

Complex Langevin and boundary terms

Manuel Scherzer^{a*}, Erhard Seiler^{b†}, Dénes Sexty^{c,d‡}, and
Ion-Olimpiu Stamatescu^{a§}

^a*Institut für Theoretische Physik, Universität Heidelberg
Heidelberg, Germany*

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

^c*Institut für Theoretische Physik, Bergische Universität Wuppertal
Wuppertal, Germany*

^d*IAS/JSC, Forschungszentrum Jülich, D-52425 Jülich, Germany*

Abstract

It is well known that the Complex Langevin (CL) method sometimes fails to converge or converges to the wrong limit. We identified one reason for this long ago: insufficient decay of the probability density either near infinity or near poles of the drift, leading to boundary terms that spoil the formal argument for correctness. To gain a deeper understanding of this phenomenon, we analyze the emergence of such boundary terms thoroughly in a simple model, allowing some analytic results for comparing with numerics. We also show how some simple modification stabilizes the CL process in such a way that it produces results agreeing with direct integration.

*email: scherzer@thphys.uni-heidelberg.de

†email: ehs@mpp.mpg.de

‡email: sext@uni-wuppertal.de

§email: I.O.Stamatescu@thphys.uni-heidelberg.de

1 Introduction

It has been known for a long time that the Complex Langevin (CL) method for simulating systems with complex action may fail by either not converging or by converging to the wrong limit. These failures were traced either to insufficient decay of the probability distribution in the complexified configuration space [1, 2] (at infinity or at poles of the drift force), or to failure of ergodicity [3, 4]. Recently Salcedo [5] has formulated interesting criteria for failure that at first sight seem to be unrelated to the ones identified by us. The most interesting ones derive support properties of the equilibrium measure which are shown to be in conflict with the correct expectation values.

In this note we will focus on one such example and show explicitly that the problems are due to slow decay, leading to the appearance of boundary terms in an integration by parts, spoiling the formal proof of correctness.

Here we consider a complex density

$$\rho(x) = \exp(-S(x)) , \quad (1)$$

periodic with period 2π and extending to an entire analytic function without zeroes.

The complex Langevin equation (CLE) in the form used here is

$$\begin{aligned} dx &= K_x dt + dw, \\ dy &= K_y dt, \end{aligned} \quad (2)$$

where dw is the Wiener process normalized as

$$\langle dw^2 \rangle = 2dt \quad (3)$$

and the drift is given by

$$\begin{aligned} K_x &= -\operatorname{Re} S'(x + iy), \\ K_y &= -\operatorname{Im} S'(x + iy). \end{aligned} \quad (4)$$

The long time asymptotic average of a generic observable \mathcal{O} is denoted by $\langle \mathcal{O} \rangle_\infty$; we say that the CL process yields correct results if this agrees with the ‘correct’ expectation value of the same observable defined as

$$\langle \mathcal{O} \rangle_c = \int dx \mathcal{O}(x) \rho(x), \quad (5)$$

i. e.

$$\langle \mathcal{O} \rangle_\infty = \langle \mathcal{O} \rangle_c. \quad (6)$$

In [1, 2] correctness was derived from the consideration of CL expectation values at finite Langevin time; it was shown that correctness is assured if a certain quantity $F_{\mathcal{O}}(t, \tau)$ is independent of an interpolation parameter $\tau \in [0, t]$, i. e.

$$\frac{\partial}{\partial \tau} F_{\mathcal{O}}(t, \tau) = 0. \quad (7)$$

Here $F_{\mathcal{O}}(t, \tau)$ interpolates between the ‘correct’ time evolution $F_{\mathcal{O}}(t, t)$ (defined in Section 3 such that $\lim_{t \rightarrow \infty} F_{\mathcal{O}}(t, t) = \langle \mathcal{O} \rangle_c$) and the time evolution of the expectation of \mathcal{O} under the Langevin process $F_{\mathcal{O}}(t, 0) = \langle \mathcal{O} \rangle_t$.

The vanishing of this quantity can be traced back to the vanishing of a boundary term arising under integration by parts; explicitly this term is

$$\lim_{Y \rightarrow \infty} B_{\mathcal{O}}(Y; t, \tau) \quad (8)$$

where

$$\begin{aligned} B_{\mathcal{O}}(Y; t, \tau) \equiv & \int [K_y(x, Y)P(x, Y; t - \tau)\mathcal{O}(x + iY; \tau) \\ & - K_y(x, -Y)P(x, -Y; t - \tau)\mathcal{O}(x - iY; \tau)]dx; \end{aligned} \quad (9)$$

$P(x, y; t)$ gives the time evolved probability density under the Langevin evolution and $\mathcal{O}(z; t)$ gives the ‘correctly evolved’ observable (see appendix A).

This form of the boundary term makes clear that correctness requires sufficient decay of the product $K_y P \mathcal{O}$.

2 The model

The model studied here is defined by the complex density

$$\rho = \frac{1}{Z(\beta)} \exp[-i\beta \cos(x)] \quad (10)$$

and has been studied already in 2007 by Stamatescu [6] and in 2008 by Berges and Sexty [7]. The ‘correct’ expectation values of exponentials (“modes”) are

$$\int dx \exp(ikx) \rho(x) = \frac{I_k(-i\beta)}{I_0(-i\beta)} = (-i)^k \frac{J_k(\beta)}{J_0(\beta)} \neq 0. \quad (11)$$

It was found in [6, 7] that the CL process does not reproduce the correct EV’s which, however, can be regained by a certain reweighting procedure (with different observables requiring sometimes different reweightings).

The remarkable fact found by Salcedo [5] is that the static probability distribution $P(x, y) = P(x, y; \infty)$ for this model can be written down explicitly by solving the time independent Fokker-Planck equation (FPE); it is

$$P(x, y) = \frac{1}{4\pi \cosh^2(y)}. \quad (12)$$

It is the only non-Gaussian example known to us for which a solution of the static FPE is known in analytic form. Three features of this solution are remarkable:

- (1) P is independent of x ,
- (2) P is independent of β ,
- (3) P decays as $\exp(-2|y|)$ for large $|y|$; this decay is not sufficient to make the integrals of the modes

$$\exp(ik(x + iy)), \quad |k| \geq 2 \quad (13)$$

absolutely convergent, in other words, already here we are faced with slow decay.

2.1 Complex Langevin results

But first let us demonstrate that (12) is indeed the distribution produced by running a CL simulation for a long time. The drift force is

$$K_x = \operatorname{Re} \frac{\rho'}{\rho} = -\beta \cos(x) \sinh(y) \quad K_y = \operatorname{Im} \frac{\rho'}{\rho} = \beta \sin(x) \cosh(y), \quad (14)$$

for the Langevin process Eq. (2).

In Fig. 1 we show the histogram of the converged marginal distribution $P_y(y; t) = \int dx P(x, y; t)$ in log scale for $\beta = 1$, overlaid with the distribution (12). The histogram is obtained from one long trajectory (Langevin time $t \approx 125000$). The agreement over about 6 orders of magnitude is convincing. The distribution can also be seen to be independent of x , cf. also Fig. 8.

We also show in Fig. 2 the histograms of $P_y(y; t)$ for various shorter times and $\beta = 0.1$, illustrating the convergence as $t \rightarrow \infty$.

As noted by Salcedo [5], it is obvious that the distribution P (12) cannot reproduce the correct expectation values of the observables $\mathcal{O}_k = \exp(ikx)$, because it is independent of x , entailing

$$\int dx P(x, y) e^{\pm ikx} = 0, \quad (15)$$

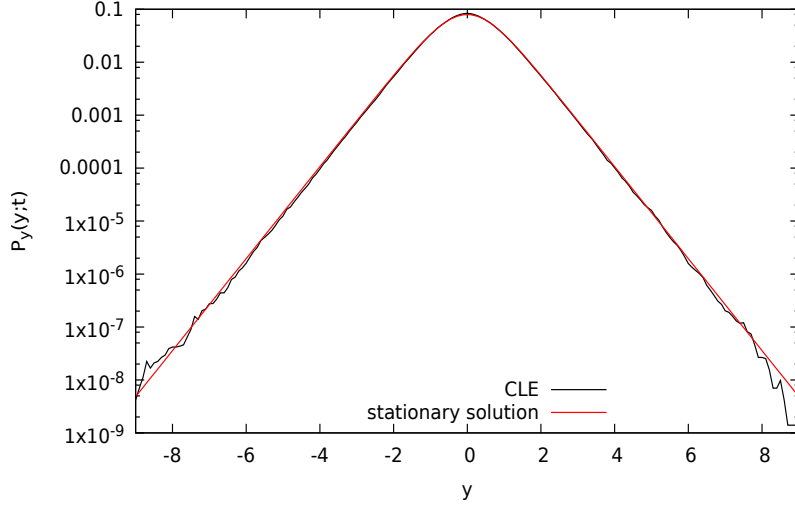


Figure 1: Comparison of the analytic expression for the marginal distribution $P_y(y; t)$ (12) (red) with the histogram of a CL simulation with $\beta = 1$.

and its slow decay makes the expectation values of \mathcal{O}_k ill defined for $|k| \geq 2$. In Table 2.1 we collect a few CL results, together with the exact expectation values determined by ρ for $\beta = 1$. The simulation used 100 independent trajectories with randomly chosen starting points, running for a Langevin time of $t = 1250$. The CL values for $|k| > 2$ are completely submerged by noise, as expected. For $k = \pm 2$, even though the results are probably not sufficiently well converged, a value near 1 is suggested.

The Schwinger-Dyson equation (SDE)

$$ik\langle e^{ikz} \rangle + \frac{\beta}{2}\langle e^{i(k+1)z} \rangle - \frac{\beta}{2}\langle e^{i(k-1)z} \rangle = 0, \quad (16)$$

arising from the identity

$$\int_{-\pi}^{\pi} \rho(x) \mathcal{O}'(x) dx = - \int_{-\pi}^{\pi} \rho'(x) \mathcal{O}(x) dx \quad (17)$$

would be satisfied for $k = 0, \pm 1$ if the modes ± 1 are 0 and the modes ± 2 are 1, even though these values are not the ‘correct’ ones.

So the CL results, where they are defined, are incorrect, but mostly – for $|k| \geq 3$, they are completely undefined due to uncontrollable fluctuations.

The last row in the table gives the correct results, as will be explained in the next subsection.

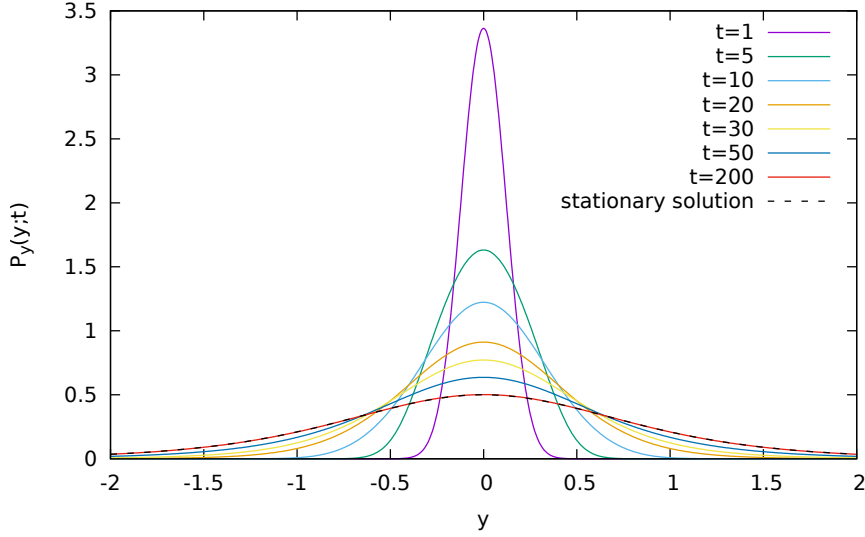


Figure 2: The marginal distributions $P_y(y; t)$ obtained by numerically solving the FPE for $\beta = 0.1$. The ordering of times corresponds to decreasing maxima; note that for $t = 200$ no difference is visible between the FPE and the analytical solution.

2.2 A puzzle

The remaining question is: how can CL fail for the first mode, i. e. observables $\mathcal{O}_{\pm 1} \equiv \exp(\pm i(x + iy))$? $\mathcal{O}_{\pm 1}P$ as well as KP decay exponentially in y .

Actually the densities of $\mathcal{O}_{\pm 1}P$ and K , if considered not as functions of y , but as functions of their actual value decay only power-like (see Eq. 37). Nagata et al [8] formulated a criterion stating that exponential decay of the distribution of K is necessary to guarantee correctness; so by this criterion correctness is not to be expected here. We will formulate, however, a more precise criterion in Section 3.

The CL simulation produces well converged, yet incorrect results, close to 0 (the result from (12), inconsistent with (11)).

The resolution lies in the nonvanishing boundary terms arising in the time dependent expectation values and persisting for arbitrarily large times; this is the mechanism described in [1, 2]. In the following section we will analyze those boundary terms in detail.

$\langle \mathcal{O} \rangle$	$\langle e^{ix} \rangle$	$\langle e^{-ix} \rangle$	$\langle e^{2ix} \rangle$	$\langle e^{-2ix} \rangle$
CL	-0.0022(16)	0.0019(19)	0.929(7)	1.205(9)
correct	0.575081 <i>i</i>	0.575081 <i>i</i>	-0.150162	-0.150162
$\exp(tL_c)\mathcal{O}$	0.575081 <i>i</i>	0.575081 <i>i</i>	-0.150162	-0.150162

Table 1: CLE and ‘correct’ results for model (10) with $\beta = 1$. The last line gives the ‘correct evolution’ for $t = 20$ (see Appendix B).

3 Boundary terms for finite Langevin time

The formal argument for correctness [1, 2] is revisited in Appendix A. It requires the choice of an initial distribution $P(x, y; 0)$; in the following we will choose for simplicity

$$P(x, y; 0) = \frac{1}{2\pi} \delta(y). \quad (18)$$

The identity (7) follows by integrating by parts, assuming that there are no boundary terms, and using the Cauchy-Riemann equations.

In order to check for the appearance of boundary terms as in (9), we need the time evolutions of both the observables (the ‘correct evolution’, see Appendix B) and of the probability density P by solving the FPE with the initial condition (18).

3.1 Indirect evidence for boundary terms

In [2] we found numerically for a somewhat different model that the correctly evolved observables $\mathcal{O}(x + iy; t)$ grow in the y direction as an iterated exponential. The same can be seen here, but we will not go into this. This growth makes the appearance of boundary terms already plausible.

In the following we show explicitly that Eq. (7) is numerically satisfied for short times (up to $t \approx 20$), choosing $\beta = 0.1$.

The ‘correct evolution’ or L_c evolution is defined by the differential equation

$$\partial_t \mathcal{O}_k(z; t) = L_c \mathcal{O}_k(z; t) \quad (t \geq 0), \quad \mathcal{O}_k(z; 0) = \exp(ikz); \quad (19)$$

with

$$L_c = [\partial_z - S'(z)] \partial_z. \quad (20)$$

We compare

$$\frac{1}{2\pi} \int dx \mathcal{O}(x; t) \rho_0(x) = F_{\mathcal{O}}(t, t) \equiv \langle \mathcal{O}(t) \rangle_c \quad (21)$$

with

$$\int dx dy \mathcal{O}(x + iy; 0) P(x, y; t) = F_{\mathcal{O}}(t, 0) \equiv \langle \mathcal{O} \rangle_{P(t)}. \quad (22)$$

Here $P(x, y; t)$ is the solution of the *real* Fokker-Planck equation (FPE)

$$\frac{\partial}{\partial t} P(x, y; t) = L^T P(x, y; t), \quad (23)$$

with

$$L^T = \partial_x [\partial_x - K_x] - \partial_y K_y, \quad (24)$$

and initial condition (18), which describes the time evolution of the probability density under the CL process. $\langle \mathcal{O} \rangle_{P(t)}$ is the t -dependent expectation value obtained in the CL process.

For our model the FPE is (23) with the drift force (14). Eq.(23) is solved numerically as well; some details are found in Appendix C.

Fig. 3 compares (21) and (22) for the Fourier modes

$$\mathcal{O}_k(z) = \exp(ikz). \quad (25)$$

for $k = 1, 2, 3$, Langevin times t between 0 and 50 and $\beta = 0.1$.

It is seen that the left hand side (21) reaches its asymptotic value already for quite short Langevin times (around $t \approx 7$). This is in accordance with the value of the smallest nonzero eigenvalue $\lambda_1 \approx -1$ of L_c (cf. Eq.(58)). For this value of β also the right hand side does the same; for $t \lesssim 20$ there is no difference visible between the left and the right hand sides (dashed and solid curve). This indicates that any boundary terms are negligible there.

So there is a “plateau” corresponding to the correct value in the solid curve, and the boundary term starts picking up around $t = 20$.

In Fig. 4 we show (in black) the evolution of the first mode up to time $t = 200$. It is seen that after the plateau it converges to zero, the value corresponding to the stationary solution (12) of the FPE. We will return to this figure in Section 5.

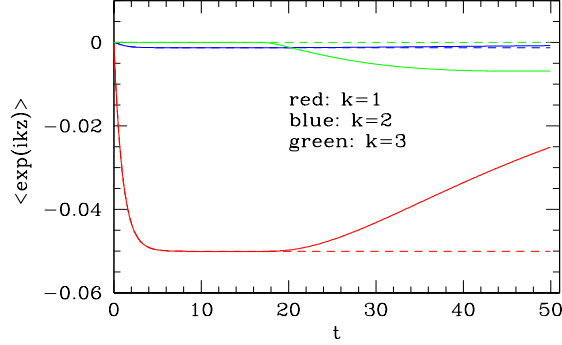


Figure 3: Comparison of expectation values using the FPE evolution of P (solid lines) Eq. (22) with the L_c evolution of the observables (dashed lines – Eq. (21)) for $\beta = 0.1$. Note that for times up to about 20 the dashed and solid lines are practically indistinguishable.

3.2 Direct study of the boundary terms

We next study explicitly the evolution of the boundary term Eq. (9) for the mode $k = 1$ and $\tau = 0$ with Langevin time t :

$$\begin{aligned} B_1(Y; t, 0) &= \frac{\partial}{\partial \tau} F_1(t, \tau)|_{\tau=0} \\ &= \beta \int_{-\pi}^{\pi} dx \sin(x) \cosh(Y) e^{ix} [P(x, Y; t) e^{-Y} - P(x, -Y; t) e^Y] . \end{aligned} \quad (26)$$

We first note that we can take the limit $t \rightarrow \infty$ of this expression, using the fact that $P(x, y; t)$ indeed converges to Eq. (12), which was verified before. We obtain

$$B_1(Y; \infty, 0) = -2\beta \int_{-\pi}^{\pi} dx \sin(x) \cosh(Y) e^{ix} \sinh(Y) \frac{1}{4\pi \cosh^2(Y)} . \quad (27)$$

or

$$B_1(Y; \infty, 0) = -\frac{i\beta}{2} \tanh(Y) , \quad (28)$$

(converging to $-i\beta/2$ for $Y \rightarrow \infty$).

In Fig. 5 we compare the numerical results for Langevin times up to $t = 200$ with the asymptotic value at $t = \infty$ for $\beta = 0.1$. Y was chosen to be 5 which is close to the asymptotic value $Y = \infty$ ($\tanh(5.) = 0.99991$). We see here

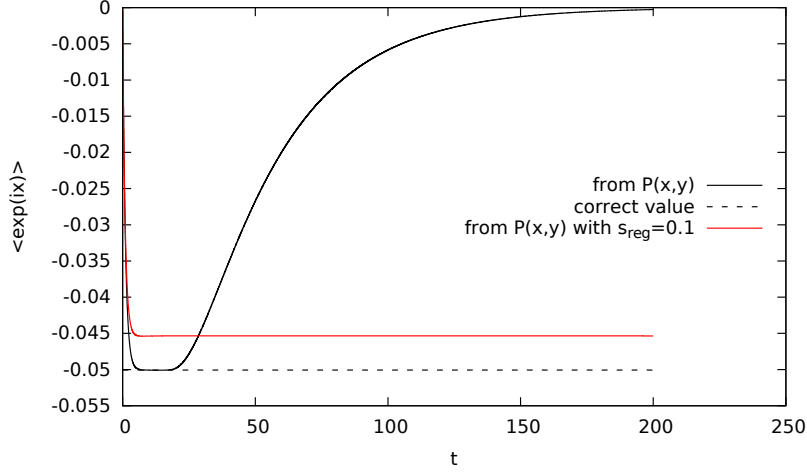


Figure 4: FPE evolution of O_1 at $\beta = 0.1$ (black). For comparison we show the evolution with regularization $K_{R,y} = -s y$, $s = 0.1$ (red), see Section 4.

directly that the boundary term stays very small up to $t \lesssim 20$, then picks up and approaches the analytically determined value $-i\beta/2$. For the value $\beta = 0.1$ it also follows closely the difference between the first mode shown in Fig. 4 and the correct value, but this cannot remain true for larger β .

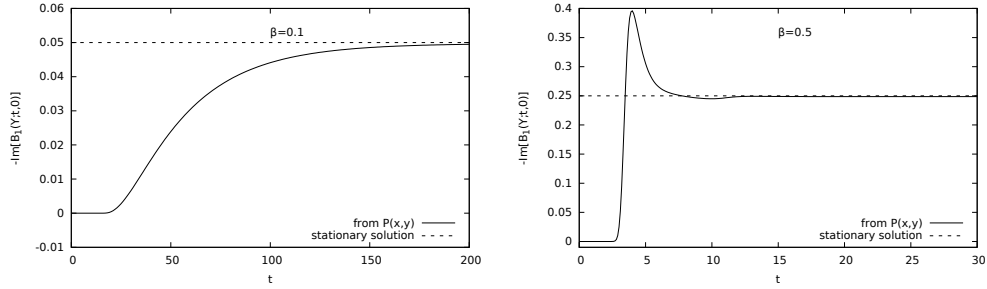


Figure 5: Numerical evolution via FPE of the imaginary part of the boundary term Eq. (26) for $\beta = 0.1$ (left panel) and $\beta = 0.5$ (right panel), $k = 1$ and $Y = 5$.

In Fig. 6 we also show the boundary term for different values of the cutoff Y , showing the fast approach to the asymptotic value. Note that in the right panel we show the boundary term as measured using the CLE alone, without making use of the Fokker-Planck evolution, which would be prohibitively costly in a lattice model.

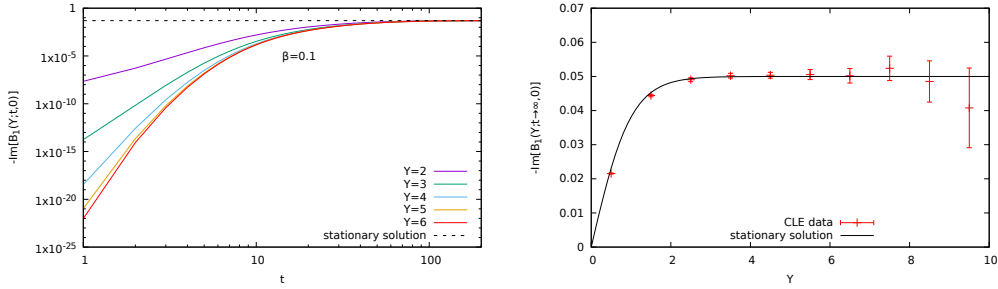


Figure 6: Left panel: numerical evolution via FPE of the imaginary part of the boundary term Eq. (26) for $\beta = 0.1$ and different values of Y vs. t . Right panel: the same boundary term evaluated via Langevin simulation at asymptotic t vs. Y .

So we established implicitly and explicitly that boundary terms appearing appreciably only after some Langevin time are responsible for invalidating the formal correctness argument.

It can also be seen by inspection of Eq. (26) that the presence of the observable \mathcal{O}_1 is essential; the distribution of the drift force alone goes to zero. Quite generally it is the product of observable, drift and probability P that decides about the presence or absence of boundary terms; this indicates some possible limitations of the criterion for correctness advocated in [8].

3.3 Boundary terms and skirts

Thinking now of Y not as a cutoff, but as a variable, and denoting it by y again, we see that the first term of the boundary term $B_1(y; \infty, 0)$, considered as a function of y Eq. (26) is just the probability density of the observable

$$v(y) \equiv \text{Im} \int dx K_y(x, y) \mathcal{O}_1(x - iy) \sim e^{2y} \quad (29)$$

for large y . The nonvanishing of the boundary term is the fact that

$$\lim_{y \rightarrow \infty} v(y) P_y(y) \neq 0. \quad (30)$$

On the other hand the distributions of v itself has a density $p(v)$, related to P by

$$p(v) = P_y(y(v)) \frac{dy}{dv}. \quad (31)$$

This can easily be worked out, but the point is that for large y

$$p(v) \sim v^{-2}, \quad (32)$$

which shows that there is no finite expectation value of v since $v p(v)$ is not integrable.

In other words: *a ‘skirt’ in the distribution of $K_y \mathcal{O}$ falling off like the power -2 or more slowly corresponds to a nonvanishing (possibly diverging) boundary term.* Note, however, such a simple reasoning is only possible because here P is independent of x .

3.4 The interpolating function

So far we have only compared $F_k(t, 0)$ and $F_k(t, t)$. But it is instructive also to look at the interpolating function $F_k(t, \tau)$

$$F_{\mathcal{O}}(t, \tau) \equiv \int P(x, y; t - \tau) \mathcal{O}(x + iy; \tau) dx dy, \quad (33)$$

which should be independent of τ for the correctness argument to hold. This is shown in Fig. 7 for $k = 1$ for $\beta = 0.1$ and $\beta = 0.5$.

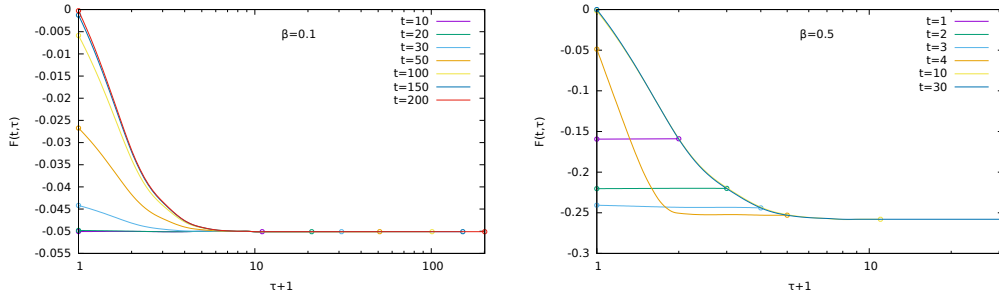


Figure 7: The interpolating function $F_1(t, \tau)$ defined in (33) for the first mode; $\beta = 0.1$ (left panel) and $\beta = 0.5$ (right panel) for various values of t ; the small circles denote the beginning and end of the respective curves.

Again it is seen that for $\beta = 0.1$ $t \lesssim 20$ the curves are flat, indicating the absence of any appreciable boundary terms. For $t > 20$ a τ dependence develops, being maximal at $\tau = 0$. This is understandable from what we have seen: the FPE evolution of P proceeds up to time $t - \tau$, which allows for the boundary terms to arise. On the other hand, for $\tau \gtrsim 7$ $\mathcal{O}_k(z; \tau)$

has practically reached its asymptotic limit (cf. Appendix B), in which only the constant mode survives; this constant can be pulled outside the integral defining F , so that for $t, \tau > 7$

$$\mathcal{O}_k(z; \tau) \approx \frac{1}{2\pi} \int_{-\pi}^{\pi} dx' \rho(x') \mathcal{O}_k(x') = \langle \mathcal{O}_k \rangle_c \quad (34)$$

and

$$F_k(t, \tau) \approx \frac{1}{2\pi} \int dx dy P(x, y; t) \langle \mathcal{O}_k \rangle_c = \langle \mathcal{O}_k \rangle_c. \quad (35)$$

i.e. the correct value (where we used the fact that the density P is always normalized).

Notice that the slope of $F_k(t, \tau)$ appears maximal at $\tau = 0$ for large t . Therefore the estimation of noxious boundary terms as defined in (26) is relevant for judging the asymptotic correctness of the CL procedure - cf. Figs. 5, 7.

Plots similar to Fig. 7 appeared in [2] for a different model.

3.5 Evolution of some marginal distributions

For $\beta = 0.1$ we saw clearly the evolution first apparently converging to the correct value and then departing from it (the ‘plateau’ in Fig. 3). Similar behavior in Langevin time was observed in a real time $SU(2)$ lattice simulation [9]. This is also reflected in some marginal distributions.

In Fig. 8 we show the evolution of $P_x(x; t) = \int dy P(x, y; t)$ for $\beta = 0.1$. It starts out flat, corresponding to our choice of initial condition; at $t = 10$ and $t = 20$ it shows maximal structure, while for larger t it approaches a flat distribution again, in agreement with (12).

The distribution of the first mode also show a similar behavior. Of interest is the imaginary part. Its density is

$$\sigma(u; t) \equiv \int dx dy P(x, y; t) \delta(\sin(x) e^{-y} - u). \quad (36)$$

We present in Fig. 9 histograms for $\sigma(u; t)$, obtained from the numerical solution of the FPE; for the limiting distribution $P(x, y; \infty) = 1/(4\pi \cosh^2(y))$ we can evaluate (36) analytically:

$$\sigma(u; \infty) = \frac{|u|}{\pi} \int_{-1}^1 dt \frac{t^2}{(t^2 + u^2)^2 \sqrt{1 - t^2}} = \frac{1}{2(1 + u^2)^{3/2}}. \quad (37)$$

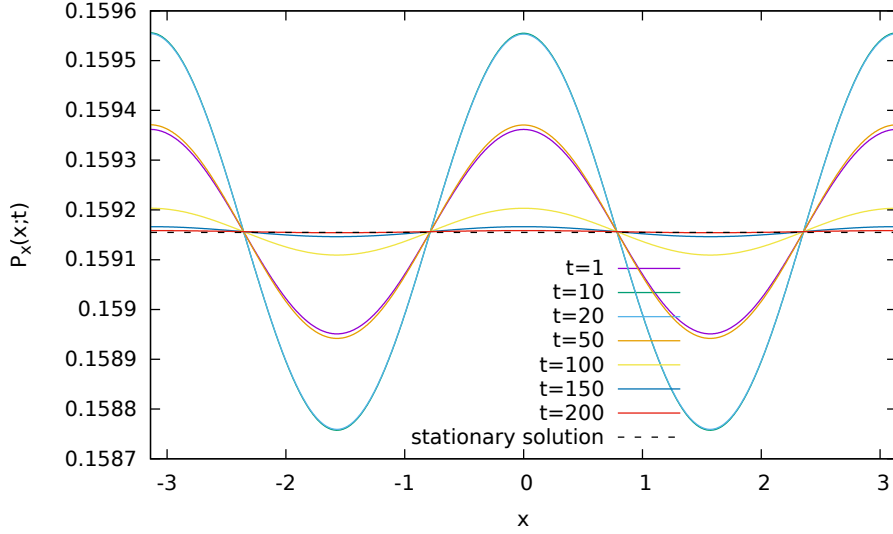


Figure 8: The marginal distribution $P_x(x;t)$ obtained from solving the Fokker-Planck equation for $\beta = 0.1$.

Fig. 9 shows first the development of an asymmetric structure with two maxima, whereas for larger t one sees clearly the approach to the symmetric analytic result (37).

4 Illustration of the effect of the boundary terms in a regularized model

In the preceding section we described how the boundary terms accumulate in the Langevin (and Fokker-Planck) evolution, spoiling the proof of convergence such that the process would lead to wrong results.

Here we want to explicitly see the effect of those terms by considering a “regularization” of the model using a damping term in the action, $S_R = \frac{s}{2} x^2$, which leads to a modification of the drift by $K_R(z) = -s z$. The philosophy of this regularization is very similar to that of dynamical stabilization [10]. In both cases, and different from modifications using symmetries, such as in the gauge cooling paradigm, the dynamics is really changed, but in a way intended to be controllable.

For $s = 0$ we regain the original model Eq. (10) (including its problems)

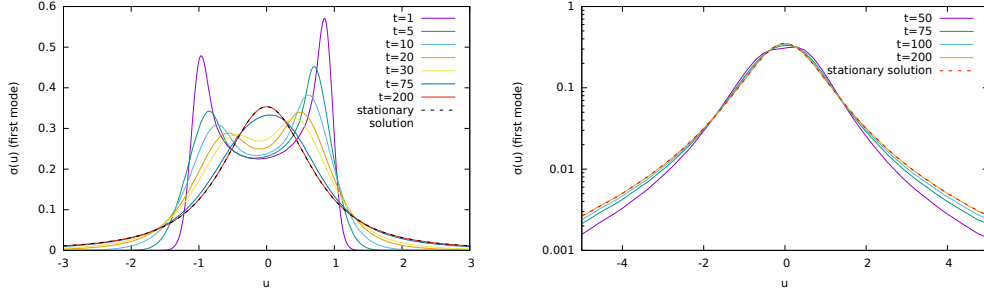


Figure 9: Evolution of the distribution of the first mode $\sigma(u)$. Left panel: times from $t = 1$ to 200 in linear scale; right panel: times from $t = 50$ to 200 in log scale. Again note that for $t = 200$ no difference is visible between the numerical results and the analytic expression.

while for $s > 0$ we should observe an interplay between the original tendency to build boundary terms and their damping in the modified model, allowing us to estimate the effect of these terms. This particular modification leads to loss of periodicity in x which becomes noncompact at $s > 0$. The CLE process was allowed to drift unbounded in the full z plane and the exact integral was correspondingly done in the infinite interval. The following plots show $O_1 = \text{Im}\langle e^{iz} \rangle$. We see from Fig. 10 that the regularization stabilises

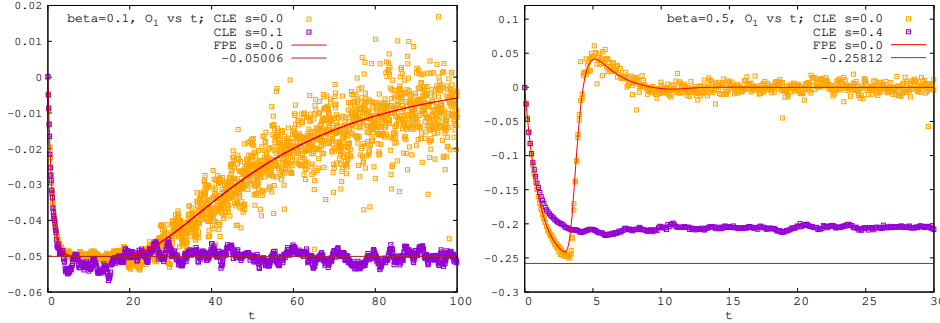


Figure 10: Comparison of the O_1 expectation values from FPE (solid line) and CLE and from regularized CLE, vs t . Left: $\beta = 0.1$, $s = 0$ and 0.1 , respectively. Right: $\beta = 0.5$, $s = 0$ and 0.4 , respectively.

the expectation values in the CLE evolution. When the non-regularized data show a plateau at the correct value for intermediary t the regularization extends this plateau into the asymptotics ($\beta = 0.1$ case). When a plateau is missing the regularization still stabilises the EV but at a value shifted from

the correct one ($\beta = 0.5$ case), since now a larger s is needed to counteract the boundary terms.

Note that an alternative regularization of the process itself is to modify *only the imaginary drift* by a damping term $K_{R,y} = -s y$. This leads to similar results (see Fig. 4, here from the FPE evolution), and has the advantage that periodicity in x is preserved. We preferred the action variant, however, also since it allows us to obtain exact correct results for the regularized model by simple numerical integration.

In Fig. 11 we show the s dependence in CLE for the regularized model for the same values of β . The plots suggest an extrapolation toward the exact expectation value (EV) for $s \rightarrow 0$, however this might not be simply linear, but depend on the particular regularization, β , etc. Therefore we mean this discussion not yet as a genuine cure but only as illustration of the effects of the boundary terms on the EV's. For $k = 1$, e.g., these effects can be estimated from the distance between the CLE regularized data and the exact values from the numerical integration of the regularized model: As can be seen from the figure at small s these effects are still present while gradually vanishing with increasing s .

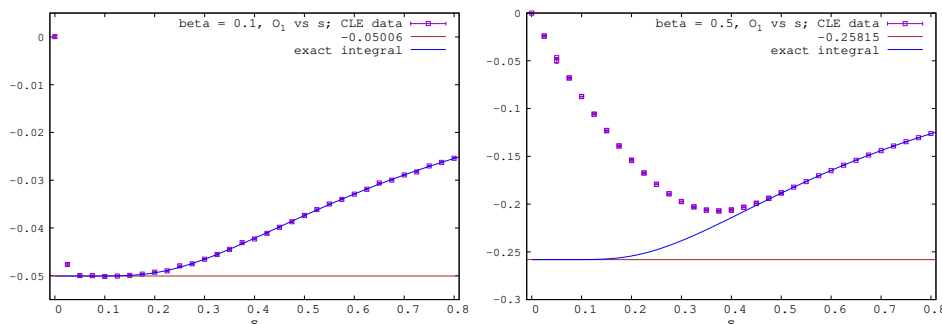


Figure 11: Dependence of O_1 on the regularization parameter s of the CLE simulation at $\beta = 0.1$ (left) and 0.5 (right). The solid lines show the exact correct values from numerical integration.

5 Conclusions

We have in great detail analyzed a simple example in which the CL fails, establishing very explicitly that the failure is due to boundary terms spoiling the correctness argument, as argued already long ago [1, 2]. The absence of

such boundary terms requires that the product of observable, drift force and probability distribution (\mathcal{OKP}) goes to zero in the noncompact (imaginary) directions. The relation between boundary terms and ‘skirts’, i. e. decay of distributions was addressed in Subsection 3.3, making clear that possible skirts in the distribution of the product \mathcal{OKP} and not just KP are relevant, so the general validity of the criterion proposed by [8] and not involving the observables is not obvious.

Generally the τ -dependent boundary term (52) cannot be estimated in a realistic (lattice) calculation. Fortunately, however, the considerations in this paper suggest that relevant for the correctness of the asymptotic (large t) EV’s is the boundary term at $\tau = 0$, $B_k(t, 0)$, as defined in (26). This term appears to maximize $B_k(t, \tau)$ and it stabilizes at large t ; it is accessible in principle to online monitoring using the CLE alone, and might provide a correctness criterion for the EV’s obtained in the CL simulation.

Acknowledgments: M. S., E. S. and I.-O. S. gratefully acknowledge kind support from DFG under Grant Sta 283/16-2. D. S. gratefully acknowledges funding by the DFG grant Heisenberg Programme (SE 2466/1-2). The authors acknowledge support by the High Performance and Cloud Computing Group at the Zentrum für Datenverarbeitung of the University of Tübingen, the state of Baden-Württemberg through bwHPC and the German Research Foundation (DFG) through grant no INST 37/935-1 FUGG.

A The argument for correctness revisited

We briefly revisit the formal proof of correctness for the CL method for our simple periodic one-dimensional models, spelling out the conditions needed for it to work as well as the mechanisms that may lead either to no convergence or ‘wrong convergence’ of the CL process (cf. [2, 3, 4]).

$P(x, y; t)$ is the time dependent probability distribution corresponding to the CL process, determined by the *real* Fokker-Planck equation (23).

We also consider the time evolution of the complex density $\rho(x; t)$ by the *complex* Fokker-Planck equation

$$\frac{\partial}{\partial t} \rho(x; t) = L_c^T \rho(x; t), \quad (38)$$

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

$$L_c^T = \partial_x [\partial_x + S'(x)] . \quad (39)$$

The initial conditions for (23) and (38) are required to be consistent, i.e.

$$P(x, y; 0) = \rho(x; 0)\delta(y) \geq 0 \quad \int \rho(x; 0)dx = 1. \quad (40)$$

The crucial point is that one can now, under conditions to be spelled out below, show that

$$\int dx \mathcal{O}(x) \rho(x; t) = \int dx dy \mathcal{O}(x + iy) P(x, y; t). \quad (41)$$

If in addition the operator L_c^T has spectrum in the left half plane with 0 a nondegenerate eigenvalue, it follows that

$$\lim_{t \rightarrow \infty} \int dx \mathcal{O}(x) \rho(x; t) = \int dx \mathcal{O}(x) \rho(x) \quad (42)$$

and by (41)

$$\lim_{t \rightarrow \infty} \int dx dy \mathcal{O}(x + iy) P(x, y; t) = \int dx \mathcal{O}(x) \rho(x). \quad (43)$$

By our choice of initial conditions, (41) holds for $t = 0$. For $t > 0$ we consider $F_{\mathcal{O}}(t, \tau)$ defined in Eq. (33), which interpolates between the two sides of (41):

$$F_{\mathcal{O}}(t, \tau) \equiv \int P(x, y; t - \tau) \mathcal{O}(x + iy; \tau) dx dy, \quad (44)$$

with $\mathcal{O}(x + iy; t)$ defined by solving the differential equation (19) L_c , the complex Langevin operator, is the transpose of L_c^T :

$$L_c = [\partial_z - S'(z)] \partial_z. \quad (45)$$

We call the solution of Eq. (19) the “ L_c evolved” or “correctly evolved” observable.

The interpolating property follows from

$$\begin{aligned} F_{\mathcal{O}}(t, 0) &= \int dx dy \mathcal{O}(x + iy) P(x, y; t) = \langle \mathcal{O} \rangle_{P(t)}, \\ F_{\mathcal{O}}(t, t) &= \int dx \mathcal{O}(x) \rho(x; t) = \langle \mathcal{O} \rangle_c, \end{aligned} \quad (46)$$

see Section 3.1, Eqs. (21) and (22). The first equality is obvious, the second one follows by integration by parts in x ; because of periodicity there are no

boundary terms. (41) would follow if we could prove

$$\begin{aligned} \frac{\partial}{\partial \tau} F_{\mathcal{O}}(t, \tau) = & - \int (L^T P(x, y; t - \tau)) \mathcal{O}(x + iy; \tau) dx dy \\ & + \int P(x, y; t - \tau) L_c \mathcal{O}(x + iy; \tau) dx dy = 0. \end{aligned} \quad (47)$$

This would again follow using integration by parts, provided there are no boundary terms. For the term ∂_x^2 of both L^T and L_c this is obvious because of periodicity, so we can drop these terms, obtaining

$$\begin{aligned} \frac{\partial}{\partial \tau} F_{\mathcal{O}}(t, \tau) = & \int ((\partial_x K_x + \partial_y K_y) P(x, y; t - \tau)) \mathcal{O}(x + iy; \tau) dx dy \\ & - \int P(x, y; t - \tau) S'(x + iy) \partial_x \mathcal{O}(x + iy; \tau) dx dy. \end{aligned} \quad (48)$$

In [1] we argued that $\mathcal{O}(x + iy; \tau)$ is holomorphic for any τ , i.e. it obeys the Cauchy Riemann equations

$$\partial_y \mathcal{O}(x + iy; \tau) = i \partial_x \mathcal{O}(x + iy; \tau). \quad (49)$$

This allows us to write the second term of the right hand side of (48) as

$$\int P(x, y; t - \tau) (K_x \partial_x + K_y \partial_y) \mathcal{O}(x + iy; \tau) dx dy. \quad (50)$$

Again the part involving ∂_x can be canceled against the corresponding term in the first term of (50) using integration by parts in x , so we only have to consider

$$\begin{aligned} \frac{\partial}{\partial \tau} F_{\mathcal{O}}(t, \tau) = & \int (\partial_y K_y P(x, y; t - \tau)) \mathcal{O}(x + iy; \tau) dx dy \\ & + \int P(x, y; t - \tau) K_y \partial_y \mathcal{O}(x + iy; \tau) dx dy. \end{aligned} \quad (51)$$

We have to interpret this as a the limit $Y \rightarrow \infty$ of the integral restricted to $|y| \leq Y$. For finite Y (51), using integration by parts, this is given by the boundary term

$$\begin{aligned} B_{\mathcal{O}}(Y; t, \tau) \equiv & \int [K_y(x, Y) P(x, Y; t - \tau) \mathcal{O}(x + iY; \tau) \\ & - K_y(x, -Y) P(x, -Y; t - \tau) \mathcal{O}(x - iY; \tau)] dx. \end{aligned} \quad (52)$$

This form of the boundary term makes clear that correctness requires sufficient decay of the products $K_y P \mathcal{O}$.

B The correct evolution

We rewrite the Langevin operator L_c in the basis of Fourier modes:

$$L_c \exp(ikx) = -k^2 \exp(ikx) - \frac{i\beta}{2} k \exp(i(k+1)x) + \frac{i\beta}{2} k \exp(i(k-1)x) \quad (53)$$

or equivalently, for a general observable

$$\mathcal{O}(x) = \sum a_k \exp(ikx) \quad (54)$$

$$(L_c a)_k = -k^2 a_k - \frac{i\beta}{2} (k-1) a_{k-1} + \frac{i\beta}{2} (k+1) a_{k+1}. \quad (55)$$

So L_c is represented on the Fourier coefficients by the sparse infinite matrix with elements

$$(L_c)_{kl} = -k^2 \delta_{kl} - \frac{i\beta}{2} (k-1) \delta_{k-1,l} + \frac{i\beta}{2} (k+1) \delta_{k+1,l}. \quad (56)$$

It is easy to compute numerically the action of $\exp(tL_c)$ on observables of the form $\mathcal{O}_k = \exp(ikx)$; cutting off the modes at $|k| \geq K$ with $K = 50$ and $K = 150$, and for $t = 50$, gave identical results, with only the constant mode surviving. Its value agrees to at least 5 digits with

$$\lim_{t \rightarrow \infty} \exp(tL_c) \mathcal{O}_k = \int dx \rho(x) \mathcal{O}_k(x) \quad (57)$$

i.e. the correct expectation value.

We also checked, using Mathematica, that the eigenvalues of the truncated matrix $(L_c)_{kl}$ have negative real part except for the unique zero eigenvalue corresponding to $a_k \propto \delta_{k0}$. All nonzero eigenvalues are real and doubly degenerate. The one with the smallest modulus determines the approach to the infinite time limit; it depends only weakly on β , e. g.

$$\lambda_1 = -0.998333 \ (\beta = 0.1); \quad \lambda_1 = -0.832189 \ (\beta = 1) \quad (58)$$

Remark: It is easy to show that by a similarity transformation L_c can be transformed into the *dissipative* operator

$$-H = \exp(S/2) L \exp(-S/2) = \frac{d^2}{dx^2} - \beta^2 \sin^2(x) - \frac{i\beta}{2} \cos(x). \quad (59)$$

Dissipativity means $-H - H^* \leq 0$, which is obvious. For such operators general theorems guarantee that the spectrum is contained in the left half of the complex plane (see for instance [11]). It is also not hard to see that there is exactly one vector with eigenvalue zero.

C Remarks on the numerical solution of the FPE

The real Fokker-Planck equation in our case is

$$\begin{aligned} \frac{\partial P(x, y; t)}{\partial t} &= [\partial_x (\partial_x - K_x) - \partial_y K_y] P(x, y; t) \\ &= [\partial_x^2 + \beta (-2 \sin x \sinh y + \cos x \sinh y \partial_x - \sin x \cosh y \partial_y)] P(x, y; t). \end{aligned} \quad (60)$$

Discretizing (60) in x and y using symmetric derivatives yields

$$\begin{aligned} P(x, y; t + dt) &= \frac{1}{dx^2} (P(x + dx, y; t) - 2P(x, y; t) + P(x - dx, y; t)) \\ &\quad - 2\beta \sin x \sinh y P(x, y; t) \\ &\quad + \frac{\beta}{2dx} \cos x \sinh y (P(x + dx, y; t) - P(x - dx, y; t)) \\ &\quad - \frac{\beta}{2dy} \sin x \cosh y (P(x, y + dy; t) - P(x, y - dy; t)). \end{aligned} \quad (61)$$

In case of a regularization term in the y -drift $K_y \rightarrow K_y - s_y y$ (see Section 4), additional terms occur

$$\begin{aligned} P(x, y; t + dt) &\rightarrow P(x, y; t + dt) + s_y P(x, y; t) \\ &\quad + s_y \frac{y}{2dy} (P(x, y + dy; t) - P(x, y - dy; t)). \end{aligned} \quad (62)$$

We solved the Fokker-Planck equation on an x - y -grid with parameters $dt = 10^{-6}$, $dx = 0.005 = dy$, a cutoff in y -direction of $Y = 5$ was found to be sufficient (compare (27), $\tanh(5) \approx \tanh(\infty)$), and a cutoff in x -direction of $X = 3.14$, which is due to the 2π periodicity of the problem. Boundary conditions in x and y were both chosen to be periodic. Initial condition were chosen according to (18), however the δ -function was smeared out slightly to avoid numerical issues; so we actually used

$$P(x, y; 0) = \frac{1}{2\pi\sqrt{2\pi\sigma_y^2}} e^{-\frac{y^2}{2\sigma_y^2}}, \quad (63)$$

where we chose $\sigma_y = 0.1$. Note that using this discretization it is hard to resolve the higher modes. This can be done more easily when solving the

Fokker-Planck equation in Fourier space, where it is given by

$$\begin{aligned}
P(k, y; t + dt) = & -k^2 P(k, y; t) \\
& - \frac{i\beta}{2} \sinh(y) ((k-1) P(k+1, y; t) + (k+1) P(k-1, y; t)) \\
& + \frac{i\beta}{4dy} \cosh(y) (P(k+1, y+dy; t) - P(k-1, y+dy; t)) \\
& + \frac{i\beta}{4dy} \cosh(y) (-P(k+1, y-dy; t) + P(k-1, y-dy; t)) .
\end{aligned} \tag{64}$$

Here we chose $dt = 0.5 \times 10^{-5}$, $k \in \{-19, \dots, 20\}$, $dy = \sqrt{dt}$, $Y \approx 2.8$ with antiperiodic boundary conditions in k for the imaginary part of $P(k, y; t)$ and periodic boundary conditions for the real part in k and for y . After $t \sim 30$ or so the result strongly depends on the choice of discretization. Hence we use the k - y discretization to resolve the plateaus in the higher modes and the x - y -discretization for everything else.

The solution to the Fokker-Planck equation shows that for $\beta = 0.1$ the evolution initially follows the CLE evolution. This is suggested by Fig. 7 and 5. By looking at $P(x, y; t)$ in Fig. 2 and 8 and the histogram of the first mode in Fig. 9, one can see that initially nontrivial structures occur. Those die out and everything approaches the asymptotic solution, which yields the wrong results. This strengthens the argument that until $t \sim 20$ or so CLE yields the correct solution but then the occurrence of boundary terms leads to wrong convergence.

References

- [1] G. Aarts, E. Seiler and I. O. Stamatescu, Phys. Rev. D **81** (2010) 054508 <https://arxiv.org/abs/0912.3360>.
- [2] G. Aarts, F. A. James, E. Seiler and I. O. Stamatescu, Eur. Phys. J. C **71** (2011) 1756 <https://arXiv.org/abs/1101.3270>.
- [3] G. Aarts, E. Seiler, D. Sexty and I. O. Stamatescu, JHEP **1705** (2017) 044 <https://arXiv.org/abs/1701.02322>.
- [4] E. Seiler, EPJ Web Conf. **175** (2018) 01019 <https://arXiv.org/abs/1708.08254>.
- [5] L. L. Salcedo, <https://arXiv.org/abs/1611.06390>.

- [6] I.-O. Stamatescu, unpublished notes (2007).
- [7] J. Berges and D. Sexty, Nucl. Phys. B **799** (2008) 306 <https://arXiv.org/abs/0708.0779>.
- [8] K. Nagata, J. Nishimura and S. Shimasaki, JHEP **1607** (2016) 073 <https://arXiv.org/abs/1604.07717>.
- [9] J. Berges, S. Borsanyi, D. Sexty and I.-O. Stamatescu, Phys. Rev. D **75** (2007) 045007 <https://arXiv.org/abs/hep-lat/0609058>.
- [10] F. Attanasio and B. Jäger, <https://arXiv.org/abs/1808.04400>.
- [11] E. B. Davies, Linear Operators and their Spectra, Cambridge University Press, Cambridge 2007.