Localised distributions in complex Langevin dynamics

Pietro Giudice[*]
Universität Münster, Institut für Theoretische Physik, Münster, Germany
E-mail: `p.giudice@uni-muenster.de`

Gert Aarts
Department of Physics, College of Science, Swansea University, Swansea, United Kingdom
E-mail: `g.aarts@swan.ac.uk`

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

Complex Langevin dynamics can be used to perform numerical simulations of theories with a complex action. In order to justify the procedure, it is important to understand the properties of the real and positive distribution, which is effectively sampled during the stochastic process. In the context of a simple model, we study this distribution and relate the results to the recently derived criteria for correctness. We demonstrate analytically that if the distribution has support only on a strip in the complexified configuration space, correct results are expected.

31st International Symposium on Lattice Field Theory LATTICE 2013
July 29 - August 3, 2013
Mainz, Germany

---

[*]Speaker.

1. Introduction

Numerical simulations of lattice field theories with a complex action are not possible with standard Monte Carlo methods. Complex Langevin (CL) dynamics provides an appealing alternative since it does not rely on importance sampling, see Refs. [1, 2] for two recent reviews. In the recent past, the important role played by the properties of the real and positive probability distribution in the complexified configuration space, which is effectively sampled during the Langevin process, has been clarified [3, 4]. An important conclusion was that this distribution should be sufficiently localised in order for CL to yield valid results. Moreover, a set of criteria for correctness, which have to be satisfied in order for CL to be reliable, has been determined. In this contribution we aim to combine the insights that can be distilled from the criteria for correctness discussed above with the explicit solution of the Fokker-Planck equation (FPE), adapting the method employed in Ref. [5]. A comprehensive version of this work has been published [6].

2. Complex Langevin dynamics and criteria for correctness

The model we have studied in this work is described by this simple partition function:

$$Z = \int_{-\infty}^{\infty} dx\, e^{-S}, \qquad S = \frac{1}{2}\sigma x^2 + \frac{1}{4}\lambda x^4, \qquad (2.1)$$

where the parameters in the action are complex-valued. This model has been studied shortly after CL was introduced [7, 8, 9], but no complete solution was given. We take λ real and positive, so that the integral exists, while σ is taken complex. Analytical results are available: a direct evaluation of the integral yields $Z = \sqrt{4\xi/\sigma}\; e^{\xi}\, K_{-\frac{1}{4}}(\xi)$, where $\xi = \sigma^2/(8\lambda)$ and $K_p(\xi)$ is the modified Bessel function of the second kind. Moments $\langle x^n \rangle$ can be obtained by differentiating with respect to σ.

We evaluate expectation values numerically, by solving a CL process. This is done by complexification of the Langevin equation:

$$\dot{z} = -\partial_z S(z) + \eta, \qquad (2.2)$$

i.e. we introduce:

$$z = x + iy, \qquad \eta = \eta_{\mathrm{R}} + i\eta_{\mathrm{I}}, \qquad \sigma = A + iB. \qquad (2.3)$$

The dot in Eq. (2.2) denotes differentiating with respect to the Langevin time t and the (Gaussian) noise satisfies

$$\langle \eta(t)\eta(t') \rangle = 2\delta(t - t'). \qquad (2.4)$$

The normalisation of the real and imaginary noise components is given by

$$\langle \eta_{\mathrm{R}}(t)\eta_{\mathrm{R}}(t') \rangle = 2N_{\mathrm{R}}\delta(t - t'), \quad \langle \eta_{\mathrm{I}}(t)\eta_{\mathrm{I}}(t') \rangle = 2N_{\mathrm{I}}\delta(t - t'), \quad \langle \eta_{\mathrm{R}}(t)\eta_{\mathrm{I}}(t') \rangle = 0, \qquad (2.5)$$

with $N_{\mathrm{R}} - N_{\mathrm{I}} = 1$. Here $N_{\mathrm{I}} \geq 0$ is a free parameter, which can be varied.

Expectation values are obtained by averaging over the noise; they evolve according to

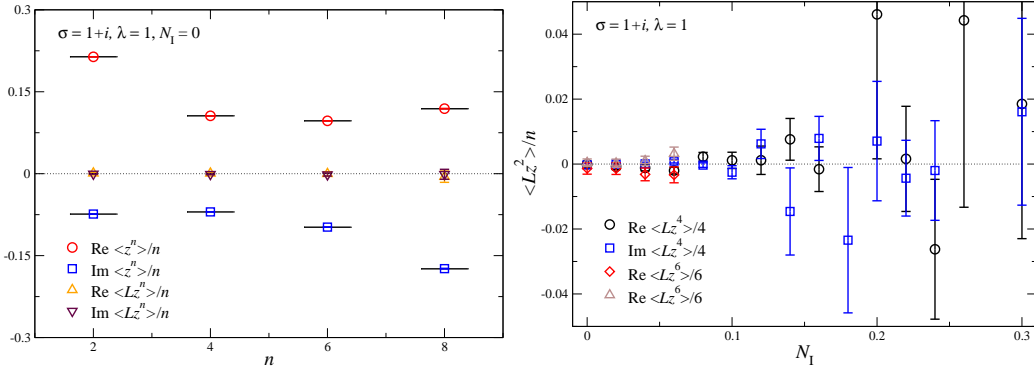$$\langle O \rangle_{P(t)} = \int dx\, dy\, P(x, y; t)\, O(x + iy), \qquad (2.6)$$

Figure 1: (Left) Real and imaginary parts of the expectation values $\frac{1}{n}\langle z^n \rangle$ and criteria for correctness $C_n = \frac{1}{n}\langle \tilde{L}z^n \rangle$ versus n at $\sigma = 1 + i$ and $\lambda = 1$ for real noise ($N_I = 0$). The horizontal lines indicate the exact value. (Right) Criteria for correctness as a function of N_I for $n = 4, 6$.

where the distribution $P(x, y; t)$ satisfies the FPE

$$\dot{P}(x, y; t) = L^T P(x, y; t), \tag{2.7}$$

with the FP operator ($K_x = -\operatorname{Re}\partial_z S(z)$ and $K_y = -\operatorname{Im}\partial_z S(z)$ are the drift terms):

$$L^T = \partial_x (N_R \partial_x - K_x) + \partial_y (N_I \partial_y - K_y). \tag{2.8}$$

In CL dynamics, convergence to the correct result is not guaranteed. A necessary condition is that the so-called *criteria for correctness* are satisfied [3, 4]:

$$C_O \equiv \langle \tilde{L}O(z) \rangle = 0, \tag{2.9}$$

in principle for a complete set of holomorphic observables $O(z)$. Here $\tilde{L}$ is the Langevin operator

$$\tilde{L} = [\partial_z - (\partial_z S(z))]\, \partial_z. \tag{2.10}$$

We consider as observables $O_n(z) = \frac{1}{n}z^n$, with n even.

In Fig. 1 (Left) CL results are shown for the real and imaginary parts of the observables $\frac{1}{n}\langle z^n \rangle$ and for the criteria for correctness $C_n = \frac{1}{n}\langle \tilde{L}z^n \rangle$, for $n = 2, 4, 6, 8$. The figure shows the result for real noise, $N_I = 0$, and parameters $\sigma = 1 + i$ and $\lambda = 1$: all expectation values agree with the exact result, denoted with the horizontal lines, and the criteria for correctness are all consistent with zero, as it should be. Moreover, we have studied how the observables and the criteria for correctness depend on the amount of complex noise. For small N_I the observables with $n = 2, 4$ appear to be consistent with the exact result, while for larger N_I they start to deviate. Problems can be detected by considering higher moments. For small N_I the observables (with $n \geq 6$) and the criteria (with $n \geq 4$) are only marginally consistent with the expected results, while for larger N_I they suffer from large fluctuations, see Fig. 1 (Right), and can no longer be sensibly determined. According to the analytical justification [3, 4], this implies that the results from CL cannot be trusted in the presence of complex noise.

3

3. Probability distributions

In this section we study the distribution $P(x,y)$ in the complexified space. It can be obtained directly collecting histograms of the distribution during the CL evolution; this method can be easily extended to multi-dimensional integrals and field theories. A different method to determine $P(x,y)$ was followed in Ref. [5]: the idea is to expand the distribution in terms of a truncated set of basis functions and solve the resulting matrix problem numerically. If we denote the eigenvalues of $-L^T$ with κ and the eigenfunctions with $P_\kappa(x,y)$, what we have to do is to solve the following eigenvalue problem: $-L^T P_\kappa(x,y) = \kappa P_\kappa(x,y)$. The time-dependent distribution can then be written as:

$$P(x,y;t) = P_0(x,y) + \sum_{\kappa \neq 0} e^{-\kappa t} P_\kappa(x,y), \tag{3.1}$$

where P_0 is the ground state, i.e. the one with eigenvalue $\kappa = 0$. Note that, in order for $P_0(x,y)$ to be the equilibrium distribution, it is necessary that for all other eigenvalues Re $\kappa > 0$.

Following Ref. [5], we expand $P(x,y)$ in a basis of Hermite functions, i.e.

$$P(x,y) = \sum_{k=0}^{N_H-1} \sum_{l=0}^{N_H-1} c_{kl} H_k \left(\sqrt{\omega}x\right) H_l \left(\sqrt{\omega}y\right), \tag{3.2}$$

where ω is a variational parameter appearing in the harmonic oscillator eigenfunctions, and N_H indicates the number of Hermite functions included in the truncated basis. The coefficients c_{kl} have to be determined; this is explained in detail in Ref. [6].

We start with the case of complex noise, $N_I = 1$. The parameters in the action are taken as $\sigma = 1+i$ and $\lambda = 1$, and we consider a basis with $30 \leq N_H \leq 150$ Hermite functions and different values of ω. In the limit of large N_H the results are expected to be independent of the value of ω. In practice however, we find that for finite N_H the parameter ω plays the role of a tuning parameter: in particular, when ω is too small, there are eigenvalues with a negative real part. We find that there is always an ω interval for which: i) there is an eigenvalue consistent with 0; ii) the other eigenvalues are in the right half-plane; iii) the reconstructed ground state distribution is stable under variation of N_H and ω. The smallest 15 eigenvalues are shown in Fig. 2 (Left): for the ω values shown here, all eigenvalues are in the right half-plane and the spectrum around the origin is to a good extent independent of ω. The reconstructed distribution $P(x,y)$, using Eq. (3.2), is shown in Fig. 2 (Right). We find a smooth distribution with a double peak structure. In Fig. 3 (Left) we present the partially integrated distributions $P_y(y) = \int_{-\infty}^{\infty} dx P(x,y)$ on a logarithmic scale. Besides presenting results for various ω values, we also show the histogram obtained during a CL simulation. We observe an acceptable agreement between the CL results and the solution of the FPE for $\omega \sim 1.5, 2$, down to a relative size of 10^{-6}, after which the FP solution can no longer cope. We interpret this as a manifestation of the truncation. When ω is taken too large, the disagreement occurs for smaller values of y. Moreover, we have verified that both partially integrated distributions $P_x(y)$ and $P_y(y)$ are characterised by a power decay with power 5. In the case of real noise, $N_I = 0$, we note that in all cases there is an eigenvalue at (or close to) the origin, but in general convergence is much harder to establish from a study of the eigenvalues alone. Convergence of P_y as ω is increased is demonstrated in Fig. 3 (Right) and we observe that a large value of ω is required, $\omega \sim 50$. It is of course expected that the chosen value of ω eventually becomes irrelevant, but for finite N_H
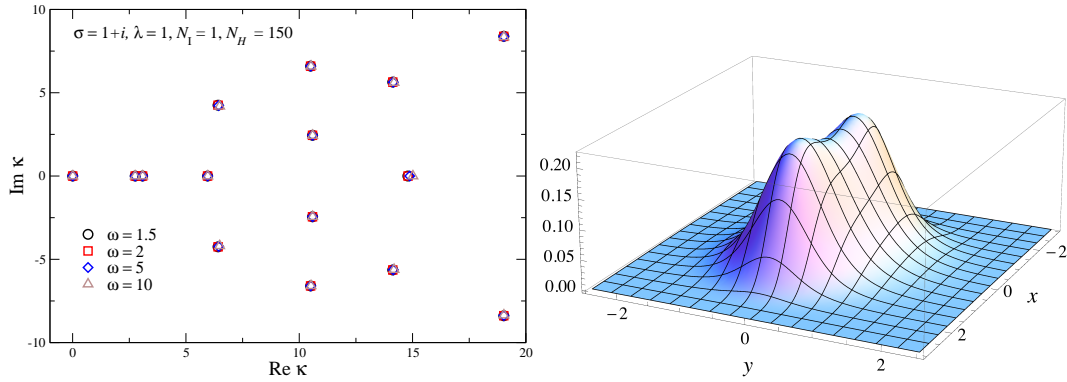
Figure 2: (Left) Eigenvalues of the FP operator $-L^T$ for complex noise, with $N_I = 1$, magnified around the smallest eigenvalues, for various values of ω, at $\sigma = 1+i$, $\lambda = 1$, and $N_H = 150$. (Right) Distribution $P(x, y)$ in the *xy*-plane for complex noise, with $N_I = 1$ and $\omega = 1.5$.
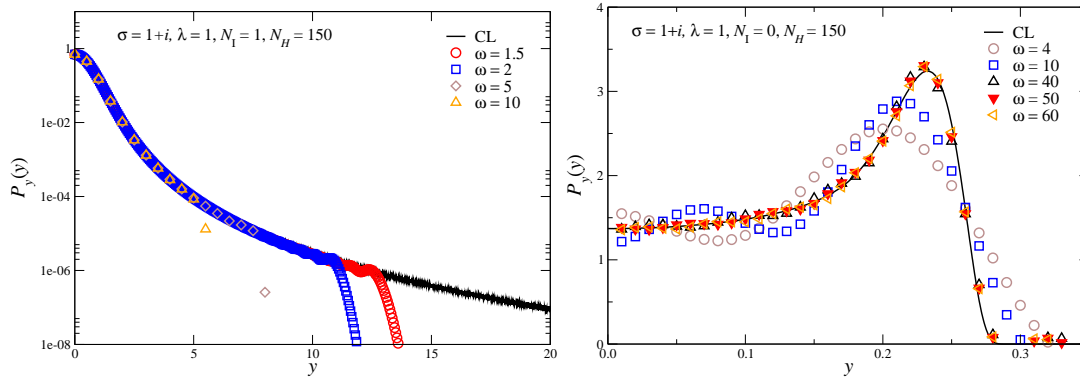

Figure 3: Partially integrated distributions $P_y(y)$ for different values of ω with complex noise, $N_I = 1$ (left plot) and real noise, $N_I = 0$ (right plot). In both cases the noisy (black) data was obtained by a CL simulation.

keeping ω as a tuning parameter is essential. The distribution is very well localised and appears to drop to 0 around $y = 0.28$. The distribution $P_x(x)$, in contrast to the case of complex noise, is now characterised by an exponential rather than a power decay.

We conclude that for this choice of parameters ($\sigma = 1+i$ and $\lambda = 1$) the decay in the case of real noise is manifestly different compared to complex noise. In the latter we found a power decay, resulting in ill-defined moments $\langle z^n \rangle$ when $n > 4$, while here we find exponential decay in the x direction and, as we will see below, in the y direction support only inside a strip. As a result there is no problem in computing higher moments, since they are all well-defined.

4. Interpretation

The classical flow diagram is shown in Fig. 4 (Left), for $\sigma = 1+i$ and $\lambda = 1$. We show the direction of the classical force by an arrow pointing in the direction $(K_x(x,y), K_y(x,y))$. The arrows are normalised to have the same length. There are three fixed points, where $K_x = K_y = 0$: an attractive point at the origin and two repulsive fixed points. The flow is directed towards the
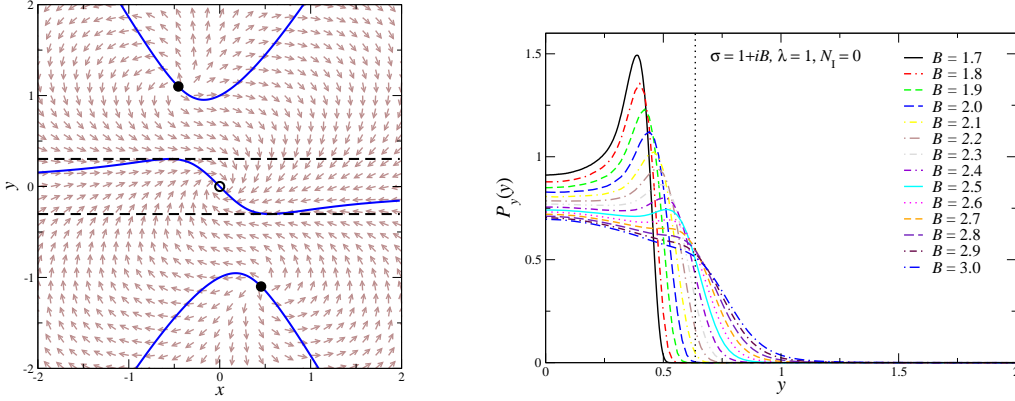
Figure 4: (Left) Classical flow in the xy-plane, for $\sigma = 1 + i$ and $\lambda = 1$. The attractive/repulsive fixed points are indicated with the open/filled circles. The full lines indicate where $K_y(x,y) = 0$. The horizontal dashed lines indicate the strip in which the CL process takes place in the case of real noise. (Right) Distribution $P_y(y)$ for different values of B, with $N_I = 0$, at $\sigma = 1 + iB$ and $\lambda = 1$, obtained with CL; the vertical line indicates the boundary of the strip for $B = 1.7$. For larger B values, there is no longer a boundary.

origin, provided that $|y|$ is not too large. This can be made more precise by studying where $K_y(x,y)$ changes sign: this is indicated in the classical flow diagram with full (blue) lines. We now realise that along the horizontal dashed lines, which are determined by the extrema of the centre curve where $K_y = 0$ ($y = \pm 0.3029$ in this case), the flow is always pointing inwards, i.e. towards the real axis. In absence of a noise component in the vertical direction, this creates a barrier for the Langevin evolution beyond which it cannot drift. We conclude therefore that in the case of real noise the process takes place in the strip determined by $-0.3029 < y < 0.3029$. This is consistent with the conclusions drawn above from the histograms and the FPE solution of the distribution $P(x,y)$. In the presence of complex noise, this conclusion no longer holds and the entire xy-plane can be explored.

As shown in Ref. [6] it is possible to make the argument based on classical flow presented above rigorous and show directly from the FPE that the equilibrium distribution $P(x,y)$ is strictly zero in strips in the xy-plane. To summarise, for real noise, we find the following:

1. when $3A^2 > B^2$, $P(x,y) = 0$ for $y^2 > \frac{A}{2\lambda}\left(1 - \sqrt{1 - \frac{B^2}{3A^2}}\right)$;

2. when $3A^2 < B^2$, there are no restrictions on $P(x,y)$.

Interestingly, the derivation in Ref. [6] demonstrates that when $3A^2 < B^2$, i.e. in absence of strips, one may therefore expect a breakdown of CL with real noise, similar as with complex noise. This is indeed what happens: as shown in Fig. 4 (Right) where the distribution $P_y(y)$ is plotted as B is increased. The delocalisation has a detrimental effect on the results of the CL process. Studying the moments and the criteria for correctness in this case, we observe that increasing B has a similar effect as increasing N_I: moreover a power law takes place again with power 5. In addition in Ref. [6], we have shown that it is possible to understand the universal power decay directly from the FPE: as a consequence, for large $|x|$ and $|y|$, the distribution decays as a power, according to $P(x,y) \sim (x^2 + y^2)^{-3}$. We conclude that in absence of strips a universal power law decay is present,

which results in a breakdown of the formal justification [3, 4] and wrong or wildly fluctuating results in practice.

5. Conclusion

In order to justify the results obtained with complex Langevin dynamics, it is necessary that the probability distribution is sufficiently localised in the complexified configuration space. Here we have studied properties of this distribution via a number of methods, in the case of a simple model. In this case the FPE can be solved explicitly, via an expansion in a truncated set of basis functions. However, it is still a nontrivial problem and perhaps the best way to find the distribution is by brute force, i.e. during the CL simulation. We have demonstrated that the essential insight can already be obtained from a combination of histograms of partially integrated distributions and the criteria for correctness, which gives a consistent picture of the dynamics. These tools are readily available in field theory.

Recently, a new approach, based on deforming the integration contour of the path integral into the complex plane and performing Monte Carlo simulations on the so-called Lefschetz thimbles, has been introduced [10, 11]. A comparison between CL and the Lefschetz thimble for the model considered here, can be found in Ref. [12].

Finally, we remark that our conclusions are also immediately applicable to nonabelian SU(N) gauge theories [13], for which gauge cooling [2, 14] provides a means to control the distribution in SL($N, \mathbb{C}$), a possibility not present in simpler models.

References

[1] G. Aarts, PoS LATTICE **2012** (2012) 017 [arXiv:1302.3028 [hep-lat]].

[2] G. Aarts, L. Bongiovanni, E. Seiler, D. Sexty and I. -O. Stamatescu, Eur. Phys. J. A **49** (2013) 89 [arXiv:1303.6425 [hep-lat]].

[3] G. Aarts, E. Seiler and I. -O. Stamatescu, Phys. Rev. D **81** (2010) 054508 [arXiv:0912.3360 [hep-lat]].

[4] G. Aarts, F. A. James, E. Seiler and I. -O. Stamatescu, Eur. Phys. J. C **71** (2011) 1756 [arXiv:1101.3270 [hep-lat]].

[5] A. Duncan and M. Niedermaier, Annals Phys. **329** (2013) 93 [arXiv:1205.0307 [quant-ph]].

[6] G. Aarts, P. Giudice and E. Seiler, Annals Phys. **337** (2013) 238 [arXiv:1306.3075 [hep-lat]].

[7] J. R. Klauder and W. P. Petersen, Journal of Stat. Phys. 39 (1985) 53, Print-85-0295 (BTL).

[8] J. Ambjorn and S. K. Yang, Phys. Lett. B **165** (1985) 140.

[9] H. Okamoto, K. Okano, L. Schulke and S. Tanaka, Nucl. Phys. B **324** (1989) 684.

[10] M. Cristoforetti *et al.* [AuroraScience Collaboration], Phys. Rev. D **86** (2012) 074506 [arXiv:1205.3996 [hep-lat]].

[11] A. Mukherjee, M. Cristoforetti and L. Scorzato, arXiv:1308.0233 [physics.comp-ph].

[12] G. Aarts, arXiv:1308.4811 [hep-lat].

[13] D. Sexty, arXiv:1307.7748 [hep-lat].

[14] E. Seiler, D. Sexty and I. -O. Stamatescu, Phys. Lett. B **723** (2013) 213 [arXiv:1211.3709 [hep-lat]].