

The path optimization for the sign problem of low dimensional QCD*

Yuto Mori[†]

Department of Physics, Faculty of Science, Kyoto University, Kyoto 606-8502, Japan

E-mail: mori.yuto.47z@st.kyoto-u.ac.jp

Kouji Kashiwa

Fukuoka Institute of Technology, Wajiro, Fukuoka 811-0295, Japan

E-mail: kashiwa@fit.ac.jp

Akira Ohnishi

Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan

E-mail: ohnishi@yukawa.kyoto-u.ac.jp

The path optimization has been proposed to weaken the sign problem which appears in some field theories such as finite density QCD. In this method, we optimize the integration path in complex plain to enhance the average phase factor. In this study, we discuss the application of this method to low dimensional QCD as a first step of finite density QCD.

37th International Symposium on Lattice Field Theory - Lattice2019

16-22 June 2019

Wuhan, China

*Report No.: KUNS-2780, YITP-19-110

[†]Speaker.

1. Introduction

The sign problem in the path integral is a numerical difficulty that appears in system with complex actions. Complex action causes rapid oscillation and serious cancellation of the Boltzmann weight in calculating the partition function, and this cancellation makes the numerical accuracy worse. In lattice QCD at finite chemical potential μ , Fermion determinant satisfies $(\det D(\mu))^* = \det D(-\mu^*)$, and generally takes a complex value. So the sign problem is one of the most important issues because this makes it difficult to investigate finite density QCD from the first principles.

Recently, in order to avoid or evade the sign problem, some complexified variable methods such as the complex Langevin method [1, 2, 3], the Lefschetz thimble method [4, 5, 6, 7] and the path optimization method [8, 9, 10, 11] have been proposed. In the Lefschetz thimble method, one complexifies the dynamical variables and give the integration path (manifold) by solving the flow equation from the fixed point or deforming original real path by the flow equation.

On the other hand, in the path optimization method, the integration path is given by minimization of the cost function, which represents the seriousness of the sign problem. This method has been applied to a one dimensional integral [8], a finite density complex scalar field theory [9], and the Polyakov loop extended Nambu-Jona-Lasinio model with and without vector type interaction [10, 11]. In these cases, the average phase factor (APF) is enhanced significantly.

In this proceedings, we apply the path optimization method to 0+1D QCD [12] and 1+1D QCD at finite chemical potentials as a first step towards the finite density QCD on four dimensional lattice. It should be noted that the low dimensional QCD at finite densities have been studied by using the Complex Langevin method (0+1D and 1+1D) [13] and by using Lefschetz thimble method [14].

2. Path Optimization method

We consider the partition function given by the integral of the complex Boltzmann weight on the complexified integration path $z(x)$,

$$\mathcal{Z} = \int_{\mathcal{C}_{\mathbb{R}}} d^N x \exp(-S(x)) = \int_{\mathcal{C}_{\mathbb{C}}} d^N z \exp(-S(z)) = \int_{\mathcal{C}_{\mathbb{R}}} d^N x J(z) \exp(-S(z(x))), \quad (2.1)$$

where $J(z) = \det(\partial z / \partial x)$ is the Jacobian. Here we note that the integral of analytic function is independent of integration path from the Cauchy integral theorem.

To obtain the optimized integration path, we first give the integration path by a trial function containing variational parameters, and next we evaluate the cost function which represents the degree of the sign problem. We adopt the following cost function,

$$\mathcal{F}[z(x)] = \int_{\mathcal{C}_{\mathbb{R}}} d^N x |J(z) \exp(-S(z(x)))| - |\mathcal{Z}| \quad (2.2)$$

$$= |\mathcal{Z}| \{ |\langle e^{i\theta} \rangle|^{-1} - 1 \}, \quad (2.3)$$

where θ is the phase of the statistical weight $\theta = \arg(J(z) \exp(-S(z)))$ and $\langle e^{i\theta} \rangle$ is APF. In the third step, we optimize the variational parameters to reduce the cost function. Hence we can increase APF which determines the precision in phase reweighting method. Finally, we generate field configurations by the Monte-Carlo method on the optimized path.

We use the neural network to represent the integration path theories with many degree of freedom. The neural network is a powerful tool to represent any functions of many inputs. Neural network with mono-hidden layer is composed of the input layer, hidden layer, and output layer. The variables given to a layer are passed to the next layer after the linear and non-linear transformation,

$$a_i = g(W_{ij}^{(1)} x_j + b_i^{(1)}), \quad y_i = \alpha_i g(W_{ij}^{(2)} a_j + b_i^{(2)}), \quad (2.4)$$

where x_i and y_i are input and output variables respectively, $g(x)$ is called activation function and we use $g(x) = \tanh(x)$, and W, b, α are the variational parameters. By using these transformation, any function is known to be represented in the large number limit of hidden layer units [15].

Inputs of the neural network are the real integral variables, and it is natural to take variables before the complexification. For example, in the scalar field theory, we input the real part of variables x_i , and obtain the imaginary part from outputs of the neural network y_i . For the SU(3) gauge theory, an SU(3) link variable U is complexified to an SL(3) matrix \mathcal{U} . So we input the matrix elements of U and obtain the complexified link variables, $\mathcal{U} \in \text{SL}(3, \mathbb{C})$ as follows,

$$\mathcal{U}(U) = U \prod_{i=1}^8 e^{\lambda_i y_i (\text{Re } U, \text{Im } U)}, \quad (2.5)$$

where λ_i are the generators of SU(3).

3. 0+1D QCD

We here discuss 0+1D lattice QCD based on the results in Ref. [12]. The action of 0+1D lattice QCD with one species of staggered fermion consists of temporal fermion hopping terms and mass terms,

$$S = \frac{1}{2} \sum_{\tau=1}^{N_\tau} (\bar{\chi}_\tau e^{\mu} U_\tau \chi_{\tau+1} - \bar{\chi}_{\tau+1} e^{-\mu} U_\tau^{-1} \chi_\tau) + m \sum_{\tau} \bar{\chi}_\tau \chi_\tau, \quad (3.1)$$

where N_τ is lattice size, m and μ are mass and chemical potential respectively. The partition function is obtained as

$$\mathcal{Z} = \int \mathcal{D}U \det(D(U)) = \int dU \det \left[X_{N_\tau} + (-1)^{N_\tau} e^{\mu/T} U + e^{-\mu/T} U^{-1} \right] \quad (3.2)$$

$$X_{N_\tau} = 2 \cosh(E/T), \quad E = \text{arcsinh } m, \quad U = U_1 U_2 \cdots U_{N_\tau}, \quad T = 1/N_\tau. \quad (3.3)$$

This theory can be reduced to a one-link variable problem by using the gauge transformation. It should be noted that no plaquette terms appear because there is no spatial direction. Fermion determinant becomes complex at finite chemical potential.

We can further reduce the number of variables from eight in the SU(3) link variable to two real variables $x_i \in [0, 2\pi]$ ($i = 1, 2$) by using the diagonalized gauge fixing, $U = \text{diag}(e^{ix_1}, e^{ix_2}, e^{-i(x_1+x_2)})$. So we demonstrate the two ways of complexifying link variables. In the first way, we complexify the variables after diagonalized gauge fixing, we complexify x_i by Eq. (2.4). In the second method, we complexify the link variable without diagonalized gauge fixing, link variable is complexified by Eq. (2.5).

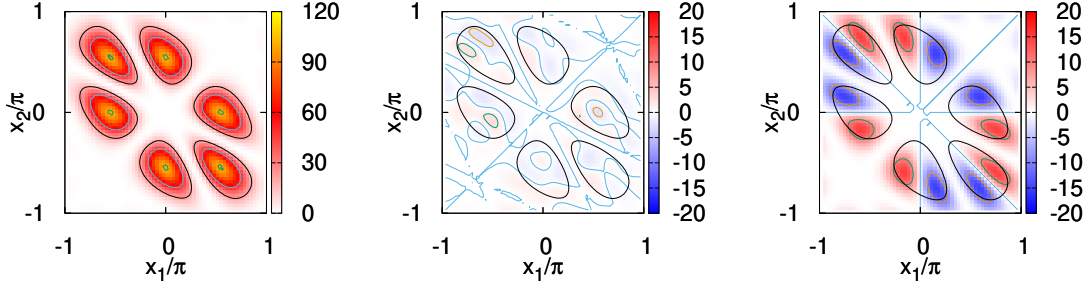


Figure 1: Absolute value of the statistical weight distribution $|JW|$ with the optimization (left), and imaginary part of the statistical weight distribution $\text{Im}(JW)$ with (middle) and without (right) the optimization. Black curves show the contour of $|JW| = 20$. The results are calculated at $\mu/T = 1.0$.

Let us first discuss the results of complexification after diagonalized gauge fixing. We optimize the function $z(x)$ by the gradient descent method to minimize the cost function. Figure 1 shows the statistical weight, the Jacobian times the Boltzmann factor Je^{-S} , as functions of x_1 and x_2 with and without the optimization [12]. We find that there are six separated regions having large statistical weights, which will be discussed later. The imaginary part of the statistical weight is suppressed by the optimization, as seen by the white color showing $\text{Im}(JW) \simeq 0$ in the middle panel.

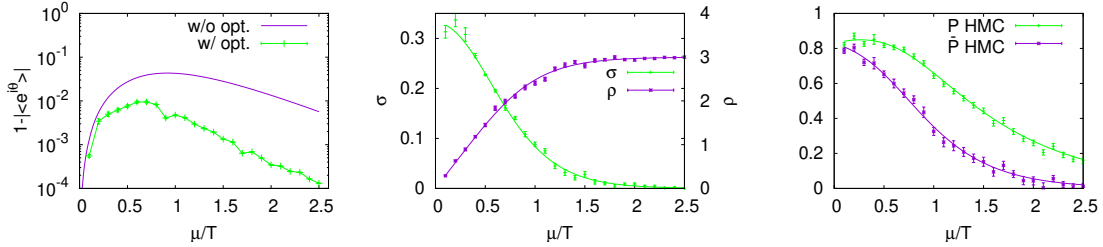


Figure 2: The average phase factor (left), the expectation value of quark condensate and quark density (middle), and that of Polyakov loop (right).

Next, we discuss the results of Hybrid Monte-Carlo (HMC) calculation without diagonalized gauge fixing. We generate the configurations by HMC and optimize the variational parameters by the stochastic gradient descent method. Specifically we use ADADELTA [16] as an optimizer to minimize the cost function evaluated using the generated configurations. Figure 2 shows APF and the expectation value of observables as a function of μ . After some steps of optimization, $1 - |\langle e^{i\theta} \rangle|$ becomes 3 to 100 times smaller than before. So we can weaken the sign problem in this theory. The expectation values are compared with the exact solutions, and results agree with exact results within the error.

Finally, in Fig. 3 we compare the eigenvalue distribution of the results from two different ways of complexifications. We diagonalize the link variables \mathcal{U} generated by HMC without diagonalized

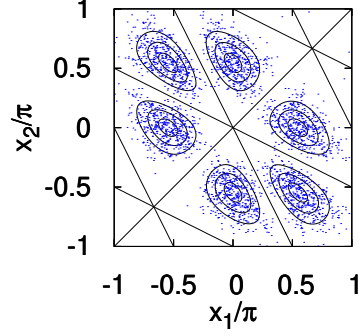


Figure 3: The scatter plot of eigenvalue of the link variable \mathcal{U} . White curves represents the contour of $|Je^{-S}|$ with diagonalized gauge fixing results. The results are calculated at $\mu/T = 1.0$.

gauge fixing,

$$\mathcal{U} = P \text{diag}(e^{iz_1}, e^{iz_2}, e^{-i(z_1+z_2)})P^{-1}, \quad (3.4)$$

where P is a regular matrix. As a result, the real part of z_1 and z_2 are distributed in six separated regions. This distribution is consistent with that with diagonalized gauge fixing which is shown by the contour. Thus the separated regions are found to be connected by the continuous transformation and do not cause the multimodal problem.

4. 1+1D QCD

Let us now proceed to discuss 1+1 dimensional QCD. We use one species of staggered fermion without rooting and take the strong coupling limit. We take the Polyakov gauge without diagonalized gauge fixing, where $U_{0,x} = I$ for $x_\tau \neq N_\tau$ with I being the identity matrix, and we complexify the link variables as in 0+1 dimensional QCD but in a more symmetric form as,

$$\mathcal{U}_{\mu,x} = U_{\mu,x} \frac{H_{\mu,x}(U)}{\sqrt[3]{H_{\mu,x}(U)}}, \quad (4.1)$$

$$H_{\mu,x}(U) = I + \sum_{i=1}^8 \lambda_i y_{\mu,x,i}(\text{Re } U, \text{Im } U), \quad (4.2)$$

except for $\mu = 0$, $x_\tau \neq N_\tau$. Here $y_{\mu,x,i}$ are the outputs of neural network, and inputs of neural network are the real and imaginary part of the link variables.

We show the results on the 2^2 lattice and the bare mass is taken to be 0.1. The left panel of Fig. 4 shows the histogram of the unitary norm defined as,

$$N = \max_{\mu,x} \text{tr}(\mathcal{U}_{\mu,x}^\dagger \mathcal{U}_{\mu,x} - 1)^2. \quad (4.3)$$

The finite unitary norm means the some deformations from $SU(3)$ are caused by the optimization.

The right panel of Fig. 4 shows APF. Original APF of this theory is smaller than that in 0+1D QCD, and have dip structure at finite chemical potential. By the optimization, APF can be enhanced. But the increase is not significant, and then the upper bound of APF may exist. In order to confirm the latter point, we need to compare the results with that of Lefschetz thimble method and to clarify the reasons of the existence of the upper bound in the future work.

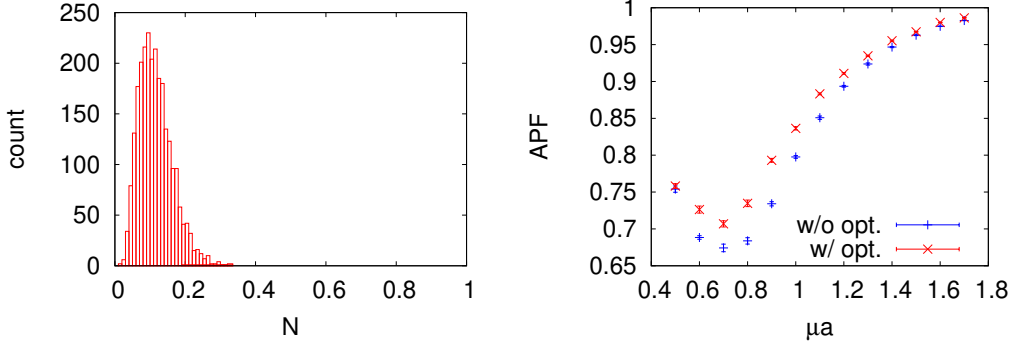


Figure 4: The unitary norm at $\mu a = 0.7$ (left), and the average phase factor with and without optimization (right).

5. Summary

In the path optimization method, we can regard the sign problem as an optimization problem. Neural network and optimization methods developed in machine learning are helpful. We apply this method to 0+1D QCD with and without diagonalized 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 the hybrid Monte-Carlo sampling has been demonstrated to work well, while the statistical weight distribution has six separated regions and multimodal problem can be expected. In 1+1D QCD, the average phase factor is found to increase by the optimization, but we have not yet found the path having a large average phase factor, e.g. larger than 0.9 in some regions of the chemical potential.

Acknowledgments

This work is supported in part by the Grants-in-Aid for Scientific Research from JSPS (Nos. 18J21251, 18K03618, 19H01898, and 19H05151), and by the Yukawa International Program for Quark-hadron Sciences (YIPQS).

References

- [1] G. Parisi, *ON COMPLEX PROBABILITIES*, *Phys. Lett.* **B131** (1983) 393.
- [2] G. Aarts, E. Seiler and I.-O. Stamatescu, *The Complex Langevin method: When can it be trusted?*, *Phys. Rev.* **D81** (2010) 054508 [0912.3360].

- [3] K. Nagata, J. Nishimura and S. Shimasaki, *Argument for justification of the complex Langevin method and the condition for correct convergence*, *Phys. Rev.* **D94** (2016) 114515 [1606.07627].
- [4] E. Witten, *Analytic Continuation Of Chern-Simons Theory*, *AMS/IP Stud. Adv. Math.* **50** (2011) 347 [1001.2933].
- [5] AURORASCIENCE collaboration, *New approach to the sign problem in quantum field theories: High density QCD on a Lefschetz thimble*, *Phys. Rev.* **D86** (2012) 074506 [1205.3996].
- [6] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu and T. Sano, *Hybrid Monte Carlo on Lefschetz thimbles - A study of the residual sign problem*, *JHEP* **1310** (2013) 147 [1309.4371].
- [7] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway and N. C. Warrington, *Sign problem and Monte Carlo calculations beyond Lefschetz thimbles*, *JHEP* **05** (2016) 053 [1512.08764].
- [8] Y. Mori, K. Kashiwa and A. Ohnishi, *Toward solving the sign problem with path optimization method*, *Phys. Rev.* **D96** (2017) 111501 [1705.05605].
- [9] Y. Mori, K. Kashiwa and A. Ohnishi, *Application of a neural network to the sign problem via the path optimization method*, *PTEP* **2018** (2018) 023B04 [1709.03208].
- [10] K. Kashiwa, Y. Mori and A. Ohnishi, *Controlling the model sign problem via the path optimization method: Monte Carlo approach to a QCD effective model with Polyakov loop*, *Phys. Rev.* **D99** (2019) 014033 [1805.08940].
- [11] K. Kashiwa, Y. Mori and A. Ohnishi, *Application of the path optimization method to the sign problem in an effective model of QCD with a repulsive vector-type interaction*, *Phys. Rev.* **D99** (2019) 114005 [1903.03679].
- [12] Y. Mori, K. Kashiwa and A. Ohnishi, *Path optimization in 0+1 dimensional QCD at finite density*, *PTEP* **2019** (2019) 113B01 [1904.11140].
- [13] G. Aarts and K. Splittorff, *Degenerate distributions in complex Langevin dynamics: one-dimensional QCD at finite chemical potential*, *JHEP* **08** (2010) 017 [1006.0332].
- [14] F. Di Renzo and G. Eruzzi, *One-dimensional QCD in thimble regularization*, *Phys. Rev.* **D97** (2018) 014503 [1709.10468].
- [15] G. Cybenko, *Approximation by superpositions of a sigmoidal function*, *Mathematics of Control, Signals, and Systems (MCSS)* **2** (1989) 303.
- [16] M. D. Zeiler, *Adadelta: an adaptive learning rate method*, 1212.5701.