

Complex Langevin simulations with a kernel

Michael Mandl,^{a,*} Erhard Seiler^b and Dénes Sexty^a

^a*Institute of Physics, NAWI Graz, University of Graz,
Universitätsplatz 5, 8010 Graz, Austria*

^b*Max-Planck-Institut für Physik (Werner-Heisenberg-Institut),
Boltzmannstraße 8, 85748 Garching bei München, Germany*

E-mail: michael.mandl@uni-graz.at, ehs@mpp.mpg.de, denes.sexty@uni-graz.at

We discuss recent developments regarding the use of kernels in complex Langevin simulations. In particular, we outline how a kernel can be used to solve the problem of wrong convergence in a simple toy model. Since conventional correctness criteria for complex Langevin results are only necessary but not sufficient, the correct convergence of complex Langevin simulations is not always straightforward to assess. Hence, we furthermore discuss a condition for correctness that we have recently derived, which is both necessary and sufficient. Finally, we outline a machine-learning approach for finding suitable kernels in lattice gauge theories and present preliminary results of its application to the heavy-dense limit of QCD.

*The 42nd International Symposium on Lattice Field Theory (LATTICE2025)
2-8 November 2025
Tata Institute of Fundamental Research, Mumbai, India*

*Speaker

1. Introduction

The complex Langevin method [1, 2] is a promising approach for solving the sign problem that plagues lattice simulations of various systems of physical interest, such as quantum chromodynamics (QCD) at non-zero baryon density or quantum field theories in real time. It is based on stochastic quantization [3], but requires the complexification of the underlying field space due to the complex actions involved. While the method has produced numerous useful results in various different theories, much is still to be learned about its drawbacks and, most importantly, how to overcome them. In this contribution, we address some of these points. In particular, we outline recent developments dealing with the use of a so-called kernel in the context of the complex Langevin approach that is used to ensure that the method produces correct results. Moreover, we also discuss a novel condition for assessing correctness in the first place.

2. The complex Langevin method

Consider a quantum field theory defined by an (in this work: Euclidean) action $S[\Phi]$ depending on some fields Φ . In the conventional approach to lattice field theory, one estimates expectation values of observables $O[\Phi]$ via approximating the path integral

$$\langle O \rangle = \frac{1}{Z} \int \mathcal{D}\Phi O[\Phi] e^{-S[\Phi]}, \quad Z = \int \mathcal{D}\Phi e^{-S[\Phi]} \quad (1)$$

by a sum over a finite number of field configurations distributed according to the path integral weight $e^{-S[\Phi]}$. This importance sampling method, however, breaks down when the action has a non-vanishing imaginary part, as is the case, for instance, in QCD with a non-zero baryon chemical potential, causing a sign problem.

The complex Langevin approach attempts to solve this sign problem via a stochastic evolution of complexified field degrees of freedom in an auxiliary time dimension τ , called the Langevin time. This evolution is governed by the complex Langevin equation, which for a theory of N real scalar fields Φ_i^R ($i = 1, \dots, N$) reads, after complexification $\Phi_i^R \rightarrow \Phi_i \equiv \Phi_i^R + i\Phi_i^I$,

$$\frac{\partial \Phi_i(x; \tau)}{\partial \tau} = -H_{ij} H_{jk}^T \frac{\delta S[\Phi]}{\delta \Phi_k(x; \tau)} + H_{ij} \eta_j(\tau). \quad (2)$$

Here, η_j is a real Gaussian noise vector satisfying $\langle \eta_i(\tau) \rangle = 0$ and $\langle \eta_i(\tau) \eta_j(\tau') \rangle = 2\delta_{ij} \delta(\tau - \tau')$ for all i and j . Moreover, we have introduced a so-called kernel H_{ij} . While, in general, this matrix-valued quantity could depend on the fields Φ_i , in which case there would be an additional term in (2), we only consider field-independent kernels here. Note that the summation over indices occurring twice in a product is assumed throughout.

The solution of (2) gives rise to a probability density $P(\Phi^R, \Phi^I; \tau)$ in complex Φ -space, whose τ -dependence is given by an associated Fokker–Planck equation. Assuming that its equilibrium limit

$$P(\Phi^R, \Phi^I) \equiv \lim_{\tau \rightarrow \infty} P(\Phi^R, \Phi^I; \tau) \quad (3)$$

exists, the complex Langevin approach provides estimators for expectation values of observables via the integration of the real density $P(\Phi^R, \Phi^I)$ over the complex domain as

$$\langle O \rangle_{\text{CL}} \equiv \int \mathcal{D}\Phi^R \mathcal{D}\Phi^I O[\Phi^R + i\Phi^I] P(\Phi^R, \Phi^I). \quad (4)$$

Thus, the complex Langevin evolution solves the sign problem if and only if

$$\langle O \rangle_{\text{CL}} = \langle O \rangle, \quad (5)$$

with $\langle O \rangle$ given in (1). We mention in passing that the holomorphicity of the observables $O[\Phi]$ as well as of the path integral density $e^{-S[\Phi]}$ is a necessary prerequisite for the method.

3. The wrong convergence problem

It is a well known problem of the complex Langevin evolution that (5) does not always hold in practise. The possible reasons for this failure have been investigated on many occasions and can be summarized as follows:

- The probability density $P(\Phi^R, \Phi^I)$ may not decay fast enough in complex field space. This causes the appearance of boundary terms at infinity when trying to perform integration by parts, spoiling a formal proof of correctness of the method [4, 5]. While such boundary terms can – in principle – be measured in a simulation [6, 7], dealing with them is a much harder task.
- In the absence of boundary terms, the complex Langevin evolution may suffer from contributions from unwanted so-called integration cycles [8–10], which are associated with spurious solutions of the theory’s Dyson–Schwinger equations [11].

Here, we shall focus on the second problem in particular. For simplicity, let us follow [8] and consider a theory with only a single degree of freedom z , for which the Dyson–Schwinger equations read

$$\langle Az^n \rangle = 0, \quad A = \frac{\partial}{\partial z} - \frac{\partial S(z)}{\partial z}, \quad (6)$$

where n is a non-negative integer. They can be derived from the generating functional

$$Z(j) = \int_{-\infty}^{\infty} dz e^{-S(z)+jz}. \quad (7)$$

depending on the source j , via the master equation

$$\left. \frac{\partial S(z)}{\partial z} \right|_{z=\frac{\partial}{\partial j}} Z(j) = jZ(j). \quad (8)$$

The important point to note is that the solution (7) to (8) is not unique. In fact, any

$$Z_\gamma(j) \equiv \int_\gamma dz e^{-S(z)+jz} \quad \text{with} \quad \left. e^{-S(z)+jz} \right|_{\partial\gamma} = 0 \quad (9)$$

solves the master equation. Such integration contours γ are referred to as integration cycles. Thus, the complexification of the underlying field space enlarges the solution space of a theory’s Dyson–Schwinger equations, which has profound consequences for complex Langevin expectation values.

In particular, it was shown for one-dimensional systems in [9] that if (6) holds in a complex Langevin simulation, the complex Langevin expectation values are linear combinations of expectation values computed along each of the theory's n_γ linearly independent integration cycles, i.e.,

$$\langle z^n \rangle_{\text{CL}} = \sum_{i=1}^{n_\gamma} a_i \langle z^n \rangle_{\gamma_i}, \quad \text{with} \quad \langle z^n \rangle_{\gamma_i} \equiv \left. \frac{\partial \ln(Z_\gamma(j))}{\partial j} \right|_{j=0} \quad \text{and} \quad a_i \in \mathbb{C}. \quad (10)$$

Thus, defining $\gamma_1 = \mathbb{R}$ without loss of generality, complex Langevin produces correct (i.e., physical) results if and only if $a_i = \delta_{i1}$. In [10], it was argued that one may – at least to a certain extent – use a kernel to control which integration cycles are being sampled in a simulation. Moreover, the same reference also established first numerical evidence for the validity of (10) for more than one degree of freedom.

The crucial observation is that conventional correctness criteria for complex Langevin do not seem to be sensitive to contributions from unwanted integration cycles: the Dyson–Schwinger equations hold and boundary terms vanish, but the results complex Langevin produces are nonetheless incorrect. We also mention that the correctness criterion based on the decay properties of the complex Langevin drift term [12] also fails in this respect, which we are planning to investigate on a quantitative level in future work. A reliable correctness criterion is thus much needed.

4. A novel condition for correctness

In [13], we have worked out a family of necessary and sufficient conditions for correctness of complex Langevin simulation results. We demonstrate them here in a simplified manner, using a particular one-dimensional toy model as an example. For a full derivation in general dimensions, a more detailed discussion of the underlying assumptions, as well as applications to a handful of other models, we refer to [13]. The theory we shall consider here is given by the action

$$S(z) = \frac{\lambda}{4} z^4, \quad \lambda = e^{\frac{5i\pi}{6}}. \quad (11)$$

It has been investigated before in the context of complex Langevin simulations, in particular regarding the application of a kernel, on multiple occasions [13–15].

In order to outline the correctness conditions, let us first define the p -norms (with $p \geq 1$) of functions $f(z)$:

$$\|f\|_p = \left(\int_{\mathcal{M}_r} dz |f(z)|^p \right)^{1/p} \quad \text{and} \quad \|f\|_\infty = \sup_{z \in \mathcal{M}_r} |f(z)|, \quad (12)$$

where \mathcal{M}_r can be chosen to be any real manifold that contains no zeros of $e^{-S(z)}$. In most applications it will either be the real line \mathbb{R} or some continuous deformation of \mathbb{R} to avoid such zeros. Next, we introduce the functions $w(z)$ and $\rho_r(z)$ as

$$w(z) = e^{-\frac{\lambda}{8} z^4}, \quad \rho_r(z) = \frac{1}{Z} e^{-\frac{\lambda}{8} z^4}, \quad (13)$$

which gives $\frac{e^{-S(z)}}{Z} = w(z)\rho_r(z)$. Note that this choice of $w(z)$ and $\rho_r(z)$ is by no means unique; see [13] for further details and the precise requirements on $w(z)$ and $\rho_r(z)$. With this, we then

define the constants $C^{(p)}$ and the w -norms $\|f\|_w^{(p)}$ as

$$C^{(p)} \equiv \|\rho_r\|_p \quad \text{and} \quad \|f\|_w^{(p)} \equiv \|w(z)f(z)\|_{\frac{p}{p-1}}, \quad (14)$$

respectively. It was shown in [13] that a necessary and sufficient condition for correctness of complex Langevin results in the model (11) is that, on the one hand, the Dyson–Schwinger equations of (11) hold and, on the other hand, the following bounds are obeyed for any choice of $p \geq 1$, both for all $f(z)$ in the space of polynomials:

$$|\langle f(z) \rangle_{\text{CL}}| \leq C^{(p)} \|f\|_w^{(p)}. \quad (15)$$

Needless to say, it is impossible in practice to verify the validity of (15) for all polynomials, which severely limits the usefulness of the proposed correctness criterion. However, we have found it helpful as a tool for detecting wrong convergence. To be clear, (15) really defines an entire family of conditions, since the choice of p , as well as that of the functions $w(z)$ and $\rho_r(z)$ and the real manifold \mathcal{M}_r , is essentially arbitrary. This means that if (15) is violated for even a single such choice and any $f(z)$, one concludes that the complex Langevin results are necessarily incorrect. It should be emphasized, however, that what is really meant by “incorrect results” is that (5) is violated for at least a single observable. Nonetheless, it is conceivable that even in a scenario in which some bounds in (15) do not hold, certain observables (for instance low powers of z) could still come out right. In other words, it is not sufficient to check (15) for a finite set of observables in order to draw conclusions about the correctness of that set.

We have successfully tested this new correctness condition in [13] using $p = 1$. Here, however, we shall investigate more general choices $p \geq 1$. As in previous works, we parametrize the kernel in (2), which is just a single complex number in the present case, as

$$H = e^{-\frac{im\pi}{48}}, \quad m = 0, 1, \dots, 47, \quad (16)$$

and we study results as a function of the kernel parameter m . As is known from [14], complex Langevin produces correct results in the model (11) for a sizeable range of m close to $m = 10$. However, there is another such plateau of equal size around $m = 34$, as well as two smaller plateaus near $m = 22$ and $m = 46$, respectively, on which the Dyson–Schwinger equations are fulfilled but the results are nonetheless incorrect due to contributions from unwanted integration cycles [10, 15]. To the best of our knowledge, the condition (15) is the only correctness criterion that exists to date that could detect this incorrectness without the need of computing exact results.

To begin with, we plot in Fig. 1 (left) the dependence of the bound $C^{(p)} \|f\|_w^{(p)}$ on p for the particular choice

$$f(z) = \frac{\lambda}{4} z^4 - \frac{\sqrt{\lambda}}{2} z^2. \quad (17)$$

Curiously, the bound attains a minimum around $p = 2$ before converging to its $p \rightarrow \infty$ limit monotonically. While each value of p gives rise to an independent condition for correctness, in practise one should of course try to minimize the bounds, in order to have as strong a condition as possible. In the present case, the optimal value is thus $p \approx 2$, while the choice $p = 1$ in [13] is actually the largest such bound. Note, however, that the bounds for different control observables $f(z)$, as well as for different choices of $w(z)$, $\rho_r(z)$, and \mathcal{M}_r will generally assume minima at

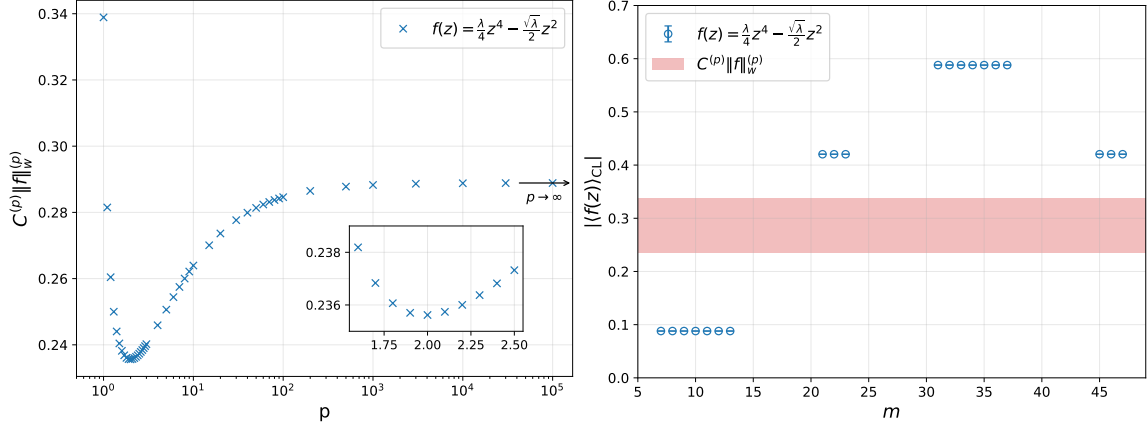


Figure 1: Validity of the bounds (15) with $f(z)$ given by (17). (left): Dependence of $C^{(p)} ||f||_w^{(p)}$ on p on a logarithmic scale. The arrow indicates the limit $p \rightarrow \infty$, while the inset shows the region around $p = 2$ on a linear scale. (right): $|(f(z))_{CL}|$ as a function of the kernel parameter m in (16). The red band indicates the range of bounds obtained by varying p . For values of m in between the four plateaus, there are non-vanishing boundary terms and the Dyson–Schwinger equations are not fulfilled [13], such that we refrain from showing any data there.

different values of p . In the following we shall nonetheless consider the full range of bounds (15) obtained by varying p . In Fig. 1 (right), we show the dependence of $|(f(z))_{CL}|$ on m and compare with (15). One finds that the bounds are violated on all plateaus except the one around $m = 10$, on which the results are known to be correct. Thus, our criterion is indeed capable of ruling out all incorrect results. Interestingly, this observation holds independently of the choice of p . We emphasize once again that these findings on their own cannot prove the correctness around $m = 10$.

5. Machine-learning kernels in lattice QCD

One of the conclusions that can be drawn from the previous section is that the use of a kernel in a complex Langevin simulation can have a large impact on the τ -evolution and the associated equilibrium distribution P in (3). Most importantly, it can potentially give rise to correct results in cases where the method would fail without a kernel. The important question for practical applications is then how to choose appropriate kernels in realistic theories of physical interest, QCD at finite baryon density being just one example. One way to find such stabilizing kernels, which has proven successful in the past [16, 17], is to employ machine-learning techniques. In the following, we briefly summarize the key idea and present a few preliminary results within the heavy-dense limit of QCD. A more detailed analysis is deferred to future work.

In the case of lattice gauge theories, the gauge degrees of freedom are represented by the group-valued link variables $U_\mu(x)$ connecting the lattice site x with its nearest neighbor in μ -direction. For QCD, the relevant gauge group is $SU(3)$. The complexification required for the application of the complex Langevin method then takes the gauge links outside of this compact domain and into $SL(3, \mathbb{C})$. More precisely, for a lattice link $X = (x, \mu)$, a naive discretization of the complex

Langevin evolution equation for U_X reads

$$U_X \rightarrow U'_X = \exp \left[i\lambda_a \left(-\varepsilon H_{Xa,Yb} H_{Yb,Zc}^T D_{Zc} S + \sqrt{\varepsilon} H_{Xa,Yb} \eta_{Yb} \right) \right] U_X . \quad (18)$$

Here, λ_a are the Gell-Mann matrices, ε denotes the discrete Langevin-time step size, $D_{Zc} S[U]$ is the usual left-derivative of the lattice action $S[U]$, η_{Yb} is Gaussian noise with the same properties as above and $H_{Xa,Yb}$ denotes the kernel matrix to be determined.

A notion of distance between the $\text{SL}(3, \mathbb{C})$ gauge configurations U_X and the target group manifold $\text{SU}(3)$ is given by the so-called unitarity norm

$$N_U \equiv \frac{1}{V} \left[\sum_X \left(U_X U_X^\dagger - \mathbb{1} \right)^2 \right] , \quad (19)$$

with V denoting the number of lattice points. A common failure mode of complex Langevin simulations in QCD at non-zero chemical potential is that the unitarity norm becomes too large, indicating the sampling of regions of complexified field space that are irrelevant for the path integral. To counteract the growth of N_U , we use it as the loss function in a gradient descent approach, with the kernel elements $H_{Xa,Yb}$ as optimization parameters. As a first step, for simplicity, we choose the kernel such that it is the same for every lattice link, i.e., $H_{Xa,Yb} = \delta_{XY} H_{ab}$, and perform the optimization with respect to the 8×8 complex matrix elements $H_{ab} = A_{ab} + iB_{ab}$, considering their real and imaginary parts A_{ab} and B_{ab} separately. Concretely, the gradient descent step reads

$$H_{ab} \rightarrow H_{ab} - l \left(\frac{\partial N_{U'}}{\partial A_{ab}} + i \frac{\partial N_{U'}}{\partial B_{ab}} \right) = H_{ab} - 2l \frac{\partial N_{U'}}{\partial H_{ab}^*} , \quad (20)$$

where $l > 0$ is a real parameter called the learning rate that we may choose freely. Notice that we are taking the derivatives of $N_{U'}$, which is (19) with the updated links (18) inserted.

Our training strategy is as follows: We start a simulation with a trivial kernel $H_{ab} = \delta_{ab}$ and let it thermalize for a Langevin-time duration τ_{therm} . We then compute the gradient on the right-hand side of (20) multiple times, namely after every τ_{grad} Langevin time for a duration τ_{update} , after which we average over these gradients and update the kernel according to (20) using this average. Before the update, the kernel is normalized to keep its (Frobenius) norm constant. With the updated kernel, we then go back to computing the gradients, but without an additional equilibration.

In the following, we test this approach for the machine-learning of kernels in the context of QCD in the heavy-dense limit [18]. In this setting, the fermionic determinant $\det(M)$ for Wilson fermions reduces to the simple form

$$\det(M) = \prod_{\mathbf{x}} \det \left(\mathbb{1} + h e^{\mu N_t} \mathcal{P}_{\mathbf{x}} \right)^2 \det \left(\mathbb{1} + h e^{-\mu N_t} \overline{\mathcal{P}_{\mathbf{x}}}^{-1} \right)^2 , \quad (21)$$

where the product runs over the spatial lattice volume, $h = (2\kappa)^{N_t}$ (κ is the usual hopping parameter), N_t denotes the number of lattice points in time direction, and μ is the chemical potential (in lattice units). Finally, the Polyakov loop and the inverse Polyakov loop are defined via the temporal gauge links $U_0(x) = U_0(t, \mathbf{x})$ as

$$\mathcal{P}_{\mathbf{x}} = \prod_{t=0}^{N_t-1} U_0(t, \mathbf{x}) , \quad \mathcal{P}_{\mathbf{x}}^{-1} = \prod_{t=N_t-1}^0 U_0^{-1}(t, \mathbf{x}) , \quad (22)$$

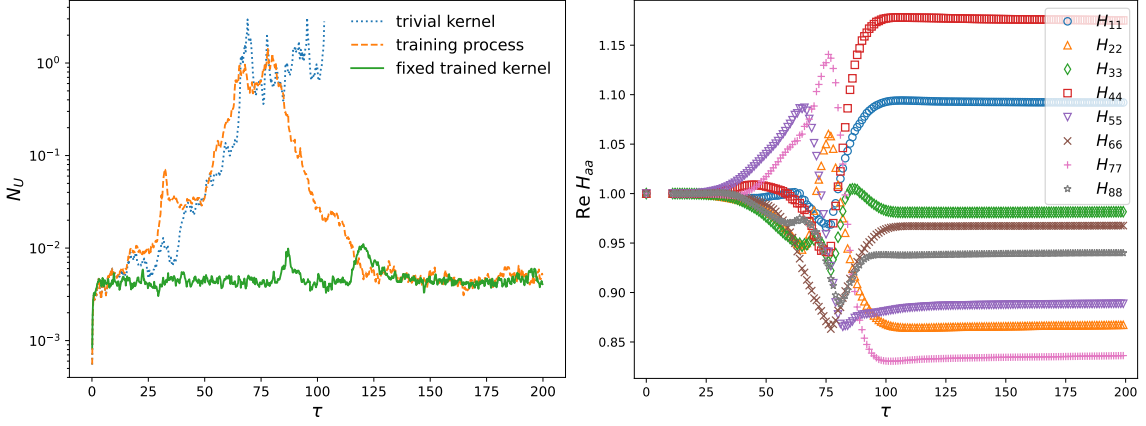


Figure 2: Langevin-time evolution of the training process for the complex Langevin kernel; see the main text for the choice of parameters. (*left*): Unitarity norm (19) in simulations with a trivial kernel (dotted line), an active training process (dashed line) and a fixed trained kernel (solid line). The initial conditions are the same for each simulation. (*right*): Real parts of the diagonal kernel elements during the training process.

respectively. In our test scenario, we employ a simple Wilson gauge action and the following set of parameters: $\kappa = 0.12$, $N_t = 4$, $N_s = 8$ (with N_s denoting the number of lattice points in the spatial directions), and $\mu = 0.8$. Moreover, for the inverse gauge coupling $\beta = 6/g^2$, we choose the value $\beta = 6.0$. Since we are conducting a proof-of-principle study, we do not perform any sort of scale setting, let alone a continuum or infinite-volume extrapolation for the time being. The hyper-parameters we employ in the training are given by $l = 10^3$, $\tau_{\text{therm}} = 10$, $\tau_{\text{grad}} = 0.01$, and $\tau_{\text{update}} = 1$, leaving us with a batch size of 100. Finally, we use a Langevin step size of $\varepsilon = 10^{-5}$. We do not employ any improved solvers [19] or adaptive step-size algorithms [20], but we do perform 16 gauge-cooling steps [21] after every Langevin update. We would like to stress at this point the possibility that the usage of gauge-cooling updates might be in conflict with the idea of using a homogeneous and isotropic kernel in the way outlined above, since our choice of kernel is not invariant under gauge transformations. The effects of this, as well as potential cures, need to be investigated and we are planning to do so in future work.

In Fig. 2 (left), we show the results of our optimization procedure. In particular, we show the evolution of N_U with the Langevin time in a typical simulation. As can be seen, with a trivial kernel the unitarity norm grows with τ rather quickly due to the constant step size and the lack of improvement (dotted curve). At a point where it becomes too large to draw meaningful conclusions from the simulation, the latter is terminated. Once we apply our training procedure (using the same initial configuration and random number seed here for better comparison), however, we observe a significant improvement in the τ -dependence of N_U : While initially it grows at about the same rate as with a trivial kernel, it then goes back to small values due to the trained kernel (dashed curve). Most strikingly, the stable behavior observed for large enough Langevin times $\tau \gtrsim 130$ indicates that the unitarity norm has reached equilibrium. This is very much desirable, as it extends the range in τ from which one may reliably draw samples and thus the statistics by a significant amount.

To get an idea about what exactly the kernel learns, we show in Fig. 2 (right) the τ -evolution of the (real parts of the) diagonal kernel elements H_{aa} . They are initialized as $H_{aa} = 1$ and at first vary

only slowly with Langevin time. However, at the point where the unitarity norm becomes large, the kernel elements also change faster. Finally, for large enough Langevin times, the H_{aa} show almost no variation as a function of τ anymore, supporting the picture of equilibration. The real parts of the non-diagonal elements exhibit an analogous behavior, but stay below ~ 0.2 in magnitude for the entire simulation. A similar statement can be made for the imaginary parts of the kernel elements. All in all, the kernel is still predominantly diagonal after the training process.

In practice, of course, one would rather work with a single fixed kernel to avoid computing gradients for the entire duration of a simulation. To this end, one may use the final kernel matrix H_{ab} that was trained in the above way. In particular, according to Fig. 2 (right) that final kernel appears to be stable as a function of τ . The results of a simulation with this fixed trained kernel are shown in Fig. 2 (left) as well (solid line). As can be seen, the unitarity norm stays below $\sim 10^{-2}$ for the entire simulation, i.e., the simulation trajectory never seems to sample irrelevant regions of the complexified field space. Thus, it appears as if there are no issues related to the use of both a (homogeneous and isotropic) kernel and gauge cooling in the present case. If such a behavior is reproducible for smaller β , corresponding to lower temperatures, and, in particular, in full QCD, it has the potential to greatly advance the range of applicability of the complex Langevin method. In particular, it could allow the study of previously unexplored regions of the QCD phase diagram, especially when combined with improved solvers and/or adaptive step size algorithms.

Acknowledgments

We are grateful to Enno Carstensen and Ion-Olimpiu Stamatescu for valuable discussions. This research was funded in whole, or in part, by the Austrian Science Fund (FWF) [[10.55776/P36875](#)]. The numerical simulations underlying this work were performed on the computing cluster of the University of Graz (GSC).

References

- [1] G. Parisi, *On complex probabilities*, *Phys. Lett. B* **131** (1983) 393.
- [2] J.R. Klauder, *A Langevin approach to fermion and quantum spin correlation functions*, *J. Phys. A* **16** (1983) L317.
- [3] G. Parisi and Y.-S. Wu, *Perturbation theory without gauge fixing*, *Sci. Sin.* **24** (1981) 483.
- [4] G. Aarts, E. Seiler and I.-O. Stamatescu, *Complex Langevin method: When can it be trusted?*, *Phys. Rev. D* **81** (2010) 054508 [[0912.3360](#)].
- [5] G. Aarts, F.A. James, E. Seiler and I.-O. Stamatescu, *Complex Langevin: etiology and diagnostics of its main problem*, *Eur. Phys. J. C* **71** (2011) 1756 [[1101.3270](#)].
- [6] M. Scherzer, E. Seiler, D. Sexty and I.-O. Stamatescu, *Complex Langevin and boundary terms*, *Phys. Rev. D* **99** (2019) 014512 [[1808.05187](#)].

- [7] M. Scherzer, E. Seiler, D. Sexty and I.-O. Stamatescu, *Controlling complex Langevin simulations of lattice models by boundary term analysis*, *Phys. Rev. D* **101** (2020) 014501 [1910.09427].
- [8] G. Guralnik and C. Pehlevan, *Complex Langevin equations and Schwinger–Dyson equations*, *Nucl. Phys. B* **811** (2009) 519 [0710.3756].
- [9] L.L. Salcedo and E. Seiler, *Schwinger–Dyson equations and line integrals*, *J. Phys. A* **52** (2019) 035201 [1809.06888].
- [10] M.W. Hansen, M. Mandl, E. Seiler and D. Sexty, *Role of integration cycles in complex Langevin simulations*, *Phys. Rev. D* **111** (2025) 074502 [2412.17137].
- [11] S. García, Z. Guralnik and G.S. Guralnik, *Theta Vacua and Boundary Conditions of the Schwinger-Dyson Equations*, [hep-th/9612079](#).
- [12] K. Nagata, J. Nishimura and S. Shimasaki, *Argument for justification of the complex Langevin method and the condition for correct convergence*, *Phys. Rev. D* **94** (2016) 114515 [1606.07627].
- [13] M. Mandl, E. Seiler and D. Sexty, *Necessary and sufficient conditions for correctness of complex Langevin*, [2508.14512](#).
- [14] H. Okamoto, K. Okano, L. Schülke and S. Tanaka, *The role of a kernel in complex langevin systems*, *Nucl. Phys. B* **324** (1989) 684.
- [15] M. Mandl, M.W. Hansen, E. Seiler and D. Sexty, *Kernels and integration cycles in complex Langevin simulations*, *PoS (LATTICE2024)* (2024) 059 [2410.13423].
- [16] D. Alvestad, R. Larsen and A. Rothkopf, *Towards learning optimized kernels for complex Langevin*, *JHEP* **2023** (2023) 057 [2211.15625].
- [17] D. Alvestad, A. Rothkopf and D. Sexty, *Lattice real-time simulations with learned optimal kernels*, *Phys. Rev. D* **109** (2023) L031502 [2310.08053].
- [18] I. Bender, T. Hashimoto, F. Karsch, V. Linke, A. Nakamura, M. Plewnia et al., *Full QCD and QED at finite temperature and chemical potential*, *Nucl. Phys. B (Proc. Suppl.)* **26** (1992) 323.
- [19] A. Ukawa and M. Fukugita, *Langevin Simulation Including Dynamical Quark Loops*, *Phys. Rev. Lett.* **55** (1985) 1854.
- [20] G. Aarts, F.A. James, E. Seiler and I.-O. Stamatescu, *Adaptive stepsize and instabilities in complex Langevin dynamics*, *Phys. Lett. B* **687** (2010) 154 [0912.0617].
- [21] E. Seiler, D. Sexty and I.-O. Stamatescu, *Gauge cooling in complex Langevin for lattice QCD with heavy quarks*, *Phys. Lett. B* **723** (2013) 213 [1211.3709].