Complex Langevin dynamics for dynamical QCD at nonzero chemical potential: A comparison with multiparameter reweighting

Z. Fodor,1,2,3 S. D. Katz,3,4 D. Sexty,1 and C. Török1,3,4

1Department of Physics, Wuppertal University, Gaussstrasse 20, D-42119 Wuppertal, Germany
2Jülich Supercomputing Centre, Forschungszentrum Jülich, D-52425 Jülich, Germany
3Institute for Theoretical Physics, Eötvös University, Pázmány P. sétány I/A, H-1117 Budapest, Hungary
4MTA-ELTE “Lendület” Lattice Gauge Theory Research Group, Pázmány P. sétány I/A, H-1117 Budapest, Hungary

(Received 26 August 2015; published 24 November 2015)

We study lattice QCD at nonvanishing chemical potential using the complex Langevin equation. We compare the results with multiparameter reweighting both from $\mu = 0$ and phase-quenched ensembles. We find a good agreement for lattice spacings below $\approx 0.15$ fm. On coarser lattices the complex Langevin approach breaks down. Four flavors of staggered fermions are used on $N_t = 4, 6$ and 8 lattices. For one ensemble we also use two flavors to investigate the effects of rooting.

DOI: 10.1103/PhysRevD.92.094516 PACS numbers: 11.15.Ha, 12.38.Gc

I. INTRODUCTION AND OVERVIEW

Dense and/or high-temperature phases of strongly interacting matter are becoming experimentally accessible nowadays due to heavy ion collision experiments at the Relativistic Heavy Ion Collider, the Large Hadron Collider, and especially the FAIR facility at GSI, as well as astrophysical observations of neutron stars. Theoretical understanding of the dense, strongly interacting phases and the first-principles determination of the phase diagram of QCD as a function of the temperature and chemical potential are still lacking. This is a consequence of the sign problem, which makes lattice calculations at nonzero baryon density challenging.

The standard nonperturbative tool for QCD, lattice QCD, is defined by the path integral

$$Z = \int DU e^{-S_{\text{YM}} \det M(\mu)} \quad (1)$$

with the Yang-Mills action $S_{\text{YM}}$ of the gluons and the fermion determinant $M(\mu)$ on a cubic space-time lattice. At nonzero chemical potential the determinant is nonreal; therefore, importance sampling methods are not applicable. For a review of ideas to circumvent the sign problem see [1–3].

One of the ways to avoid the sign problem is using the analyticity of the action and complexifying the field manifolds of the theory with the complex Langevin equation (CLE) [4,5]. (See also the related but distinct approach of the Lefschetz thimbles, where the integration contours are pushed into the complex plane [6].)

After promising initial results, it was noticed that the complex Langevin equation can also deliver convergent but wrong results in some cases [7,8]. Also technical problems could arise which are avoided using adaptive step sizes for the Langevin equation [7,9]. In the last decade the method has enjoyed increasing attention related to real-time systems [10–14], as well as finite-density problems [15–27]. The method showed remarkable success in the case of finite density Bose gas [16] or the SU(3) spin model [18,19], but the breakdown of the method was also observed a few times [17,22]. The theoretical understanding of the successes and the failures of the method has improved: it has been proved that provided a few requirements (some “offline” such as the holomorphicity of the action and the observables, and some “online” such as the quick decay of the field distributions at infinity) the method will provide correct results [28,29].

It has been recently demonstrated that complex Langevin simulations of gauge theories are made feasible using the procedure of gauge cooling [30,31] (see also [32]), which helps to reduce the fluctuations corresponding to the complexified gauge freedom of the theory. This method was first used to solve heavy dense QCD (HDQCD) where the quarks are kept static (their spatial hopping terms are dropped) [30,31], and it has been also extended to full QCD using light quarks in the staggered [33] as well as the Wilson formulation [34]. Gauge cooling makes the investigation of QCD with a theta term also possible [35].

In this paper we compare results of the reweighting approach and the complex Langevin approach for $N_F = 4$ and $N_F = 2$ QCD using staggered fermions.

In Sec. II we give a brief overview of the complex Langevin method. In Sec. III we summarize the reweighting method. In Sec. IV we present our numerical results comparing the reweighting and complex Langevin simulations. Finally, we conclude in Sec. V.

II. THE COMPLEX LANGEVIN EQUATION

The complex Langevin equation [4,5] is a straightforward generalization of the real Langevin equation [36]. For the link variables $U_{a,b}$ of lattice QCD an update with Langevin time step $\epsilon$ reads [37]
with \( \lambda_a \) the generators of the gauge group, i.e. the Gell-Mann matrices, and the Gaussian noise \( \eta_{axv} \). The drift force \( K_{axv} \) is determined from the action \( S[U] \) by

\[
K_{axv} = -D_{axv}S[U]
\]

with the left derivative

\[
D_{axv}f(U) = \frac{\partial f(e^{iK_{axv}} U_{x,v})}{\partial \alpha_{x,v}}|_{\alpha_{x,v}=0}.
\]

In case the drift term is nonreal the manifold of the link variables is complexified to \( SL(3, \mathbb{C}) \). The original theory is recovered by taking averages of the observables analytically continued to the complexified manifold.

For the case of QCD the action of the theory involves the fermionic determinant through the complex logarithm

\[
S_{\text{eff}} = S_{YM} - \ln \det M(\mu). \tag{5}
\]

The drift term in turn is given by

\[
K_{axv} = -D_{axv}S_{YM}[U] + \frac{N_F}{4} \text{Tr}[M^{-1}(\mu, U)D_{axv}M(\mu, U)], \tag{6}
\]

where the second term is calculated using one conjugate gradient inversion per update using noise vectors [33]. The action we are interested in is thus nonholomorphic, and in turn this results in a drift term which has singularities where the fermionic measure \( \det M(\mu, U) \) is vanishing.

The theoretical understanding of the behavior of the theory with a meromorphic drift term is still lacking, but we have some observations as detailed below. Such a drift term seems to lead to incorrect results in toy models if the trajectories encircle the origin frequently [24,38]. In other cases the simulations yield a correct result in spite of a logarithm in the action [19,20]. In [27] an explicit example is presented where the simulations give correct results in spite of the frequent rotations of the phase of the measure. The condition for correctness is that the distribution of configurations vanishes sufficiently fast (faster than linearly) near the pole.

For QCD itself we have a few indications that the poles do not affect the simulations at high temperatures: observing the spectrum of the Dirac operator [39], comparisons with expansions which use a holomorphic action [34], and the results presented in this paper. It remains to be seen whether simulations in the confined phase are affected.

The “distance” of a configuration from the original SU(3) manifold can be monitored with the unitarity norm

\[
1 \frac{1}{4\Omega} \sum_{x,v} \text{Tr}((U_{x,v} U_{x,v}^{-1} - 1)^2), \tag{7}
\]

where \( \Omega = N^3 \tau \) is the volume of the lattice. In naive complex Langevin simulations, this distance grows exponentially, and the simulation breaks down because of numerical problems if it gets too large. This behavior can be countered with gauge cooling, which means that several gauge transformations of the enlarged manifold are performed in the direction of the steepest descent of the unitarity norm (7) [30,31]. With gauge cooling, the unitarity norm remains bounded at a safe level as long as the \( \beta \) parameter of the action is not too small. The value \( \beta_{\text{min}} \) corresponds to a maximal lattice spacing, which seems to depend weakly on the lattice size, as can be checked easily for the cheaper HDQCD theory [40].

### III. REWEIGHTING

In the multiparameter reweighting approach one rewrites the partition function as [41]

\[
Z = \int DU e^{-S_{YM}(\beta)} \det M(\mu) = \int DU e^{-S_{YM}(\beta_0)} \det M(\mu_0) \left\{ e^{-S_{YM}(\beta) + S_{YM}(\beta_0)} \det M(\mu) \right\} \det M(\mu_0), \tag{8}
\]

where \( \mu_0 \) is chosen such that the second line contains a positive definite measure which can be used to generate the configurations and the terms in the curly bracket in the last line are taken into account as an observable. The expectation value of any observable can be then written in the form

\[
\langle O \rangle_{\beta, \mu} = \frac{\sum O(\beta, \mu) w(\beta, \beta_0, \mu, \mu_0)}{\sum w(\beta, \beta_0, \mu, \mu_0)} \tag{9}
\]

with \( w(\beta, \beta_0, \mu, \mu_0) \) being the weights of the configurations defined by the curly bracket of Eq. (8). Note that gauge observables do not explicitly depend on \( \mu \); therefore, their \( \mu \) dependence comes entirely from the weight factors. Fermionic observables, on the other hand, also explicitly depend on the chemical potential.

In this paper we use two choices for the original, positive measure ensemble. The first choice is to use \( \mu_0 = 0 \), i.e. reweighting from zero chemical potential. For any choice of the target \( \beta, \mu \) parameters one can find the optimal \( \beta_0 \) for which the fluctuation of the weights \( w(\beta, \mu) \) is minimal. This corresponds to the best reweighting line as discussed in [42–44]. We generated configurations at \( \mu = 0 \) for \( \beta \) in the range 4.9–5.5. These were then used to reach the entire \( \mu, \beta \) plane via multiparameter reweighting. Our second choice is to use the phase-quenched ensemble, i.e. replacing \( \det M(\mu_0) \) by \( | \det M(\mu) | \) in Eqs. (8) and (10). In this
case the reweighting factor contains only the phase of the determinant.

For staggered fermions an additional rooting is required; for \( N_F \) flavors the weights become

\[
w(\beta, \mu_0, \mu, \phi_0) = e^{-S_{\text{YM}}(\beta) + S_{\text{YM}}(\phi_0)} \left( \frac{\det M(\mu)}{\det M(\mu_0)} \right)^{N_F/4}. \tag{10}
\]

Since for \( N_F < 4 \) a fractional power is taken which has cuts on the complex plane it is important to choose these cuts such that the weights are analytic for real \( \mu \) values. This can be achieved by expressing \( \det M(\mu) \) analytically as a function of \( \mu \) as discussed in [45, 46].

**IV. RESULTS**

We use the Wilson plaquette action for the gauge sector of the theory and unimproved staggered fermions with \( N_F = 4 \) flavors if not otherwise noted. We have used three different lattice sizes for this study: \( 8^3 \times 4, 12^3 \times 6 \) and \( 16^3 \times 8 \), all having the aspect ratio \( L_s/L_t = 2 \).

Our main observables are the plaquette averages, the spatial average of the trace of the Polyakov loop

\[
\sum_x \text{Tr} P(x)/N_x^3, \quad P(x) = \prod_{i=1...N_t} U_4(x, i) \tag{11}
\]

and its inverse \( \sum_x \text{Tr} P^{-1}(x)/N_x^3 \), the chiral condensate \( \langle \bar{\psi} \psi \rangle \) and the fermionic density \( n \) defined as

\[
\langle \bar{\psi} \psi \rangle = \frac{1}{\Omega} \left( \frac{\partial \ln Z}{\partial m} \right), \quad n = \frac{1}{\Omega} \left( \frac{\partial \ln Z}{\partial \mu} \right), \tag{12}
\]

with \( \Omega \) the volume of the space-time lattice. We are also interested in the average phase of the fermion determinant, which measures the severity of the sign problem

\[
\langle e^{2i\psi} \rangle = \left( \frac{\det M(\mu)}{\det M(-\mu)} \right). \tag{13}
\]

We perform the complex Langevin simulations using adaptive step size, with a control parameter which puts the typical step sizes in the range \( \epsilon \approx 10^{-5} - 5 \times 10^{-5} \). Using such small step sizes allows us to avoid having to take the \( \epsilon \to 0 \) limit as the results are in the zero Langevin step limit within errors. We use initial conditions on the SU(3) manifold and allow \( \tau = 10 - 30 \) Langevin time for thermalization, after which we perform the measurements for another \( \tau = 10 - 30 \) Langevin time. We checked that proper thermalization is reached by observing that halving the thermalization time leads to consistent results.

We have determined the pion masses as well as the lattice spacing using the \( w_0 \) scale as proposed in [47] for several quark masses; see Table I. One sees that choosing the quark masses \( ma = 0.05 \) for the \( N_t = 4 \) lattice, \( ma = 0.02 \) for the \( N_t = 6 \) lattice and \( ma = 0.01 \) for the \( N_t = 8 \) lattice, in the vicinity of the critical temperature we have \( m_\pi/T_c \approx 2.2 - 2.4 \). We have additionally investigated the \( N_t = 8 \) lattice with \( am = 0.05 \), which corresponds to the rather heavy pion mass of \( m_\pi/T_c \approx 4.8 \).

**A. Reweighting from \( \mu = 0 \)**

First we have tested the theory at a fixed \( \beta = 5.4 \) at \( N_t = 4 \) as a function of \( \mu \), which is well above the deconfinement transition which at \( \mu = 0 \) and \( m = 0.05 \) is at \( \beta_c \approx 5.04 \).

In Fig. 1 we show the comparison of the gauge observables: plaquette averages and Polyakov loops. We generated \( O(10^3) \) independent configurations in the \( \mu = 0 \) ensemble with the usual hybrid Monte carlo algorithm (using every 50th configuration of the Markov chain), and we calculated the reweighting as detailed in Sec. III.
notes that the reweighting performs well for small chemical potentials $\mu/T < 1 - 1.5$, where there is a nice agreement between reweighting and CLE. The errors of the reweighting approach start to grow large as one increases $\mu$ above $1.5T$, where the average of the reweighting is dominated by a few configurations. This is the manifestation of the overlap problem: the ensemble we have sampled has typical configurations which are not the typical configurations of the ensemble we wish to study.

Next we turn to the fermionic observables: chiral condensate, fermionic density in Fig. 2. One notes that the reweighting of these quantities is possible to much higher values of $\mu/T$. This is the consequence of their explicit dependence on $\mu$, which dominates their change as the chemical potential is changed. This is in contrast to the gauge observables in Fig. 1, where the change is given entirely by the change in the measure of the path integral. The downward turn of the Polyakov loop and its inverse around $\mu/T = 3$ is the result of the phenomenon of saturation: at this chemical potential half of all of the available fermionic states on the lattice are filled, as visible in Fig. 2. This lattice artifact can also be observed with static quarks [30] and even in the strong coupling expansion [48].

Finally, the average phase factor in Fig. 2 is a good indicator of the severeness of the sign problem in the theory. One sees that the average phase in the region $\mu/T > 1.5 - 2$ indeed gets very small. Note that to see agreement between CLE and reweighting one has to be careful to choose the observable to be the analytic continuation of an observable on the SU(3) manifold. In this case we define the phase factor from the analytic continuation of the determinants, as written in (13).

In Fig. 3 we show the histogram of the absolute value of the weights of the configurations normalized by the biggest weight in the ensemble. This illustrates the overlap problem: the “further” one tries to reweight from the original ensemble, the less and less will be the contribution of an average configuration to the average, which becomes dominated by very few configurations. Thus the
fluctuations of the result become larger, and even the error bars are not reliable as the distribution of the observables becomes non-Gaussian. As we show below, this situation improves if one chooses an ensemble “closer” to the target ensemble: in this case taking the phase-quenched ensemble $\left[ \left| \det M(\mu) \right| \right]$ instead of the zero $\mu$ ensemble.

In Fig. 4 we use a theory with $N_F = 2$ flavors of fermions, by taking the square root of the staggered fermion determinant. We perform reweighting from the $\mu = 0$ ensemble using $\approx 1700$ configurations. To maintain analyticity, in the reweighting procedure one must make sure that no cut of the complex square root function is crossed while the chemical potential is changed. In the complex Langevin simulations the rooting is implemented simply by multiplying the fermion drift terms with an appropriate factor. We observe good agreement for small values $\mu/T$, similarly to the case of the $N_F = 4$ theory, indicating that the effect of rooting is the same in these different approaches.

B. Reweighting from the phase-quenched ensemble

We have investigated the efficiency of reweighting from the “phase-quenched” ensemble. In Fig. 5 we show the comparison of the plaquette averages as well as the Polyakov loop averages. We have used about 4000–5000 independent configurations at $N_t = 4$ for each $\mu$ value. One
notes that the agreement is much better when compared to the reweighting from the $\mu = 0$ ensemble and also for higher $\mu/T$ values (compare with Fig. 1). Note that this comparison is in the deconfined phase; therefore, no phase transition corresponding to the pion condensation is expected in the phase-quenched ensemble, making reweighting easier. For the $\beta = 5.4$ value used for these plots, the complex Langevin simulation breaks down in the saturation region $\mu/T > 5$ (not shown in the plots), also signaled by a large "skirt" of the distributions (meaning a slow, typically power law decay) and the disagreement of the reweighting and CLE simulations, most detectable in the plaquette averages.

C. Comparisons as a function of $\beta$

Next we have investigated the appearance of a discrepancy of the CLE and reweighting results at smaller $\beta$ values arising from a skirt of the complexified distributions [30,31]. In Fig. 6 we compare reweighting and CLE as a function of the $\beta$ parameter at fixed $\mu/T = 1$ on an $8^3 \times 4$ lattice. One observes that the reweighting is nicely reproduced by the complex Langevin simulations as long as $\beta > 5.10^{-5}$. Below this limit the distributions develop a long skirt and CL simulations become instable, also signaled by large unitarity norm and the conjugate gradient algorithm (needed for the calculation of the drift terms in the CLE) failing to converge. Similar behavior is detected on the fermionic observables in Fig. 7. This behavior has been observed also in HDQCD simulations [30,40], where a limit value $\beta_{\text{min}} = 5.6 - 5.7$ was seen independent of the value of $N_f \geq 6$, and $\beta_{\text{min}}$ was slightly smaller for $N_f = 4$. This minimal $\beta$ parameter corresponds to a maximal lattice spacing $a_{\text{max}} \approx 0.2$ fm in HDQCD. Apparently the limiting $\beta$ value is different in full QCD, but it turns out that the corresponding lattice spacing is roughly equal for $N_F = 4$ with $am = 0.05$: $a_{\text{max}} \approx 0.2 - 0.25$ fm. This breakdown is also visible on histograms

![FIG. 6 (color online). Comparison of the plaquette averages and Polyakov loops as a function of the $\beta$ parameter at a fixed $\mu/T = 1$ calculated with CLE and reweighting from the $\mu = 0$ ensemble.](image1)

![FIG. 7 (color online). Comparison of the chiral condensate and the fermionic density as a function of the $\beta$ parameter at a fixed $\mu/T = 1$ calculated with CLE and reweighting from the $\mu = 0$ ensemble.](image2)

![FIG. 8 (color online). The histograms of the spatial plaquette variable in the CLE simulation, measured at various $\beta$ values corresponding to Figs. 6 and 7.](image3)
of various observables. In Fig. 8 we show the histograms of the spatial plaquettes at various $\beta$ values. One notices that the skirt of the distribution is indeed large at $\beta = 5.1$, where the CLE breaks down. Although a small skirt is also present at $\beta = 5.2$, it is not visited frequently enough to change the averages noticeably.

A similar behavior is observed on the finer $12^3 \times 6$ lattice, as depicted in Fig. 9. We used 200–300 configurations for the reweighting procedure on $N_t = 6$ lattices at every $\beta$ value. We observe a limiting $\beta_{\text{min}} \approx 5.15$ corresponding to $a_{\text{max}} \approx 0.15$ fm which at $N_t = 6$ allows simulations right down to the transition temperature, but not below.

Finally we investigated $N_t = 8$ lattices. In Fig. 10 we show the behavior of the gauge observables, in Fig. 11 the fermionic density. We used 200–300 independent configurations at each $\beta$ value to perform the reweighting. At small betas the complex Langevin simulations become instable also on these lattices, which can be observed in Figs. 10

FIG. 9 (color online). Comparison of the plaquette averages and Polyakov loops as a function of the $\beta$ parameter at a fixed $\mu/T = 0.72$ calculated with CLE and reweighting from the $\mu = 0$ ensemble on $12^3 \times 6$ lattices.

FIG. 10 (color online). Comparison of the plaquette averages and Polyakov loops as a function of the $\beta$ parameter at a fixed $\mu/T = 0.96$, using $m = 0.01$ and $m = 0.05$ as indicated, calculated with CLE and reweighting from the $\mu = 0$ ensemble on $16^3 \times 8$ lattices.

FIG. 11 (color online). Comparison of the fermionic density as a function of the $\beta$ parameter at a fixed $\mu/T = 0.96$, using $m = 0.01$ and $m = 0.05$ as indicated, calculated with CLE and reweighting from the $\mu = 0$ ensemble on $16^3 \times 8$ lattices.
and 11 by the absence of results. One observes that the CLE breaks down above the lattice spacing $a \approx 0.15$ fm.

V. CONCLUSIONS

In this paper we have compared complex Langevin simulations of finite density QCD with reweighting from the positive ensembles of the phase-quenched theory and $\mu = 0$.

Both methods have a limited region of parameter space where they are applicable. The complex Langevin method fails for too small $\beta$ parameters, as noted earlier, but this still allows the exploration of the whole phase diagram in HDQCD. Reweighting from zero $\mu$ breaks down because of the overlap and sign problems around $\beta_\ell \approx 1 - 1.5$. In contrast, the reweighting from the phase-quenched ensemble in the deconfined phase performs better also for large $\mu$, suggesting that the sign problem is not that severe.

We observe good agreement of these two methods in the region where they are both applicable. The failure of both methods can be assessed independently of the comparison: the complex Langevin simulations develop skirted distributions as the gauge cooling loses its effectiveness, and the errors of the reweighting start to grow large signaling sign and overlap problems.

An important question for the applicability of the complex Langevin method to explore the phase diagram of QCD is the behavior of $\beta_{\text{mat}}$, the lattice parameter below which gauge cooling is not effective. In this study we have determined that using $N_t = 4$, $N_s = 6$ and $N_s = 8$ lattices (with pion mass $m_\pi/T_c \approx 2.2 - 2.4$) this breakdown prevents the exploration of the deconfinement transition and the location of a possible critical point.

ACKNOWLEDGMENTS

This project was funded by the DFG Grant No. SFB/TR55. S. D. K. is funded by the “Lendület” program of the Hungarian Academy of Sciences (LP2012-44/2012). The authors gratefully acknowledge the Gauss Centre for Supercomputing e.V. [50] for funding this project by providing computing time on the GCS Supercomputer SuperMUC at Leibniz Supercomputing Centre (LRZ) [51]. Some parts of the numerical calculation were done on the GPU cluster at Eotvos and Wuppertal Universities.