The path optimization for the sign problem of low dimensional QCD

Yuto Mori (Kyoto Univ.)

Kouji Kashiwa (Fukuoka Inst. Tech.)
Akira Ohnishi (YITP)
Content

1. Sign problem
2. Path Optimization
3. 0+1 dimensional QCD
   1. w/ diagonal gauge fixing
   2. w/o diagonal gauge fixing (one link variable)
4. 1+1 dimensional QCD (preliminary)
5. Summary
Sign problem

- When $S \in \mathbb{C}$, serious cancellation occurs in integration at large volume. And, integrals cannot be obtained precisely.

\[ Z = \int \mathcal{D}x e^{-\Re S - i \Im S} \ll \int \mathcal{D}x e^{-\Re S} \]

- Seriousness of the sign problem
  - average phase factor (APF)

\[
\text{APF} = \frac{\int \mathcal{D}x e^{-\Re S} e^{-i \Im S}}{\int \mathcal{D}x e^{-\Re S}} \\
\sim e^{-\beta V \Delta f} \sim 0
\]

ex.) $e^{-S(x)} = (x + 10i)^{50} e^{-x^2/2}$

Path Optimization

Optimize the integral path in the complexified variable space to weaken the sign problem. (Integral of holomorphic (analytic) function is independent of integral path.)

We can regard the sign problem as an optimization problem.

(ex.) One variable case
- Trial function (integral path)
  \[ z(\cdot) : \mathbb{R} \to \mathbb{C} \]
- Cost function (function to minimize)
  \[ \mathcal{F}[z(t)] = |Z| \{|\text{APF}|^{-1} - 1\} \]

oscillation of the integrand becomes weaker.

Path Optimization

Optimize the integral path in the complexified variable space to weaken the sign problem. (Integral of holomorphic(analytic) function is independent of integral path.)

We can regard the sign problem as an optimization problem.

(ex.) One variable case

- Trial function (integral path)
  \[ z(\cdot) : \mathbb{R} \to \mathbb{C} \]
- Cost function (function to minimize)
  \[ F[z(t)] = |Z|\{|\text{APF}|-1\} \]

oscillation of the integrand becomes weaker.
Neural Network for field theories

Neural Network (NN) is powerful to represent any functions of many inputs.

- Combination of linear and non-linear transformation
  
  \[ a_i = g(W^{(1)}_{ij} t_j + b^{(1)}_i) \]
  \[ f_i = g(W^{(2)}_{ij} a_j + b^{(2)}_i) \]
  \[ g(x) : \text{Activation fn. (ex. tanh)} \]

※ \( W, b, \alpha, \beta \) : parameters

- Any fn. can be reproduced at (\# of units of hidden layer) \( \to \infty \) (Universal approximation theorem)
  
  G. Cybenko, MCSS 2, 303 (1989)
  K. Hornik, Neural networks 4, 251 (1991)

- We input the real part of variables, and obtain the imaginary part from outputs.
  \[ z_i(t) = t_i + i(\alpha_i f_i(t) + \beta_i) \]
0+1 dim. QCD

Application to gauge theories

-1-spiecies of Staggered fermion

\[ S = \frac{1}{2} \sum_{\tau} (\bar{\chi}_{\tau} e^{\mu U_{\tau}} \chi_{\tau+\hat{0}} - \bar{\chi}_{\tau+\hat{0}} e^{-\mu U_{\tau}^{-1}} \chi_{\tau}) + m \sum_{\tau} \bar{\chi}_{\tau} \chi_{\tau} \]

\[ Z = \int D\bar{U} \det D[U] = \int d\bar{U} \det[X_N + (-1)^N e^{\mu/T} U + e^{-\mu/T} U^{-1}] \]

\[ X_N = 2 \cosh(E/T), \quad E = \arcsinh m, \quad T = 1/N \]

One link variable, No plaquette
0+1 dim. QCD

There are two ways of complexifying link variables.

1) Complexification after diagonal gauge fixing

\[ U = \text{diag}(e^{ix_1}, e^{ix_2}, e^{ix_3}), \ (x_3 = -(x_1 + x_2)) \]

\[ Z = \int dx_1 dx_2 H(x) \exp(-S(x)) \]

Haar measure

\[ H(x) = \frac{8}{3\pi^2} \prod_{a<b} \sin^2 \left( \frac{x_a - x_b}{2} \right) \]

\[ e^{-S(x)} = \prod (X_N + 2 \cos(x_a - i\mu)) \]

\[ x_a \in \mathbb{R} \rightarrow z_a \in \mathbb{C} \quad z_a(x) = x_a + iy_a(x), \ y_a^q(x) : \text{trial function} \]

2) Complexification without diagonal gauge fixing

\[ U \in SU(3) \rightarrow U \in SL(3, \mathbb{C}) \]

\[ \text{ex.} \quad U(U) = U \prod_a \exp(y_a \lambda_a) \quad y_a(U) : \text{trial function} \]
1 dim. QCD diagonal gauge

Complexification after diagonal gauge fixing

\[ Z = \int dz_1 dz_2 H(z)e^{-S(z)} = \int dx_1 dx_2 JH(z(x))e^{-S(z(x))} \]

Optimization of \( z \) by gradient descent method

\[ m = 0.05, \ T = 0.5, \ \mu/T = 1.0 \]

\[ W = H e^{-S} \]

Phase of \( J \) cancels that of \( W \)

---

\( Y, M, Kashiwa, Ohnishi, \)
\( arXiv:1904.11140 \)
1 dim. QCD diagonal gauge

① Complexification after diagonal gauge fixing

\[ Z = \int dz_1 dz_2 H(z)e^{-S(z)} = \int dx_1 dx_2 J H(z(x))e^{-S(z(x))} \]

Optimization of \( z \) by gradient descent method

\[ m = 0.05, \ T = 0.5, \ \mu/T = 1.0 \]

\[ H(x) = \frac{8}{3\pi^2} \prod_{a < b} \sin^2 \left( \frac{x_a - x_b}{2} \right) \]

→ Separation of probability distribution. Difficulty in HMC.
1 dim. QCD one link

2 Complexification without diagonal gauge fixing

\[ U(U) = U \prod\limits_{a} e^{y_{a} \lambda_{a}/2} \]

\( y_{a}(U) \) : trial function

We generate the configurations by HMC and optimize NN parameters by stochastic gradient descent (SGD, Adadelta).

- 1 – average phase factor (APF)

After some steps of optimization, 1-APF becomes 3~100 times smaller than before.
1 dim. QCD one link

- **Eigenvalue distribution**

\[ PUP^{-1} = \text{diag}(e^{iz_1}, e^{iz_2}, e^{-i(z_1+z_2)}) \]

Six separated region are produced by HMC

- **Expectation value**

Consistent with exact solutions.

\[ m = 0.05, \quad T = 0.5, \quad \mu/T = 1.0 \]
1D QCD

- Action with staggered fermion (strong coupling)

\[ S = \bar{\chi}_x D_{xy} \chi_y \]

\[ D_{xy} = m\delta_{x,y} + \frac{1}{2} (-1)^{x_0} \{ \delta_{x+\hat{1},y} U_{1,x} - \delta_{x,y+\hat{1}} U_{1,y}^{-1} \} \]

\[ + \frac{1}{2} \{ \delta_{x+\hat{0},y} e^\mu U_{0,x} - \delta_{x,y+\hat{0}} e^{-\mu} U_{0,y}^{-1} \} \]

- Polyakov gauge (without diagonal gauge fixing)

\[ U_{0,x} \begin{cases} = 1 & (x_0 \neq N_\tau) \\ \in SU(3, \mathbb{C}) & (x_0 = N_\tau) \end{cases} \]

- Complexification

\[ U_{\nu,x} \in SU(3) \rightarrow U_{\nu,x}(U) \in SL(3, \mathbb{C}) \]
$2\times 2$ Lattice, $N_f = 2, m = 0.1$

**Unitary norm after opt.**

$$N = \max_{x,\nu} \text{tr}(U_{\nu,x}^\dagger U_{\nu,x} - 1)^2$$

Average phase factor can be slightly enhanced by the deformation of manifold.

…Is there upper bound of APF?

---

**Average phase factor**

![Graph showing the change in APF with varying $\mu$.](image)
Summary

- In the path optimization method, we can regard the sign problem as an optimization problem.
  - Neural Network and optimization methods (SGD) developed in machine learning are helpful.

- We apply this method to 0+1 dim. QCD with and without diagonal gauge fixing.
  - Average phase factor becomes large and exact results are reproduced in observable calculations.
  - Without gauge fixing, eigenvalue distribution agrees with gauge fixed results, and HMC works well.

- In 1+1D QCD, works in progress.