PROCEEDINGS OF SCIENCE

Complex Langevin dynamics: criteria for correctness

Gert Aarts
Department of Physics, Swansea University, Swansea, UK
E-mail: g.aarts@swan.ac.uk

Frank A. James*
Department of Physics, Swansea University, Swansea, UK
E-mail: pyfj@swan.ac.uk

Erhard Seiler
Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München, Germany
E-mail: ehs@mppmu.mpg.de

Ion-Olimpiu Stamatescu
Institut für Theoretische Physik, Universität Heidelberg and FEST, Heidelberg, Germany
E-mail: I.O.Stamatescu@thphys.uni-heidelberg.de

The complex Langevin method is a leading candidate for solving the sign problem occurring in various physical situations, notably QCD at finite chemical potential. Its most vexing problem is 'convergence to the wrong limit', where the simulation gives a well defined, but incorrect, result. Here, we first outline a formal justification of the method and identify points at which it might fail. From these we derive a condition that must be satisfied in order for correct results to be obtained. We then apply these ideas to the three-dimensional SU(3) spin model at finite chemical potential and show strong indications that complex Langevin dynamics yields correct results in this theory.

XXIX International Symposium on Lattice Field Theory
July 10 - 16 2011
Squaw Valley, Lake Tahoe, California

*Speaker.

PoS(Lattice 2011)197

1. Introduction

Many systems of physical relevance have a complex action, QCD with a finite chemical potential is a particularly interesting example. Such theories are difficult to study numerically because the weight is complex and therefore standard Markov chain Monte Carlo techniques based on a probability interpretation cannot be applied; this is often called the sign problem [1]. The complex Langevin equation (CLE) enjoyed a certain limited popularity after being proposed in the early 1980s [2, 3] because it can, in principle, avoid the sign problem since it does not rely on importance sampling. In some cases, complex Langevin simulations have been found to be numerically unstable due to runaway solutions but this can be cured by use of an adaptive stepsize [4]. A more serious and vexing issue is convergence to the wrong limit, where the simulation gives well defined results, but when compared with known values are found to be incorrect. It is this problem that we shall focus on here.

In Section 2 we briefly cover a formal argument for the correctness of the CLE and identify points at which it might fail. By studying the long time evolution of observables with respect to real and complex measures, we derive a criterion which must be satisfied in order for correct results to be obtained [5, 6]. In Section 3 these ideas are applied to the SU(3) spin model at finite chemical potential [7, 8, 9]. A comparison with imaginary chemical potential shows that the CLE works correctly in the region $\mu^2 \sim 0$. An analysis of the criterion with a larger chemical potential supports the claim that in this case complex Langevin dynamics can be trusted [10].

2. Criteria for correctness

The central object of interest is the expectation value of a particular observable, given by

$$\langle O \rangle = \frac{\int O(x) e^{-S(x)} dx}{\int e^{-S(x)} dx}, \tag{2.1}$$

where for notational simplicity we use a single real degree of freedom, x. The action $S(x)$ is complex, preventing a probability interpretation of the measure and ruling out methods based on importance sampling. The complex Langevin equation is

$$\frac{dx}{dt} = K_x + \sqrt{N_{\mathrm{R}}}\eta_{\mathrm{R}}, \qquad \frac{dy}{dt} = K_y + \sqrt{N_{\mathrm{I}}}\eta_{\mathrm{I}}, \tag{2.2}$$

where the real variable is analytically continued as $x \to z = x + iy$. The drift terms are given by

$$K_x = \mathrm{Re}\ \frac{dS(x)}{dx}\bigg|_{x\to x+iy}, \qquad K_y = \mathrm{Im}\ \frac{dS(x)}{dx}\bigg|_{x\to x+iy}, \tag{2.3}$$

and the two noise terms η_{R}, η_{I} are independent Gaussian random numbers with variance 2 and normalisation $N_{\mathrm{I}} \geq 0$ and $N_{\mathrm{R}} - N_{\mathrm{I}} = 1$. A numerical simulation can then be implemented by integrating these equations to large times $t \to \infty$.

The resulting dynamics is described by a dual Fokker-Planck equation for the evolution of the probability density $P(x,y;t)$,

$$\frac{\partial}{\partial t} P(x,y;t) = L^T P(x,y;t), \tag{2.4}$$

with the operator

$$L^T = \nabla_x[N_{\rm R}\nabla_x - K_x] + \nabla_y[N_{\rm I}\nabla_y - K_y]. \tag{2.5}$$

To understand the time evolution of the real density $P(x,y;t)$ one must also examine the evolution of the complex density $\rho(x;t)$, determined by

$$\frac{\partial}{\partial t}\rho(x;t) = L_0^T\rho(x;t). \tag{2.6}$$

Here, the complex Fokker-Planck operator L_0^T is

$$L_0^T = \nabla_x[\nabla_x + \nabla_x S(x)]. \tag{2.7}$$

This equation has $\rho(x;\infty) \propto \exp[-S(x)]$ as a stationary solution, which is expected to be unique. Numerical studies (where feasible) of Eq. (2.6) confirm this to be true; in fact, convergence to this distribution seems exponentially fast.

Expectation values with respect to the two densities can now be defined as

$$\langle O\rangle_{P(t)} = \frac{\int O(x+iy)P(x,y;t)dxdy}{\int P(x,y;t)dxdy}, \qquad \langle O\rangle_{\rho(t)} = \frac{\int O(x)\rho(x;t)dx}{\int \rho(x;t)dx}. \tag{2.8}$$

The result that one would like to show is

$$\langle O\rangle_{P(t)} = \langle O\rangle_{\rho(t)}, \tag{2.9}$$

if the initial conditions $\langle O\rangle_{P(0)} = \langle O\rangle_{\rho(0)}$ match, which is assured provided

$$P(x,y;0) = \rho(x;0)\delta(y). \tag{2.10}$$

One expects the dependence on the initial conditions to vanish in the limit $t \to \infty$ by ergodicity.

To establish a connection between the expectation values with respect to P and ρ, one moves the time evolution from the densities to the observables. Since we are only interested in functions of $z = x + iy$ (holomorphic functions), we may act with the Langevin operator

$$\tilde{L} = [\nabla_z - (\nabla_z S(z))]\nabla_z,$$

whose action on holomorphic functions agrees with that of L.

We now use $\tilde{L}$ to evolve observables according to the equation

$$\frac{\partial}{\partial t}O(z;t) = \tilde{L}O(z;t), \tag{2.11}$$

which is formally solved by

$$O(z;t) = \exp[t\tilde{L}]O(z).$$

Due to the fact that L and $\tilde{L}$ agree on holomorphic functions, the tilde may be dropped.

To examine the evolution we define the function

$$F(t,\tau) = \int P(x,y;t-\tau)O(x+iy;\tau)dxdy, \tag{2.12}$$

and observe that $F(t,\tau)$ interpolates between the two expectation values:

$$F(t,0) = \langle O \rangle_{P(t)}, \qquad F(t,t) = \langle O \rangle_{\rho(t)}. \tag{2.13}$$

The first can be seen easily, while the second makes use of the initial conditions and

$$\begin{aligned} F(t,t) &= \int P(x,y;0) \left(e^{tL} O\right)(x+iy;0) dxdy \\ &= \int \rho(x;0) \left(e^{tL_0} O\right)(x;0) dx = \int O(x;0) \left(e^{tL_0^T} \rho\right)(x;0) dx = \langle O \rangle_{\rho(t)}, \end{aligned} \tag{2.14}$$

where it is only necessary to assume that integration by parts in x does not produce any boundary terms.

The desired result (2.9) follows if $F(t,\tau)$ is independent of τ. To check this, we need the τ derivative to vanish,

$$\frac{\partial}{\partial \tau} F(t,\tau) = -\int \left(L^T P(x,y;t-\tau)\right) O(x+iy;\tau) dxdy + \int P(x,y;t-\tau) LO(x+iy;\tau) dxdy. \tag{2.15}$$

Integration by parts, *if applicable without boundary terms at infinity*, then shows that the two terms cancel and therefore $F(t,\tau)$ is independent of τ, irrespective of N_{I}.

This is therefore a point at which the formal argument might fail: if the decay of the product

$$P(x,y;t-\tau) O(x+iy;\tau)$$

and its derivatives is insufficient for integration by parts without boundary terms.

In Ref. [6] a study of the U(1) one-link model found that the τ-derivative is largest at $\tau = 0$. This motivates the superficially weaker condition

$$\lim_{t\to\infty} \frac{\partial}{\partial t} F(t,\tau)\bigg|_{\tau=0} = 0. \tag{2.16}$$

This modification to the condition is still sufficient for correctness if it holds for a sufficiently large set of suitably chosen observables. Taking the limit $t \to \infty$ in Eq. (2.15) causes the first contribution to vanish because of the equilibrium condition, $L^T P(x,y;\infty) = 0$. Therefore, the criterion for correctness reduces to

$$E_O \equiv \int P(x,y;\infty) LO(x+iy,0) dxdy = \langle LO \rangle = 0. \tag{2.17}$$

This is fairly simple to check for a given observable, but it is in fact a strong statement since it must hold *for all observables*. Therefore, Eq. (2.17) really represents an infinite tower of identities which must all be satisfied. In practice it can be checked for a small number of observables, still yielding a necessary criterion [6].

3. SU(3) spin model

Motivated by recent work [9] (for related models see e.g. Ref. [11]), we re-examine the three-dimensional SU(3) spin model at finite chemical potential, for which promising results with complex Langevin dynamics have been obtained in earlier studies [7, 8]. The action is formed from three contributions

$$S = S_B + S_F + S_H, \tag{3.1}$$

which read

$$S_B = -\beta \sum_{x,\nu} \left(\mathrm{Tr} U_x \mathrm{Tr} U^{-1}_{x+\hat{\nu}} + \mathrm{Tr} U_x^{-1} \mathrm{Tr} U_{x+\hat{\nu}} \right), \tag{3.2}$$

$$S_F = -h \sum_x \left(e^{\mu} \mathrm{Tr} U_x + e^{-\mu} \mathrm{Tr} U_x^{-1} \right), \tag{3.3}$$

$$S_H = -\sum_x \log \left[\sin^2 \left(\frac{\phi_{1x} - \phi_{2x}}{2} \right) \sin^2 \left(\frac{2\phi_{1x} + \phi_{2x}}{2} \right) \sin^2 \left(\frac{\phi_{1x} + 2\phi_{2x}}{2} \right) \right]. \tag{3.4}$$

The final component, S_H, originates from the Haar measure introduced by diagonalising U in terms of angles,

$$U = \mathrm{diag}\,(e^{i\phi_1}, e^{i\phi_2}, e^{-i(\phi_1+\phi_2)}).$$

The action is complex due to the "heavy fermion" contribution, $S_{\mathrm{F}}^*(\mu) = S_{\mathrm{F}}(-\mu^*)$. For small h the theory has a confined and deconfined phase separated by a first order transition line at small μ, turning into a crossover at larger μ (see Figure 1, left) [7].

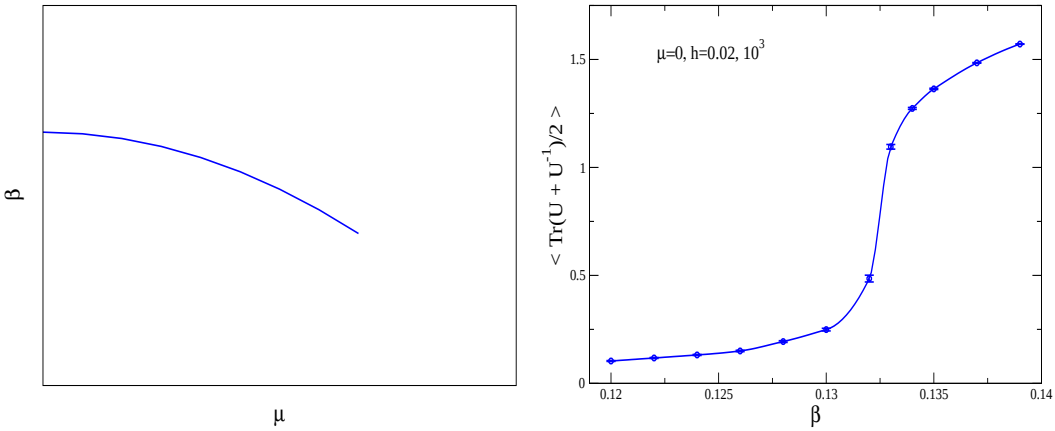


Figure 1: Left: sketch of the phase diagram at small h. Right: $\langle \mathrm{Tr}\,(U+U^{-1})/2 \rangle$ as a function of β at $\mu = 0$, $h = 0.02$ on a 10^3 lattice.

Figure 1 (right) shows the transition as a function of β at $\mu = 0$, using the observable $\langle \mathrm{Tr}\,(U + U^{-1})/2 \rangle$. Note that at $\mu = 0$, $\langle \mathrm{Tr} U \rangle = \langle \mathrm{Tr} U^{-1} \rangle$, but at $\mu^2 > 0$ they differ. Since observables that are even in μ should be continuous across $\mu^2 = 0$, we also perform simulations with an imaginary chemical potential, for which the action is real and simulations using standard techniques can be used (we use real Langevin dynamics). A comparison with complex Langevin data should therefore show even observables to be continuous across $\mu^2 = 0$. In Figure 2 the observable $\langle \mathrm{Tr}\,(U+U^{-1})/2 \rangle$ is plotted. The data from complex Langevin dynamics with $\mu^2 \geq 0$ is consistent with those from real Langevin dynamics with $\mu^2 \leq 0$ in both phases, including the critical region, suggesting that complex Langevin dynamics is working correctly here. This is in contrast to the case of the XY model recently studied using complex Langevin dynamics, where correct results were obtained in only part of the phase diagram [12].

At larger values of μ^2, one can no longer rely on analytical continuation to justify the results. Instead, we assess them using the criteria developed above and compute $\langle LO \rangle$ for $O = \mathrm{Tr} U$. The outcome is shown in Figure 3 for real chemical potential (left) and also in the phase-quenched

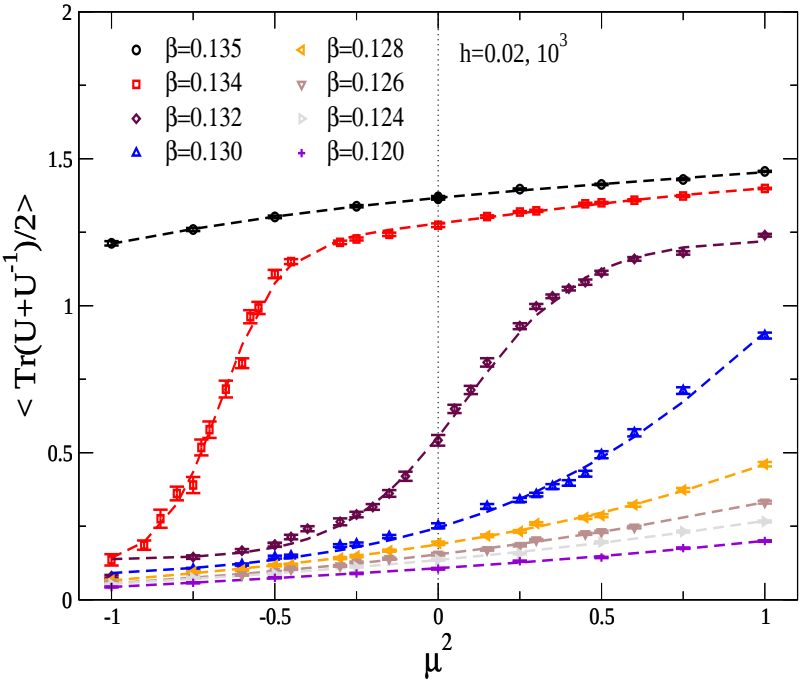
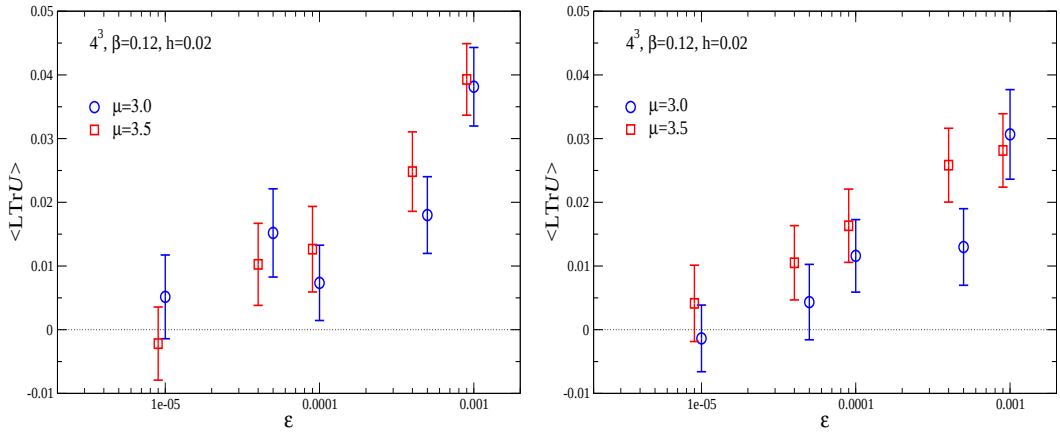


Figure 2: $\langle \mathrm{Tr}\,(U+U^{-1})/2\rangle$ as a function of μ^2 for various β values. For $\mu^2<0$ (imaginary μ) the action is real and real Langevin dynamics is used; when $\mu^2>0$ complex Langevin is necessary.



(a) Complex Langevin dynamics (b) Phase-quenched using real Langevin dynamics

Figure 3: Langevin stepsize dependence of the correctness criterion for $\mathrm{Tr}U$ with complex Langevin dynamics (left) and in the phase-quenched approximation (right) using real Langevin dynamics. The points for $\mu=3.5$ are shifted horizontally for clarity ($\beta=0.12$, $h=0.02$ and lattice volume 4^3).

theory (right), using $\mu=3$ and 3.5. In the phase-quenched theory, the action is real and we use real Langevin dynamics. The figures indicate that at nonzero stepsize the criterion is not satisfied, both for real and complex Langevin dynamics, but that in the limit $\varepsilon\to 0$ it is satisfied.

We find therefore that in this model the test is passed successfully. Moreover, finite stepsize corrections can be quantified by the deviation of $\langle LO\rangle$ from zero. We are currently extending this analysis and have implemented a higher-order algorithm to eliminate the effects of finite step-

size [10].

4. Conclusions

Complex Langevin dynamics can in principle be applied where the sign problem prevents the use of importance sampling. An analysis of the long-time evolution of real and complex measures using the Fokker-Planck equation shows that the correct stationary solution exists. However, if the decay of the distribution $P(x,y;t)$ is insufficient to allow integration by parts without boundary terms, convergence to the wrong limit can occur. In practice, this can be diagnosed by testing the condition that $\langle LO \rangle = 0$ for a suitably large set of observables in the limit of vanishing stepsize. In the case that the criterion is not satisfied, complex Langevin dynamics fails.

An analysis of the SU(3) spin model shows that the criterion for correctness is satisfied, justifying the claim that complex Langevin dynamics works correctly with this model. This is corroborated by analytic continuation from imaginary chemical potential in the region of small chemical potential $\left|\mu^2\right| \lesssim 1$, which shows that data for observables even in μ are continuous between results from a real action when $\mu^2 \leq 0$ and a complex action with $\mu^2 > 0$. A further and more detailed study of the criterion and stepsize dependence is in progress [10].

Acknowledgments

This work is supported by STFC.

References

[1] P. de Forcrand, PoS **LAT2009** (2009) 010 [arXiv:1005.0539 [hep-lat]].

[2] G. Parisi, Phys. Lett. **B131** (1983) 393-395.

[3] J. R. Klauder, *Stochastic quantization*, H. Mitter, C. B. Lang eds., in *Recent Developments in High-Energy Physics*, Springer-Verlag, Wien Austria (1983), pg. 351.

[4] G. Aarts, F. A. James, E. Seiler, I. -O. Stamatescu, Phys. Lett. **B687** (2010) 154 [arXiv:0912.0617 [hep-lat]].

[5] G. Aarts, E. Seiler, I. -O. Stamatescu, Phys. Rev. **D81** (2010) 054508 [arXiv:0912.3360 [hep-lat]].

[6] G. Aarts, F. A. James, E. Seiler, I. -O. Stamatescu, Eur. Phys. J. **C71** (2011) 1756 [arXiv:1101.3270 [hep-lat]].

[7] F. Karsch, H. W. Wyld, Phys. Rev. Lett. **55** (1985) 2242.

[8] N. Bilic, H. Gausterer, S. Sanielevici, Phys. Rev. **D37** (1988) 3684.

[9] C. Gattringer, Nucl. Phys. **B850** (2011) 242 [arXiv:1104.2503 [hep-lat]].

[10] G. Aarts and F. A. James, in preparation

[11] S. Lottini, O. Philipsen, J. Langelage, arXiv:1105.5284 [hep-lat].

[12] G. Aarts, F. A. James, JHEP **1008** (2010) 020 [arXiv:1005.3468 [hep-lat]].