New criterion for correctness in the complex Langevin method - an application to finite density QCD

Keitaro Nagata (KEK)
Hideo Matsufuru (KEK), Jun Nishimura (KEK, Sokendai), Shinji Shimasaki (Keio Univ.)

Reference)
Matsufuru, KN, Nishimura, Shimasaki, work in progress => this talk
KN, Nishimura, Shimasaki, 1606.07627 => Shimasaki’s talk

see also KN, Nishimura, Shimasaki, 1604.07717, 1508.02377, 1511.08580
1. Introduction
Recent progress of complex Langevin method (CLM) enables us to study systems of complex action, e.g. QCD at finite density.

[Aarts, Seiler, Stamatescu (‘10), Seiler, Sexty, Stamatescu (‘13), Sexty (‘13), Fodor, Katz, Sexty, Torok (‘14), and talks in lat’16]
Crucial question:

How do we distinguish whether results in CLM are correct or wrong?
Revisiting the argument for justification of CLM leads us to

a new criterion of correctness
= probability of drift terms

[KN, Nishimura, Shimasaki, 1606.07627, Shimasaki’s talk]
New criterion - probability of drift terms

✓ legitimate
  - obtained from theoretical argument

✓ clear
  - signal is qualitative

✓ cheap
  - no additional calculation required
Purpose of this talk

- Demonstration of effectiveness of the new criterion in a solvable system
  - 2d SU(2) Yang-Mills

- Applications of the criterion to finite density QCD
  - we study QCD at finite density using CLM and determine the reliable range of chemical potential using the new criterion.
2. Framework
CLM for gauge theory, gauge cooling and probability of drift
CLM for gauge theory with $S \in \mathbb{C}$

- Complex Langevin method
  - complexify link variables $U_{n,\mu} \in SU(N) \rightarrow U_{n,\mu} \in SL(N, \mathbb{C})$
  - action and observables $S(U) \rightarrow S(U)$
  - Langevin equation for link variables

$$U_{n,\mu}(t + \epsilon) = e^{iX}U_{n,\mu}(t),$$

$$X = \sum_{a} \lambda_{a}[(\mathcal{D}_{an\mu}S)\epsilon + \sqrt{2\epsilon}\eta_{an\mu}]$$

t : Langevin time

Gaussian white noise (we consider only real noise)

**Drift terms**

generators: e.g. Gell-Mann matrices for SU(3)
Gauge cooling to stabilize CL simulations

• complexified gauge transformation after each Langevin step

\[ U_{n,\mu} \rightarrow g_n U_{n,\mu} g_{n+\hat{\mu}}^{-1} \quad g_n \in \text{SL}(N, \mathbb{C}) \]

• chose \( g_n \) so that it suppresses a norm
  – unitarity norm (we concentrate on u. norm in this work)

\[ N_u = \frac{1}{N_V} \sum_{n,\mu} \text{tr}[U_{n,\mu}^{\dagger} U_{n,\mu} + (U_{n,\mu}^\dagger)^{-1} U_{n,\mu}^{-1} - 2] \]

  H.c. is taken after complexification

  – Justification of the complex Langevin method with the gauge cooling procedure [KN, Nishimura, Shimasaki, 1508.02377]
The probability of the drift terms

- We introduce the probability of drift terms $p(u)$

$$p(u; t) = \int D\mathcal{U} \sum_{n,\mu} \delta (u - u_{n,\mu}(\mathcal{U})) \, P(\mathcal{U}; t)$$

where

$$u_{n,\mu} = \sqrt{\frac{1}{N_c^2 - 1} \sum_{a} v_{an\mu} v_{an\mu}^\dagger}$$

$$v_{an\mu} = D_{an\mu} S, \, (a = 1, 2, \cdots N_c^2 - 1)$$

If the probability $p(u; t)$ falls off exponentially or faster, then the equality holds between the Langevin time average of an observable and its physical expectation value.
3. Application to Solvable system
2d SU(2) YM theory
2d SU(2) Yang-Mills theory

- Framework

\[ Z = \int \mathcal{D}U e^{-S} \]
\[ S = -\frac{\beta}{2N} \sum_x \text{tr}[U_{01}(x) + U_{01}^{-1}(x)] \]

- exactly solvable using character expansion

- sign problem for complex \( \beta \)

- CLM works for small \( \text{Im}[\beta] \), while it fails for large \( \text{Im}[\beta] \)
  [Makino et al.\text{('15)]}
Plaquette in 2d SU(2) YM with $\beta=1.5 \exp(i \theta)$.

**Setup**
- lattice size $4^4$
- $\varepsilon = 10^{-5}$, $t=100$
- one g.c. with unitarity norm
- \( \theta \lesssim 0.6 \): agreement with exact result
- \( \theta \gtrsim 0.6 \): deviation from exact result

\[ \text{Plaquette in 2d SU(2) YM w/ } \beta=1.5 \exp(i \theta). \]

Setup
- lattice size 4\(^4\)
- \( \varepsilon = 10^{-5}, t=100 \)
- one g.c. with unitarity norm

\[ \text{Re}[P], \text{CLM} \]
\[ \text{Re}[P], \text{Exact} \]
\[ \text{Im}[P], \text{CLM} \]
\[ \text{Im}[P], \text{Exact} \]

[Makino et al (‘15)]
• CLM can fail even if the unitarity norm is under control

[Makino et al (’15)]
Probability of drift terms

Plaquette in 2d SU(2) YM w/ $\beta = 1.5 \exp(i \theta)$.
• θ = 0.2, 0.4 : CLM successful : \( p(u) \) falls off exponentially or faster

• θ \( \gtrsim 0.6 \) : CLM fails : power law
Correctness of the results can be distinguished by the probability of drift terms.
Remark: large drift and spikes in norms

- Large drift is correlated with spikes of unitarity norm.
- If the unitarity norm has spikes frequently, results may be wrong. => The probability of drift tell you if it is reliable or not!
- Langevin time should be sufficiently larger than the auto correlation of norms to use the criterion.
4. Application to QCD at finite density
Setup

- We consider finite density QCD at low T with light quarks
  - lattice size: $4^3 \times 8$
  - $Nf=4$ staggered fermion with $m a=0.05$ (we keep mass small to see singular drift problem)

- Langevin setup
  - Langevin time: $t = 10\sim20$ with fixed $\epsilon = 10^{-4}$
  - gauge cooling: $10\sim20$ times
  - we use bilinear noise method with Kogut-Sinclair type improvement [Sinclair’s talk Lat’15]

- Results are preliminary: $t=20$ may not be sufficient to confirm the tail part of the probability of the drift terms.
• Results in CLM agree with those in PQ, but deviation found at $\mu=0.5$.

• Are the results reliable?
- Unitarity norm is almost under control for all the cases.
Dirac Eigenvalues (\(ev(D+m)\))

\(\mu=0.3\)

\(\mu=0.5\)

\(\mu=0.7\)

\(\mu=0.9\)

\(\mu=1.3\)
It is difficult to judge correctness of the results from the unitarity norm and Dirac evs.
• $\mu \leq 0.3$ : fall-off exponentially or faster $\Rightarrow$ reliable
• $\mu = 0.5$ : fall-off exponentially $\Rightarrow$ reliable
• $0.7 \leq \mu$ : power law
Data in semi-log plot
Probability of drift: (L) gauge part, (R) fermion part

\[ \nu = D S_G \]

\[ \nu = D S_F \]

- \( \mu = 0.7, 0.9 \Rightarrow \) singular drift problem
- \( \mu = 1.3 \Rightarrow \) excursion problem
- From the criterion, we can identify the origin of problems
Conclusion

• We proposed a new criterion of correctness of results in CLM by revisiting the argument for justification.
  – The criterion works well for 2d SU(2) (and also cRMT [Shimasaki’talk]).

• We study QCD at low T with light quarks using CLM, we determine the reliable range of chemical potential.
  – singular drift problem occurs at intermediate values of $\mu$.

• [work in progress]
  – application of new norms with Dirac operators to avoid the singular drift problem.
The probability of the drift terms is a reliable criterion of correctness.

It is a legitimate, cheap and clear way of judging correctness.
Back up for 2d SU(2)
Probability of drift – semi-log plot

\[ p(u) \]

\[ u \]

\[ \theta = 0.4 \quad \text{red} \]

\[ \theta = 0.2 \quad \text{green} \]
intermediate value of $\theta$

Plaquette in 2d SU(2) YM w/ $\beta=1.5 \exp(i \theta)$.

- Our criterion: CLM is correct for the fall off exponentially or faster.
- It is possible that the CLM is correct even if it has power law.
Back up slide data at $\mu=0.5$
The new criterion suggests the data at $\mu = 0.5$ is reliable.

- At $\mu=0.5$: the probability falls off faster than power law.
Tail part of the distribution

$p(u)$

$log(u)$

$\mu = 0.1$

$\mu = 0.3$

$\mu = 0.5$

$\mu = 0.7$

$\mu = 0.9$

$\mu = 1.3$
• The data in semi-log plot seems to fit with linear
• The probability falls off exponentially.
Data at $\mu = 0.7$ and $0.9$ in semi-log plot

- The data in semi-log plot seems to fall off slower than linear.
- The probability falls off slower than exponentially.
The data in semi-log plot seems to fall off slower than linear.
The probability falls off slower than exponentially.
$T_{\text{therm}}$ dep-Data at $\mu = 0.5$ in semi-log-plot

- Data for $\mu = 0.5$ with $t_{\text{therm}} = 4$ and 6
- The plot implies the exponential damp.
  - the data is reliable.
Unitariry norm and anti-hermiticity

- We found the similarity between unitarity norm and anti-hermiticity norm

\[ N_u \equiv \frac{1}{4N_V} \sum_{x,\nu} \text{tr}[(U_{x\nu})^\dagger U_{x\nu} + (U^{-1}_{x\nu})^\dagger U^{-1}_{x\nu} - 2]. \]

\[ \mathcal{N}_{\text{a.h.}} = \frac{1}{4N_V} \text{tr} (D + D^\dagger)(D + D^\dagger)^\dagger \]

\[ = \frac{2}{4N_V} \sum_{x,\sigma} \text{tr}_c \left[ e^{2\mu a \delta_{4,\sigma}} U_{\sigma}^\dagger(x) U_{\sigma}(x) + e^{-2\mu a \delta_{4,\sigma}} (U_{\sigma}^{-1}(x - \hat{\sigma}))^\dagger U_{\sigma}^{-1}(x - \hat{\sigma}) - 2 \mathbf{1}_{3\times3} \right] \]

At \( \mu=0 \), they are equivalent.
Cooling for unitarity norm has an effect to reduce the anti-hermiticity norm.

 Cooling for new norm may extend the applicable range further, work in progress.
Back up for $4^3 \times 8$
Anti-Hermiticity norm

- $4^3 \times 8$

![Graph showing the dependence of a system on different µa values over Langevin time. The graph is a log-log plot with axes labeled as follows:

- Y-axis: logarithmic scale ranging from $10^{-2}$ to $10^2$
- X-axis: linear scale ranging from 0 to 25

Legend:
- $\mu_a=0.1$
- $\mu_a=0.3$
- $\mu_a=0.5$
- $\mu_a=0.7$
- $\mu_a=0.9$
- $\mu_a=2.0$]
Back up – framework
Sign problem

\[ Z = \int \prod_{k} dx_k e^{-S(x)} \]

- Monte Carlo method with importance sampling
  - powerful tool to solve path-integrals non-perturbatively

- Importance sampling breaks down if \( S \) is complex
  - QCD at finite density or with theta-term
  - Chern-Simons gauge theories
  - Hubbard model away from half-filling
(real) Langevin Method (LM)

\[ Z = \int \prod_k dx_k \exp(-S(x)), \quad (x_k, S \in \mathbb{R}) \]

- Generation of ensemble using Langevin eq. (stochastic quantization)

\[ \frac{dx_k^{(\eta)}}{dt} = -\frac{\partial S}{\partial x_k^{(\eta)}} + \eta_k(t) \]

- Ensemble is generated by the stochastic equation rather than importance sampling

\[ \text{\textit{t}: parameter (Langevin time)} \]
\[ \text{\textit{\eta}: Gaussian white noise} \]

Parisi-Wu 1981
proof of LM

- Average of an observable in LM is given by

\[
\langle O(x^{(\eta)}(t)) \rangle_\eta = \int dx\ O(x)P(x; t)
\]

\[
P(x; t) = \left\langle \prod_k \delta(x_k - x_k^{(\eta)}(t)) \right\rangle_\eta
\]

\[
\langle \cdots \rangle_\eta = \frac{\int \mathcal{D}\eta \cdots e^{-\frac{1}{4} \int d\tau \eta^2}}{\int \mathcal{D}\eta e^{-\frac{1}{4} \int d\tau \eta^2}}
\]

- According to the Fokker-Planck equation, P converges to

\[
\lim_{t \to \infty} P(x; t) \propto e^{-S(x)}
\]

- Average of the observable converges to

\[
\lim_{t \to \infty} \langle O(x(t)) \rangle_\eta = \lim_{t \to \infty} \int \prod dx_k\ O(x)P(x; t),
\]

\[
\propto \int \prod dx_k\ O(x)e^{-S(x)}\quad \text{physical expectation value}
\]
Complex Langevin method (CLM)

• Stochastic quantization is available for complex action
  – LM is free from the probability interpretation of $\exp(-S)$
  – however, complexification is inevitable

  [Parisi('83), Klauder('83)]

• CLM
  – extend originally real variables to complex
    \[ x \in \mathbb{R} \rightarrow z = x + iy \in \mathbb{C} \]
  – extend also action and observables in a holomorphic manner
    \[ S(x) \rightarrow S(z) = S(x + iy) \]
  – Langevin equation
    \[ \frac{\partial z}{\partial t} = -\frac{\partial S}{\partial z} + \eta(t) \]
    (noise term can be complex. However, we prefer to use real noise throughout this talk)
Problem of convergence

• In CLM, $P(x; t)$ in equilibrium is not ensured to converge to correct limit

$$\int dx dy \ O(x + iy) P(x, y; t) = \int dx \ O(x) \rho(x; t)$$

$$\lim_{t \to \infty} \rho(x; t) = e^{-S(x)}$$

– CLM works well for some cases, but fails for other.
– there had been no criteria to distinguish if results in CLM are correct or not.
Justification of CLM / criteria of correctness

- CLM is justified if some conditions are satisfied
  
  \[ \int dxdy O(x + iy) P(x, y; t) = \int dx O(x) \rho(x; t) \]

  \[ \lim_{t \to \infty} \rho(x; t) = e^{-S(x)} \]

  - fast fall-off of the probability distribution in the imaginary direction
  - holomorphy of action and observables

This argument also tells what causes the failure of the CLM.

(“Revisit the argument of justification”, KN, Nishimura, Shimasaki in preparation.)
Advantages of CLM

- CLM overcomes several points which are serious difficulties for approaches based on importance sampling
  - CLM is possible even if the phase fluctuation is very large
  - exponential increase of numerical cost does not occur

\[ \mu = 0 \] (or imaginary-, isospin-\( \mu \))

```
[\text{e.g. KN, Nakamura, PRD91, (2015) no9. 094507,}
```
Back up for $4^4$
Chiral condensate and number density

![Graph showing chiral condensate and number density vs. \( \mu/T \).]
Unitarity norm

Langevin time

$\mu a = 0.1$
$\mu a = 0.3$
$\mu a = 0.5$
$\mu a = 0.7$
$\mu a = 0.9$
$\mu a = 2.0$
New types of norm for singular drift problem

• CLM fails when such Dirac eigenvalues appear that the fermion drift becomes singular [Mollgaard & Splittorff, Greensite]

• We showed that the singular drift problem can be avoided by choosing suitable norm in RMT
  – norms including Dirac operator
  – e.g. anti-hermiticity norm

\[ N_{\text{a.h.}} = \frac{1}{N_V} \text{tr}[(D + D^\dagger)^2] \]