Path optimization for $U(1)$ gauge theory with complexified parameters

Kouji Kashiwa[1,*] and Yuto Mori[2,†]

[1]*Fukuoka Institute of Technology, Wajiro, Fukuoka 811-0295, Japan*
[2]*Department of Physics, Faculty of Science, Kyoto University, Kyoto 606-8502, Japan*

In this article, we apply the path optimization method to handle the complexified parameters in the $1 + 1$ dimensional pure $U(1)$ gauge theory on the lattice. Complexified parameters make it possible to explore the Lee-Yang zeros which helps us to understand the phase structure and thus we consider the complex coupling constant with the path optimization method in the theory. We clarify the gauge fixing issue in the path optimization method; the gauge fixing helps to optimize the integration path effectively. With the gauge fixing, the path optimization method can treat the complex parameter and control the sign problem. It is the first step to directly tackle the Lee-Yang zero analysis of the gauge theory by using the path optimization method.

I. INTRODUCTION

Exploring the phase structure of theories and models with finite external parameters such as the temperature (T), the chemical potential (μ) and the external magnetic field (B) is an important subject to understand our Universe. For example, the phase structure of quantum chromodynamics (QCD) at finite T, μ and B is directly related to the early Universe, current heavy ion collision experiments, neutron star physics and so on; see Ref. [1] as an example.

One of the interesting approaches to investigate the phase structure is the Lee-Yang zero analysis [2,3]. In the analysis, we complexify external parameters and search zeros of the partition function in the complex plane of the external parameters. Then, an approaching tendency of zeros to the real axis indicates the existence of the phase transition because singularities of the partition function are the origin of the ordinary phase transition. Particularly, the experimental observation [4] and the quantum computation by using a quantum computer [5] for the zeros are currently possible and thus the analysis has attracted much more attention recently.

There have been some attempts to perform the Lee-Yang zero analysis in the gauge theory; an interesting example is QCD with the complexified μ [6–8]. In the calculation, one first gathers numerical data at finite imaginary chemical

potential (μ_I) and after they construct the grand canonical partition function with the complex μ by using the Fourier transformation and the fugacity expansion; see Ref. [9]. However, the Fourier transformation becomes much more difficult with decreasing T because the oscillation becomes severer and thus this approach cannot tell us the phase structure at low T; for example, see Ref. [10]. The reason why we use the imaginary chemical potential to perform the Lee-Yang zero analysis in QCD is that there is the sign problem at complexified external parameters and then the Monte-Carlo calculation sometimes fails; see Ref. [11] for details of the sign problem and Ref. [12] for details of the imaginary chemical potential. If we can directly perform the Monte-Carlo calculation with finite complexified parameters, there is the possibility that we can better understand properties of the phase structure via the Lee-Yang zero analysis. In addition, such complexification of chemical potential may be related to the investigation of the confinement-deconfinement transition at finite density [13,14].

In the Lefschetz thimble and path optimization methods, dynamical variables are complexified and then the integral path and/or the configurations are generated such that those obeying the manifold on which the sign problem is weak. Since these approaches can weaken the sign problem and thus it is natural to expect that these approaches can be applied to explore the system with complexified external parameters. In this study, we concentrate on the application of the path optimization method to the system with complexified parameters.

The path optimization method is based on the standard path integral formulation with the complexification of dynamical variables [15,16]; the actual procedure is performed as follows:

(1) The cost function, which reflects the seriousness of the sign problem, is prepared.

*[*]kashiwa@fit.ac.jp
[†]mori@ruby.scphys.kyoto-u.ac.jp

(2) Dynamical variables are complexified.
(3) The integral path in the complex domain is modified to minimize the cost function.

After taking the prescription, we can have a better integral path (manifold) which has larger $|e^{i\theta}|$; $0 \le |e^{i\theta}| \le 1$ is the average phase factor which is responsible for the seriousness of the sign problem. Thanks to Cauchy's integral theorem, the modified integral path provides us the same result as that obtained on the original integral path if there are no poles or cuts between the modified and original paths and the infinite regions of the integral path do not contribute to the results. There are some attempts to apply the method to various quantum field theories and models, e.g., the complex $\lambda\phi^4$ theory [16], the Polyakov-loop extended Nambu–Jona-Lasinio model [17,18] and the $0+1$ dimensional QCD [19].

In this article, we apply the path optimization method to deal with the complexified parameters. We here employ one plaquette system in the $1+1$ dimensional $U(1)$ gauge theory with complexified coupling constant on the lattice; some results for this theory are obtained by using a modification of the integral path, see Refs. [20,21]. In Sec. II, we show the formulation of the theory on the lattice and the explanation of the path optimization method. Numerical results are shown in Sec. III. Section IV is devoted to the summary.

II. FORMULATION

In this section, we summarize detailed formulation of the $1+1$ dimensional $U(1)$ lattice gauge theory and explain how we apply the path optimization method to the theory. We here consider the one plaquette system and thus the following formulation is corresponding to the system.

A. Action and partition function

The Wilson's plaquette action [22] for only one plaquette in the case of the $U(1)$ gauge theory is written as

$$S_{\mathrm{G}} = \frac{\beta}{2}\{P + P^{-1}\}, \tag{1}$$

where $\beta = 1/g^2$ is the lattice gauge coupling constant and $P\,(P^{-1})$ denotes the plaquette (its inverse). The definition of P is given by

$$P := U_1 U_2 U_3^{-1} U_4^{-1}, \tag{2}$$

where U_n are the $U(1)$ link variables defined as

$$U_n := e^{igA_n} \in \mathrm{U}(1), \tag{3}$$

here A_n denotes the $U(1)$ gauge field, $g^2 = 1/\beta$ and $n = 1, 2, 3, 4$.

The present theory can be analytically solved as

$$\mathcal{Z} = \int \prod_n dU_n e^{-S_{\mathrm{G}}} = I_0(\beta), \tag{4}$$

where I_0 denotes the modified Bessel function of the first kind. It should be noted that we can obtain an analytic result of the partition function for any system volume with periodic or open boundary conditions on the lattice [20,21,23]. For real β, $I_0(\beta)$ is always positive and thus there are no zeros of $\mathcal{Z}$ in Eq. (4), but the gauge coupling constant is now complex, $\beta \in \mathbb{C}$, and thus the partition function can be 0 which is nothing but the Lee-Yang zeros. These zeros play an important role in understanding the phase structure.

For the gauge theory, the action is invariant under the gauge transformation. In this case, to use the gauge transformation, one can reduce the degree of freedom to $n_{\mathrm{deg}} = 1, \ldots, 4$;

$$U_n = \begin{cases} U_n & n = 1, \ldots, n_{\mathrm{deg}} \\ \mathbf{I} & \text{otherwise.} \end{cases} \tag{5}$$

B. Path optimization method

In the path optimization method, we extend dynamical variables from real $(t \in \mathbb{R})$ to complex $(z \in \mathbb{C})$. In the present case, we need to extend the plaquette and the link variable as

$$\mathcal{P} = \mathcal{U}_1 \mathcal{U}_2 \mathcal{U}_3^{-1} \mathcal{U}_4^{-1}, \tag{6}$$

$$\mathcal{U}_n = e^{ig\mathcal{A}_n} =: U_n e^{z_n}, \tag{7}$$

where $\mathcal{A}_n \in \mathbb{C}$ and then $z_n \in \mathbb{R}$ represents the modification of the integral path. To represent z_n, we employ the artificial neural network as the model to generate the integral path. We here use the simple neural network which has the mono input, hidden and output layers. In this network, the variables in the hidden layer nodes (y_j) and the output variables (z_n) are given as follows:

$$\begin{aligned} z_n &= \omega_n F(w_{jn}^{(2)} y_j + b_j), \\ y_j &= F(w_{ij}^{(1)} t_i + b_j), \end{aligned} \tag{8}$$

where t_i denotes the parametric variable which is set to $\mathrm{Re}\,U_{n'}$, $\mathrm{Im}\,U_{n'}$, $i = 1, \ldots, 2 \times n_{\mathrm{deg}}$, w, b and ω are parameters of the neural network (weight and bias) and F is the so-called activation function. In this study, we chose the tangent hyperbolic function as the activation function.

To perform the path optimization, we need the cost function $(\mathcal{F})$; we here use

$$\mathcal{F}[z(t)] = \int d^n t |e^{i\theta(t)} - e^{i\theta_0}|^2 \times |J(t)e^{-S_G(t)}|, \quad (9)$$

where $J(t)$ is the Jacobian, θ_0 denotes the phase of the average phase factor and $\theta(t)$ is the phase of $J(t)e^{-S_G(t)} = e^{i\theta(t)}|J(t)e^{-S_G(t)}|$. Of course, we do not know the actual value of θ_0 except with $\beta \in \mathbb{R}$, $\theta_0 = 0$, and thus we replace θ_0 with $\langle\theta_{\text{pre}}\rangle_{\text{EMA}}$, where $\langle\theta_{\text{pre}}\rangle_{\text{EMA}}$ is the exponential moving average (EMA) of the phase obtained in the previous optimization steps. Minimization of the above cost function makes phases of e^{-S_G} as a function of z take similar values on the modified integral path when the regions are relevant to the final result. In other words, there is no need to care for the phase of the Boltzmann weight in irrelevant regions which should be automatically suppressed.

Since there is the sign problem in the case of the complex coupling constant by definition, we use the phase reweighting to perform the Monte-Carlo calculation as

$$\langle\mathcal{O}\rangle = \frac{\langle\mathcal{O}e^{i\theta}\rangle_{\text{pq}}}{\langle e^{i\theta}\rangle_{\text{pq}}}, \quad (10)$$

where $\mathcal{O}$ represents any operator and $\langle\cdots\rangle_{\text{pq}}$ means the phase quenched expectation values where $|Je^{-S_G}|$ is used as the Boltzmann weight. Since the Boltzmann weight, $|Je^{-S_G}|$, is now real, we can perform the Monte-Carlo calculation exactly. It should be noted that we are not restricted to the original integral path in the estimation of Eq. (10) unlike the ordinary reweighting calculation.

The machine learning technique was first introduced to the path optimization method in Ref. [16] to represent the modified integral path with a weaker sign problem. The machine learning technique was also introduced to the generalized Lefschetz thimble method [24] to learn the integral manifold where the sign problem is mild in Ref. [25] a few days before Ref. [16].

C. Setting of numerical calculation

Numbers of the unit in the input, hidden and output layers are $N_{\text{input}} = 2 \times n_{\text{deg}}$, $N_{\text{hidden}} = 10$ and $N_{\text{output}} = n_{\text{deg}}$, respectively. To determine the parameters in the neural network, we optimize these by using the ADADELTA [26], one of the stochastic gradient methods, as an optimizer with the Xavier initialization [27], the batch normalization [28], the minibatch training and the exponential moving average; see Ref. [16] for details of the optimization.

Actual configurations are generated by using the path optimization method with the hybrid Monte-Carlo (HMC) method [29] in the systems which includes the single plaquette with the open boundary condition. It should be noted that we here use the HMC with the replica exchange method (exchange HMC) [30,31] because there is the global sign problem even on the modified integral path [32] which means that there are some separated regions on

the modified integral path relevant to the integration. Integration over these separated regions is quite difficult to pick up by using ordinary HMC: We prepare the N_{rep} replicas characterized by the parameters in neural network as

$$C_x = \frac{x}{N_{\text{rep}}} C, \quad (11)$$

where x means the replica number, $x = 1, \ldots, N_{\text{rep}}$, and C represents the parameters of the optimized neural network [$C = w, b, \omega$ in Eq. (8)]. We set $N_{\text{rep}} = 5$ in the numerical calculation. We use the exchange probability of the replicas $P(U_x \leftrightarrow U_{x'})$ as

$$P(U_x \leftrightarrow U_{x'}) = \min\left(1, \frac{P(U_x; C_{x'})P(U_{x'}; C_x)}{P(U_x; C_x)P(U_{x'}; C_{x'})}\right),$$
$$P(U_x; C_x) = |J(U_x; C_x)e^{-S_G(\mathcal{U}(U_x; C_x))}|. \quad (12)$$

In Ref. [16], we used the sampling of configurations based on the symmetry of the modified integral path, but the method is only available if we know the good modified integral path which can have the symmetry. In this paper,
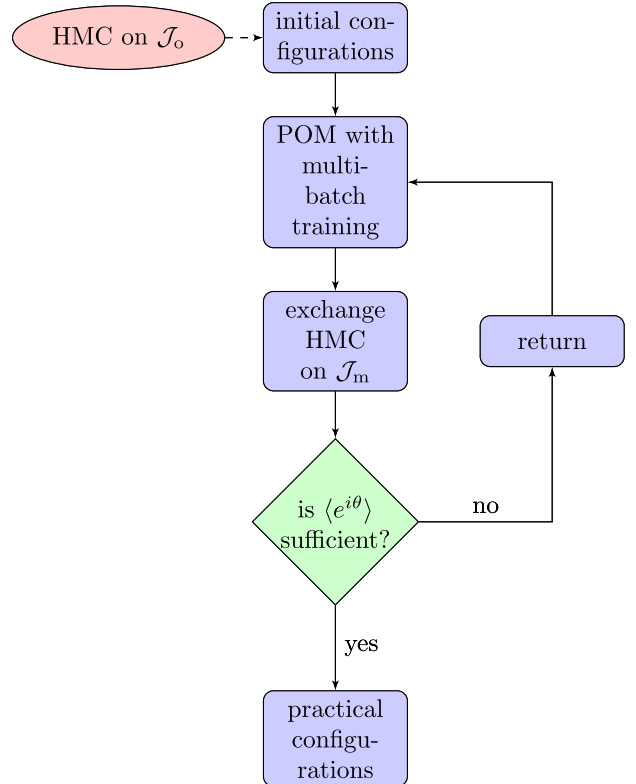

FIG. 1. The flowchart of the algorithm to generate practical configurations in this work. Symbols, $\mathcal{J}_o$ and $\mathcal{J}_m$, denote the original and the modified integral path, respectively. The closed loop in the flowchart is the one cycle of the optimization procedure of the integral path.

we employ the exchange HMC method to generally perform the path optimization.

The expectation values are calculated with 2500 configurations and then the corresponding errors are estimated by using the jackknife method with the bin size 50. For the parameter of the theory, we set $\beta = \beta_R + i\beta_I$ with $\beta_R, \beta_I \in \mathbb{R}$.

Figure 1 shows the flowchart of the algorithm to generate practical configurations where $\mathcal{J}_o$ and $\mathcal{J}_m$ are the original and modified integral path, respectively.

III. NUMERICAL RESULTS

In this section, we show the numerical results of the $1 + 1$ dimensional $U(1)$ gauge theory on the lattice with the complex coupling by using the path optimization method. Here, we show the results of the $1 + 1$ dimensional $U(1)$ gauge theory only with the single plaquette; i.e., the simplest setting of the theory. Actually, it is nothing but the n_{deg}-dimensional integral.

Figure 2 shows the scatter plot on the $\mathrm{Re}\mathcal{P} - \mathrm{Im}\mathcal{P}$ plane with $\beta = 2i$. Here, we show the results with the gauge fixing:

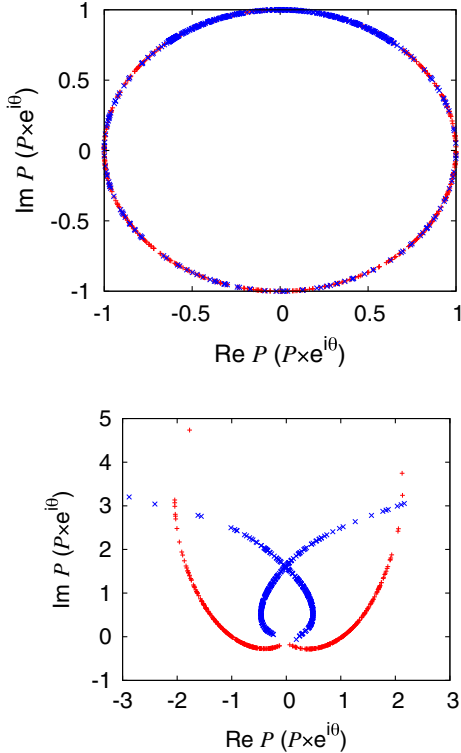$$U_n = \begin{cases} U_n & n = 1 \\ \mathbf{I} & n \neq 1. \end{cases} \quad (13)$$

We can clearly see the modification of the integral path from Fig. 2. In addition, we can see the bias of the distribution in $\mathcal{P} \times e^{i\theta}$ which plays a crucial role in the calculation of $\langle \mathcal{P} \rangle$; this bias makes $\langle \mathcal{P} \rangle$ finite because the phase quenched expectation value of $\mathcal{P}$ becomes 0. The histogram of the phase for the case of Fig. 2 is shown in Fig. 3. On the original integral path, θ distribution is widely spread, but the distribution on the modified integral path is well localized; we can generate configurations strongly localized around two separated regions. The replica exchange method works well in both cases.

Figure 4 shows the average phase factor, $\langle e^{i\theta} \rangle_{\mathrm{EMA}}$, as a function of the optimization step in one epoch with $\beta = i$ and $2i$; one epoch is defined so that one sequence of the minibatch training is finished. Here, we estimate the average phase factor by using EMA. From the figure, we can clearly see that the average phase factor cannot be enhanced without the gauge fixing. With the gauge fixing, the average phase factor approaches 1 with $\beta = i$. In the case with $\beta = 2i$, there is the serious global sign problem as shown in Fig. 3 and thus we have the upper limit of the improvement, but we can well enhance the average phase factor via the path optimization. It should be noted that the modified integral path sometimes provides the expectation
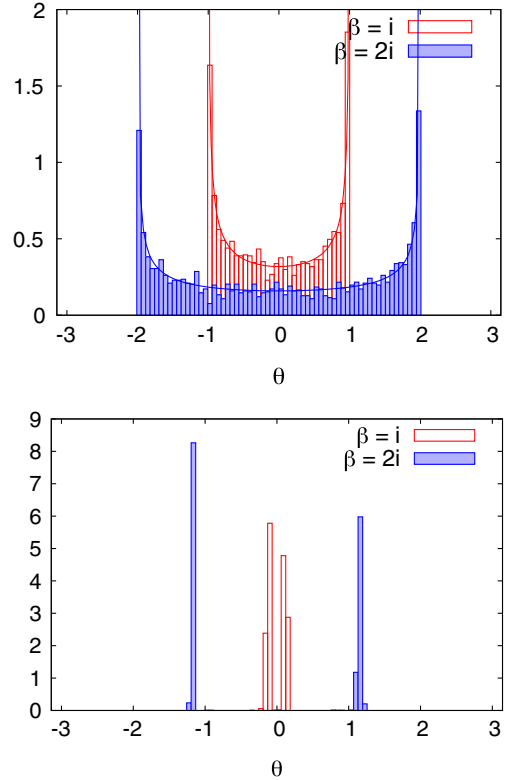

FIG. 2. The scatter plot of $\mathcal{P}$ and $\mathcal{P} \times e^{i\theta}$ on the X-Y plane with $\beta = 2i$, where $X = \mathrm{Re}\mathcal{P}$ and $Y = \mathrm{Im}\mathcal{P}$. The top and bottom panels show the results without and with the path optimization. Plus signs and crosses indicate $\mathcal{P}$ and $\mathcal{P} \times e^{i\theta}$, respectively.

FIG. 3. The normalized histogram of the phase, θ, for $\beta = i$ and $2i$. The top and bottom panels show the results without and with the path optimization. The line in the top panel shows the probability distribution on the original path, $P(\theta) = [\pi\beta_I \sin\{\arccos(\theta/\beta_I)\}]^{-1}$.
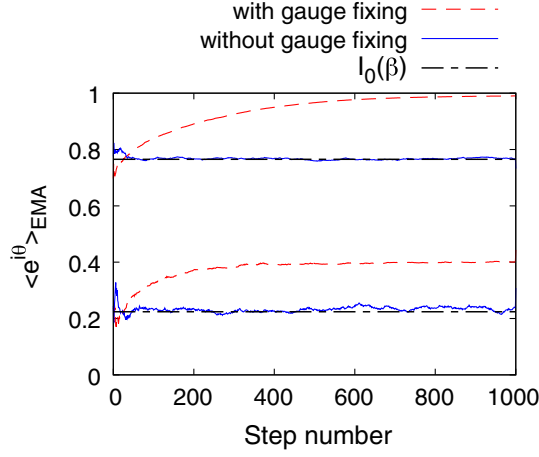
FIG. 4. The average phase factor, $\langle e^{i\theta} \rangle_{\mathrm{EMA}}$, during the optimization with and without the gauge fixing in one epoch. The dash-dotted line shows the exact average phase factor on the original path. In the case with the gauge fixing, we here perfectly fix the gauge degree of freedom; i.e., $n_{\mathrm{deg}} = 1$. Upper- and bottom-side lines are results with $\beta = i$ and $2i$, respectively. In the case with $\beta = 2i$, the serious global sign problem exists.

value with the larger error bar compared with the original one even if the average phase factor is enhanced. This indicates that there is the competition between the improvement of the original sign problem and the modification of the integral path which is responsible to the statistical error via the path optimization method. In addition, we can see from Fig. 5 that the average phase factor becomes larger with reduction of n_{deg} by using the gauge fixing; it may be related to the fact that we have a larger degree of freedom without the gauge fixing and then the neural network cannot show sufficient performance to optimize the suitable integral path.

For the reader's convenience, we finally show the expectation value of the plaquette as a function of β_{I} with $\beta_{\mathrm{R}} = 0$ where zeros exist in Fig. 6. It is clearly seen that the
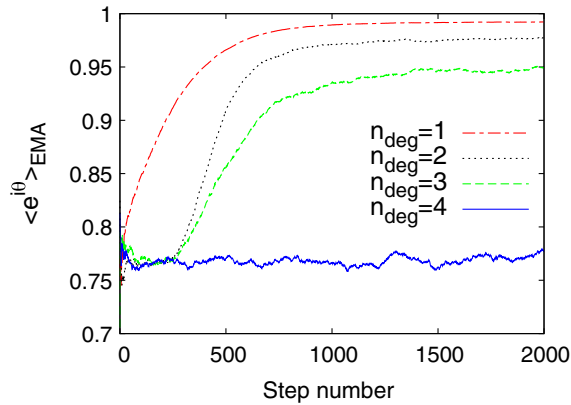

FIG. 5. The average phase factor, $\langle e^{i\theta} \rangle_{\mathrm{EMA}}$, during the optimization with and without the gauge fixing in two epochs with $\beta = i$.
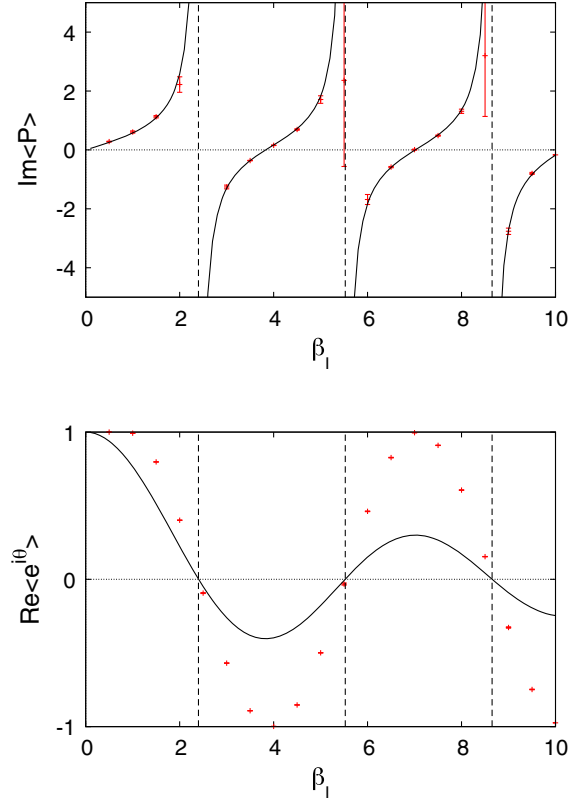

FIG. 6. The top and bottom panels show the β_{I} dependence of the expectation value of the plaquette and the average phase factor, respectively. Symbols show numerical results obtained via the path optimization with the gauge fixing and solid lines denote the analytic results; the solid line in the bottom panel is corresponding to the results on the original integral path.

modified integral path reproduces the analytic result except the region near the partition function zeros. At $\beta_{\mathrm{I}} \sim 2.4$, 5.5, 8.6, we have the zeros and then there should be the strong modification of the integral path and/or the serious global sign problem.

IV. SUMMARY

In this study, we have considered the $1 + 1$ dimensional $U(1)$ lattice gauge theory which has the single plaquette with the complex coupling constant as a first step to directly investigate the Lee-Yang zeros in the gauge theory. We have estimated how the average phase factor is improved via the modification of integral variables. Since there is the sign problem when the coupling constant is complex, we employ the path optimization method to perform the Monte-Carlo calculation. To represent the modified integral path in the complex domain of the integral variables, we employ the artificial neural network.

We have shown that the modification of the integral path represented by using the neural network can well enhance the average factor if we impose the gauge fixing to the theory, but we cannot without the gauge fixing; this suggests the importance of the gauge fixing to control

the sign problem in the path optimization method on the lattice. This issue is demonstrated in the system of the single plaquette. Also, we have checked that the replica exchange method works well to generate configuration localized in the well separated regions which are realized via the path optimization. From these results, we have clarified how to use the path optimization method in the gauge theory. It should be important in the application of the path optimization method to the more complicated gauge theory such as QCD.

In the present study, we have restricted the system size to be small because we are interested in the possibility of applying the path optimization method to the system with complexified external parameters. Usually, the sign problem becomes exponentially worse in terms of the system volume; there is the competition between the exponential suppression of it from the system volume and the enhancement of it from the optimization. In the larger volume case, we should introduce some methods to reduce the numerical cost to calculate the Jacobian, whose numerical cost is proportional to the square of the system volume. Examples of promising methods are the diagonal ansatz of the Jacobian [33] and the nearest-neighbor lattice-sites ansatz [34]. We will revisit this issue in our future work.

This study is a first step in the path optimization method to explore the phase structure of the gauge theory in the complexified parameter space which is important to understand properties of the phase transition; e.g., for investigation of the distribution of the Lee-Yang zeros. In the present work, we employ the simple gauge theory, but we believe that it sheds light on the complexification of the integral variables and also the parameters. In the future, we will apply the path optimization method to the $SU(2)$ gauge theory with the complex coupling constant. It was reported that the complex Langevin method fails in some parameter regions; see Ref. [35].

[1] K. Fukushima and T. Hatsuda, Rep. Prog. Phys. **74,** 014001 (2011).

[2] T. D. Lee and C.-N. Yang, Phys. Rev. **87,** 410 (1952).

[3] M. Biskup, C. Borgs, J. T. Chayes, L. J. Kleinwaks, and R. Kotecký, Phys. Rev. Lett. **84,** 4794 (2000).

[4] X. Peng, H. Zhou, B.-B. Wei, J. Cui, J. Du, and R.-B. Liu, Phys. Rev. Lett. **114,** 010601 (2015).

[5] A. Krishnan, M. Schmitt, R. Moessner, and M. Heyl, Phys. Rev. A **100,** 022125 (2019).

[6] A. Nakamura and K. Nagata, Prog. Theor. Exp. Phys. **2016,** 033D01 (2016).

[7] K. Nagata, K. Kashiwa, A. Nakamura, and S. M. Nishigaki, Phys. Rev. D **91,** 094507 (2015).

[8] M. Wakayama, V. G. Bornyakov, D. L. Boyda, V. A. Goy, H. Iida, A. V. Molochkov, A. Nakamura, and V. I. Zakharov, Phys. Lett. B **793,** 227 (2019).

[9] A. Roberge and N. Weiss, Nucl. Phys. **B275,** 734 (1986).

[10] R. Fukuda, A. Nakamura, and S. Oka, Phys. Rev. D **93,** 094508 (2016).

[11] P. de Forcrand, Proc. Sci., LAT2009 (**2009**) 010.

[12] K. Kashiwa, Symmetry **11,** 562 (2019).

[13] K. Kashiwa and A. Ohnishi, Phys. Lett. B **750,** 282 (2015).

[14] K. Kashiwa and A. Ohnishi, Phys. Lett. B **772,** 669 (2017).

[15] Y. Mori, K. Kashiwa, and A. Ohnishi, Phys. Lett. B **781,** 688 (2018).

[16] Y. Mori, K. Kashiwa, and A. Ohnishi, Phys. Rev. D **96,** 111501 (2017).

[17] K. Kashiwa, Y. Mori, and A. Ohnishi, Phys. Rev. D **99,** 014033 (2019).

[18] K. Kashiwa, Y. Mori, and A. Ohnishi, Phys. Rev. D **99,** 114005 (2019).

[19] Y. Mori, K. Kashiwa, and A. Ohnishi, Prog. Theor. Exp. Phys. **2019,** 113B01 (2019).

[20] J. M. Pawlowski, M. Scherzer, C. Schmidt, F. P. G. Ziegler, and F. Ziesch, Proc. Sci., LATTICE2019 (**2019**) 223.

[21] W. Detmold, G. Kanwar, M. L. Wagman, and N. C. Warrington, Phys. Rev. D **102,** 014514 (2020).

[22] K. G. Wilson, Phys. Rev. D **10,** 2445 (1974); **10,** 45 (1974); **10,** 319 (1974).

[23] R. Balian, J. M. Drouffe, and C. Itzykson, Phys. Rev. D **10,** 3376 (1974).

[24] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, J. High Energy Phys. 05 (2016) 053.

[25] A. Alexandru, P. F. Bedaque, H. Lamm, and S. Lawrence, Phys. Rev. D **96,** 094505 (2017).

[26] M. D. Zeiler, arXiv:1212.5701.

[27] X. Glorot and Y. Bengio, in *Thirteenth International Conference on Artificial Intelligence and Statistics* (Proceedings of Machine Learning Research, 2010), Vol. 9, pp. 249–256.

[28] S. Ioffe and C. Szegedy, arXiv:1502.03167.

[29] S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195,** 216 (1987).

[30] R. H. Swendsen and J.-S. Wang, Phys. Rev. Lett. **57**, 2607 (1986).

[31] M. Fukuma and N. Umeda, Prog. Theor. Exp. Phys. **2017**, 073B01 (2017).

[32] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu, and T. Sano, J. High Energy Phys. 10 (2013) 147.

[33] A. Alexandru, P. F. Bedaque, H. Lamm, and S. Lawrence, Phys. Rev. D **97**, 094510 (2018).

[34] F. Bursa and M. Kroyter, J. High Energy Phys. 12 (2018) 054.

[35] H. Makino, H. Suzuki, and D. Takeda, Phys. Rev. D **92**, 085020 (2015).

[36] S. Choe, S. Muroya, A. Nakamura, C. Nonaka, T. Saito, and F. Shoji, Nucl. Phys. B, Proc. Suppl. **106–107**, 1037 (2002).