

On the condition for correct convergence in the complex Langevin method*

Shinji Shimasaki[†]

*Research and Education Center for Natural Sciences, Keio University, 4-1-1 Hiyoshi,
Yokohama, Kanagawa 223-8521, Japan
E-mail: shinji.shimasaki@keio.jp*

Keitaro Nagata

*KEK, High Energy Accelerator Research Organization, 1-1 Oho,
Tsukuba, Ibaraki 305-0801, Japan
E-mail: knagata@post.kek.jp*

Jun Nishimura

*KEK, High Energy Accelerator Research Organization, 1-1 Oho,
Tsukuba, Ibaraki 305-0801, Japan
Graduate University for Advanced Studies (SOKENDAI), 1-1 Oho,
Tsukuba, Ibaraki 305-0801, Japan
E-mail: jnishii@post.kek.jp*

The complex Langevin method (CLM) provides a promising way to perform the path integral with a complex action using a stochastic equation for complexified dynamical variables. It is known, however, that the method gives wrong results in some cases, while it works, for instance, in finite density QCD in the deconfinement phase or in the heavy dense limit. Here we revisit the argument for justification of the CLM and point out a subtlety in using the time-evolved observables, which play a crucial role in the argument. This subtlety requires that the probability distribution of the drift term should fall off exponentially or faster at large magnitude. We demonstrate our claim in some examples such as chiral Random Matrix Theory and show that our criterion is indeed useful in judging whether the results obtained by the CLM are trustable or not.

*34th annual International Symposium on Lattice Field Theory
24-30 July 2016
University of Southampton, UK*

*KEK-TH/1946

[†]Speaker.

1. Introduction

The complex Langevin method (CLM) [1, 2] is a promising approach to the complex action problem based on a stochastic process for complexified variables. Although it is well-known that the method does not always give correct results, the range of applicability has been substantially enlarged thanks to the recent development of a new technique such as gauge cooling [3]. In particular, the gauge cooling has enabled the application of the method to finite density QCD in the deconfined phase [4, 5].

The argument for justification of the CLM was given in ref. [6, 7]. There, a crucial step was to shift the time evolution of the probability distribution of the complexified variables to that of observables. This is possible only if the integration by parts used there is valid. For this reason, the CLM fails when the probability distribution does not fall off fast enough in the asymptotic region [6, 7] (the excursion problem) or in the region near singularities of the drift term when they exist [8] (the singular drift problem). It was shown that not only the excursion problem [3] but also the singular drift problem [9] may be cured by the gauge cooling.

Here we revisit the argument for justification of the CLM and point out a subtlety in the use of time-evolved observables [10]. In the previous argument, it was implicitly assumed that the time-evolved observables can be used for an infinitely long time. We point out that this is a too strong assumption, which is not necessarily satisfied even in cases where the CLM gives correct results. In fact, what is needed for justification is the use of the time-evolved observables for a finite but nonzero time. This still requires that the probability distribution of the drift term should be suppressed exponentially at large magnitude, which is slightly stronger than the condition for the validity of the integration by parts.

The rest of this article is organized as follows. In section 2, we discuss the condition for justification of the CLM taking into account the subtlety in the time-evolved observables. In section 3, we demonstrate our condition in simple models. Section 4 is devoted to a summary.

2. New argument for justification of the CLM

Let us consider a partition function $Z = \int dx w(x)$ written in terms of a real variable x with a complex weight $w(x)$. In the CLM, we consider the Langevin equation for the complexified variable $x \rightarrow z = x + iy$, which takes the form

$$z^{(\eta)}(t + \varepsilon) = z^{(\eta)}(t) + \varepsilon v(z) + \sqrt{\varepsilon} \eta(t) \quad (2.1)$$

in its discretized version. Here, $v(z)$ represents the drift term, which can be obtained by analytically continuing $v(x) = w(x)^{-1} \partial w(x) / \partial x$, and $\eta(t)$ is a real noise, which obeys the probability distribution $\propto e^{-\frac{1}{4} \sum_t \eta^2(t)}$. Below, we denote the expectation value with respect to η as $\langle \cdots \rangle_\eta$. Let us also extend the observable $\mathcal{O}(x)$ in the original model to a holomorphic function $\mathcal{O}(z)$ of z by analytic continuation and define the expectation value of $\mathcal{O}(z)$ as

$$\Phi(t) = \langle \mathcal{O}(x^{(\eta)}(t) + iy^{(\eta)}(t)) \rangle_\eta = \int dx dy \mathcal{O}(x + iy) P(x, y; t), \quad (2.2)$$

where $P(x, y; t)$ is the probability distribution of $x^{(\eta)}(t)$ and $y^{(\eta)}(t)$ defined by $P(x, y; t) = \langle \delta(x - x^{(\eta)}(t)) \delta(y - y^{(\eta)}(t)) \rangle_\eta$. The crucial issue for the CLM is whether this quantity $\Phi(t)$, after taking

the $t \rightarrow \infty$ and $\varepsilon \rightarrow 0$ limits, reproduces the expectation value of $\mathcal{O}(x)$ with respect to the original path integral with a complex weight, namely whether the identity

$$\lim_{t \rightarrow \infty} \lim_{\varepsilon \rightarrow 0} \Phi(t) = \frac{1}{Z} \int dx \mathcal{O}(x) w(x) \quad (2.3)$$

holds or not.

In order to address this issue, we consider the time-evolution of the expectation value $\Phi(t)$. Using the Langevin equation (2.1), we can write it as

$$\Phi(t + \varepsilon) = \int dx dy \mathcal{O}_\varepsilon(x + iy) P(x, y; t), \quad (2.4)$$

$$\mathcal{O}_\varepsilon(z) = \frac{1}{\mathcal{N}} \int d\eta e^{-\frac{1}{4}\eta^2} \mathcal{O}\left(z + \varepsilon v(z) + \sqrt{\varepsilon} \eta\right), \quad (2.5)$$

where \mathcal{N} is a normalization constant. Since $\mathcal{O}(z)$ and $v(z)$ are holomorphic, so is $\mathcal{O}_\varepsilon(z)$. By “holomorphic”, we actually mean that the functions are holomorphic in the region visited by the Langevin process. In particular, we allow the case in which the functions have singularities, which are measure zero in the configuration space. Expanding (2.5) with respect to ε and using the holomorphy of $\mathcal{O}(z)$, we can rewrite (2.4) as

$$\Phi(t + \varepsilon) = \sum_{n=0}^{\infty} \frac{1}{n!} \varepsilon^n \int dx dy \left(\tilde{L}^n : \mathcal{O}(z) \right) P(x, y; t), \quad (2.6)$$

where $\tilde{L} = (\partial/\partial z + v(z)) \partial/\partial z$ and the symbol $: \dots :$ implies that the derivative operators are moved to the right as in $:(f(x) + \partial)^2: = f(x)^2 + 2f(x)\partial + \partial^2$.

If the ε -expansion (2.6) is valid, one can take the $\varepsilon \rightarrow 0$ limit and get

$$\frac{d}{dt} \Phi(t) = \int dx dy \left\{ \tilde{L} \mathcal{O}(z) \right\} P(x, y; t). \quad (2.7)$$

In fact, it is known from the previous argument [6, 7] that (2.7) does not hold when the integration by parts, which is necessary in showing it, becomes invalid. In the present argument, on the other hand, the possible failure of (2.7) should be attributed to the breakdown of the ε -expansion (2.6). Indeed, \tilde{L}^n involves the n -th power of the drift term, which may become infinitely large. Therefore, the integral in (2.6) can be divergent for large enough n . What we have done so far is just an alternative presentation of the known problem that (2.7) can be violated. However, we will see below that a similar argument for a finite time-evolution of $\Phi(t)$ gives rise to a condition, which is stronger than the one needed for the validity of (2.7).

The time-evolution of the expectation value $\Phi(t)$ for a finite τ can be obtained formally by repeating the above argument for $\tilde{L}^n \mathcal{O}(z)$ as

$$\Phi(t + \tau) = \sum_{n=0}^{\infty} \frac{1}{n!} \tau^n \int dx dy \left\{ \tilde{L}^n \mathcal{O}(z) \right\} P(x, y; t). \quad (2.8)$$

In order for this expression to be valid, however, it is not sufficient to require that the integral appearing in the infinite series is convergent. What also matters is the convergence radius of the infinite series. In the previous argument, (2.3) was proved by assuming implicitly that the convergence radius is infinite. Actually, this assumption is too strong and can be relaxed if we employ

the induction with respect to the Langevin time [10]. What is needed to prove (2.3) then is that the convergence radius, which depends on t in general, is bounded from below as a function of t .

In what follows, we discuss the explicit condition for the expression (2.8) to be valid [10]. Let us define the probability distribution of the magnitude of the drift $u(z) = |v(z)|$ by

$$p(u; t) \equiv \int dx dy \delta(u(z) - u) P(x, y; t) . \quad (2.9)$$

Then, the most dominant contribution for each n in (2.6) and (2.8) may be written as

$$\int dx dy u(z)^n P(x, y; t) = \int_0^\infty du u^n p(u; t) . \quad (2.10)$$

In order for this to be finite for arbitrary n , $p(u; t)$ should fall off at large u faster than any power law. This condition is required for the ε -expansion (2.6) and the τ -expansion (2.8) to be valid. In order for the τ -expansion (2.8) to be valid for a finite τ , we need to require further that the convergence radius of the infinite series should be non-zero. For instance, in cases where the probability distribution of the drift term is suppressed exponentially as $p(u; t) \sim e^{-\kappa u}$ for some $\kappa > 0$, the convergence radius is estimated as $\tau \sim \kappa$. This implies that, in order for the τ -expansion (2.8) to have a nonzero convergence radius, $p(u; t)$ has to fall off exponentially or faster. Note that this condition is slightly stronger than the one obtained from the validity of the ε -expansion, which is equivalent to the validity of the integration by parts discussed in refs. [6, 7]. Therefore, our condition may be viewed as a necessary and sufficient condition for justification of the CLM.

While our argument above is given in a single variable case for simplicity, we can extend it to more general cases with multiple variables including the lattice gauge theory. We can also include the gauge cooling in the argument as in ref. [11]. For a comprehensive presentation of the argument including such generalizations, see ref. [10].

3. Demonstration of our condition

In this section, we demonstrate our new condition for justification of the CLM in some models. First, we discuss two one-variable models, in which the CLM fails in some parameter region due to the singular drift problem or the excursion problem [10]. According to our argument above, these failures should be attributed to the appearance of a large drift and hence they are understood in a unified manner. Next, we discuss the chiral Random Matrix Theory (cRMT) [13], which suffers from the singular drift problem [12, 9] at small quark mass. This example clearly shows that our condition is valid also in a multi-variable case.

3.1 A model with a singular drift

Here we consider a model with a singular drift term, whose partition function is given by

$$Z = \int dx w(x) , \quad w(x) = (x + i\alpha)^p e^{-x^2/2} , \quad (3.1)$$

where x is a real variable and α and p are real parameters. We perform simulations for $p = 4$ and various α with the step-size $\varepsilon = 10^{-5}$. The initial configuration is taken to be $z = 0$ and the first 3×10^5 steps are discarded for thermalization. After thermalization, we make 10^{10} steps and

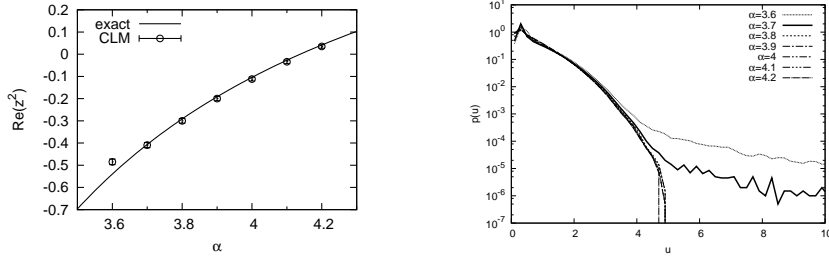


Figure 1: (Left) The real part of the expectation value of $\theta(z) = z^2$ obtained by the CLM is plotted against α for $p = 4$. The solid line represents the exact result. (Right) The probability distribution $p(u)$ of the magnitude $u = |v|$ of the drift term is shown for various α within $3.6 \leq \alpha \leq 4.2$ in a semi-log plot.

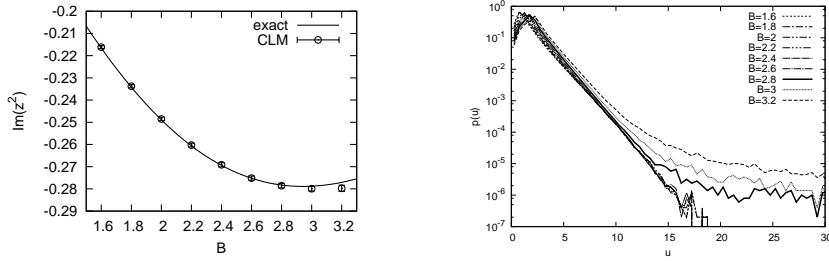


Figure 2: (Left) The imaginary part of the expectation value of $\theta(z) = z^2$ is plotted against B for $A = 1$. The solid line represents the exact result. (Right) The probability distribution $p(u)$ for the magnitude $u = |v|$ of the drift term is shown for various B within $1.6 \leq B \leq 3.2$ in a semi-log plot.

perform measurements every 10^3 steps. In Fig. 1 (Left), we plot the real part of the expectation value of $\theta(z) = z^2$ against α , which shows that the CLM reproduces the correct results for $\alpha \gtrsim 3.7$.

According to our new argument, the CLM fails when the probability distribution of the drift term is not suppressed exponentially at large magnitude. This is confirmed in Fig. 1 (Right), in which we plot the probability distribution of the magnitude of the drift term for various α . We find that the fall-off of the distribution is faster than exponential for $\alpha \geq 3.8$, while it is a power law for $\alpha \leq 3.7$ [10]. The result for $\alpha = 3.7$ seems to agree with the exact result presumably because the discrepancy is too small to be measured.

3.2 A model with a possibility of excursions

Here we consider a model with the excursion problem, whose partition function is [14]

$$Z = \int dx w(x), \quad w(x) = e^{-\frac{1}{2}(A+iB)x^2 - \frac{1}{4}x^4}, \quad (3.2)$$

where x is a real variable and A and B are real parameters. The simulation parameters are the same as those in the previous example except for the use of an adaptive step-size with the step-size ε being replaced by $\varepsilon = 0.01/|v(z)|$ when the magnitude of the drift $|v(z)|$ exceeds 10^3 . This is needed to avoid the runaway problem, which occurs for $B \geq 3.0$. In Fig. 2 (Left), we plot the imaginary part of the expectation value of $\theta(z) = z^2$ against B . It shows that the CLM reproduces the correct results for $B \leq 2.8$. Our new argument is confirmed in Fig. 2 (Right), in which the probability distribution of the magnitude of the drift term is plotted for various B . We find that the

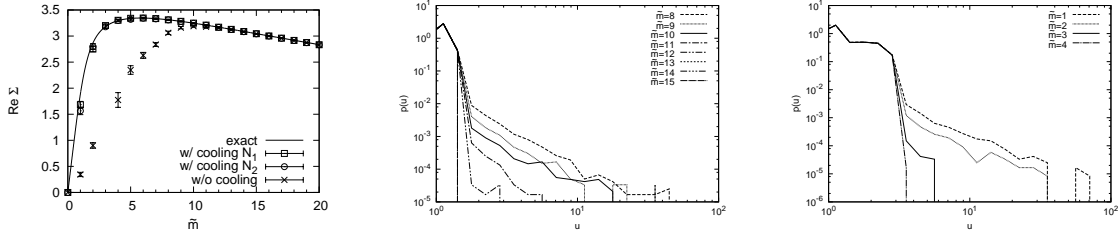


Figure 3: (Left) The real part of the expectation value of the chiral condensate is plotted against \tilde{m} . The results are obtained by the CLM with or without gauge cooling. The solid line represents the exact result. The probability distribution of the magnitude of the drift $p(u)$ (3.5) is shown in log-log plots for various \tilde{m} in the cases without gauge cooling (Middle) and with the gauge cooling using the norm N_1 (Right).

distribution falls off exponentially for $B \leq 2.6$, while it falls off by a power for $B \geq 2.8$ [10]. The result at $B = 2.8$ agrees with the exact result presumably because the discrepancy is too small to be measured.

3.3 chiral Random Matrix Theory

In order to demonstrate our condition in a multi-variable case, we consider the cRMT for N_f quarks with the degenerate mass m and the chemical potential μ . The partition function is given by

$$Z = \int d\Phi_1 d\Phi_2 [\det(D + m)]^{N_f} e^{-S_b}, \quad (3.3)$$

where Φ_k ($k = 1, 2$) are $N \times N$ general complex matrices. The bosonic action S_b in (3.3) is given by $S_b = 2N \sum_{k=1}^2 \text{Tr}(\Phi_k^\dagger \Phi_k)$ and the $2N \times 2N$ matrix D is given as

$$D = \begin{pmatrix} 0 & e^\mu \Phi_1 + e^{-\mu} \Phi_2 \\ -e^{-\mu} \Phi_1^\dagger - e^\mu \Phi_2^\dagger & 0 \end{pmatrix}. \quad (3.4)$$

We apply the CLM for $N = 30$, $N_f = 2$, $\tilde{\mu} \equiv \mu\sqrt{N} = 2$ and various $\tilde{m} \equiv mN$. In Fig. 3 (Left), we plot the real part of the chiral condensate $\Sigma = N^{-1} \partial \log Z / \partial m$ obtained with or without gauge cooling as a function of \tilde{m} [9]. In this Figure, N_1 and N_2 stand for the types of norm used for gauge cooling. See ref. [9] for the details. We find that the CLM reproduces the exact results for $\tilde{m} \gtrsim 10$ without gauge cooling and for $\tilde{m} \gtrsim 1$ with gauge cooling using the norm N_1 .

Next, we discuss the probability distribution of the magnitude of the drift term. Let us denote the drift terms for Φ_i and Φ_i^\dagger by F_i and \bar{F}_i ($i = 1, 2$), respectively. Then, the probability distribution may be defined as

$$p(u) = \frac{1}{2N} \sum_{i=1}^2 \sum_{a=1}^N \langle (\delta(u - v_i^{(a)}) + \delta(u - \bar{v}_i^{(a)})) \rangle, \quad (3.5)$$

where $v_i^{(a)}$ and $\bar{v}_i^{(a)}$ ($a = 1, \dots, N$) are the eigenvalues of $(F_i^\dagger F_i)^{1/2}$ and $(\bar{F}_i^\dagger \bar{F}_i)^{1/2}$, respectively. Note that this definition respects the $U(N) \times U(N)$ symmetry of the original cRMT. In Fig. 3 (Middle) and (Right), we plot the probability distribution of the magnitude of the drift term (3.5) against various \tilde{m} in cases with or without gauge cooling. We observe a power-law fall-off of the distribution for $\tilde{m} \lesssim 12$ without gauge cooling and for $\tilde{m} \lesssim 3$ with gauge cooling using the norm N_1 . These regions agree with the regions where the CLM gives wrong results.

4. Summary

We revisited the argument for justification of the CLM, which was first given in ref. [6, 7]. In particular, we pointed out that the assumption made in the previous argument that time-evolved observables can be used for an infinitely long time is too strong. All we need to show (2.3) is the use of time-evolved observables for a finite but nonzero time. This still requires that the probability distribution of the drift term should fall off exponentially or faster at large magnitude. Our new condition can be used to probe the two possible problems in the CLM, namely the excursion problem and the singular drift problem, in a unified manner and to judge whether the results obtained by the CLM are trustable or not. This was demonstrated in two one-variable models and the cRMT, where it was shown that the CLM reproduces the exact results when the probability distribution of the drift term falls off exponentially or faster at large magnitude. Obviously, our condition should be of particular use in applying the CLM to cases in which the results are not known a priori. See ref. [15] for an application to finite density QCD.

Acknowledgments

This work was supported by Grant-in-Aid for Scientific Research (No. 26800154 for K.N. and No. 23244057, 16H03988 for J.N.) from Japan Society for the Promotion of Science. S.S. was supported by the MEXT-Supported Program for the Strategic Research Foundation at Private Universities “Topological Science” (Grant No. S1511006).

References

- [1] G. Parisi, Phys. Lett. B **131** (1983) 393.
- [2] J. R. Klauder, Phys. Rev. A **29** (1984) 2036.
- [3] E. Seiler, D. Sexty and I. O. Stamatescu, Phys. Lett. B **723** (2013) 213
- [4] D. Sexty, Phys. Lett. B **729** (2014) 108
- [5] Z. Fodor, S. D. Katz, D. Sexty and C. Török, Phys. Rev. D **92**, no. 9, 094516 (2015)
- [6] G. Aarts, E. Seiler and I. O. Stamatescu, Phys. Rev. D **81** (2010) 054508
- [7] G. Aarts, F. A. James, E. Seiler and I. O. Stamatescu, Eur. Phys. J. C **71** (2011) 1756
- [8] J. Nishimura and S. Shimasaki, Phys. Rev. D **92** (2015) 1, 011501
- [9] K. Nagata, J. Nishimura and S. Shimasaki, JHEP **1607**, 073 (2016)
- [10] K. Nagata, J. Nishimura and S. Shimasaki, arXiv:1606.07627 [hep-lat].
- [11] K. Nagata, J. Nishimura and S. Shimasaki, PTEP **2016** (2016) no.1, 013B01
- [12] A. Mollgaard and K. Splittorff, Phys. Rev. D **88**, no. 11, 116007 (2013)
- [13] K. Nagata, J. Nishimura and S. Shimasaki, in preparation.
- [14] G. Aarts, P. Giudice and E. Seiler, Annals Phys. **337** (2013) 238
- [15] K. Nagata, H. Matsufuru, J. Nishimura and S. Shimasaki, arXiv:1611.08077[hep-lat]