On the condition for correct convergence in the complex Langevin method

Shinji Shimasaki (Keio U.)

based on the work done in collaboration with Keitaro Nagata (KEK), Jun Nishimura (KEK)

arXiv:1606.07620 [hep-lat] and more
Complex Langevin Method (CLM)

- Path integral for a theory with complex $S$

$$Z = \int_{-\infty}^{\infty} dx e^{-S(x)} \quad \langle \mathcal{O}(x) \rangle = \frac{1}{Z} \int dx \mathcal{O}(x) e^{-S(x)}$$

- Langevin eq. for complexified variables $x \to z = x + iy \in \mathbb{C}$

$$\frac{dz(t)}{dt} = v(z) + \eta(t)$$

- drift term $v(z) = -\frac{dS(z)}{dz}$, gaussian noise $\eta(t)$

$$\lim_{t \to \infty} \langle \mathcal{O}(z \eta(t)) \rangle_\eta = \langle \mathcal{O}(x) \rangle \quad \langle \cdots \rangle_\eta = \int \mathcal{D}\eta \cdots e^{-\frac{1}{4} \int \eta(t')^2 dt'}$$
Key identity

[Aarts, Seiler, Stamatescu 09][Aarts, James, Seiler, Stamatescu 11]

\[
\int dx dy \mathcal{O}(x + iy) P(x, y; t) = \int dx \mathcal{O}(x) \rho(x; t)
\]

\[
P(x, y; 0) = \rho(x) \delta(y), \quad \rho(x; 0) = \rho(x)
\]

- Probability distribution \( P(x, y; t) = \langle \delta(x - x_\eta(t)) \delta(y - y_\eta(t)) \rangle_\eta \)

- Complex weight \( \rho(x; t) \)
  \[
  \frac{\partial \rho(x; t)}{\partial t} = \frac{\partial}{\partial x} \left( \frac{dS}{dx} + \frac{\partial}{\partial x} \right) \rho(x; t)
  \]
  (Fokker-Planck eq)
  \( \rightarrow e^{-S} \) is a stationary solution

- If the key identity holds and \( \lim_{t \to \infty} P(x, y; t) \) converges to a unique function, then \( \lim_{t \to \infty} \rho(x; t) = e^{-S}/Z \) and

\[
\lim_{t \to \infty} \int dx dy \mathcal{O}(x + iy) P(x, y; t) = \frac{1}{Z} \int dx \mathcal{O}(x) e^{-S}
\]
Argument on the conditions for the CLM to work

- The previous argument for the justification of the CLM uses the continuous Langevin equation from the outset.

- Starting from a discretized Langevin equation, the $\epsilon \to 0$ limit turns out to be subtle.

- The expectation value of the time-evolved observable $\mathcal{O}(z; t)$, which plays an important role in the argument, can be ill-defined.

- We fix these subtleties and establish the argument for the justification. This leads to a simple condition, which tells us whether the results are reliable or not.
Discretized Langevin equation

\[ z^{(\eta)}(t + \epsilon) = z^{(\eta)}(t) + \epsilon z^{(\eta)}(t) + \sqrt{\epsilon} \eta(t) \]

- \( \eta(t) = \eta^{(R)}(t) + i\eta^{(I)}(t) \) is a gaussian noise with probability

\[ \langle \cdots \rangle_\eta \sim \int d\eta^{(R)}(t)d\eta^{(I)}(t) \cdots e^{-\frac{1}{4} \sum_t \left( \frac{1}{N_R} \eta^{(R)}(t)^2 + \frac{1}{N_I} \eta^{(I)}(t)^2 \right)} \]

\[ N_R - N_I = 1 \]

- Expectation value

\[ \langle \mathcal{O}(z^{(\eta)}(t)) \rangle_\eta = \int dxdy \mathcal{O}(x + iy) P(x, y; t) \]
Infinitesimal ($\varepsilon$) time-evolution of the expectation value

$$\langle \mathcal{O}(z^{(\eta)}(t + \varepsilon)) \rangle_{\eta} = \int dxdy \mathcal{O}_\varepsilon(x + iy) P(x, y; t)$$

$$\mathcal{O}_\varepsilon(z) = \int d\eta_R d\eta_I e^{-\frac{1}{4N_R} \eta_R^2 - \frac{1}{4N_I} \eta_I^2} \mathcal{O}(z + \varepsilon v(z) + \sqrt{\varepsilon} (\eta_R + i\eta_I))$$

$\varepsilon$-expansion and $\eta$-integration

$$\mathcal{O}_\varepsilon(z) = \sum_{n=0}^{\infty} \frac{1}{n!} \varepsilon^n : L^n : \mathcal{O}(z)$$

$$L = \left\{ \text{Re}v(z) + N_R \frac{\partial}{\partial x} \right\} \frac{\partial}{\partial x} + \left\{ \text{Im}v(z) + N_I \frac{\partial}{\partial y} \right\} \frac{\partial}{\partial y}$$

$$(f(x) + \partial_x)^2 := f(x)^2 + 2f(x)\partial_x + \partial_x^2$$
For holomorphic $O(z)$

$$L = \left\{ \Re v(z) + N_R \frac{\partial}{\partial x} \right\} \frac{\partial}{\partial x} + \left\{ \Im v(z) + N_I \frac{\partial}{\partial y} \right\} \frac{\partial}{\partial y}$$

$$\frac{\partial}{\partial x} \rightarrow \frac{\partial}{\partial z}, \quad \frac{\partial}{\partial y} \rightarrow i \frac{\partial}{\partial z} \quad N_R - N_I = 1$$

$$\tilde{L} = \left( v(z) + \frac{\partial}{\partial z} \right) \frac{\partial}{\partial z}$$

$$LO(z) \rightarrow \tilde{L}O(z) \quad \text{holomorphic}$$

$$O_\epsilon(z) = \sum_{n=0}^{\infty} \frac{1}{n!} \epsilon^n : \tilde{L}^n : O(z)$$
Subtlety in the $\epsilon$ time-evolution

$$\langle O(\epsilon^{(n)}(t+\epsilon))\rangle_n = \int dxdy O_\epsilon(x + iy) P(x, y; t)$$

$$= \sum_{n=0}^{\infty} \frac{1}{n!} \epsilon^n \int dxdy \left\{ : \tilde{L}^n : O(x + iy) \right\} P(x, y; t)$$

- The integrand involves the n-th power of the drift term
  → The prob. dist. of the drift should fall off faster than any power law at large magnitude

- If the integrals are convergent, we can take the $\epsilon \to 0$ limit

$$\frac{d}{dt} \langle O(\epsilon^{(n)}(t))\rangle_n = \int dxdy \left\{ \tilde{L} O(x + iy) \right\} P(x, y; t)$$
finite ($\tau$) time-evolution and its subtlety

- Repeat the $\varepsilon$ time-evolution

$$\langle \mathcal{O}(z^{(n)}(t + \tau)) \rangle_{\eta} = \sum_{n=0}^{\infty} \frac{1}{n!} \tau^n \int dxdy \left\{ \tilde{L}^n \mathcal{O}(x + iy) \right\} P(x, y; t)$$

- The infinite series should have a finite convergence radius

- The prob. dist. of the drift: $p(u; t) = \int dxdy \delta(u - |v(z)|) P(x, y; t)$

- i.e. $p(u; t) \sim e^{-\kappa u}$

$$\sum_{n=0}^{\infty} \frac{1}{n!} \tau^n \int_0^{\infty} du u^n p(u; t) = \sum_{n=0}^{\infty} \frac{1}{\kappa^n} \left( \frac{\tau}{\kappa} \right)^n$$

- For the finite $t$-evolution of $O(x)$ to be well-defined, the prob. dist. of the magnitude of the drift should fall off exponentially or faster
Proof of the key identity

- We prove the following identity for any $k$ by induction wrt $t$

\[
\int dxdy \left\{ \tilde{L}^k \mathcal{O}(x + iy) \right\} P(x, y; t) = \int dx \left\{ L_0^k \mathcal{O}(x) \right\} \rho(x; t)
\]

\[
\tilde{L} = \left( v(z) + \frac{\partial}{\partial z} \right) \frac{\partial}{\partial z} \quad L_0 = \left( v(x) + \frac{\partial}{\partial x} \right) \frac{\partial}{\partial x}
\]

- At $t=0$, it is trivially satisfied because $P(x, y; 0) = \rho(x)\delta(y), \quad \rho(x; 0) = \rho(x)$

- Assume this holds at some $t$, then for $\tau < \tau_{\text{conv}}$

\[
\sum_{n=0}^{\infty} \frac{1}{n!} \tau^n \int dxdy \left\{ \tilde{L}^{n+k} \mathcal{O}(x + iy) \right\} P(x, y; t) = \sum_{n=0}^{\infty} \frac{1}{n!} \tau^n \int dx \left\{ (L_0)^{n+k} \mathcal{O}(x) \right\} \rho(x; t)
\]

integration by parts & FP eq for $\rho(x; t)$
Demonstration of our condition in simple examples

- A model with a singular drift

\[ Z = \int dx (x + i\alpha)^p e^{-\frac{1}{2}x^2} \]

\[ v(z) = \frac{p}{z + i\alpha} - z \]

- A model with a possibility of excursions

\[ Z = \int dx e^{-\frac{1}{2}(A+iB)x^2 - \frac{1}{4}x^4} \]

\[ v(z) = (A + iB)z - z^3 \]
A model with a singular drift

\[ Z = \int dx (x + i\alpha)^p e^{-\frac{1}{2}x^2} \]

\[ \text{drift: } v(z) = \frac{p}{z + i\alpha} - z \]

singularity at \( z = -i\alpha \)

\[ p = 4 \]

\[ \text{CLM fails for } \alpha \lesssim 3.6 \]
As a model with a singular drift, we consider the partition function \[15\]

3.1 A model with a singular drift

CLM fails, whereas it is exponentially suppressed when the CLM works. Thus the failures of the drift term and show that it is only power-law suppressed at large magnitude when the fixed points, and the filled triangles represent the singular points.

Figure 1: (Left) The real part of the expectation value of \(z^\alpha\) with \(x = 5\) (Left) and \(x = 4\) (Right) with \(\alpha = 6\).

\[\int_{x}^{\infty} \text{exact} \quad \int_{x}^{\infty} \text{CLM}\]

The solid line represents the exact result. (Right) Zoom-up of the same plot in the region 3.

\[\alpha \text{ is a real variable and } \alpha \text{ and } \alpha Sngl. pt. \]

\[\text{Fxd. pt.} \quad \text{Sngl. pt.}\]

\[\text{CLM} \]

\[6\]

\[p = 4, \alpha = 3\]

\[p = 4, \alpha = 5\]

\[\text{Fxd. pt.} \quad \text{Sngl. pt.}\]

\[\text{CLM} \]

\[\alpha = 3\]

\[\alpha = 5\]

many configurations near the singularity

\(\rightarrow\) Large drifts appear

no configurations near the singularity
Demonstration of our condition

\[ u = |v(z)| = \left| \frac{p}{z + i\alpha} - z \right| \]
Demonstration of our condition

[\text{Nagata-Nishimura-SS 16}]

semi-log plot
prob. dist. of the drift

\[ u = |v(z)| = \left| \frac{p}{z + i\alpha} - z \right| \]
Demonstration of our condition

\[ u = |v(z)| = \left| \frac{p}{z + i\alpha} - z \right| \]
Demonstration of our condition

$\alpha$-3d step-size

The step-size $\alpha$ is singular at $\alpha = 3$.

Figure 2: The scatter plot of thermalized configurations (red dots) and the flow diagram (black circles) for $\alpha = 5$ (Left) and $\alpha = 5$ (Right). The data points appear around the singular point.

log-log plot

prob. dist. of the drift

power law fall off

$u = |v(z)| = \left| \frac{p}{z + i\alpha} - z \right|$
A model with a possibility of excursions

\[ Z = \int dx e^{-\frac{1}{2}(A+iB)x^2 - \frac{1}{4}x^4} \]

\[
\text{drift: } v(z) = -(A + iB)z - z^3
\]

\[ \Phi \]

\[ \text{potential danger of excursions in large } |z| \]

\[ \bullet A=1 \]

\[ \text{Nagata-Nishimura-SS 16} \]

\[ \text{CLM fails for } B \geq 3.0 \]
Problem Distribution

[Probable-Nishimura-SS 16]

<table>
<thead>
<tr>
<th>Figure 5:</th>
<th>Figure 6:</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Left) The imaginary part of the expectation value of $A_{\chi}$ as a function of $1 +</td>
<td>B</td>
</tr>
</tbody>
</table>

$B=2$ configurations are restricted in $|y|<C$ $B=4$ configurations spread out into large $|y|$ region $\rightarrow$ Large drifts appear
Demonstration of our condition

\[ u = \left| v(z) \right| = \left| (1 + iB)z - z^3 \right| \]
Demonstration of our condition

\[ u = |v(z)| = |(1 + iB)z - z^3| \]

semi-log plot

prob. dist. of the drift
Demonstration of our condition

\[ \text{[Nagata-Nishimura-SS 16]} \]

\[ \mathbf{u} = |v(z)| = |(1 + iB)z - z^3| \]
Demonstration of our condition

\[ u = |v(z)| = |(1 + iB)z - z^3| \]
Demonstration of our condition in Chiral Random Matrix Theory
CLM for Chiral Random Matrix Theory

[Bloch-Bruckmann-Kieburg-Splitterff-Verbaarschot 13]

\[ Z = \int d\Phi_1 d\Phi_2 d\Psi_1 d\Phi_2 \left[ \det(D + m) \right]^{N_f} e^{-S} \]  

[Mollgaard-Splitterff 13, 15]

[Nagata-Nishimura-SS 16]

- For small masses, naive application of the CLM fails due to the singular drift problem (large drift)  
  [Mollgaard-Splitterff 13]

- By using a gauge cooling, we can avoid the problem and make the CLM work even at small masses  
  [Nagata-Nishimura-SS 16]

![Graphs showing chiral condensate and baryon number density](image-url)
Demonstration of our condition in CRMT

\[ \Phi_1, \Phi_2, \Psi_1, \Psi_2 \rightarrow \text{drift } F_i \quad (i=1,2,3,4) \]

eigenvalues of \( \sqrt{F_i^\dagger F_i} \) \quad (i=1,2,3,4)

without cooling (blue points)
Demonstration of our condition in CRMT

$\Phi_1, \Phi_2, \Psi_1, \Psi_2 \rightarrow \text{drift } F_i \quad (i=1,2,3,4)$

eigenvalues of $\sqrt{F_i^\dagger F_i} \quad (i=1,2,3,4)$

without cooling (blue points)
Demonstration of our condition in CRMT

\[ \Phi_1, \Phi_2, \Psi_1, \Psi_2 \rightarrow \text{drift } F_i \quad (i=1,2,3,4) \]

eigenvalues of \( \sqrt{F_i^\dagger F_i} \quad (i=1,2,3,4) \)

with a cooling (red points)
Demonstration of our condition in CRMT

\( \Phi_1, \Phi_2, \Psi_1, \Psi_2 \rightarrow \text{drift } F_i \) \ (i=1,2,3,4)

eigenvalues of \( \sqrt{F_i^\dagger F_i} \) \ (i=1,2,3,4)

with a cooling (red points)
Summary

- We have established the argument for justification of the CLM by starting with a finite step-size $\epsilon$.

- We find that in order for the CLM to work, the probability distribution of the drift should fall off exponentially or faster.

- We demonstrate the validity of our condition in simple models involving the singular drift problem or the excursion problem and Chiral Random Matrix Theory.

- Importantly, our condition turns out to be practically useful in the CLM for lattice QCD $\rightarrow$ Nagata’s talk (next)