10^{-1}}$
$B_{1,1}$	8.11×10^{-4}	-1.24×10^{-3}	$B_{1,5}$	-5.92×10^{-4}	1.86×10^{-3}
$\boldsymbol{A_{2,1}}$	$\mathbf{1.37 \times 10^{-2}}$	$\mathbf{1.36 \times 10^{-2}}$	$\boldsymbol{A_{2,5}}$	$\mathbf{1.19 \times 10^{-2}}$	$\mathbf{1.30 \times 10^{-2}}$
$B_{2,1}$	-1.26×10^{-3}	-5.67×10^{-4}	$B_{2,5}$	-5.27×10^{-4}	-1.33×10^{-4}
$A_{0,2}$	-3.66×10^{-3}	3.18×10^{-3}	$A_{0,6}$	2.53×10^{-3}	2.42×10^{-3}
$\boldsymbol{A_{1,2}}$	$\mathbf{-1.11 \times 10^{-1}}$	$\mathbf{-1.19 \times 10^{-1}}$	$\boldsymbol{A_{1,6}}$	$\mathbf{-1.20 \times 10^{-1}}$	$\mathbf{-1.14 \times 10^{-1}}$
$B_{1,2}$	-7.33×10^{-4}	-4.09×10^{-4}	$B_{1,6}$	8.31×10^{-4}	-2.00×10^{-3}
$\boldsymbol{A_{2,2}}$	$\mathbf{1.26 \times 10^{-2}}$	$\mathbf{1.28 \times 10^{-2}}$	$\boldsymbol{A_{2,6}}$	$\mathbf{1.44 \times 10^{-2}}$	$\mathbf{1.23 \times 10^{-2}}$
$B_{2,2}$	2.64×10^{-4}	-7.20×10^{-4}	$B_{2,6}$	-1.36×10^{-3}	4.47×10^{-5}
$A_{0,3}$	3.88×10^{-3}	2.35×10^{-3}	$A_{0,7}$	-4.80×10^{-3}	4.61×10^{-3}
$\boldsymbol{A_{1,3}}$	$\mathbf{-1.16 \times 10^{-1}}$	$\mathbf{-1.18 \times 10^{-1}}$	$\boldsymbol{A_{1,7}}$	$\mathbf{-1.09 \times 10^{-1}}$	$\mathbf{-1.20 \times 10^{-1}}$
$B_{1,3}$	-2.13×10^{-3}	-2.14×10^{-3}	$B_{1,7}$	-2.58×10^{-3}	1.24×10^{-5}
$\boldsymbol{A_{2,3}}$	$\mathbf{1.22 \times 10^{-2}}$	$\mathbf{1.40 \times 10^{-2}}$	$\boldsymbol{A_{2,7}}$	$\mathbf{1.35 \times 10^{-2}}$	$\mathbf{1.39 \times 10^{-2}}$
$B_{2,3}$	2.10×10^{-4}	8.39×10^{-4}	$B_{2,7}$	-1.68×10^{-3}	-1.04×10^{-3}

TABLE I. Optimal value of the Fourier coefficients obtained on $\Omega = 8 \times 4^2$ and 8×6^2 lattices for $\beta = 0.4$, $\mu^2 = 0.15$, $K = 2$ with parametrization given by Eq. (13). The most relevant values are written in bold while the others are negligible.

term	value, 8^3	value, 8×10^2	term	value, 8^3	value, 8×10^2
$A_{0,0}$	-1.47×10^{-4}	3.77×10^{-3}	$A_{0,4}$	4.15×10^{-4}	-2.28×10^{-3}
$\boldsymbol{A_{1,0}}$	$\mathbf{-1.23 \times 10^{-1}}$	$\mathbf{-1.21 \times 10^{-1}}$	$\boldsymbol{A_{1,4}}$	$\mathbf{-1.19 \times 10^{-1}}$	$\mathbf{-1.24 \times 10^{-1}}$
$B_{1,0}$	-3.05×10^{-4}	1.10×10^{-3}	$B_{1,4}$	1.70×10^{-3}	2.63×10^{-3}
$\boldsymbol{A_{2,0}}$	$\mathbf{1.54 \times 10^{-2}}$	$\mathbf{1.37 \times 10^{-2}}$	$\boldsymbol{A_{2,4}}$	$\mathbf{1.38 \times 10^{-2}}$	$\mathbf{1.29 \times 10^{-2}}$
$B_{2,0}$	5.83×10^{-4}	-1.43×10^{-4}	$B_{2,4}$	-7.15×10^{-5}	2.83×10^{-4}
$A_{0,1}$	-1.33×10^{-3}	-3.80×10^{-3}	$A_{0,5}$	-2.43×10^{-4}	-2.34×10^{-3}
$\boldsymbol{A_{1,1}}$	$\mathbf{-1.18 \times 10^{-1}}$	$\mathbf{-1.21 \times 10^{-1}}$	$\boldsymbol{A_{1,5}}$	$\mathbf{-1.20 \times 10^{-1}}$	$\mathbf{-1.23 \times 10^{-1}}$
$B_{1,1}$	1.05×10^{-3}	2.84×10^{-3}	$B_{1,5}$	-1.58×10^{-3}	6.28×10^{-4}
$\boldsymbol{A_{2,1}}$	$\mathbf{1.66 \times 10^{-2}}$	$\mathbf{1.58 \times 10^{-2}}$	$\boldsymbol{A_{2,5}}$	$\mathbf{1.21 \times 10^{-2}}$	$\mathbf{1.13 \times 10^{-2}}$
$B_{2,1}$	1.56×10^{-3}	1.78×10^{-3}	$B_{2,5}$	-1.64×10^{-3}	-1.46×10^{-5}
$A_{0,2}$	9.76×10^{-3}	8.90×10^{-3}	$A_{0,6}$	-2.74×10^{-3}	-4.49×10^{-3}
$\boldsymbol{A_{1,2}}$	$\mathbf{-1.20 \times 10^{-1}}$	$\mathbf{-1.21 \times 10^{-1}}$	$\boldsymbol{A_{1,6}}$	$\mathbf{-1.19 \times 10^{-1}}$	$\mathbf{-1.22 \times 10^{-1}}$
$B_{1,2}$	2.15×10^{-3}	1.31×10^{-3}	$B_{1,6}$	9.43×10^{-5}	-4.86×10^{-4}
$\boldsymbol{A_{2,2}}$	$\mathbf{1.47 \times 10^{-2}}$	$\mathbf{1.61 \times 10^{-2}}$	$\boldsymbol{A_{2,6}}$	$\mathbf{1.69 \times 10^{-2}}$	$\mathbf{1.42 \times 10^{-2}}$
$B_{2,2}$	2.10×10^{-3}	-7.80×10^{-4}	$B_{2,6}$	6.84×10^{-4}	-6.65×10^{-4}
$A_{0,3}$	-4.54×10^{-3}	-2.31×10^{-3}	$A_{0,7}$	5.67×10^{-3}	3.58×10^{-3}
$\boldsymbol{A_{1,3}}$	$\mathbf{-1.21 \times 10^{-1}}$	$\mathbf{-1.20 \times 10^{-1}}$	$\boldsymbol{A_{1,7}}$	$\mathbf{-1.22 \times 10^{-1}}$	$\mathbf{-1.23 \times 10^{-1}}$
$B_{1,3}$	-2.61×10^{-3}	-2.21×10^{-3}	$B_{1,7}$	1.49×10^{-3}	7.52×10^{-4}
$\boldsymbol{A_{2,3}}$	$\mathbf{1.35 \times 10^{-2}}$	$\mathbf{1.62 \times 10^{-2}}$	$\boldsymbol{A_{2,7}}$	$\mathbf{1.27 \times 10^{-2}}$	$\mathbf{1.29 \times 10^{-2}}$
$B_{2,3}$	-3.84×10^{-4}	-2.73×10^{-4}	$B_{2,7}$	6.48×10^{-4}	4.98×10^{-5}

TABLE II. Optimal value of the Fourier coefficients obtained on an $\Omega = 8^3$ lattice for $\beta = 0.4$, $\mu^2 = 0.15$, $K = 2$ with parametrization given by Eq. (13). The most relevant values are written in bold while the others are negligible.

term	value, 8×4^2	value, 8×6^2	value, 8^3	value, 8×10^2
$\boldsymbol{\bar{A}_{1,0}}$	$\mathbf{-1.15 \times 10^{-1}}$	$\mathbf{-1.15 \times 10^{-1}}$	$\mathbf{-1.18 \times 10^{-1}}$	$\mathbf{-1.20 \times 10^{-1}}$
$\bar{B}_{1,0}$	2.44×10^{-3}	-3.63×10^{-4}	-1.85×10^{-3}	-2.40×10^{-3}
$\boldsymbol{\bar{A}_{2,0}}$	$\mathbf{1.35 \times 10^{-2}}$	$\mathbf{1.48 \times 10^{-2}}$	$\mathbf{1.41 \times 10^{-2}}$	$\mathbf{1.39 \times 10^{-2}}$
$\bar{B}_{2,0}$	8.16×10^{-5}	-1.26×10^{-3}	-1.50×10^{-3}	1.57×10^{-3}
$\bar{A}_{1,1}$	5.44×10^{-3}	6.13×10^{-3}	3.68×10^{-3}	3.60×10^{-3}
$\bar{B}_{1,1}$	2.57×10^{-3}	-5.47×10^{-4}	-2.61×10^{-4}	2.42×10^{-3}
$\bar{A}_{2,1}$	-5.32×10^{-4}	-1.92×10^{-3}	-5.14×10^{-5}	1.42×10^{-3}
$\bar{B}_{2,1}$	5.13×10^{-4}	6.43×10^{-5}	1.06×10^{-3}	1.56×10^{-4}
$\bar{A}_{1,2}$	6.04×10^{-3}	5.62×10^{-3}	5.16×10^{-3}	4.43×10^{-3}
$\bar{B}_{1,2}$	2.22×10^{-3}	6.81×10^{-4}	-3.57×10^{-4}	-1.23×10^{-3}
$\bar{A}_{2,2}$	2.48×10^{-4}	-1.13×10^{-3}	-5.08×10^{-4}	-1.40×10^{-3}
$\bar{B}_{2,2}$	-3.67×10^{-4}	-7.64×10^{-4}	3.33×10^{-4}	1.15×10^{-3}

TABLE III. Optimal values of the Fourier coefficients obtained on $\Omega = 8 \times 4^2, 8 \times 6^2, 8^3$ and 8×10^2 lattices for $\beta = 0.4$, $\mu^2 = 0.15$ and $K = 2$ with parametrization given by Eq. (16). The most relevant values are written in bold while the others are negligible.

term	modified	simple	term	modified	simple
$A_{0,0}$	-1.48×10^{-4}	-5.66×10^{-3}	$A_{0,4}$	4.15×10^{-4}	-3.06×10^{-2}
$\boldsymbol{A_{1,0}}$	$\mathbf{-1.23 \times 10^{-1}}$	$\mathbf{-1.11 \times 10^{-1}}$	$\boldsymbol{A_{1,4}}$	$\mathbf{-1.19 \times 10^{-1}}$	$\mathbf{-1.19 \times 10^{-1}}$
$B_{1,0}$	-3.06×10^{-4}	1.66×10^{-3}	$B_{1,4}$	1.71×10^{-3}	3.67×10^{-3}
$\boldsymbol{A_{2,0}}$	$\mathbf{1.55 \times 10^{-2}}$	$\mathbf{1.03 \times 10^{-2}}$	$\boldsymbol{A_{2,4}}$	$\mathbf{1.38 \times 10^{-2}}$	$\mathbf{1.03 \times 10^{-2}}$
$B_{2,0}$	5.83×10^{-4}	2.40×10^{-3}	$B_{2,4}$	-7.15×10^{-5}	-4.33×10^{-3}
$A_{0,1}$	-1.33×10^{-3}	-5.65×10^{-3}	$A_{0,5}$	-2.43×10^{-4}	4.74×10^{-3}
$\boldsymbol{A_{1,1}}$	$\mathbf{-1.19 \times 10^{-1}}$	$\mathbf{-1.09 \times 10^{-1}}$	$\boldsymbol{A_{1,5}}$	$\mathbf{-1.20 \times 10^{-1}}$	$\mathbf{-1.21 \times 10^{-1}}$
$B_{1,1}$	1.05×10^{-3}	2.51×10^{-3}	$B_{1,5}$	-1.58×10^{-3}	-6.40×10^{-3}
$\boldsymbol{A_{2,1}}$	$\mathbf{1.66 \times 10^{-2}}$	$\mathbf{1.44 \times 10^{-2}}$	$\boldsymbol{A_{2,5}}$	$\mathbf{1.22 \times 10^{-2}}$	$\mathbf{1.63 \times 10^{-2}}$
$B_{2,1}$	1.57×10^{-3}	-5.24×10^{-3}	$B_{2,5}$	-1.65×10^{-3}	3.31×10^{-4}
$A_{0,2}$	9.76×10^{-3}	-3.26×10^{-3}	$A_{0,6}$	-2.75×10^{-3}	-1.30×10^{-3}
$\boldsymbol{A_{1,2}}$	$\mathbf{-1.21 \times 10^{-1}}$	$\mathbf{-1.27 \times 10^{-1}}$	$\boldsymbol{A_{1,6}}$	$\mathbf{-1.20 \times 10^{-1}}$	$\mathbf{-1.04 \times 10^{-1}}$
$B_{1,2}$	2.15×10^{-3}	-1.21×10^{-2}	$B_{1,6}$	9.44×10^{-5}	1.26×10^{-3}
$\boldsymbol{A_{2,2}}$	$\mathbf{1.47 \times 10^{-2}}$	$\mathbf{1.81 \times 10^{-2}}$	$\boldsymbol{A_{2,6}}$	$\mathbf{1.70 \times 10^{-2}}$	$\mathbf{6.71 \times 10^{-3}}$
$B_{2,2}$	2.11×10^{-3}	8.61×10^{-3}	$B_{2,6}$	6.84×10^{-4}	1.36×10^{-4}
$A_{0,3}$	-4.55×10^{-3}	-2.27×10^{-2}	$A_{0,7}$	5.68×10^{-3}	1.61×10^{-3}
$\boldsymbol{A_{1,3}}$	$\mathbf{-1.22 \times 10^{-1}}$	$\mathbf{-1.22 \times 10^{-1}}$	$\boldsymbol{A_{1,7}}$	$\mathbf{-1.22 \times 10^{-1}}$	$\mathbf{-1.18 \times 10^{-1}}$
$B_{1,3}$	-2.62×10^{-3}	9.84×10^{-3}	$B_{1,7}$	1.49×10^{-3}	-1.48×10^{-3}
$\boldsymbol{A_{2,3}}$	$\mathbf{1.36 \times 10^{-2}}$	$\mathbf{1.28 \times 10^{-2}}$	$\boldsymbol{A_{2,7}}$	$\mathbf{1.28 \times 10^{-2}}$	$\mathbf{1.36 \times 10^{-2}}$
$B_{2,3}$	-3.84×10^{-4}	7.24×10^{-4}	$B_{2,7}$	6.49×10^{-4}	6.42×10^{-3}

TABLE IV. Comparison of the values of Fourier coefficients obtained with the simple and the reweighted optimization method on an $\Omega = 8^3$ lattice for $\beta = 0.4$, $\mu^2 = 0.15$ and $K = 2$ with parametrization given by Eq. (13).

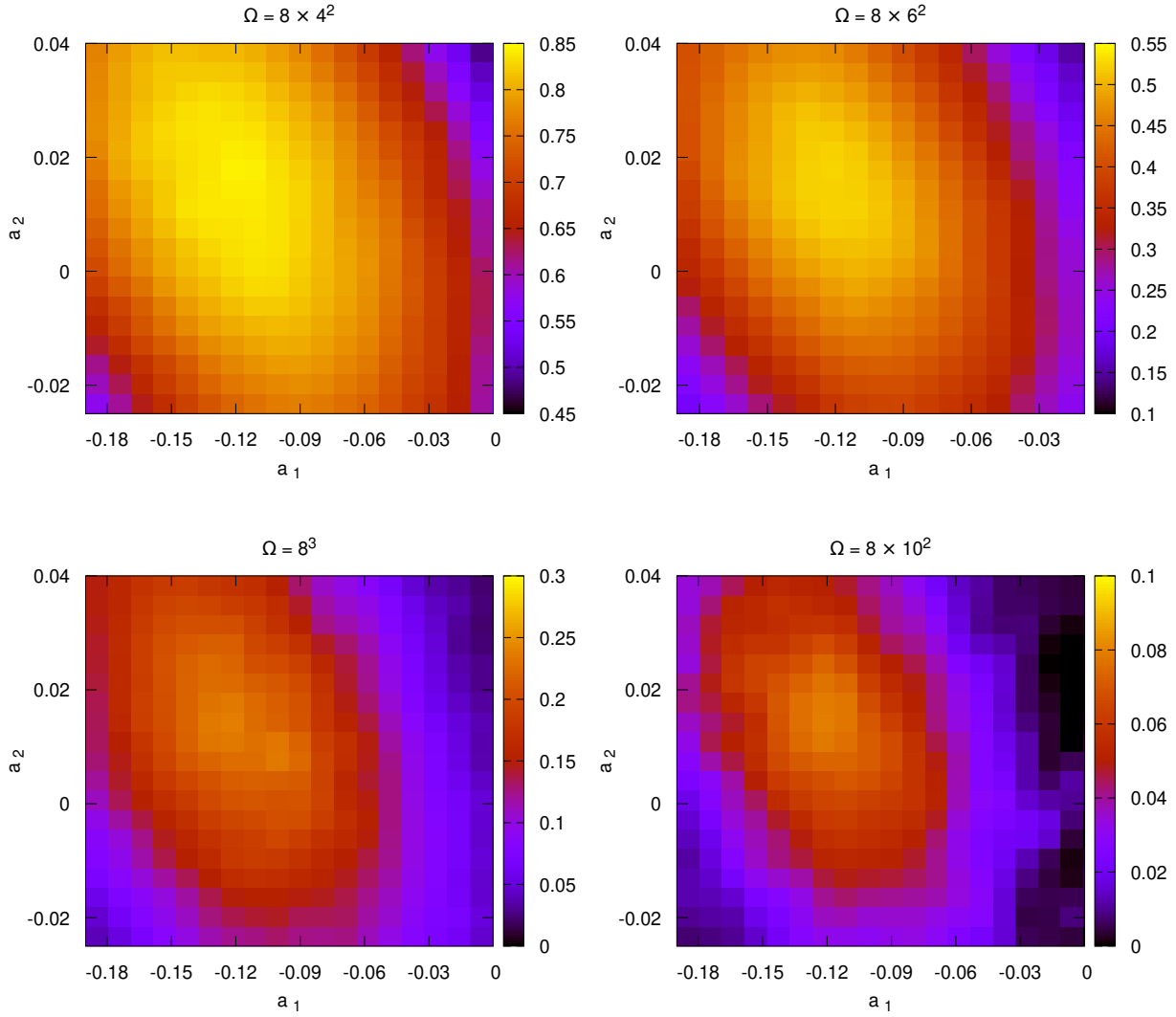

FIG. 7. Average sign over the space of Fourier coefficients a_1 and a_2 of Eq. (16) with the (a_1, a_2) setup. Scans were carried out on $\Omega = 8 \times 4^2, 8 \times 6^2, 8^3$ and 8×10^2 lattices at $\beta = 0.4$ and $\mu^2 = 0.15$.

Appendix B: Scans in coefficient space

Figure 7 shows the average sign over the space of coefficients a_1 and a_2 of the parametrization described by Eq. (18) in the (a_1, a_2) setup with $\beta = 0.4$ and $\mu^2 = 0.15$. Each scan was performed at different spatial sizes with temporal size fixed at $N_0 = 8$.

[1] S. Chandrasekharan and U.-J. Wiese, "Meron cluster solution of a fermion sign problem," *Phys. Rev. Lett.*, vol. 83, pp. 3116–3119, 1999.

[2] M. G. Alford, S. Chandrasekharan, J. Cox, and U. J.

Wiese, "Solution of the complex action problem in the Potts model for dense QCD," *Nucl. Phys. B*, vol. 602, pp. 61–86, 2001.

[3] M. G. Endres, "Method for simulating O(N) lattice models at finite density," *Phys. Rev. D*, vol. 75, p. 065012, 2007.

[4] F. Bruckmann, C. Gattringer, T. Kloiber, and T. Sulejmanpasic, "Dual lattice representations for O(N) and CP(N−1) models with a chemical potential," *Phys. Lett. B*, vol. 749, pp. 495–501, 2015. [Erratum: Phys.Lett.B 751, 595–595 (2015)].

[5] C. Gattringer, T. Kloiber, and V. Sazonov, "Solving the sign problems of the massless lattice Schwinger model with a dual formulation," *Nucl. Phys. B*, vol. 897, pp. 732–748, 2015.

[6] Z. Fodor, S. D. Katz, and C. Schmidt, "The Density of states method at non-zero chemical potential," *JHEP*, vol. 03, p. 121, 2007.

[7] G. Endrődi, Z. Fodor, S. D. Katz, D. Sexty, K. K. Szabó, and C. Török, "Applying constrained simulations for low temperature lattice QCD at finite baryon chemical potential," *Phys. Rev. D*, vol. 98, no. 7, p. 074508, 2018.

[8] M. Giordano, K. Kapás, S. D. Katz, D. Nógrádi, and A. Pásztor, "New approach to lattice QCD at finite density; results for the critical end point on coarse lattices," *JHEP*, vol. 05, p. 088, 2020.

[9] S. Borsányi, Z. Fodor, M. Giordano, S. D. Katz, D. Nógrádi, A. Pásztor, and C. H. Wong, "Lattice simulations of the QCD chiral transition at real baryon density," 8 2021.

[10] G. Parisi, "On complex probabilities," *Phys. Lett. B*, vol. 131, pp. 393–395, 1983.

[11] G. Aarts, E. Seiler, and I.-O. Stamatescu, "The Complex Langevin method: When can it be trusted?," *Phys. Rev. D*, vol. 81, p. 054508, 2010.

[12] E. Seiler, D. Sexty, and I.-O. Stamatescu, "Gauge cooling in complex Langevin for QCD with heavy quarks," *Phys. Lett. B*, vol. 723, pp. 213–216, 2013.

[13] G. Aarts, E. Seiler, D. Sexty, and I.-O. Stamatescu, "Complex Langevin dynamics and zeroes of the fermion determinant," *JHEP*, vol. 05, p. 044, 2017. [Erratum: JHEP 01, 128 (2018)].

[14] M. Scherzer, E. Seiler, D. Sexty, and I.-O. Stamatescu, "Complex Langevin and boundary terms," *Phys. Rev. D*, vol. 99, no. 1, p. 014512, 2019.

[15] M. Golterman, Y. Shamir, and B. Svetitsky, "Breakdown of staggered fermions at nonzero chemical potential," *Phys. Rev. D*, vol. 74, p. 071501, 2006.

[16] M. Giordano, K. Kapás, S. D. Katz, D. Nógrádi, and A. Pásztor, "Radius of convergence in lattice QCD at finite μ_B with rooted staggered fermions," *Phys. Rev. D*, vol. 101, no. 7, p. 074511, 2020.

[17] A. Alexandru, G. Başar, P. F. Bedaque, and N. C. Warrington, "Complex Paths Around The Sign Problem," 7 2020.

[18] M. Cristoforetti, F. Di Renzo, and L. Scorzato, "New approach to the sign problem in quantum field theories: High density QCD on a Lefschetz thimble," *Phys. Rev. D*, vol. 86, p. 074506, 2012.

[19] M. Cristoforetti, F. Di Renzo, A. Mukherjee, and L. Scorzato, "Monte Carlo simulations on the Lefschetz thimble: Taming the sign problem," *Phys. Rev. D*, vol. 88, no. 5, p. 051501, 2013.

[20] A. Alexandru, G. Başar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, "Sign problem and Monte Carlo calculations beyond Lefschetz thimbles," *JHEP*, vol. 05, p. 053, 2016.

[21] M. Fukuma and N. Matsumoto, "Worldvolume approach to the tempered Lefschetz thimble method," *PTEP*, vol. 2021, no. 2, p. 023B08, 2021.

[22] F. Di Renzo, S. Singh, and K. Zambello, "Taylor expansions on Lefschetz thimbles," *Phys. Rev. D*, vol. 103, no. 3, p. 034513, 2021.

[23] F. Di Renzo and K. Zambello, "Solution of the Thirring model in thimble regularization," *Phys. Rev. D*, vol. 105, no. 5, p. 054501, 2022.

[24] S. Lawrence, "Beyond Thimbles: Sign-Optimized Manifolds for Finite Density," *PoS*, vol. LATTICE2018, p. 149, 2018.

[25] Y. Mori, K. Kashiwa, and A. Ohnishi, "Toward solving the sign problem with path optimization method," *Phys. Rev. D*, vol. 96, no. 11, p. 111501, 2017.

[26] F. Bursa and M. Kroyter, "A simple approach towards the sign problem using path optimisation," *JHEP*, vol. 12, p. 054, 2018.

[27] K. Kashiwa, Y. Mori, and A. Ohnishi, "Controlling the model sign problem via the path optimization method: Monte Carlo approach to a QCD effective model with Polyakov loop," *Phys. Rev. D*, vol. 99, no. 1, p. 014033, 2019.

[28] Y. Mori, K. Kashiwa, and A. Ohnishi, "Path optimization in 0+1D QCD at finite density," *PTEP*, vol. 2019, no. 11, p. 113B01, 2019.

[29] Y. Mori, K. Kashiwa, and A. Ohnishi, "Application of a neural network to the sign problem via the path optimization method," *PTEP*, vol. 2018, no. 2, p. 023B04, 2018.

[30] A. Alexandru, P. F. Bedaque, H. Lamm, and S. Lawrence, "Finite-Density Monte Carlo Calculations on Sign-Optimized Manifolds," *Phys. Rev. D*, vol. 97, no. 9, p. 094510, 2018.

[31] A. Alexandru, P. F. Bedaque, H. Lamm, S. Lawrence, and N. C. Warrington, "Fermions at Finite Density in 2+1 Dimensions with Sign-Optimized Manifolds," *Phys. Rev. Lett.*, vol. 121, no. 19, p. 191602, 2018.

[32] F. Bursa and M. Kroyter, "Optimisation of complex integration contours at higher order," *JHEP*, vol. 04, p. 181, 2021.

[33] W. Detmold, G. Kanwar, M. L. Wagman, and N. C. Warrington, "Path integral contour deformations for noisy observables," *Phys. Rev. D*, vol. 102, no. 1, p. 014514, 2020.

[34] W. Detmold, G. Kanwar, H. Lamm, M. L. Wagman, and N. C. Warrington, "Path integral contour deformations for observables in $SU(N)$ gauge theory," *Phys. Rev. D*, vol. 103, no. 9, p. 094517, 2021.

[35] K. Kashiwa and Y. Mori, "Path optimization for $U(1)$ gauge theory with complexified parameters," *Phys. Rev. D*, vol. 102, no. 5, p. 054519, 2020.

[36] A. Roberge and N. Weiss, "Gauge Theories With Imaginary Chemical Potential and the Phases of QCD," *Nucl. Phys. B*, vol. 275, pp. 734–745, 1986.

[37] G. Aarts, F. A. James, E. Seiler, and I.-O. Stamatescu, "Adaptive stepsize and instabilities in complex Langevin dynamics," *Phys. Lett. B*, vol. 687, pp. 154–159, 2010.

[38] G. Aarts and F. A. James, "On the convergence of complex Langevin dynamics: The Three-dimensional XY model at finite chemical potential," *JHEP*, vol. 08, p. 020,

2010.

[39] M. Scherzer, E. Seiler, D. Sexty, and I. O. Stamatescu, "Controlling Complex Langevin simulations of lattice models by boundary term analysis," *Phys. Rev. D*, vol. 101, no. 1, p. 014501, 2020.

[40] D. Banerjee and S. Chandrasekharan, "Finite size effects in the presence of a chemical potential: A study in the classical non-linear O(2) sigma-model," *Phys. Rev. D*, vol. 81, p. 125007, 2010.

[41] K. Langfeld, "Phase diagram of the quantum O(2)-model in 2+1 dimensions," *Phys. Rev. D*, vol. 87, no. 11, p. 114504, 2013.

[42] J. Bloch, R. G. Jha, R. Lohmayer, and M. Meister, "Tensor renormalization group study of the three-dimensional O(2) model," *Phys. Rev. D*, vol. 104, no. 9, p. 094517, 2021.

[43] P. de Forcrand, S. Kim, and T. Takaishi, "QCD simulations at small chemical potential," *Nucl. Phys. B Proc. Suppl.*, vol. 119, pp. 541–543, 2003.

[44] A. Alexandru, M. Faber, I. Horváth, and K.-F. Liu, "Lattice QCD at finite density via a new canonical approach," *Phys. Rev. D*, vol. 72, p. 114513, 2005.