Kernel controlled real-time Complex Langevin simulation

Daniel Alvestad[1],[*], *Rasmus* Larsen[1],[**], and *Alexander* Rothkopf[1],[***]

[1]Faculty of Science and Technology, University of Stavanger, NO-4021 Stavanger, Norway

Abstract. This study explores the utility of a kernel in complex Langevin simulations of quantum real-time dynamics on the Schwinger-Keldysh contour. We give several examples where we use a systematic scheme to find kernels that restore correct convergence of complex Langevin. The schemes combine prior information we know about the system and the correctness of convergence of complex Langevin to construct a kernel. This allows us to simulate up to 2β on the real-time Schwinger-Keldysh contour with the $0 + 1$ dimensional anharmonic oscillator using $m = 1, \lambda = 24$, which was previously unattainable using the complex Langevin equation.

1 Introduction

The notorious sign problem [1, 2] prevents the computation of the real-time dynamics of strongly correlated quantum fields by means of established Monte-Carlo methods. It thus constitutes a major roadblock in the study of phenomenologically relevant transport phenomena in various field of physics, be it in the quark-gluon plasma created in relativistic heavy-ion collisions or in the functional materials studied in condensed matter systems.

One promising strategy to enable direct simulations of the real-time dynamics is the complex Langevin approach [3], based on stochastic quantization [4] (for other approaches see e.g. [5]) of Feynman's path integral $\langle O \rangle = \int \mathcal{D}\phi\, O \exp[iS[\phi]]$. It proposes to complexify the degrees of freedom of the system $\phi \to \phi^R + i\phi^I$ and evolve them stochastically in an artificial additional time, so called Langevin time τ_L with Gaussian noise η

$$\frac{d\phi^R}{d\tau_L} = \text{Re}\left[i\frac{\delta S[\phi]}{\delta\phi}\bigg|_{\phi=\phi^R+i\phi^I} \right] + \eta(\tau_L), \quad \frac{d\phi^I}{d\tau_L} = \text{Im}\left[i\frac{\delta S[\phi]}{\delta\phi}\bigg|_{\phi=\phi^R+i\phi^I} \right]. \tag{1}$$

Instead of describing the individual realizations of the complexified system with a Langevin equation we can consider the evolution of the distribution of the d.o.f. via the associated Fokker-Planck equations. There exists two of those, a genuine FP equation, which describes the real distribution $P[\phi^R, \phi^I]$ of the complexified ϕ and a complex FP equation, which describes the distribution $\rho(\phi)$ and whose late time limit must be $\lim_{\tau_L\to\infty} \rho(\phi) \propto \exp[iS]$ for complex Langevin to succeed.

This condition is one among several (for the boundary criterion see [6]), which over the past years have been explored to verify whether complex Langevin dynamics converge to the

---

[*]e-mail: daniel.alvestad@uis.no
[**]e-mail: rasmus.n.larsen@uis.no
[***]e-mail: alexander.rothkopf@uis.no

correct solution. The possibility to converge to incorrect solutions, together with the occurrence of so-called runaway trajectories constitute the two main drawbacks of this approach (see e.g. [7]).

In a previous publication [8] we have shown how to avoid runaway trajectories, by introducing inherently stable solvers to the solution of the complex Langevin dynamics. In contrast to the adaptive step size prescription [9] discussed in the literature, the implicit solvers also provide a novel mechanism to regularize the underlying path integral we wish to simulate. The simplest extension of the often deployed forward Euler scheme is the generalized Euler-Maruyama (EM) scheme. It evolves each discretized field ϕ_j via time step ε according to

$$\phi_j^{n+1} = \phi_j^n + i\varepsilon\Big[\theta\frac{\delta S^{n+1}}{\delta\phi_j} + (1-\theta)\frac{\delta S^{n+1}}{\delta\phi_j}\Big] + \sqrt{\varepsilon}\eta_j^n, \tag{2}$$

where the implicitness parameter θ governs the stability properties. For $\theta \geq 1/2$ the approach is unconditionally stable. The stability and inherent ability of this solver to regularize the path integral, allowed us in ref. [8] to simulate, for the first time, the real-time dynamics of the strongly coupled anharmonic oscillator on the canonical Schwinger-Keldysh (SK) contour without tilt.

The real-time extent of these simulations was very small $t_{max}m = 0.5$ in order for the naive complex Langevin equations to converge to the correct solution. Of course to obtain phenomenologically relevant insight these simulations must be extended to larger real-times, which is the aim of the study presented in this contribution.

As one extends the real-time extent of the Schwinger-Keldysh contour it is known that naive complex Langevin fails to converge correctly [10]. This phenomenon has been investigated in details in the community and it was found that the appearance of so called boundary terms [6] signals the failure of convergence. In the context of complex Langevin for gauge theories the gauge cooling approach [11] and more recently dynamical stabilization [12] have been proposed as means to avoid boundary terms and have led to impressive improvements in the simulations of QCD at finite Baryo-chemical potential [13]. The combination of the two techniques is also deployed to study the real-time dynamics of gauge fields [14].

One way to tame the sign problem and the associated incorrect convergence of CL in the past was to construct coordinate transformations or field redefinitions [15]. They can all be summarized by a general modification of the complex Langevin equations in terms of a so called *kernel*. The role of kernels for the convergence of CL has been explored early on in simple models in refs. [16].

In our study we focus on kernelled complex Langevin and will introduce a novel strategy based on prior knowledge that allows us to learn optimal kernels to restore correct convergence. Note that the sign problem is known to be NP-hard [17] and thus no generic solution strategy is believed to exist. In order to make progress we thus believe system-specific information needs to be incorporated into the CL simulations and our learning strategy offers exactly that.

2 Kernelled Complex Langevin

It is known [3] that real-valued Langevin equations offer the freedom to introduce a real-valued kernel in the stochastic dynamics (dW denotes a Wiener process)

$$d\phi = -K\frac{\delta S}{\delta\phi}d\tau_L + \frac{\delta K}{\delta\phi}d\tau_L + \sqrt{K}dW. \tag{3}$$
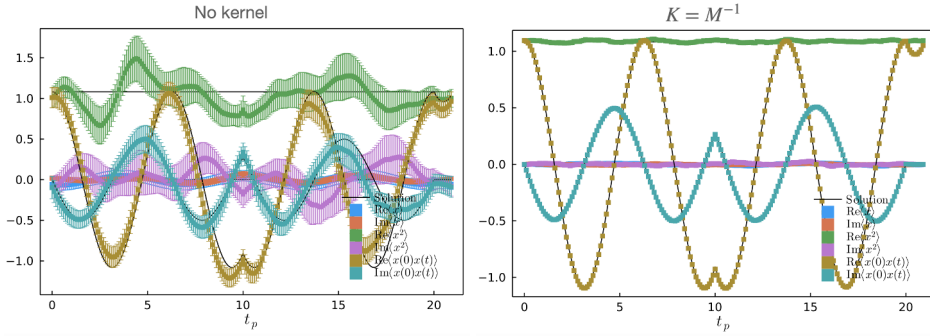
Figure 1. (left) The CL dynamics of the harmonic oscillator on the SK contour with real-time extent $mt_{max} = 10$ in the absence of a kernel. (right) Same system after introducing the manually constructed kernel of eq. (6).

It will change the approach to the unique stationary distribution but does not change its form. In complex Langevin we may introduce a complex valued kernel (making sure that it factorizes $K = H^\dagger H$ so that one can take the square root), which amounts to a non-neutral modification of the dynamics

$$\frac{d\phi^R}{d\tau_L} = \mathrm{Re}\left[iK[\phi]\frac{\delta S[\phi]}{\delta\phi} + \frac{\delta K}{\delta\phi}\right]_{\phi=\phi^R+i\phi^I} + \mathrm{Re}[H[\phi^R+i\phi^I]]\eta(\tau_L), \qquad (4)$$

$$\frac{d\phi^I}{d\tau_L} = \mathrm{Im}\left[iK[\phi]\frac{\delta S[\phi]}{\delta\phi} + \frac{\delta K}{\delta\phi}\right]_{\phi=\phi^R+i\phi^I} + \mathrm{Im}[H[\phi^R+i\phi^I]]\eta(\tau_L). \qquad (5)$$

Through a tuning of such a kernel it is thus fathomable that the correct late time distribution can be recovered. One example where a manually constructed kernel can be used to restore correct convergence is in the harmonic oscillator on the SK contour. It was shown in ref. [16] that by turning the drift term of the CL equation into $-\phi$ through the application of a kernel, correct convergence can be achieved in a simple one d.o.f. model. We now understand this result from connecting complex Langevin to the thimble structure of the theory (see section section 3.1). The construction for the harmonic oscillator proceeds similarly

$$S = \int_{SK} dt\, \phi(t)M\phi(t), \quad \rightarrow \quad K = iM^{-1} \quad \rightarrow \quad K\delta S/\delta\phi = -\phi. \qquad (6)$$

Using the implicit EM scheme with $\theta = 0.6$ at $\beta = 1$ and collecting statistics over hundred different trajectories up to $m\tau_L = 100$, we show in fig. 1 how the introduction of the manual kernel of eq. (6) (right) tames the large noise and incorrect convergence of the naive CL dynamics (left). The colored data points show the real- and imaginary part of different correlators evaluated along the SK contour. For $mt_p \leq mt_{max} = 10$ the quantities are evaluated on the forward contour, for $10 \leq mt_p \leq 20$ on the backward contour and for the remaining values of t_P on the Euclidean branch (blue and orange: $\langle x \rangle$, green and pink $\langle x^2 \rangle$ and brown and turquoise $\langle x(0)x(t_P)\rangle$). The solid lines indicate the known values obtained from the Schrödinger equation.

Interestingly the kernel manually constructed in the free theory also allows us to improve convergence for the strongly coupled anharmonic oscillator with $\lambda = 24$ a common benchmark system in the literature. Deploying K of eq. (6) naively allows us to extend correct
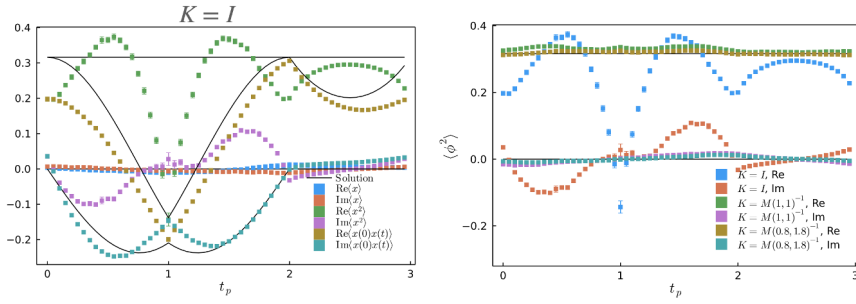
Figure 2. (left) The CL dynamics of the strongly coupled anharmonic oscillator on the SK contour with real-time extent $mt_{max} = 1$ in the absence of a kernel. Note the incorrect convergence in all observables. (right) Same system after introducing the manually constructed kernel of eq. (6).

convergence up to $mt_{max} = 0.75$. Modifying the relative weight between the kinetic and potential term in the kernel matrix M by multiplying the derivative term by $g = 0.8$ and the potential term by $m_g = 1.8$, values that were heuristically found, we even manage to get to $mt_{max} = 1$. In the right panel of fig. 2 we plot the values of $\langle x^2 \rangle$ for different choices of kernel and in the left panel show the anharmonic oscillator results wihthout the presence of a kernel.

These results are encouraging as they show that a well chosen kernel indeed can help restore correct convergence. The central challenge however lies in the fact that no systematic prescription exists on how to modify the kernel further to restore convergence at later real-time, i.e. for a larger extent of the SK contour.

3 Learning optimal kernels for correct convergence

Our proposal therefore is to take inspiration from the machine learning community and attempt to systematically find optimal kernels based on prior knowledge available on the system. Such prior information comes the form of e.g. the symmetries of the system, Euclidean correlators accessible in conventional simulations and also in the form of the correctness criterion that requires that boundary terms are absent. For each type of information we propose to construct a cost functional L^{sym}, L^{Eucl} and L^{BT} whose value taken together $L^{prior} = L^{sym} + L^{Eucl} + L^{BT}$ shall be minimized by tuning the entries of the kernel matrix K. As a first step we will consider only field-independent kernels for simplicity, which can be seen as the lowest order in an expansion in basis functions depending on the field ϕ. In this study we will only include L^{sym} and L^{Eucl}, where we exploit that the two-point functions are known in Euclidean time and that time-translation invariance enforces that the equal-time correlation functions must remain constant over the whole of the SK contour.

To minimize the loss we have to compute gradients of the cost functionals and in turn evaluate the dependence of observables on the values of the kernel K. This entails taking the derivative w.r.t. the whole of the CL simulation. Such operation in principle can be accomplished by the use of modern autodifferentiation techniques [18], but we have found that due to the stiffness of the CL equation in real-time many of these direct methods in their standard implementation are too costly in practice. One path forward in the future is to implement the adjoint method specific to real-time CL.

Instead we constructed a heuristic low-cost gradient with which we manged to significantly reduce the values of the true cost functional L^{prior}. This low-cost gradient is obtained from approximating the true gradient of the functional $L_D = \left\langle \left| D(x) \cdot (-x) - \|D(x)\| \, \|x\| \right|^{\xi} \right\rangle$,
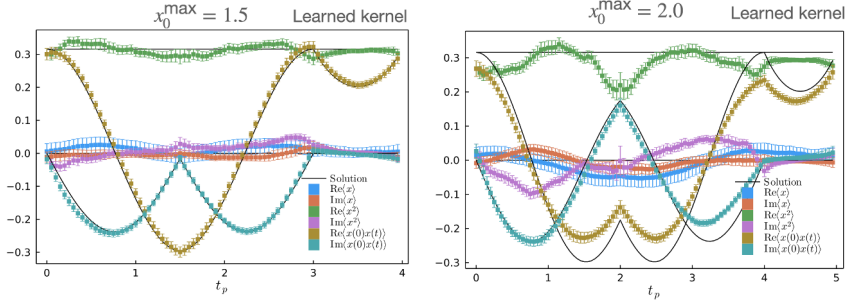
Figure 3. (left) CL dynamics of the anharmonic oscillator in the presence of our learned field-independent optimal kernel at $mt_{\mathrm{max}} = 1.5$. Correct convergence is restored (c.f. left panel of fig. 2) (right) Results of the CL simulations at $mt_{\mathrm{max}} = 2$, where we fail to learn a field-independent kernel to restore correct convergence.

with $D = K\partial S/\partial\phi$, which is designed to favor the absence of boundary terms by pulling the field degrees towards the origin. Best results were obtained when choosing $1 < \xi < 2$. The efficient evaluation of the low-cost gradient from L_D was achieved by using autodifferentiation.

By iteratively finding the optimal values for the kernel using the low-cost gradient and monitoring the success of the optimization via L^{prior} we managed to extend correct convergence of CL for the strongly coupled anharmonic oscillator up to $mt_{\mathrm{max}} = 1.5$ as shown in the left panel of fig. 3. This constitutes a threefold improvement over the previously achieved record for the real-time simulation of the anharmonic oscillator in complex Langevin.

Our approach of course has its limitations. Using a field independent kernel, we fail to restore convergence when extending the SK contour to $mt_{\mathrm{max}} = 2$, as shown on the right of fig. 3. There are several possible reasons for this failure: the use of the low-cost gradient may simply fail to locate the true minimum of L^{prior}. This is related to the fact that from simple models it is known that pulling the complexified fields towards the origin has a beneficial effect if the theory has its critical points at the origin. In general the critical points of the theory may lie away from the origin, rendering the idea behind the gradient based on L_D moot. We will most probably require a field dependent kernel to restore correct convergence for larger real-time extent a direction of future work.

3.1 Connection to Thimbles

Let us briefly mention an interesting connection between the introduction of a kernel and the thimble structure of the simulated theory. As was shown in ref. [16] in the simple complex Gaussian model $S = \frac{1}{2}\phi^2$, the kernel that turns the drift term into the simple form $iK\delta S/\delta\phi = -\phi$ restores correct convergence. We now understand this behavior through the thimble structure of the model. There exists a single thimble oriented downward 45 degrees in the complex plane passing through the origin. While naive complex Langevin samples preferentially parallel to the real ϕ^R axis, the kernel rotates the sampling exactly onto the thimble. This restores correct convergence, as sampling along the thimble has a well-defined statistical meaning.

In systems where multiple thimbles are present rotating the sampling with a field independent kernel only works in certain parameter ranges and in general a field dependent kernel is required to make sure that the physics of different thimbles is captured.

4 Conclusion

Bringing together the concept of kernelled complex Langevin with a machine learning inspired strategy to locate their optimal values based on prior information we propose to systematically find kernels that restore correct convergence in simulations of strongly correlated quantum systems on the real-time SK contour. We have shown that the strategy is viable by extending correct convergence for the anharmonic oscillator threefold beyond the previous state-of-the-art. Exploration of field dependent kernels and the implementation of custom autodifferentiation methods, such as adjoint methods for real-time complex Langevin are work in progress.

5 Acknowledgments

References

[1] C. Gattringer, K. Langfeld, International Journal of Modern Physics A **31**, 1643007 (2016), number: 22 arXiv: 1603.09517

[2] G. Pan, Z.Y. Meng (2022), 2204.08777

[3] M. Namiki, I. Ohba, K. Okano, Y. Yamanaka, A.K. Kapoor, H. Nakazato, S. Tanaka, Lect. Notes Phys. Monogr. **9**, 1 (1992)

[4] P.H. Damgaard, H. Hüffel, Physics Reports **152**, 227 (1987), number: 5

[5] C.E. Berger, L. Rammelmüller, A.C. Loheac, F. Ehmann, J. Braun, J.E. Drut, Physics Reports **892**, 1 (2021)

[6] G. Aarts, F.A. James, E. Seiler, I.O. Stamatescu, Eur. Phys. J. C **71**, 1756 (2011)

[7] E. Seiler, EPJ Web Conf. **175**, 01019 (2018), 1708.08254

[8] D. Alvestad, R. Larsen, A. Rothkopf, JHEP **08**, 138 (2021), 2105.02735

[9] G. Aarts, F.A. James, E. Seiler, I.O. Stamatescu, Physics Letters B **687**, 154 (2010)

[10] J. Berges, S. Borsányi, D. Sexty, I.O. Stamatescu, Physical Review D **75** (2007), hep-lat/0609058

[11] E. Seiler, D. Sexty, I.O. Stamatescu (2012), 1211.3709

[12] G. Aarts, F. Attanasio, B. Jäger, D. Sexty, Acta Phys. Polon. Supp. **9**, 621 (2016), 1607.05642

[13] F. Attanasio, B. Jäger, F.P.G. Ziegler (2022), 2203.13144

[14] K. Boguslavski, P. Hotzy, D.I. Müller, *A stabilizing kernel for complex Langevin simulations of real-time gauge theories*, in *39th International Symposium on Lattice Field Theory* (2022), 2210.08020

[15] G. Aarts, F.A. James, J.M. Pawlowski, E. Seiler, D. Sexty, I.O. Stamatescu, JHEP **03**, 073 (2013), 1212.5231

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

[17] M. Troyer, U.J. Wiese, Phys. Rev. Lett. **94**, 170201 (2005), cond-mat/0408370

[18] A.G. Baydin, B.A. Pearlmutter, A.A. Radul, J.M. Siskind, Journal of Marchine Learning Research **18**, 1 (2018)