

Complex Langevin method: When can it be trusted?

Gert Aarts*

Department of Physics, Swansea University, Swansea, United Kingdom

Erhard Seiler†

Max-Planck-Institut für Physik (Werner-Heisenberg-Institut), München, Germany

Ion-Olimpiu Stamatescu‡

Institut für Theoretische Physik, Universität Heidelberg and FEST, Heidelberg, Germany

(Received 18 December 2009; published 22 March 2010)

We analyze to what extent the complex Langevin method, which is in principle capable of solving the so-called sign problems, can be considered as reliable. We give a formal derivation of the correctness and then point out various mathematical loopholes. The detailed study of some simple examples leads to practical suggestions about the application of the method.

DOI: [10.1103/PhysRevD.81.054508](https://doi.org/10.1103/PhysRevD.81.054508)

PACS numbers: 11.15.Ha, 11.15.Bt

I. INTRODUCTION

The complex Langevin method solves in principle the sign problems arising in simulations of various systems, in particular, in QCD with finite chemical potential. After being proposed in the early 1980s by Klauder [1] and Parisi [2], it enjoyed a certain limited popularity [3,4], but very quickly certain problems were found. The first one was instability of the simulations with absence of convergence (runaways), the second one convergence to a wrong limit [5–8]. Nevertheless in recent years the method has been revived with sometimes impressive success [9–16]. In particular the use of adaptive step size has eliminated the problem of runaways [17].

But nagging problems remained due to the lack of clear criteria to decide when an apparently convergent simulation actually represented the truth. This was linked to the lack of a clear mathematical basis for the method, that would at the same time also provide criteria for its applicability. The purpose of the present paper is to clarify the situation at least to some extent. While we are still not able to close certain mathematical gaps and reach a complete analytic solution to the problems that have plagued the method, we give some strong numerical evidence that the method is correct in some cases and also suggest a plausible explanation for the failure in other cases; this leads to some pragmatic conclusions suggesting how to proceed in practice in a way that promises credible results.

The paper is organized as follows. In Sec. II we give a formal justification of the method, highlighting the assumptions underlying the derivation. In Sec. III three main questions raised by the formal arguments are listed. We then focus on one particular issue, boundary effects, in

Sec. IV, and present detailed case studies in Sec. V. Tentative conclusions are given in Sec. VI.

II. FORMAL DERIVATIONS

For simplicity we concentrate here on models in which the fields take values in flat manifolds $\mathcal{M} = \mathbb{R}^n$ or $\mathcal{M} = T^n$, where T^n is the n dimensional torus $(S^1)^n$ with coordinates (x_1, \dots, x_n) . The complications that arise when the fields live in nontrivial manifolds, as is of course the case in QCD, have been successfully dealt with in the literature (see for instance, Ref. [18] for real, Refs. [10,12,19,20] for complex Langevin dynamics). But these complications are not really relevant for our discussion.

As is well known, the idea is to simulate a complex measure $\exp(-S)dx$, with S a holomorphic function on a real manifold \mathcal{M} , by setting up a stochastic process on the complexification \mathcal{M}_c of \mathcal{M} , such that the expectation values of *entire holomorphic observables* \mathcal{O} in this stochastic process converge to the ones with respect to the complex measure $\exp(-S)dx$.

The complex Langevin equation (CLE) on \mathcal{M}_c is

$$dz = -\nabla S dt + dw, \quad (1)$$

where dw denotes the increment of the Wiener process and the equation is to be interpreted as a real stochastic process, namely,

$$dx = K_x dt + dw, \quad dy = K_y dt, \quad (2)$$

with

$$K_x = -\text{Re}\nabla_x S(x + iy), \quad K_y = -\text{Im}\nabla_x S(x + iy). \quad (3)$$

A slight generalization of Eq. (2) that has been considered and will play a role in this investigation is

$$dx = K_x dt + \sqrt{N_R} dw_R, \quad dy = K_y dt + \sqrt{N_I} dw_I, \quad (4)$$

*g.aarts@swan.ac.uk

†ehs@mppmu.mpg.de

‡I.O.Stamatescu@thphys.uni-heidelberg.de

where dW_R and dW_I are independent Wiener processes, $N_I \geq 0$ and $N_R = N_I + 1$. This is usually referred to as complex noise. The introduction of a nonzero N_I makes it possible to solve the Fokker-Planck equation (see below) numerically and also allows a random walk discretization of the complex Langevin process:

$$\delta x(t) = \pm \omega_x, \quad P_{x,\pm} = \frac{1}{2} \left(1 \pm \tanh \left[\frac{\omega_x}{2N_R} K_x \right] \right), \quad (5)$$

$$\delta y(t) = \pm \omega_y, \quad P_{y,\pm} = \frac{1}{2} \left(1 \pm \tanh \left[\frac{\omega_y}{2N_I} K_y \right] \right), \quad (6)$$

$$\omega_x = \sqrt{2N_R \delta t}, \quad \omega_y = \sqrt{2N_I \delta t}, \quad (7)$$

where P_{\pm} are the transition probabilities and we have defined the steps ω such as to have the same δt in both subprocesses, to ensure correct evolution.

By Itô calculus, if f is a twice differentiable function on \mathcal{M}_c and

$$z(t) = x(t) + iy(t) \quad (8)$$

is a solution of the complex Langevin equation (4), we have

$$\frac{d}{dt} \langle f(x(t), y(t)) \rangle = \langle Lf(x(t), y(t)) \rangle, \quad (9)$$

where L is the Langevin operator

$$L = [N_R \nabla_x + K_x] \nabla_x + [N_I \nabla_y + K_y] \nabla_y, \quad (10)$$

and $\langle f \rangle$ denotes the noise average of f corresponding to the stochastic process described by Eq. (4). In the standard way Eq. (4) leads to its dual Fokker-Planck equation (FPE) for the evolution of the probability density $P(x, y; t)$,

$$\frac{\partial}{\partial t} P(x, y; t) = L^T P(x, y; t), \quad (11)$$

with

$$L^T = \nabla_x [N_R \nabla_x - K_x] + \nabla_y [N_I \nabla_y - K_y]. \quad (12)$$

L^T is the formal adjoint (transpose) of L with respect to the bilinear (not Hermitian) pairing

$$\langle f, P \rangle = \int f(x, y) P(x, y) dx dy, \quad (13)$$

i.e.,

$$\langle Lf, P \rangle = \langle f, L^T P \rangle. \quad (14)$$

Note that the FPE has the form of a continuity equation

$$\frac{\partial}{\partial t} P(x, y; t) = \nabla_x J_x + \nabla_y J_y, \quad (15)$$

where (J_x, J_y) is the probability current in the $2n$ dimensional space \mathcal{M}_c , given by

$$J_x = (N_R \nabla_x - K_x)P, \quad J_y = (N_I \nabla_y - K_y)P. \quad (16)$$

We will also consider the evolution of a complex density $\rho(x)$ on \mathcal{M} under the following complex FPE:

$$\frac{\partial}{\partial t} \rho(x; t) = L_0^T \rho(x; t), \quad (17)$$

where now the complex Fokker-Planck operator L_0^T is

$$L_0^T = \nabla_x [\nabla_x + (\nabla_x S(x))]. \quad (18)$$

A slight generalization will be useful: For any $y_0 \in \mathcal{M}$ we consider a complex Fokker-Planck operator $L_{y_0}^T$ given by

$$L_{y_0}^T = \nabla_x [\nabla_x + (\nabla_x S(x + iy_0))]. \quad (19)$$

$L_{y_0}^T$ is the formal adjoint of

$$L_{y_0} = [\nabla_x - (\nabla_x S(x + iy_0))] \nabla_x. \quad (20)$$

The operators $L_{y_0}^T$ act on suitable complex valued distributions (measures) on \mathcal{M} , parametrized by the real variables (x_1, \dots, x_n) . But they do not allow a probabilistic interpretation, because they do not preserve positivity.

For any y_0 Eq. (17) with L_0^T replaced by $L_{y_0}^T$ has the complex density

$$\rho_{y_0}(x; \infty) \propto \exp[-S(x + iy_0)] \quad (21)$$

as its (hopefully unique) stationary solution.

We next consider expectation values. Let \mathcal{O} be an entire holomorphic observable with at most exponential growth; then we set

$$\langle \mathcal{O} \rangle_{P(t)} \equiv \frac{\int \mathcal{O}(x + iy) P(x, y; t) dx dy}{\int P(x, y; t) dx dy} \quad (22)$$

and

$$\langle \mathcal{O} \rangle_{\rho(t)} \equiv \frac{\int \mathcal{O}(x) \rho(x; t) dx}{\int \rho(x; t) dx}. \quad (23)$$

We would like to show that

$$\langle \mathcal{O} \rangle_{P(t)} = \langle \mathcal{O} \rangle_{\rho(t)}, \quad (24)$$

provided the initial conditions agree,

$$\langle \mathcal{O} \rangle_{P(0)} = \langle \mathcal{O} \rangle_{\rho(0)}, \quad (25)$$

which is true if we choose

$$P(x, y; 0) = \rho(x; 0) \delta(y - y_0) \quad (26)$$

(for any y_0). In the limit $t \rightarrow \infty$ the dependence on the initial condition should of course disappear by ergodicity.

The goal is to establish a connection between the ‘‘expectation values’’ with respect to ρ and P for the class of observables chosen (entire holomorphic with at most exponential growth). The idea is to move the time evolution from the densities to the observables and make use of the Cauchy-Riemann (CR) equations. Formally (i.e., without worrying about boundary terms and existence questions), this works as follows: first we use the fact that we want to apply the complex FP operators L_{y_0} only to functions that

have analytic continuations to all of \mathcal{M}_c . On those analytic continuations we may act with the Langevin operator

$$\tilde{L} \equiv [\nabla_z - (\nabla_z S(z))]\nabla_z, \quad (27)$$

whose action on holomorphic functions agrees with that of L , since on such functions $\nabla_y = i\nabla_x$ and $\Delta_x = -\Delta_y$ so that the difference $L - \tilde{L}$ vanishes.

The proliferation of Langevin/Fokker-Planck operators may be somewhat bewildering, but it is important to realize that L, \tilde{L}, L_{y_0} are really all different operators: while L and \tilde{L} act on functions on \mathcal{M}_c [i.e., functions of $(x_1, y_1, \dots, x_n, y_n)$], agreeing on holomorphic functions, but disagreeing on general functions, L_{y_0} acts on functions on \mathcal{M} , i.e., functions of (x_1, \dots, x_n) .

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

$$\partial_t \mathcal{O}(z; t) = \tilde{L} \mathcal{O}(z; t) \quad (t \geq 0) \quad (28)$$

with the initial condition $\mathcal{O}(z; 0) = \mathcal{O}(z)$, which is formally solved by

$$\mathcal{O}(z; t) = \exp[t\tilde{L}]\mathcal{O}(z). \quad (29)$$

In Eqs. (28) and (29), because of the CR equations, the tilde may be dropped, and we will do so now. So we will have

$$\partial_t \mathcal{O}(z; t) = L \mathcal{O}(z; t) \quad (t \geq 0), \quad (30)$$

with its formal solution

$$\mathcal{O}(z; t) = \exp[tL]\mathcal{O}(z). \quad (31)$$

In fact Eq. (28) is also equivalent to the family of equations

$$\partial_t \mathcal{O}(x + iy_0; t) = L_{y_0} \mathcal{O}(x + iy_0; t) \quad (t \geq 0) \quad \forall y_0. \quad (32)$$

The first thing to notice is that $\mathcal{O}(z; t)$ will be holomorphic if $\mathcal{O}(z; 0)$ is. This can be seen as follows: Let $\bar{\partial}_j$ ($i = j, \dots, n$) be the Cauchy-Riemann operators defined by

$$\bar{\partial}_j \equiv \partial_{x_j} + i\partial_{y_j}. \quad (33)$$

Applying this to both sides of Eq. (28) and using the fact that by the holomorphy of S , $\bar{\partial}_j$ commutes with L , we find

$$\partial_t \bar{\partial}_j \mathcal{O}(z; t) = L \bar{\partial}_j \mathcal{O}(z; t) \quad (t \geq 0); \quad (34)$$

this is just Eq. (28) again with $\mathcal{O}(z; t)$ replaced by $\bar{\partial}_j \mathcal{O}(z; t)$. Under the assumption that L generates a semigroup acting on $\bar{\partial}_j \mathcal{O}(z; t)$, Eq. (34) has a unique solution; since the initial condition is $\bar{\partial}_j \mathcal{O}(z; 0) = 0$ we conclude that $\bar{\partial}_j \mathcal{O}(z; t) = 0$ for all $t \geq 0$, i.e., $\mathcal{O}(z; t)$ satisfies the CR equations for all $t \geq 0$ and all $j = 1, \dots, n$. So $\mathcal{O}(z; t)$ is holomorphic in each component z_j separately. By Hartogs' theorem [21,22] this implies joint holomorphy.

We now consider, for $0 \leq \tau \leq t$,

$$F(t, \tau) \equiv \int P(x, y; t - \tau) \mathcal{O}(x + iy; \tau) dx dy, \quad (35)$$

and claim that it interpolates between the ρ and the P expectations:

$$F(t, 0) = \langle \mathcal{O} \rangle_{P(t)}, \quad F(t, t) = \langle \mathcal{O} \rangle_{\rho(t)}. \quad (36)$$

The first equality is obvious, while the second one can be seen as follows, using Eqs. (26) and (31)

$$\begin{aligned} F(t, t) &= \int P(x, y; 0) \exp(tL) \mathcal{O}(x + iy; 0) dx dy \\ &= \int \rho(x; 0) (\exp(tL_{y_0}) \mathcal{O}(x + iy_0; 0)) dx \\ &= \int \mathcal{O}(x + iy_0; 0) (\exp(tL_{y_0}^T) \rho(x; 0)) dx \\ &= \langle \mathcal{O} \rangle_{\rho(t)}, \end{aligned} \quad (37)$$

where we only had to assume that we can integrate by parts in x without worrying about boundary terms.

Our desired result follows if we can show that $F(t, \tau)$ is independent of τ . To see this, we differentiate

$$\begin{aligned} \frac{\partial}{\partial \tau} F(t, \tau) &= - \int (L^T P(x, y; t - \tau)) \mathcal{O}(x + iy; \tau) dx dy \\ &\quad + \int P(x, y; t - \tau) \tilde{L} \mathcal{O}(x + iy; \tau) dx dy. \end{aligned} \quad (38)$$

Integration by parts then shows that the two terms cancel, hence $\frac{\partial}{\partial \tau} F(t, \tau) = 0$ and thus

$$\langle \mathcal{O} \rangle_{P(t)} = \langle \mathcal{O} \rangle_{\rho(t)}. \quad (39)$$

It is important to notice that this holds for all N_j ; whereas the left-hand side seems to depend on N_I , the right-hand side is manifestly independent of it.

If we knew in addition that

$$\lim_{t \rightarrow \infty} \langle \mathcal{O} \rangle_{\rho(t)} = \langle \mathcal{O} \rangle_{\rho(\infty)}, \quad (40)$$

with $\rho(\infty)$ given by Eq. (21) with $y_0 = 0$, we could now conclude that the expectation values of the Langevin process relax to the desired values; this convergence would follow if we knew that the spectrum of $L_{y_0}^T$ lies in a half plane $\text{Re} z \leq 0$ and 0 is a nondegenerate eigenvalue. But note that we do not really need convergence of $P(x, y; t)$ for Eq. (40) to hold, since it will only be tested against analytic observables \mathcal{O} .

Nevertheless the numerical evidence in many cases points to the existence of a unique stationary probability density $P(x, y; \infty)$. The corresponding probability currents are divergenceless, but unlike the situation in the real Langevin process, they cannot vanish. A general feature of the stationary distribution that can be read off the FPE is the following: Assume that (x_0, y_0) is a local stationary point of P , then

$$(N_R \Delta_x + N_I \Delta_y) P = \nabla_x (K_x P) + \nabla_y (K_y P), \quad (41)$$

for $x = x_0, y = y_0$. So if $N_I > 0$, a local maximum of P can only occur where the divergence of the drift force is

negative and a minimum where it is positive. For $N_I = 0$ the conclusion is even stronger: where the divergence is negative (positive), there cannot be a local minimum (maximum) in x for fixed y . These properties provide some checks on numerical solutions.

III. QUESTIONS

There are three main questions raised by the formal arguments in the previous section:

- (1) Can the operators L , \tilde{L} , L_{y_0} and their transposes be exponentiated; in more mathematical language: do these operators generate semigroups on some suitable space of functions?
- (2) Are the various integrations by parts justified, which underlie the shifting of the time evolution from the measure to the observables and back, or are there boundary terms to worry about?
- (3) Are the spectra of L , L_{y_0} and their transposes entirely in the left half-plane and is 0 a nondegenerate eigenvalue?

Concerning the first question, there are treatises (see for instance Refs. [23,24]) giving rather general sufficient conditions for the existence of a semigroup generated by differential operators of the general type considered here. Unfortunately it seems that the cases we have to deal with here are not covered by those general results; the main difficulties are (1) the strong growth of the drift given by the gradients of the action in some complex directions and (2) the fact that the drift is not always “restoring.”

Question (1) for L , L^T is intimately related to the question whether the stochastic process given by the complex Langevin equation exists for arbitrary long times. This is not obvious because typically in the classical (no noise) limit there are trajectories that go to infinity in finite time. While those trajectories occur only for a subset of measure zero of initial conditions, it is not obvious what happens after adding the noise. On the one hand, the noise will typically kick the process away from the unstable trajectories; on the other hand it may also kick it near the unstable trajectory, inducing very large excursions of the process. We will later illustrate this with some examples.

But let us say that the accumulated numerical evidence points not only to the existence of the process for arbitrarily large times, but also to the existence of a unique equilibrium measure for the process; unfortunately we could neither find results in the mathematical literature that would imply this, nor could we prove ourselves that this is the case.

Concerning the exponentiation of \tilde{L} , L_{y_0} , which should be easier, we still could not establish mathematically that it is possible on a space of functions containing the most obvious observables, such as exponentials.

There is a useful criterion for the existence of a bounded semigroup generated by an operator A on a Hilbert space [23]: A generates a bounded semigroup if it is *dissipative*,

i.e., if $A + A^\dagger \leq 0$. Unfortunately even in the simplest cases of a quadratic S the corresponding Langevin and FPE operators are *not* dissipative. So they can at best generate exponentially bounded semigroups; if in addition the spectrum is in the left half-plane, convergence to the equilibrium should still take place.

For the second question it would be necessary to have good control over the falloff of the solutions of the FPE in the imaginary directions: if we insert the observables $\mathcal{O}_k(z) \equiv \exp(ikz)$ into Eq. (39), we get for the Fourier transform (Fourier coefficients in the compact case) $\hat{\rho}(k; t)$ of the complex density $\rho(x; t)$

$$\hat{\rho}(k; t) = \frac{1}{(2\pi)^n} \int P(x, y; t) e^{ikx - ky} dx dy. \quad (42)$$

This makes sense for all k *only* if $P(x, y; t)$ decays more strongly than any exponential in imaginary direction. Our case studies described in the following sections indicate that this does not seem to be the case: in our first example the decay is probably exponential, but not stronger; in our second example the decay seems to be even weaker [$O(|y|^{-r})$ with $r \approx 2$], so that exponentials cannot be used as observables. The remainder of the paper is mainly devoted to studying question (2) in some toy models.

Finally let us remark that the third question is more difficult than the first one, and again the answer is not known rigorously. But again the numerical evidence strongly suggests a positive answer, depending on the model and the parameter values, in many interesting and relevant cases (see, e.g., Refs. [12,14]).

IV. BOUNDARY EFFECTS

In this paper we are mainly concerned with question (2). Even though we have very little analytic control, careful numerical studies reveal that, as remarked, the answer is generally *no*. In our case studies we find indications that the probability density $P(x, y; t)$ indeed relaxes to an equilibrium density $P(x, y; \infty)$, but that limiting density decays at best like an exponential for $|y| \rightarrow \infty$, in other cases only powerlike. This limits the class of observables for which the integrations by parts can be performed without boundary terms.

But let us first take a closer look at the integrations by parts that occur in the formal arguments of Sec. II. The danger lies in a possibly insufficient falloff in the imaginary (y) directions, whereas in the real (x) direction we have either compactness or sufficient falloff due to the behavior of the action S .

Let us remark that for $N_I > 0$ the operators L and L^T are uniformly strictly elliptic; this is important, because it implies regularity for any solution of the stationary FPE [25]. It is also to be expected that the semigroup $\exp(tL^T)$ has a smooth (even real analytic) kernel so that for $t > 0$ $P(x, y; t)$ will be smooth (real analytic). This is supported by our numerical studies. So the problematic point is to

show that the two terms on the right-hand side of Eq. (38) actually cancel. This may fail because the observables \mathcal{O} typically grow in imaginary direction, whereas the decay of $P(x, y; t)$ (always assuming it exists) may be insufficient to compensate for it.

Let us see in a little more detail how the argument for the independence of N_I may fail. For simplicity of presentation we consider the one-dimensional case ($n = 1$). We write the Langevin operator (10) as

$$L = L_{N_I=0} + N_I \Delta, \quad (43)$$

where $L_{N_I=0}$ is the Langevin operator for $N_I = 0$, $N_R = 1$. Then let us consider

$$\begin{aligned} & \frac{\partial}{\partial N_I} \int P(x, y; t) \mathcal{O}(x + iy) dx dy \\ &= \frac{\partial}{\partial N_I} \int (e^{t(L_{N_I=0} + N_I \Delta)} P(x, y; 0)) \mathcal{O}(x + iy) dx dy. \end{aligned} \quad (44)$$

We use the formula

$$\frac{\partial}{\partial N_I} e^{tL^T} = \int_0^t d\tau e^{\tau L^T} \Delta e^{(t-\tau)L^T}, \quad (45)$$

and integration by parts to rewrite Eq. (44) as

$$\int_0^t d\tau \int P(x, y; t - \tau) \Delta \mathcal{O}(x + iy; \tau) dx dy + X. \quad (46)$$

The term denoted by X collects possible boundary terms arising in the integration by parts. It vanishes only if the decay of $P(x, y; t - \tau)$ is strong enough to offset any possible growth of $\mathcal{O}(x, y; \tau)$; otherwise it may either converge to a finite nonzero value or diverge. By the CR equations the first term vanishes, but the uncontrolled boundary term X remains.

Let us look at a simple example that shows how and when the formal argument fails. Consider

$$\begin{aligned} X &= \int (\Delta P(x, y; t)) \mathcal{O}(x + iy) dx dy \\ &\quad - \int P(x, y; t) \Delta \mathcal{O}(x + iy) dx dy \end{aligned} \quad (47)$$

(where the second term vanishes on account of the CR equations). By the formal argument this would be zero, being just a boundary term, but careful application of integration by parts, at first over the finite domain $-Y_- < y < Y_+$, gives

$$\begin{aligned} X &= \lim_{Y_{\pm} \rightarrow \infty} \int dx [(\nabla_y P(x, y; t)) \mathcal{O}(x + iy) \\ &\quad - P(x, y; t) \nabla_y \mathcal{O}(x + iy)] \Big|_{-Y_-}^{Y_+}. \end{aligned} \quad (48)$$

Here it is clear that what matters is the combined asymptotic behavior of P and \mathcal{O} , and depending on the observable, X may be zero, finite, or divergent. Of course the form

of the boundary terms is less simple when using integration by parts in Eqs. (38) and (44), but we expect that it is still the decay of the products like $P\mathcal{O}$, $\mathcal{O}\nabla_y P$, $P\nabla_y \mathcal{O}$ that is relevant.

When trying to investigate the effects of the boundary numerically, one would in principle like to use the probability density obtained without a cutoff. In practice this is, however, not feasible, and we therefore introduce a cutoff in the imaginary direction, which is not sent to infinity. Such a device is necessary for the solution of the FPE, and even though the CLE does not require it, for the purpose of comparison we also introduce it there. This will, however, introduce additional problems with the formal arguments relying on the CR equations as well as integration by parts.

Concretely we proceed as follows: We restrict each y_j to lie between $-Y_-$ and Y_+ and impose periodic boundary conditions on both the observables and the probability densities. This has a number of consequences. First, observables $\mathcal{O}(x + iy)$ will in general not be continuous across the ‘‘seam,’’ where we identify $y_j = -Y_-$ with $y_j = Y_+$. They can therefore not be interpreted as continuous functions and *a priori* the Itô formula (9) does not hold (it may still hold in the sense of distributions, which should be sufficient for our purposes). Furthermore, the jump across the seam will mean that the CR equations are no longer satisfied everywhere. For the evolved observables the CR equations cannot be expected to hold anywhere exactly, as the violation that occurred initially only at the boundary gets propagated everywhere by the Langevin evolution. Similarly, the drift in the FPE is expected to be discontinuous at the seam. Therefore we have to expect that P has a jump there as well; this in turn forces us to interpret the FPE in the sense of distributions.

One might wonder whether it would not be better to limit the fluctuations in the imaginary direction by introducing a smooth cutoff; but since such a smooth cutoff function will necessarily be nonholomorphic it will destroy the formal arguments even more; we therefore stick with the simplest choice of a periodic cutoff.

We conclude therefore that the introduction of a cutoff $-Y_- < y_j < Y_+$ and imposing periodic boundary conditions leads to a breakdown of the formal arguments given in Sec. II. Although it is difficult to quantify precisely the effect of this, it seems reasonable to expect that it is still the behavior of $P\mathcal{O}$ and similar products at large $|y_j|$ that determines what is happening. For the CLE it is also clear that a very large cutoff will practically not be felt, because the system very rarely will make contact with it. This is borne out by our numerics which clearly shows convergence to the limit of infinite cutoff. For the FPE, on the other hand the issue is less clear, because there are very large boundary terms arising from the gradients of the drift across the ‘‘seam.’’ In any case, for the FPE we cannot directly compare with the cutoff-free results, because these do not exist.

V. CASE STUDIES

A. The $U(1)$ one-link model

To understand in more detail how boundary terms affect the behavior of complex Langevin simulations, we studied in some detail the $U(1)$ one-link model in the hopping approximation that was already discussed in Ref. [12] for $N_I = 0$.

The action is

$$S = -\beta \cos z - \kappa \cos(z - i\mu) = -a \cos(z - ic), \quad (49)$$

with

$$a = \sqrt{(\beta + \kappa e^\mu)(\beta + \kappa e^{-\mu})}, \quad (50)$$

$$c = \frac{1}{2} \ln \frac{\beta + \kappa e^\mu}{\beta + \kappa e^{-\mu}}. \quad (51)$$

The complex drift force is correspondingly

$$K = -S' = -\beta \sin z - \kappa \sin(z - i\mu) = -a \sin(z - ic), \quad (52)$$

and the two components of the drift read

$$K_x = -\text{Re}S' = -a \sin x \cosh(y - c), \quad (53)$$

$$K_y = -\text{Im}S' = -a \cos x \sinh(y - c). \quad (54)$$

As discussed in Ref. [12] there are two fixed points at $(x, y) = (0, c)$ and $(x, y) = (\pi, c)$; the first one is attractive, the second one repulsive.

A special feature of this model is that for $y = c$ the drift is purely in x direction. If in addition $N_I = 0$, the Langevin process will never leave the line $y = c$ if it starts there (we emphasize that the properties discussed in this paragraph do not hold for the full $U(1)$ one-link model, which was studied in detail in Ref. [12]). It is therefore straightforward to find an explicit solution to the stationary FPE,

$$P(x, y; \infty) \propto e^{-S(x+ic)} \delta(y - c). \quad (55)$$

It follows that this model is actually equivalent to one with a real action, once we shift $y \rightarrow y + c$ and replace β by a (a and c now embody the dependence on the parameters of the model). The numerics presented below show that the line $y = c$ is an attractor for the Langevin process; this indicates that the solution (55) is unique (with proper normalization). These properties imply that for $N_I = 0$ the dynamics is completely understood.

When $N_I > 0$, the presence of the repulsive fixed point is responsible for the occurrence of large excursions, which are well known to be the scourge of complex Langevin simulations. For large y the drift terms dominate over the noise, and the Langevin process is essentially just a deterministic motion; the ‘‘classical’’ trajectories are given by

$$z(t) = i \ln \frac{1 - iC e^{-at}}{1 + iC e^{-at}} + ic, \quad (56)$$

where the complex integration constant C is related to the starting point by

$$C = \tan\left(\frac{z(0) - ic}{2}\right). \quad (57)$$

It is easy to see that all trajectories, except those starting on the unstable trajectories ($\text{Im}z(0) = \pi$) are attracted to the stable fixed point at $z = ic$.

As an illustration we show in Fig. 1 a scatter plot of a Langevin simulation clearly exhibiting the classical orbits. There were 500 update steps between consecutive points and we used $N_I = 1$. To enhance the classical features, this simulation was done with the noise terms in the CLE [see Eqs. (59) and (60) below] suppressed by a factor 10; this is of course equivalent to replacing ϵ by $\epsilon/100$ while multiplying the force terms by a factor of 100. This scatter plot should capture some generic features that will also be present for other choices of the model parameters.

In order to numerically study the role of boundary terms and large $|y|$ values at nonzero N_I , we introduce a cutoff in imaginary direction, placed symmetrically around $y = c$, i.e., $-Y + c < y < Y + c$, and impose periodic boundary conditions. To see the effect of this cutoff, we compute numerically the expectation values of the observables $\exp(ikz)$ for $k = \pm 1, \pm 2$, for various values of N_I and Y , both by simulating the Langevin process and by numerical solution of the FPE. Note that these observables grow exponentially at large y . The exact values are given by

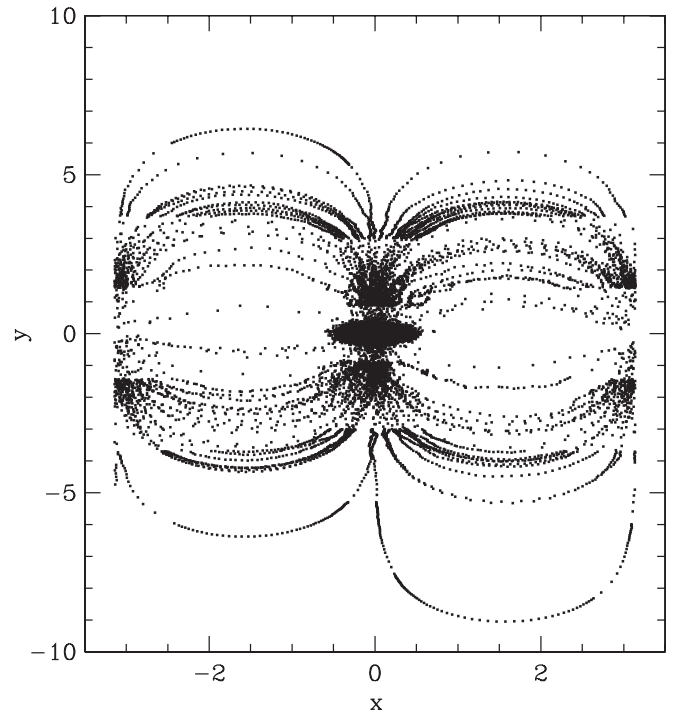


FIG. 1. Scatter plot for the $U(1)$ one-link model at $\beta = 1$, $\kappa = 0$, $N_I = 1$ with reduced noise (see text).

$$\langle e^{ikx} \rangle = \frac{I_k(a)}{I_0(a)} e^{-kc}, \quad (58)$$

where $I_k(a)$ are the modified Bessel functions of the first kind.

The Langevin process is discretized in the usual way,

$$x_{n+1} = x_n + \epsilon K_x(x_n, y_n) + \sqrt{\epsilon N_R} \eta_{x,n} \quad (59)$$

$$y_{n+1} = y_n + \epsilon K_y(x_n, y_n) + \sqrt{\epsilon N_I} \eta_{y,n} \quad (60)$$

where $\eta_{x,n}$ and $\eta_{y,n}$ are pseudorandom numbers with zero mean and variance 2. We use periodic boundary conditions in the imaginary direction, as stated above. We also use an adaptive step size [17], choosing ϵ such that the product of $\epsilon |K_x + iK_y| \leq 0.001$. To estimate the statistical error we run 100 trajectories with independent random starting points.

To solve the FPE numerically, we employ the periodicity in x and consider the Fourier decompositions

$$\hat{\rho}(k; t) = \int \frac{dx}{2\pi} e^{ikx} \rho(x; t), \quad (61)$$

$$\hat{P}(k, y; t) = \int \frac{dx}{2\pi} e^{ikx} P(x, y; t), \quad (62)$$

with the inverse transformations given by

$$\rho(x; t) = \sum_{k=-\infty}^{\infty} e^{-ikx} \hat{\rho}(k; t), \quad (63)$$

$$P(x, y; t) = \sum_{k=-\infty}^{\infty} e^{-ikx} \hat{P}(k, y; t). \quad (64)$$

The FPE can be rewritten in terms of these modes as

$$\begin{aligned} \partial_t \hat{P}(k, y; t) = & (-N_R k^2 + N_I \partial_y^2) \hat{P}(k, y; t) \\ & + \frac{a}{2} \cosh(y - c) [(k + 1) \hat{P}(k - 1, y; t) \\ & - (k - 1) \hat{P}(k + 1, y; t)] \\ & + \frac{a}{2} \sinh(y - c) \partial_y [\hat{P}(k - 1, y; t) \\ & + \hat{P}(k + 1, y; t)]. \end{aligned} \quad (65)$$

This equation is solved numerically with a discretized time step $\epsilon = 10^{-5}$ and a spatial discretization in y of $\delta = \sqrt{\epsilon}$. We vary the cutoff Y in the y direction from 10δ up to 500δ (1500δ in some cases). In the x direction, with $-\pi < x \leq \pi$, we use 20–30 points. We found that convergence was reached after 10^6 time steps, corresponding to a Langevin time $t \sim 10$. We found convergence for all the values of N_I and Y studied. For $N_I = 0$, we were not able to solve the

FPE numerically, due to the singular behavior, but the solution is known analytically, see Eq. (55).

We fix the parameters of the model to be

$$\beta = 1.0, \quad \kappa = 0.25, \quad \mu = 0.5. \quad (66)$$

In Fig. 2 we show some examples of the real probability distribution $P(x, y, t)$ at large Langevin time $t \sim 20$ for various N_I and a given cutoff $Y = 0.474$. The y direction goes from left to right and the compact x direction from

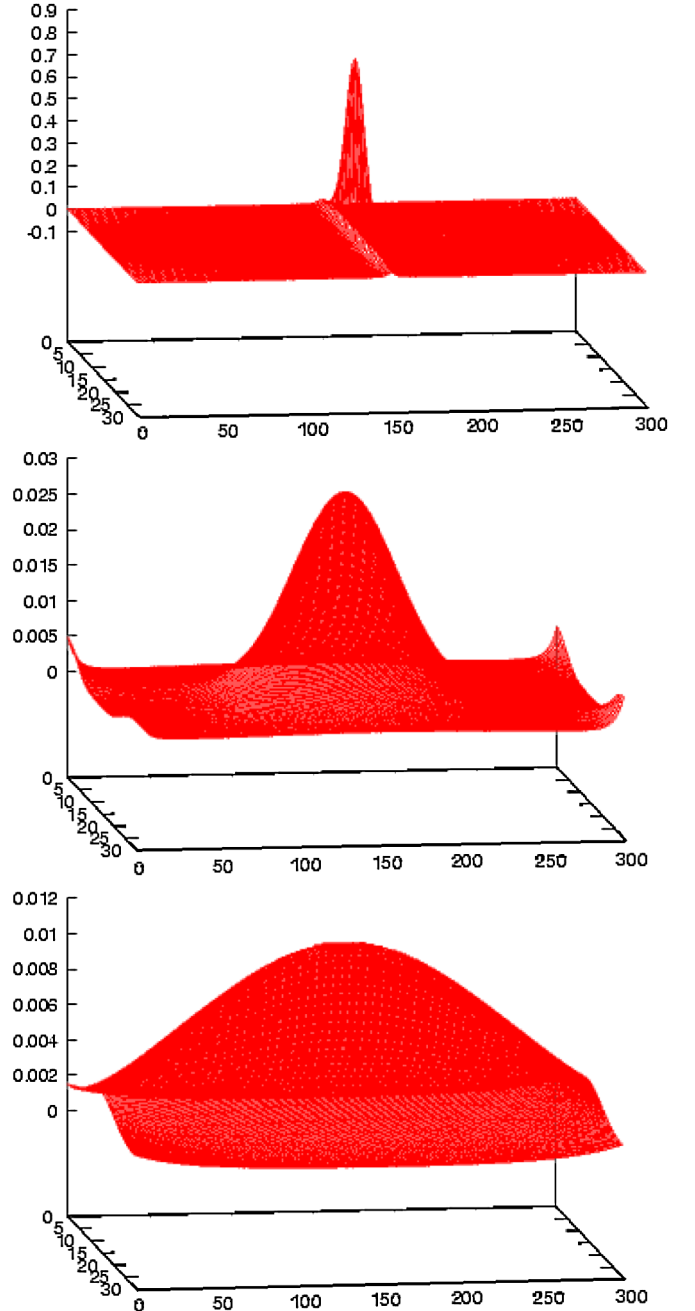


FIG. 2 (color online). Distributions $P(x, y; t \rightarrow \infty)$ in the $U(1)$ one-link model, obtained from a numerical solution of the real FPE, for various values of N_I : $N_I = 0.0001$ (top), 0.01 (middle), 0.1 (bottom). See the main text for further details.

back to front. We observe that at small $N_I = 0.0001$, the distribution resembles the analytic result at $N_I = 0$. The distribution is very narrow in the y direction and boundary effects are not expected to play a role. Increasing N_I results in a wider distribution, and boundary effects become clearly visible. The apparent nonsmooth behavior at the edge $|y - c| = Y$ is to be expected from the discontinuity across the seam, discussed at the end of the previous section.

After obtaining the distribution, expectation values of observables follow from Eq. (22). In Tables I, II, III, IV, V, VI, and VII (see end of paper) we compare the results of the Langevin simulation and the FPE for increasing values of N_I from 0 up to $N_I = 0.1$. Note that the Langevin equation can be solved without a cutoff ($Y = \infty$). The imaginary parts of the observables are consistent with zero. The data of the tables are also summarized in

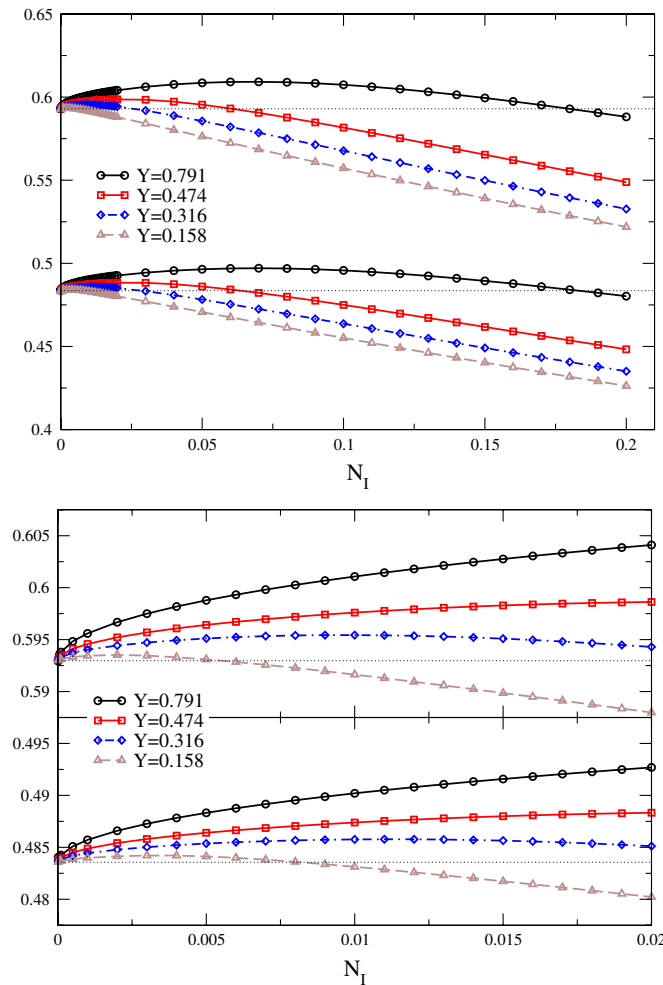


FIG. 3 (color online). N_I dependence of $\text{Re}\langle e^{iz} \rangle$ (lower points) and $\text{Re}\langle e^{-iz} \rangle$ (higher points) from FPE for various values of the cutoff Y . The bottom figure zooms in on smaller values of N_I . The lines are guides to the eye, the horizontal dotted lines indicate the correct results (0.483 564 and 0.592 966, respectively).

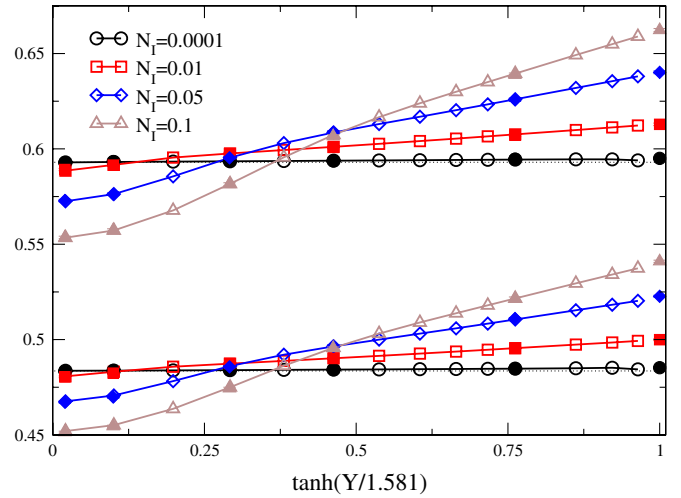


FIG. 4 (color online). Cutoff (Y) dependence of $\text{Re}\langle e^{iz} \rangle$ (lower data) and $\text{Re}\langle e^{-iz} \rangle$ (upper data) for various values of N_I , for FPE (open symbols) and CLE (full symbols, note that the errorbars are much smaller than the points). The lines are guides to the eye. The horizontal dotted lines indicate the correct results.

Figs. 3 and 4 (the FPE and the CLE data are indistinguishable at the scale of the figures; the rightmost points in Fig. 4 correspond to $Y = 4.74$ for FPE and ∞ for CLE).

- The following facts can be inferred from these results:
- (1) CLE and FPE give rather similar results, but sometimes they differ by several σ (statistical error of the CLE simulation).
 - (2) All the data show a clear dependence on N_I , in contrast to the conclusion of the formal arguments. For larger N_I values both CLE and FPE give results different from the exact values.
 - (3) The best results are generally obtained for the smallest N_I . In this case there is also the weakest Y dependence, in fact no Y dependence whatsoever for $N_I = 0$.

Obviously the presence of the cutoff and periodic boundary conditions affects the CLE and FPE in a similar way. But the N_I dependence shows the failure of the formal argument, even for $Y = \infty$. At least for $k = \pm 1$ one has a clear case of “convergence to a wrong limit.” The data of the CLE with $Y = \infty$ and $k = \pm 2$ are actually not really converged, except for $N_I = 0$. Observing the scatter plots for different Langevin times it appears unclear whether the observables $\exp(\pm 2iz)$ really reach an equilibrium distribution—they tend to drift out to infinity, whereas their averages, while remaining small, suffer from huge fluctuations. On the other hand the observables $\exp(\pm iz)$ seem to reach a stable distribution for $Y = \infty$ and any N_I .

B. The model of Guralnik and Pehlevan

Guralnik and Pehlevan [15] studied an instructive toy model on $\mathcal{M} = \mathbb{R}$ and $\mathcal{M}_c = \mathbb{C}$, called GP model henceforth. Its action is

$$S = -i\beta(z + \frac{1}{3}z^3), \quad (67)$$

and was studied in connection with PT invariant but non-Hermitian Hamiltonians (where PT indicates the combined action of parity and time reversal) [26].

The action (67) leads to the drift forces

$$K_x = -2\beta xy, \quad K_y = \beta(1 + x^2 - y^2). \quad (68)$$

There is a stable fixed point at $z = i$ and an unstable one at $z = -i$. The “classical trajectories” obtained by leaving out the noise are given by

$$z(t) = \frac{z_0 + i \tanh(\beta t)}{1 - iz_0 \tanh(\beta t)}. \quad (69)$$

Since this is a Möbius transformation from $w = \tanh(\beta t)$ to $z(t)$ the trajectories are circles. They can be imagined to emerge from the unstable fixed point $z = -i$ at $t = -\infty$ and go to the stable fixed point $z = i$ as $t \rightarrow \infty$. Those classical trajectories again can be seen clearly in the large excursions, since there the noise becomes negligible. Figure 5 shows the result from a Langevin simulation at $N_I = 1.0$ and $\beta = 1.0$ (in this case there are 50 000 update steps between two consecutive points).

Guralnik and Pehlevan give the exact results of the first three moments at $\beta = 1$, they are

$$\begin{aligned} \langle z \rangle &= -i \frac{\text{Ai}'(1)}{\text{Ai}(1)} = 1.1763i, & \langle z^2 \rangle &= -1, \\ \langle z^3 \rangle &= i - \langle z \rangle = -0.1763i. \end{aligned} \quad (70)$$

They solved the discretized Langevin equation numerically, using $N_I = 0.001$, and obtained good agreement

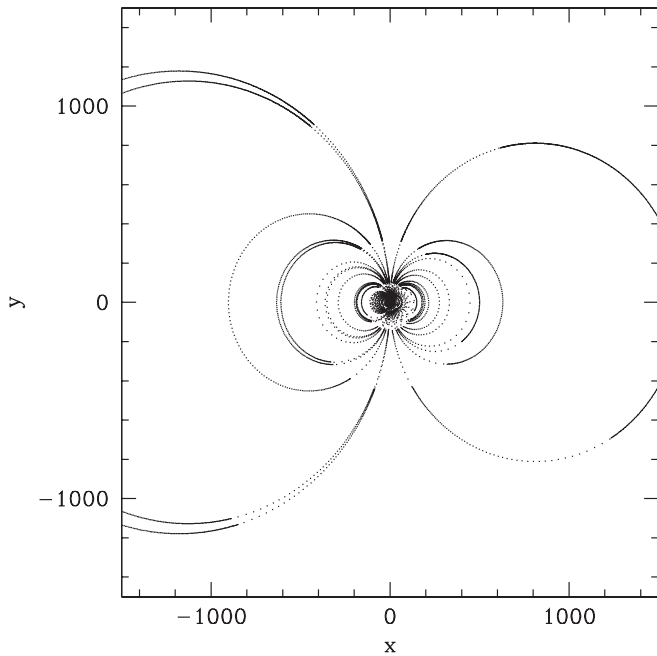


FIG. 5. Scatter plot in the GP model with $\beta = 1$, $N_I = 1$.

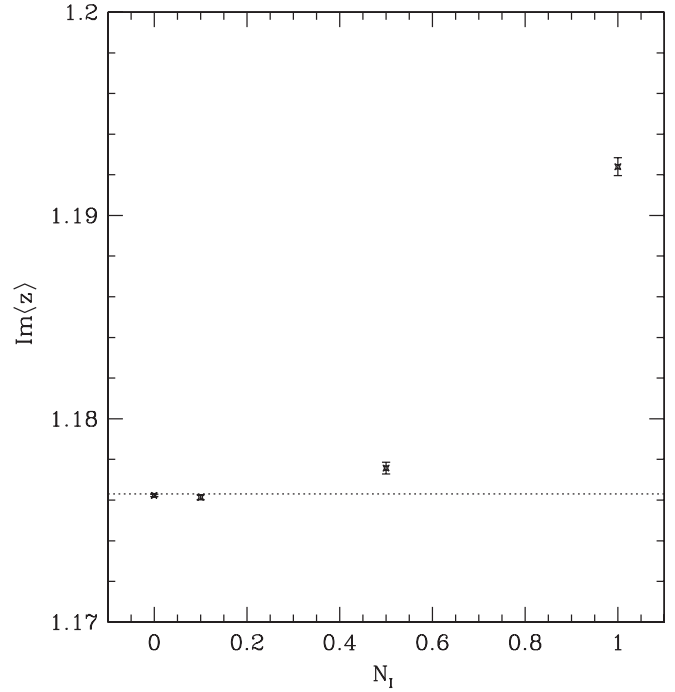


FIG. 6. N_I dependence of $\text{Im}\langle z \rangle$ in the GP model at $\beta = 1$.

with those exact results. We did some more and probably longer simulations at $N_I = 0, 0.1, 0.5, 1.0$. The results for $\text{Im}\langle z \rangle$ are shown in Fig. 6; again they show a clear dependence on N_I , in conflict with the formal reasoning. While for small N_I there is agreement with the exact result, for larger N_I we again have convergence to the wrong limit.

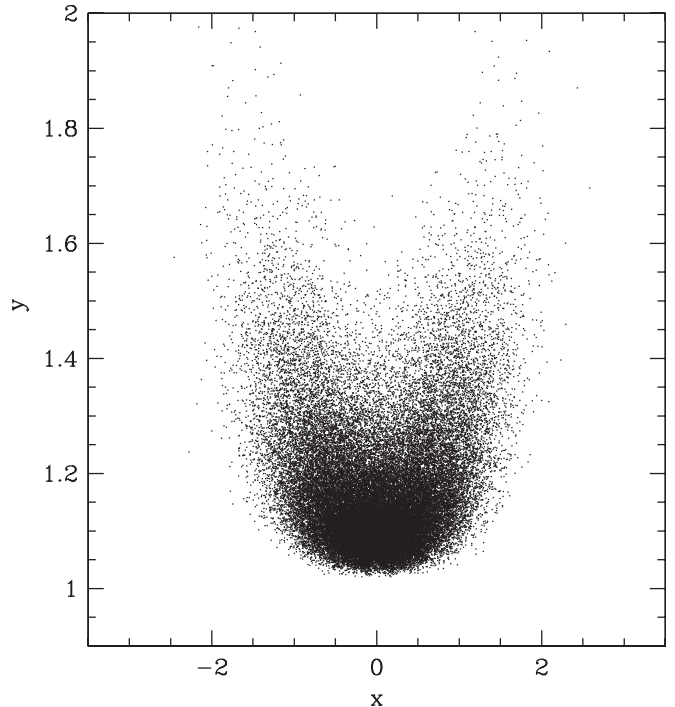


FIG. 7. Scatter plot in the GP model for $\beta = 1$, $N_I = 0$.

The results for $\text{Re}\langle z^2 \rangle$ show a similar behavior, except that at $N_I = 1.0$ the result is $-0.951(44)$, with a statistical error that is huge compared to the error 0.0007 found at $N_I = 0.5$. The data for $\text{Im}\langle z^3 \rangle$ show an even more dramatic failure for larger N_I : they diverge for $N_I = 1.0$, whereas for $N_I = 0.5$ fluctuations are large. Finally we measured $\text{Im}\langle e^{iz} \rangle$, which has an exact value of 2.2624 . Here diver-

gence becomes manifest already for $N_I = 0.5$ (but it might occur for all $N_I > 0$).

In order to have an idea of the equilibrium distribution for $N_I = 0$ we show in Fig. 7 a scatter plot of 50 000 configurations in the complex plane. Non-Gaussian behavior is quite clear from this plot. Noticeable is the appearance of sharp edges of the distribution, possibly indicating jumps, but no δ functions like in the $U(1)$ case in the hopping expansion. To obtain these pictures we sampled over 60 000 points taken at equal intervals of 0.5 in Langevin time. Similar distributions have been observed in Ref. [12] for the full $U(1)$ one-link model. The non-Gaussian character is further demonstrated in Fig. 8: the histograms for both $\text{Re}z$ and $\text{Im}z$ deviate strongly from a Gaussian distribution.

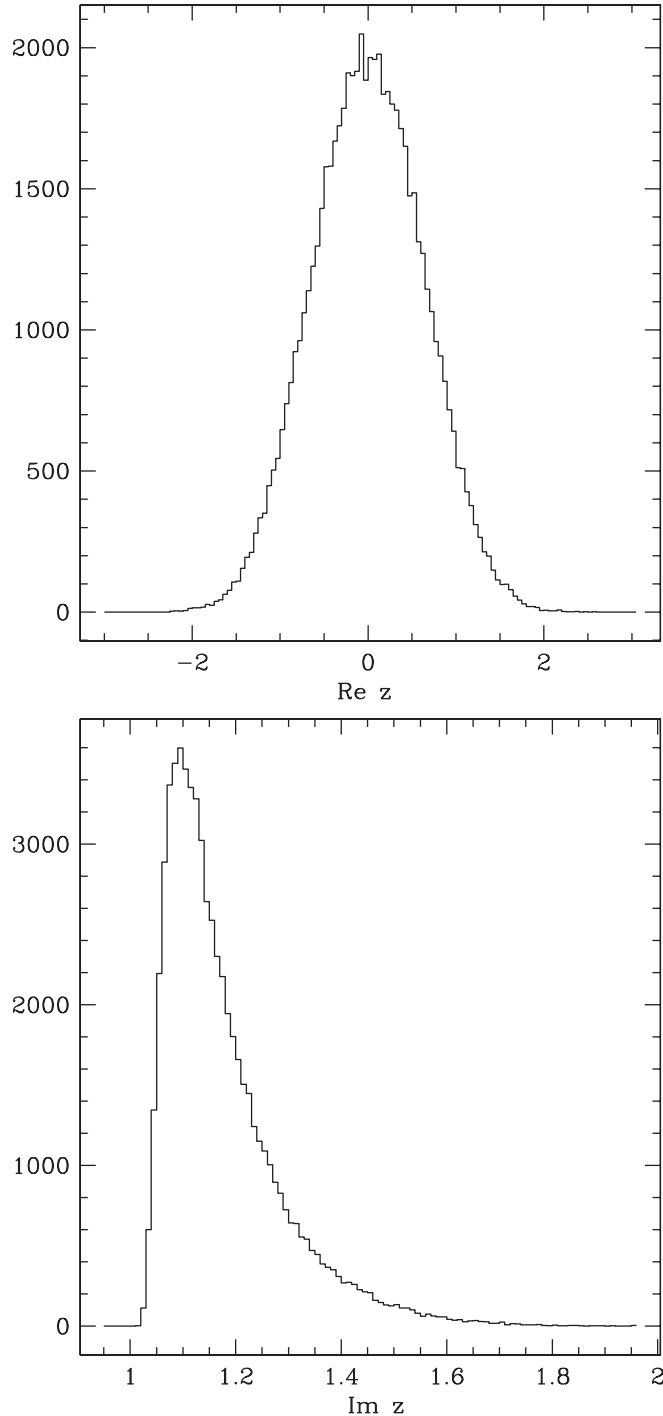


FIG. 8. Distribution of $\text{Re}z$ (top) and $\text{Im}z$ (bottom) in the GP model with $\beta = 1$, $N_I = 0$.

VI. TENTATIVE CONCLUSIONS

A. Explanation of wrong results

We will now try to interpret these findings. They are quite analogous in the two model cases and we will try to reach some conclusions that can be generalized to more realistic models.

The question of convergence vs divergence apparently depends on the values of N_I , but it is more plausible that there is no qualitative difference between the different positive values of N_I , only the time needed to observe the asymptotic behavior is different. On the other hand for $N_I = 0$ there seems to be really a qualitative difference: the distributions develop discontinuities or even δ functions; but more important is the fact that they seem to drop very rapidly in the imaginary direction.

We tentatively conclude that for $N_I = 0$ the systems relax to equilibrium measures that show at least exponential decay in imaginary direction [in our simple $U(1)$ model this decay is of course much stronger—the measure is zero for $|y| > c$].

For $N_I > 0$ the situation is less clear, but it seems that in the $U(1)$ model we get a decay at least like $\exp(-|y|)$, but probably not stronger than any exponential. In the model of Guralnik and Pehlevan the data suggest a powerlike decay (the power appears to be near 2).

To sum up our tentative conclusions: There is a unique equilibrium distribution $P(x, y; \infty)$ for the Langevin processes for any $N_I \geq 0$, but for $N_I > 0$ it shows limited decay, especially in the imaginary direction.

The type of decay for $N_I > 0$ depends on the model considered; but for some observables their growth in imaginary direction may conspire with the falloff of the equilibrium measure in such a way that we obtain convergence to a wrong limit. In any case one should not expect convergence of the mean value for *all* holomorphic observables.

For very small N_I and limited simulation time, the behavior is of course indistinguishable from the one at

$N_I = 0$, and one may reach a quasiconvergence to the right limit, even if an infinitely long simulation would diverge.

If we try to generalize boldly from our toy $U(1)$ model to real lattice gauge theories, we expect that for $N_I > 0$ and for most interesting observables, such as Wilson loops, Polyakov loops, etc., we have to expect boundary terms contributing even as the boundary is sent to infinity. For multiply charged loops and $N_I > 0$ the situation could be

even worse: those boundary contributions may diverge as the boundary is moved to infinity. But these problems probably will not occur for $N_I = 0$ and do not show up at very small values of N_I either, at least if simulations are not run excessively long.

We realize that definite conclusions about the falloff of the probability density $P(x, y)$ in the y direction have not

TABLE I. $N_I = 0.0$.

Y		$\langle U \rangle$	$\langle U^{-1} \rangle$	$\langle U^2 \rangle$	$\langle U^{-2} \rangle$
0.032	CLE	0.483 31(047)	0.592 65(058)	0.123 11(037)	0.198 65(055)
0.158	CLE	0.483 31(047)	0.592 65(058)	0.132 11(037)	0.198 65(055)
0.474	CLE	0.483 31(047)	0.592 65(058)	0.132 11(037)	0.198 65(055)
0.790	CLE	0.483 31(047)	0.592 65(058)	0.132 11(037)	0.198 65(055)
1.581	CLE	0.483 31(047)	0.592 65(058)	0.132 11(037)	0.198 65(055)
∞	CLE	0.483 31(047)	0.592 65(058)	0.132 11(037)	0.198 65(055)
Exact		0.483 56	0.592 97	0.130 65	0.196 46

TABLE II. $N_I = 0.00001$.

Y		$\langle U \rangle$	$\langle U^{-1} \rangle$	$\langle U^2 \rangle$	$\langle U^{-2} \rangle$
0.032	CLE	0.483 44(046)	0.592 81(057)	0.131 54(036)	0.197 78(054)
	FPE	0.483 58	0.592 97	0.130 65	0.196 46
0.158	CLE	0.483 69(047)	0.593 13(057)	0.131 94(036)	0.198 35(055)
	FPE	0.483 61	0.593 02	0.130 65	0.196 45
0.474	CLE	0.483 70(047)	0.593 13(058)	0.131 88(036)	0.198 22(054)
	FPE	0.483 76	0.593 07	0.130 66	0.196 39
0.790	CLE	0.483 72(048)	0.593 19(059)	0.131 86(036)	0.198 12(055)
	FPE	0.484 05	0.592 92	0.130 76	0.196 18
1.581	CLE	0.483 82(047)	0.593 30(058)	0.131 83(037)	0.197 92(056)
	FPE	0.486 47	0.590 33	0.133 18	0.194 49
∞	CLE	0.483 96(047)	0.593 45(058)	0.130 89(042)	0.196 85(056)
Exact		0.483 56	0.592 97	0.130 65	0.196 46

TABLE III. $N_I = 0.0001$.

Y		$\langle U \rangle$	$\langle U^{-1} \rangle$	$\langle U^2 \rangle$	$\langle U^{-2} \rangle$
0.032	CLE	0.483 45(050)	0.592 80(062)	0.131 14(035)	0.197 14(052)
	FPE	0.483 66	0.592 90	0.130 66	0.196 46
0.158	CLE	0.483 85(054)	0.593 35(067)	0.131 53(039)	0.197 71(059)
	FPE	0.483 71	0.593 13	0.130 64	0.196 44
0.474	CLE	0.484 02(052)	0.593 57(064)	0.131 48(037)	0.197 62(056)
	FPE	0.483 99	0.593 48	0.130 59	0.196 37
0.790	CLE	0.484 25(053)	0.593 77(065)	0.131 46(038)	0.197 64(057)
	FPE	0.484 25	0.593 79	0.130 56	0.196 27
1.581	CLE	0.484 85(051)	0.594 54(063)	0.131 15(036)	0.197 25(054)
	FPE	0.484 76	0.594 37	0.123 63	0.195 83
∞	CLE	0.485 27(051)	0.595 04(063)	0.129 44(047)	0.193 87(079)
Exact		0.483 56	0.592 97	0.130 65	0.196 46

yet been reached; we intend to return to a more detailed study of this question in a future paper.

B. Practical conclusions

The conclusions for practical applications are quite straightforward:

- (1) In complex Langevin simulations one should use $N_I = 0$. If one wants to do a random walk simulation or an iterate of the Fokker-Planck operator, this is not possible, but one should make sure that $N_I \ll 1$.

TABLE IV. $N_I = 0.001$.

Y		$\langle U \rangle$	$\langle U^{-1} \rangle$	$\langle U^2 \rangle$	$\langle U^{-2} \rangle$
0.032	CLE	0.483 42(049)	0.592 93(060)	0.130 95(038)	0.196 92(057)
	FPE	0.483 73	0.592 26	0.130 74	0.196 02
0.158	CLE	0.484 10(051)	0.593 46(062)	0.130 68(037)	0.196 32(056)
	FPE	0.483 99	0.593 44	0.130 65	0.196 47
0.474	CLE	0.485 38(051)	0.594 97(064)	0.130 90(038)	0.197 02(058)
	FPE	0.484 88	0.594 57	0.130 50	0.196 23
0.790	CLE	0.486 19(051)	0.596 05(061)	0.130 66(041)	0.196 55(063)
	FPE	0.485 72	0.059 559	0.130 27	0.195 89
1.581	CLE	0.487 62(052)	0.597 61(062)	0.129 34(044)	0.195 03(064)
	FPE	0.487 29	0.597 53	0.129 34	0.194 49
∞	CLE	0.488 90(050)	0.599 42(063)	0.123 92(068)	0.186 02(086)
Exact		0.483 56	0.592 97	0.130 65	0.196 46

TABLE V. $N_I = 0.01$.

Y		$\langle U \rangle$	$\langle U^{-1} \rangle$	$\langle U^2 \rangle$	$\langle U^{-2} \rangle$
0.032	CLE	0.480 47(049)	0.589 17(060)	0.129 35(035)	0.194 50(053)
	FPE	0.480 73	0.588 56	0.128 94	0.193 28
0.158	CLE	0.482 40(051)	0.591 50(063)	0.130 39(037)	0.196 13(056)
	FPE	0.483 13	0.591 63	0.130 55	0.195 84
0.474	CLE	0.487 47(053)	0.597 53(067)	0.130 80(041)	0.196 45(059)
	FPE	0.487 39	0.597 60	0.130 64	0.196 46
0.790	CLE	0.490 12(055)	0.600 78(069)	0.129 52(045)	0.194 37(069)
	FPE	0.490 20	0.601 07	0.129 84	0.195 25
1.581	CLE	0.495 32(052)	0.607 14(068)	0.126 47(054)	0.190 56(085)
	FPE	0.495 50	0.607 58	0.126 68	0.190 51
∞	CLE	0.499 95(049)	0.612 87(063)	0.107 33(202)	0.161 32(376)
Exact		0.483 56	0.592 97	0.130 65	0.196 46

TABLE VI. $N_I = 0.05$.

Y		$\langle U \rangle$	$\langle U^{-1} \rangle$	$\langle U^2 \rangle$	$\langle U^{-2} \rangle$
0.032	CLE	0.467 20(051)	0.572 90(063)	0.121 50(034)	0.182 69(051)
	FPE	0.467 63	0.572 53	0.121 18	0.181 65
0.158	CLE	0.469 85(054)	0.576 20(066)	0.123 35(037)	0.185 57(055)
	FPE	0.470 70	0.576 29	0.123 63	0.185 33
0.474	CLE	0.485 70(057)	0.595 33(071)	0.131 67(042)	0.197 66(065)
	FPE	0.486 08	0.595 42	0.131 84	0.197 93
0.790	CLE	0.496 64(055)	0.608 71(068)	0.131 66(046)	0.197 31(074)
	FPE	0.496 44	0.608 58	0.131 66	0.197 99
1.581	CLE	0.511 04(056)	0.625 80(080)	0.123 66(074)	0.183 70(114)
	FPE	0.510 55	0.626 01	0.123 63	0.185 95
∞	CLE	0.522 72(051)	0.640 13(069)	0.075 53(204)	0.113 59(334)
Exact		0.483 56	0.592 97	0.130 65	0.196 46

TABLE VII. $N_f = 0.1$.

Y		$\langle U \rangle$	$\langle U^{-1} \rangle$	$\langle U^2 \rangle$	$\langle U^{-2} \rangle$
0.032	CLE	0.451 44(051)	0.553 58(062)	0.112 67(035)	0.169 41(053)
	FPE	0.452 03	0.553 43	0.112 36	0.168 43
0.158	CLE	0.454 56(055)	0.557 42(067)	0.114 53(038)	0.172 24(054)
	FPE	0.455 13	0.557 22	0.114 79	0.172 07
0.474	CLE	0.474 43(057)	0.581 51(071)	0.127 38(044)	0.191 17(065)
	FPE	0.475 01	0.581 65	0.127 59	0.191 36
0.790	CLE	0.495 60(053)	0.606 68(073)	0.134 48(052)	0.201 20(081)
	FPE	0.495 75	0.607 40	0.134 46	0.201 98
1.581	CLE	0.521 41(093)	0.638 83(093)	0.123 45(098)	0.185 69(144)
	FPE	0.521 57	0.639 48	0.124 26	0.186 92
∞	CLE	0.541 04(059)	0.662 30(082)	0.050 49(354)	0.077 23(522)
Exact		0.483 56	0.592 97	0.130 65	0.196 46

- (2) Always validate the simulations by comparison with other trusted computations for parameter values where other methods are available.
- (3) Wilson or Polyakov loops of higher charge will generally have much larger fluctuations. In general they are not needed for physics applications, but they may be worth looking at because they can give information about the probability distribution.

ACKNOWLEDGMENTS

We thank Frank James and Denes Sexty for discussion. I.-O. S. thanks the MPI for Physics München and Swansea University for hospitality. G. A. is supported by STFC.

- [1] J. Klauder, Acta Phys. Austriaca Suppl. **XXXV**, 251 (1983); J. Phys. A **16**, L317 (1983); Phys. Rev. A **29**, 2036 (1984).
- [2] G. Parisi, Phys. Lett. **131B**, 393 (1983).
- [3] F. Karsch and H. W. Wyld, Phys. Rev. Lett. **55**, 2242 (1985).
- [4] P.H. Damgaard and H. Hüffel, Phys. Rep. **152**, 227 (1987).
- [5] J. Ambjorn and S. K. Yang, Phys. Lett. **165B**, 140 (1985).
- [6] J. Klauder and W. P. Petersen, J. Stat. Phys. **39**, 53 (1985).
- [7] H. Q. Lin and J. E. Hirsch, Phys. Rev. B **34**, 1964 (1986).
- [8] J. Ambjorn, M. Flensburg, and C. Peterson, Nucl. Phys. **B275**, 375 (1986).
- [9] J. Berges and I.O. Stamatescu, Phys. Rev. Lett. **95**, 202003 (2005).
- [10] J. Berges, S. Borsanyi, D. Sexty, and I.O. Stamatescu, Phys. Rev. D **75**, 045007 (2007).
- [11] J. Berges and D. Sexty, Nucl. Phys. **B799**, 306 (2008).
- [12] G. Aarts and I.O. Stamatescu, J. High Energy Phys. 09 (2008) 018.
- [13] G. Aarts, Phys. Rev. Lett. **102**, 131601 (2009).
- [14] G. Aarts, J. High Energy Phys. 05 (2009) 052.
- [15] C. Pehlevan and G. Guralnik, Nucl. Phys. **B811**, 519 (2009).
- [16] G. Guralnik and C. Pehlevan, Nucl. Phys. **B822**, 349 (2009).
- [17] G. Aarts, F. James, E. Seiler, and I.O. Stamatescu, arXiv:0912.0617.
- [18] G. G. Batrouni *et al.*, Phys. Rev. D **32**, 2736 (1985).
- [19] H. Gausterer and H. Thaler, J. Phys. A **31**, 2541 (1998).
- [20] H. Gausterer, Nucl. Phys. **A642**, c239 (1998).
- [21] R. C. Gunning and H. Rossi, *Analytic Functions of Several Complex Variables* (Prentice-Hall, Englewood Cliffs, New Jersey, 1965).
- [22] V. S. Valdimirov, *Methods of the Theory of Functions of Many Complex Variables* (MIT Press, Cambridge, Massachusetts, 1966).
- [23] E. B. Davies, *Linear Operators and Their Spectra* (Cambridge University Press, Cambridge, England, 2007).
- [24] B. Helffer and F. Nier, *Hypoelliptic Estimates and Spectral Theory for Fokker-Planck Operators and Witten Laplacians* (Springer, Berlin, Germany, 2005).
- [25] D. Gilbarg and N.S. Trudinger, *Elliptic Partial Differential Equations of Second Order* (Springer, Berlin, 1977).
- [26] C.M. Bender, Rep. Prog. Phys. **70**, 947 (2007).