

Application of the path optimization method to the sign problem in an effective model of QCD with a repulsive vector-type interaction

Kouji Kashiwa,^{1,*} Yuto Mori,^{2,†} and Akira Ohnishi^{3,‡}

¹*Fukuoka Institute of Technology, Wajiro, Fukuoka 811-0295, Japan*

²*Department of Physics, Faculty of Science, Kyoto University, Kyoto 606-8502, Japan*

³*Yukawa Institute for Theoretical Physics, Kyoto University, Kyoto 606-8502, Japan*



(Received 18 March 2019; published 6 June 2019)

The path optimization method is applied to a QCD effective model with a Polyakov loop and a repulsive vector-type interaction at finite temperature and density to circumvent the model sign problem. We show how the path optimization method can increase the average phase factor and control the model sign problem. This is the first study which correctly treats the repulsive vector-type interaction in the QCD effective model with a Polyakov loop via the Markov-chain Monte Carlo approach. It is shown that we can evade the model sign problem within the standard path-integral formulation by complexifying the temporal component of the gluon field and the vector-type auxiliary field.

DOI: [10.1103/PhysRevD.99.114005](https://doi.org/10.1103/PhysRevD.99.114005)

I. INTRODUCTION

Understanding the confinement-deconfinement transition at finite temperature (T) and chemical potential (μ) in quantum chromodynamics (QCD) is one of the important and interesting subjects in elementary particle, nuclear, and hadron physics. To investigate nonperturbative properties of QCD such as chiral symmetry breaking and the confinement mechanism, Monte Carlo simulations of lattice QCD have been utilized as a powerful tool for studying nonperturbative properties of QCD such as chiral symmetry breaking and confinement at zero baryon density. Unfortunately, lattice QCD simulations have the sign problem at a nonzero real quark chemical potential. To circumvent the sign problem, several methods have been proposed so far such as the Taylor expansion method [1–3], the reweighting method [4,5], the analytic continuation method [6–8], and the canonical approach [9–14]. However, we cannot access the cold dense region, $\mu/T > 1$, in these methods at present [15].

The QCD effective models are widely used to investigate the QCD phase structure at finite real chemical potential. We can sometimes avoid the sign problem in simple models such as the Nambu–Jona-Lasinio (NJL) model without a

repulsive vector-type interaction. In more realistic models, however, the sign problem arises again. For example, the Polyakov-loop extended NJL (PNJL) model [16] has the sign problem even in the mean-field treatment. The sign problem appearing in the QCD effective model is called the model sign problem [17,18]. Practically, one can avoid the model sign problem by using some prescriptions which may not have a clear theoretical foundation.

One of the model sign problems arises from the vector-type interaction. In the mean-field approximation for the NJL-type model, we usually neglect the Wick rotation of the vector-type auxiliary field (ω_0), and then the stationary point of the action is considered to be the solution. The stationary point corresponds to the maxima of the thermodynamic potential along the ω_0 direction, and thus it is not stable, in principle. While this treatment cannot be justified from the standard path-integral formulation, it can be acceptable in the mean-field approximation. Actually, this problem has been discussed in Ref. [19] by using the Lefschetz-thimble method [20–22]. We can clearly understand that the standard mean-field treatment implicitly employs the complexification of the vector-type auxiliary field based on the Cauchy-(Poincaré) theorem.

In this article, we use the path optimization method [23–26] to formally tackle the model sign problem induced by the Polyakov loop and also the repulsive vector-type interaction in a QCD effective model. In Ref. [26], we have shown that the complexification of the temporal component of the gluon field is sufficient to control the model sign problem in the PNJL model without a repulsive vector-type interaction. In addition, we have shown that the complexification of the vector-type auxiliary field should be responsible to control the model sign problem in the

*kashiwa@fit.ac.jp

†mori@ruby.scphys.kyoto-u.ac.jp

‡ohnishi@yukawa.kyoto-u.ac.jp

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/). Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

NJL model with a repulsive vector-type interaction [19]. It should be noted that the flow equation of the Lefschetz thimble blows up at a small value of the vector-type auxiliary field, and we failed to obtain the Lefschetz thimbles in the auxiliary-field space in Ref. [19]. Therefore, in this article, we apply the path optimization method to the PNJL model with a repulsive vector-type interaction to control its model sign problem. This study is the first attempt to treat the model sign problem correctly and systematically within the standard path-integral formulation via the complexification of the integral variables in the QCD effective model with a Polyakov loop and a repulsive vector-type interaction.

This article is organized as follows. In the next section, we explain the path optimization method and the PNJL model with a repulsive vector-type interaction. Section III shows numerical results by using the hybrid Monte Carlo method. Section IV is devoted to the summary and discussions.

II. FORMULATION

We investigate the model sign problem appearing in the PNJL model with a repulsive vector-type interaction via the path optimization method. Details are explained below.

A. Polyakov-loop extended Nambu–Jona-Lasinio model

The Euclidean Lagrangian density of the two-flavor PNJL model [16] with a repulsive vector-type interaction is given as

$$\mathcal{L} = \bar{q}(\not{D} + m_0)q - G[(\bar{q}q)^2 + (\bar{q}i\gamma_5\vec{\tau}q)^2] + G_v(\bar{q}\gamma_\mu q)^2 + \mathcal{V}_g(\Phi, \bar{\Phi}), \quad (1)$$

where m_0 denotes the current quark mass, $D_\nu = \partial_\nu - igA_\nu\delta_{\nu 4}$ is the covariant derivative, Φ ($\bar{\Phi}$) represents the Polyakov loop (its conjugate), and $\mathcal{V}_g$ expresses the gluonic contribution. The coupling constants G and G_v take positive values, as understood from the QCD one-gluon exchange interaction; see Ref. [27] as an example.

We employ the homogeneous auxiliary-field ansatz, as adopted in previous works using the Monte Carlo PNJL model [26,28], and thus our numerical results converge to mean-field results in the infinite volume limit. The homogeneous ansatz corresponds to the momentum truncation to $\mathbf{k} = \mathbf{0}$. After the bosonization and complexification of auxiliary fields, the grand-canonical partition function is given as

$$\mathcal{Z} = \int \prod_{\mathbf{k}} dz_{\mathbf{k}} e^{-\Gamma[z_{\mathbf{k}}]}, \quad (2)$$

where Γ is the effective action and $z_{\mathbf{k}}$ represents the dynamical variables in the momentum space. With the

homogeneous field ansatz, we truncate the auxiliary fields to $\mathbf{k} = \mathbf{0}$ components. Then the effective action becomes $\Gamma = \beta V \mathcal{V}$, where $\mathcal{V}$ is the effective potential and β is the inverse temperature. Thus, our Monte Carlo results should agree with the mean-field results in the infinite volume limit, where the configuration at the minimum of $\mathcal{V}$ dominates.

After the Hubbard-Stratonovich transformation (bosonization), the thermodynamic potential is obtained as

$$\mathcal{V} = \mathcal{V}_{\text{NJL}} + \mathcal{V}_g, \quad (3)$$

where $\mathcal{V}_{\text{NJL}}$ and $\mathcal{V}_g$ are the fermionic and gluonic parts of the effective potential, respectively. The actual form of $\mathcal{V}_{\text{NJL}}$ is given as

$$\mathcal{V}_{\text{NJL}} = -2N_f \int_{\Lambda} \frac{d^3 p}{(2\pi)^3} \left[N_c E_{\mathbf{p}} - N_c \sqrt{\mathbf{p}^2 + m_0^2} + T \ln(f^- f^+) \right] + G(\sigma^2 + \vec{\pi}^2) + G_v \omega_4^2, \quad (4)$$

where $N_f = 2$ ($N_c = 3$) is the number of flavor (color) and Λ is the three-dimensional momentum cutoff. We set the same momentum cutoff in the vacuum and the thermal parts. We here introduce auxiliary fields as $\sigma = \bar{q}q$, $\vec{\pi} = \bar{q}i\gamma_5\vec{\tau}q$, and $\omega_4 = -\bar{q}i\gamma_0 q$. The Fermi-Dirac distribution functions are given as

$$\begin{aligned} f^- &= 1 + 3(\Phi + \bar{\Phi}e^{-\beta E_{\mathbf{p}}^-})e^{-\beta E_{\mathbf{p}}^-} + e^{-3\beta E_{\mathbf{p}}^-}, \\ f^+ &= 1 + 3(\bar{\Phi} + \Phi e^{-\beta E_{\mathbf{p}}^+})e^{-\beta E_{\mathbf{p}}^+} + e^{-3\beta E_{\mathbf{p}}^+}, \\ E_{\mathbf{p}}^\pm &= E_{\mathbf{p}} \mp \tilde{\mu} = \sqrt{\varepsilon_{\mathbf{p}}^2 + 2N^+ N^-} \mp \tilde{\mu}, \\ \varepsilon_{\mathbf{p}} &= \sqrt{\mathbf{p}^2 + M^2 + N^2}, \end{aligned} \quad (5)$$

where M , N , $N^\pm$, and $\tilde{\mu}$ are functions of the auxiliary fields:

$$\begin{aligned} M &= m_0 - 2G\sigma, & N &= -2G\pi^0, \\ N^\pm &= -2G\pi^\pm, & \tilde{\mu} &= \mu - 2iG_v\omega_4, \end{aligned} \quad (6)$$

with $\pi^0 = \pi_3$ and $\pi^\pm = (\pi_1 \pm i\pi_2)/\sqrt{2}$. Because of the $2iG_v\omega_4$ term in $\tilde{\mu}$, the repulsive vector-type interaction induces the model sign problem in addition to that from the Polyakov loop. For $\mathcal{V}_g$, we employ the logarithmic-type Polyakov-loop potential proposed in Ref. [29]:

$$\frac{\mathcal{V}_g}{T^4} = -\frac{1}{2}a_T \bar{\Phi}\Phi + b_T \ln(h), \quad (7)$$

$$h = 1 - 6\bar{\Phi}\Phi + 4(\bar{\Phi}^3 + \Phi^3) - 3(\bar{\Phi}\Phi)^2, \quad (8)$$

$$a_T = a_0 + a_1 \left(\frac{T_0}{T}\right) + a_2 \left(\frac{T_0}{T}\right)^2, \quad b_T = b_3 \left(\frac{T_0}{T}\right)^3. \quad (9)$$

The parameters are usually set to reproduce the lattice QCD data in the pure gauge limit. The basic setup to compute the PNJL model with the Markov-chain Monte Carlo method is shown in Refs. [26,28].

Cuts in the logarithm of Eq. (7) may induce the numerical problem, but it may be a model artifact, and thus we do not consider any additional care for the singularities in this study as in Ref. [26]. One of the promising ways to avoid the problem is the modification of the functional form of the Polyakov-loop potential. The logarithmic term in the Polyakov-loop potential appears as $VT_0^3 b_3 \times \ln(h)$ in the Boltzmann weight. If $VT_0^3 b_3$ is set to be a positive integer, the singularity does not matter. In the present potential, $VT_0^3 b_3$ is not an integer. It is well known that there is another functional form of the Polyakov-loop potential that is the polynomial one [30], which also reproduces the lattice QCD data in the pure gauge limit at finite T and does not have singularities. Nevertheless, sampled configurations are found to be well localized; then the path optimization method works well practically in the present setup as shown later.

B. Path optimization method

In the path optimization method, we first complexify the integral variables, $x_i \in \mathbb{R} \rightarrow z_i \in \mathbb{C}$, where $i = 1, \dots, n$ with n being the dimension of integration. To construct the new (and good) integral path in the complex space, we use the cost function, which represents the seriousness of the sign problem. We vary the integral path in the direction to decrease the cost function. This method has a similarity from the viewpoint of the complexification of dynamical variables with the complex Langevin method [31,32] and the Lefschetz-thimble method [20–22,33]. Especially, the path optimization method belongs to the category of the off-thimble integral methods, which allow the integral path to deviate from the thimble, as proposed in the generalized Lefschetz-thimble method [33]. See Refs. [34–40] for recent progress in these methods.

The path optimization method was first proposed in Ref. [23]. Machine learning (feedforward neural network) was introduced to describe and to optimize the modified integral path in Refs. [24,25]. A few days before Ref. [25] was submitted, machine learning was introduced to learn the integral manifold in the generalized Lefschetz-thimble method in Ref. [41]. This method uses supervised learning, because we must teach the relevant integral path (manifold) to the neural network, and the results of the generalized Lefschetz-thimble method have been used as the teacher data; it is the first paper which employs supervised learning to evade the sign problem as far as we know. Also, the same group applied machine learning to optimize the integral path by using the average phase factor in Ref. [40] after our paper [25] appeared. This method has a similarity with our path optimization method which uses unsupervised

learning. Machine learning can be applied to various optimization problems, and, thus, it is quite useful in physics.

The functional form of the new integral path is represented by using the feedforward neural network [23,25]. Then, the parameters in the feedforward neural network are optimized via the minimization of the cost function. The largest advantage of using the feedforward neural network in the path optimization method is in the universal approximation theorem; the neural network even with a mono-hidden layer can approximate any kind of continuous function on the compact subset as long as we prepare a sufficient number of units in the hidden layer [42,43].

To use the feedforward neural network, we represent z_i by using a parametric quantity (t_i) as

$$\begin{aligned} a_i(t) &= g(w_{ij}^{(1)} t_j + b_i^{(1)}), & f_i(t) &= g(w_{ij}^{(2)} a_j + b_i^{(2)}), \\ z_i(t) &= t_i + i[\alpha_i f_i(t) + \beta_i], \end{aligned} \quad (10)$$

where w_{ij} , b_i , α_i , and β_i are parameters. In particular, w and b are called the weight and the bias, respectively. Thus, we have the map $\text{Re}(z_i) \rightarrow \text{Im}(z_i)$. The function $g(x)$ is called the activation function, and we use the hyperbolic tangent function. We use the back-propagation algorithm in the actual optimization of parameters. It should be noted that the path optimization method reproduces the same results as the original theory because of the Cauchy(-Poincaré) theorem as long as the integral path does not go across singular points and the contribution at $\text{Re} z \rightarrow \pm\infty$ vanishes.

To obtain the good integral path, we use the following form of the cost function:

$$\begin{aligned} \mathcal{F}[z(t)] &= \frac{1}{2} \int d^n t |e^{i\theta(t)} - e^{i\theta_0}|^2 \times |J(t) e^{-\Gamma(z(t))}| \\ &= \int d^n t |J(t) e^{-\Gamma(z(t))}| - \left| \int d^n t J(t) e^{-\Gamma(z(t))} \right|, \end{aligned} \quad (11)$$

where

$$\begin{aligned} \theta(t) &= \arg(J(t) e^{-\Gamma(z(t))}), & \theta_0 &= \arg(\mathcal{Z}), \\ J(t) &= \det\left(\frac{\partial z_i}{\partial t_j}\right); \end{aligned} \quad (12)$$

see Ref. [23] for details.

In applying the path optimization method to the PNJL model with a vector-type interaction, we complexify temporal gluon components (A_3 and A_8) and the Wick-rotated vector-type auxiliary field (ω_4), while the scalar-type and pseudoscalar-type auxiliary fields (σ and $\bar{\pi}$) are still treated as real variables. Thus, we have seven dynamical variables (σ , π^0 , $\pi^\pm$, $\text{Re}\omega_4$, $\text{Re}A_3$, and $\text{Re}A_8$) and three dependent variables ($\text{Im}\omega_4$, $\text{Im}A_3$, and $\text{Im}A_8$), where the latter three variables are given via Eq. (10), since it is

known that the model sign problem can be resolved by the complexification of the temporal gluon fields in the Lefschetz-thimble method at least in a system without a repulsive vector-type interaction [18]. Also, ω_4 can induce the model sign problem even in the NJL model, and then we must consider the complexification of ω_4 [19].

In the present study, we directly complexify A_3, A_8 , and ω_4 , but this treatment may violate the periodicity along the $\text{Re}A_3$ and $\text{Re}A_8$ directions. If we wish to take care of the periodicity, we may use the periodic functional form in the neural network as in Ref. [41]. In particular, the violation of the periodicity may become serious when the configurations are spread widely in the $\text{Re}A_3$ and $\text{Re}A_8$ variables. For example, see Ref. [44] for the issue of the periodicity. As shown later, however, the present calculation agrees well with the mean-field approximation and configurations are well localized. Thus, we do not introduce the periodic form of inputs at present.

It should be noted that the path optimization with the feedforward neural network is unsupervised learning, because we do not need teacher data. The settings of the feedforward neural network in the path optimization method such as the optimizer are the same as Ref. [26], and, thus, we skip the explanation here.

III. NUMERICAL RESULTS

In the actual numerical calculation, we have generated 80 000 configurations by using the hybrid Monte Carlo method. Then, the expectation values are estimated after a few times of optimization. We employ the simple neural network which contains the input, mono-hidden, and output layers. The number of unit in the hidden layer is set to $4N_{\text{dof}} = 12$, where N_{dof} is the number of dependent variables. The expectation value of an operator ($\mathcal{O}$) is obtained via the phase reweighting as

$$\langle \mathcal{O} \rangle = \frac{\int d^n t \mathcal{O} e^{i\theta} |J(t) e^{-\Gamma(z(t))}|}{\int d^n t e^{i\theta} |J(t) e^{-\Gamma(z(t))}|} = \frac{\langle e^{i\theta} \mathcal{O} \rangle_{\text{pq}}}{\langle e^{i\theta} \rangle_{\text{pq}}}, \quad (13)$$

where $\langle \cdots \rangle_{\text{pq}}$ means the phase quenched average and

$$e^{i\theta} = \frac{J(t) e^{-\Gamma(z(t))}}{|J(t) e^{-\Gamma(z(t))}|}. \quad (14)$$

The parameters in the NJL part are the same as in Ref. [26], and we newly introduce G_v as $G_v = 0.5G$.

Figure 1 shows the average phase factor, $\text{Re}\langle e^{i\theta} \rangle_{\text{pq}}$, at $T = 0.1$ GeV with $k = VT^3 = 8$ and 64. In some regions of μ , the average phase factor becomes almost 0 before the optimization, as shown by the dashed line in the figure. By comparison, we can successfully increase the average phase factor after the optimization. It suggests that there is no need to complexify σ and $\vec{\pi}$ auxiliary fields in the path optimization method to investigate the PNJL model with a

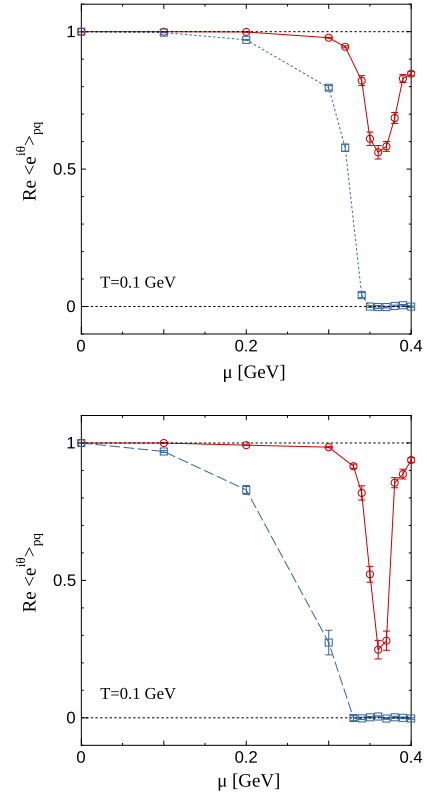


FIG. 1. The μ dependence of $\text{Re}\langle e^{i\theta} \rangle_{\text{pq}}$ at $T = 0.1$ GeV. The top and bottom panels show results with $k = 8$ and 64, respectively. The circle and square symbols are results after and before the optimization, respectively.

repulsive vector-type interaction. Also, this would be true in the Lefschetz-thimble method and other complexified integral-path approaches. Compared with the PNJL model without a repulsive vector-type interaction [26], the average phase factor becomes worse, because the ω_4 field additionally induces the sign problem at a finite density. Around $\mu = 0.36$ GeV, the optimization is neither sufficient nor automatic in the case with $k = 64$. With naive initial conditions of dynamical variables, the average phase factor stays very small. Then, various initial conditions have been examined, and we finally obtain the optimized path with a reasonably large average phase factor as shown in Fig. 1. This result indicates that the present neural network in the case with $k = 64$ does not have enough performance of the approximation to overcome the exponential suppression of the average phase factor.

Figure 2 shows the μ dependence of the order parameters at $T = 0.1$ GeV after the optimization. We also show the mean-field results based on the CK symmetry ansatz in the fermion determinant [45,46], where C and K are the charge and complex conjugations, respectively. This ansatz can be justified by using the Lefschetz-thimble method [18]. Under the CK symmetry condition, we solve gap equations

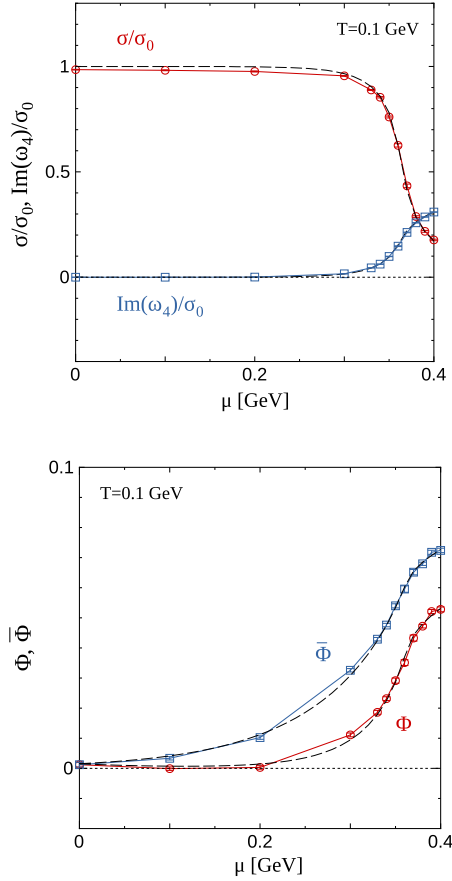


FIG. 2. The top and bottom panels show the μ dependence of σ and ω_4 , and Φ and $\bar{\Phi}$ at $T = 0.1$ GeV with $k = 64$, where σ and ω_4 are normalized by σ at $T = \mu = 0$ ($\sigma_0 < 0$) in the infinite volume limit. The thin dashed lines are eye guides which are the mean-field results with the CK symmetry ansatz in the fermion determinant and the un-Wick-rotated calculation.

$$\frac{\partial \Gamma}{\partial \sigma} = 0, \quad \frac{\partial \Gamma}{\partial \omega_0} = 0, \quad \frac{\partial \Gamma}{\partial \Phi} = 0, \quad \frac{\partial \Gamma}{\partial \bar{\Phi}} = 0, \quad (15)$$

where we do not use the Wick rotation of the vector-type auxiliary field. This treatment cannot be justified in the standard path integral formulation, but practically it reproduces the correct result in the leading order of the large N_c expansion, because ω_0 corresponds to the quark number density in the mean-field approximation. From the figure, we find that the numerical errors are well controlled and the difference between Φ and $\bar{\Phi}$ at a finite density, $\bar{\Phi} > \Phi$, is correctly reproduced. Compared with the results without a vector-type interaction [26], the chiral condensate decreases more slowly. This is reasonable, since the repulsive vector-type interaction is known to weaken the chiral phase transition. We find that $|\text{Im}\omega_4|$ strongly increases above $\mu = 0.3$ GeV. Since the quark number density and ω_4 are related to each other via $\omega_0 = \langle q^\dagger q \rangle = i\omega_4$, this sudden increase indicates the absence of the silver-blaze problem at $T = 0$; the quark number density should start to increase at

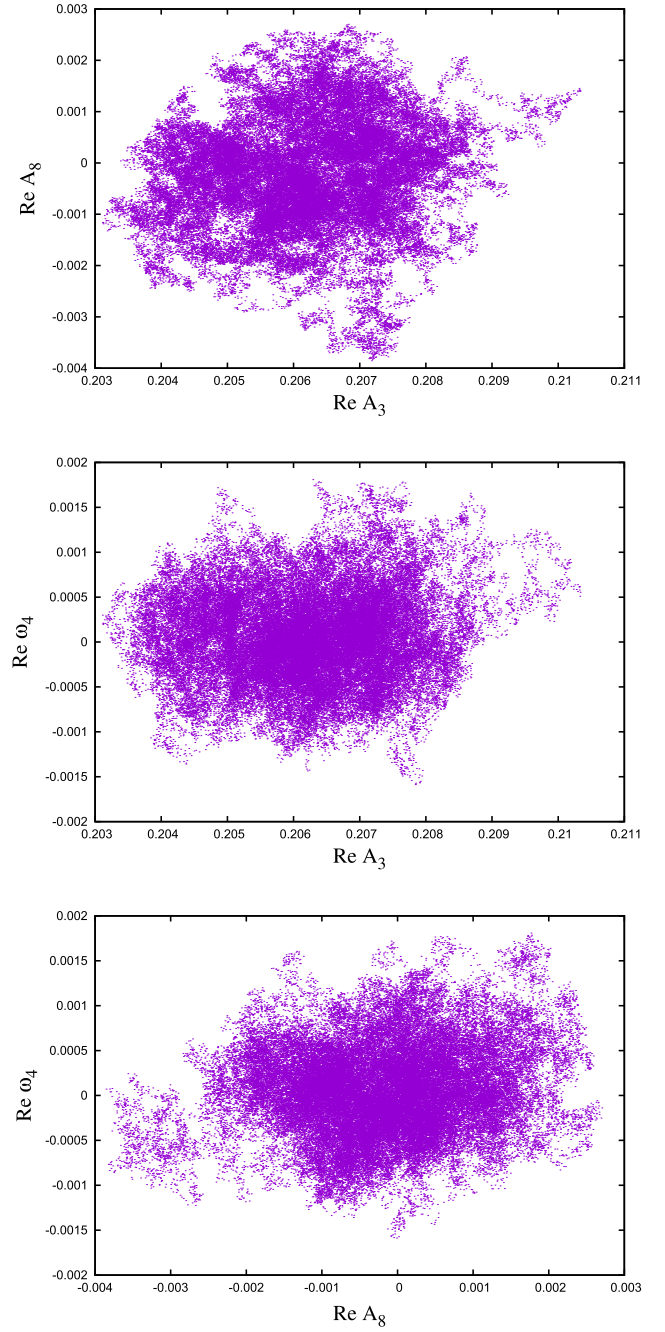


FIG. 3. The scatter plot at $T = 0.1$ GeV and $\mu = 0.3$ GeV on the $\text{Re } A_3$ – $\text{Re } A_8$, $\text{Re } A_3$ – $\text{Re } \omega_4$, and $\text{Re } A_8$ – $\text{Re } \omega_4$ planes. Here we use 80 000 configurations.

$\mu = M(\mu = 0)$. By comparison, the results at small μ are almost the same as those without a vector-type interaction [26], since the quark number density and, thus, the vector potential are small.

It should be noted that the real part of ω_4 is consistent with zero within the error bar, and, thus, the consequence obtained in the analyses in the Lefschetz-thimble method [19] is naturally understood in the path optimization method. We must consider the Wick rotation of ω_0 , and

then the model sign problem can be resolved by complexifying ω_4 . The present results imply that the ω_4 field has almost only the imaginary part; then the flow equation of the Lefschetz thimble can stall at a small value of $|\text{Re}\omega_4|$. In Fig. 3, we show the scatter plot of the hybrid Monte Carlo configurations at $T = 0.1$ GeV and $\mu = 0.3$ GeV on the $\text{Re}A_3$ - $\text{Re}A_8$, $\text{Re}A_3$ - $\text{Re}\omega_4$, and $\text{Re}A_8$ - $\text{Re}\omega_4$ planes. We can see the localized configurations around $\text{Re}A_8 = 0$ and $\text{Re}A_3 \neq 0$. Thus, this study implies that the standard PNJL model computation with a repulsive vector-type interaction under the CK symmetry ansatz and the un-Wick-rotated ω_0 in the mean-field approximation is systematically and numerically justified via the path optimization method. The scatter plot also supports the direct complexification of A_3 and A_8 without taking care of the periodicity. The Monte Carlo configurations are not spread but localized; then it is not necessary to take account of the periodic boundary condition.

IV. SUMMARY

In this study, we have applied the path optimization method to the QCD effective model with a Polyakov loop and a repulsive vector-type interaction. The feedforward neural network with the mono-hidden layer is employed to describe the good integral path in the complexified space of integral variables. The temporal components of the gluon field and the vector-type auxiliary field are complexified, and then the path is optimized via the path optimization method.

By optimizing the path (manifold), we can successfully improve the average phase factor, and calculated results of observables show reasonable behavior and have small error bars. It is not easy to optimize the integral path in the rapidly changing region of the order parameters, but we can finally improve the average phase factor by examining various initial conditions of dynamical variables.

After a few optimization steps, we can well reproduce the mean-field results at a large volume as we expect. Since we use the homogeneous ansatz of the integral variables, our numerical simulation should give the mean-field result in the large volume limit. The imaginary part of the vector-type auxiliary field starts to rapidly increase in strength above $\mu = 0.3$ GeV at $T = 0.1$ GeV. This indicates the absence of the silver-blaze problem at $T = 0$, and, thus, the

path optimization method can pick up the correct properties of the theory.

In the standard mean-field approximation, we do not perform the Wick rotation of the vector auxiliary field (ω_0). While such a treatment cannot be justified within the standard path integral formulation, it can be justified by employing a complexified theory such as the Lefschetz-thimble method and the path optimization method. In this article, we have demonstrated that the path optimization method correctly resolves the model sign problem, and then the ω_4 field takes almost a pure imaginary value which is required from the fact that the grand-canonical partition function is real. This study provides the correct numerical treatment of a repulsive vector-type interaction in the QCD effective model with a Polyakov loop.

Finally, we comment on the problem of the numerical cost. The degree of freedom is enlarged in the present calculation compared with the case without a vector-type interaction [26], the improvement of the average phase factor becomes slow, and, thus, we need more optimization steps (epochs) and/or some other extensions. One of the possible extensions to circumvent such an optimization problem is introducing a deep neural network, and it is our future work. Also, the sign problem becomes exponentially severe with an increasing system size. Then, it is important to know that the improvement of the average phase factor via the path optimization method can overcome the exponential suppression. Therefore, we need further investigation of the competition in the average phase factor between the suppression from the system size and the improvement from the path optimization. In particular, this problem becomes serious when we consider the lattice calculation. One promising approach is the reduction of the Jacobian computation cost; the diagonal ansatz of the Jacobian [40] and the nearest-neighbor lattice-cites ansatz [39] are promising examples. It will be reported elsewhere.

ACKNOWLEDGMENTS

This work is supported in part by the Grants-in-Aid for Scientific Research from JSPS (No. 15H03663, No. 16K05350, No. 18J21251, No. 18K03618, and No. 19H01898) and by the Yukawa International Program for Quark-hadron Sciences (YIPQS).

-
- [1] O. Miyamura, S. Choe, Y. Liu, T. Takaishi, and A. Nakamura, *Phys. Rev. D* **66**, 077502 (2002).
 - [2] C. Allton, M. Doring, S. Ejiri, S. Hands, O. Kaczmarek, F. Karsch, E. Laermann, and K. Redlich, *Phys. Rev. D* **71**, 054508 (2005).

- [3] R. Gaii and S. Gupta, *Phys. Rev. D* **78**, 114503 (2008).
- [4] Z. Fodor and S. Katz, *Phys. Lett. B* **534**, 87 (2002); *J. High Energy Phys.* **03** (2002) 014; **04** (2004) 050.
- [5] Z. Fodor, S. Katz, and K. Szabo, *Phys. Lett. B* **568**, 73 (2003).

- [6] P. de Forcrand and O. Philipsen, *Nucl. Phys.* **B642**, 290 (2002); **B673**, 170 (2003).
- [7] M. D’Elia and M.-P. Lombardo, *Phys. Rev. D* **67**, 014505 (2003); **70**, 074509 (2004).
- [8] H.-S. Chen and X.-Q. Luo, *Phys. Rev. D* **72**, 034504 (2005).
- [9] A. Hasenfratz and D. Toussaint, *Nucl. Phys.* **B371**, 539 (1992).
- [10] A. Alexandru, M. Faber, I. Horvath, and K.-F. Liu, *Phys. Rev. D* **72**, 114513 (2005).
- [11] S. Kratochvila and P. de Forcrand, *Phys. Rev. D* **73**, 114512 (2006).
- [12] P. de Forcrand and S. Kratochvila, *Nucl. Phys. B, Proc. Suppl.* **153**, 62 (2006).
- [13] A. Li, A. Alexandru, K.-F. Liu, and X. Meng, *Phys. Rev. D* **82**, 054502 (2010).
- [14] A. Nakamura, S. Oka, and Y. Taniguchi, *J. High Energy Phys.* **02** (2016) 054.
- [15] P. de Forcrand, *Proc. Sci.*, LAT2009 (2009) 010 [arXiv:1005.0539].
- [16] K. Fukushima, *Phys. Lett. B* **591**, 277 (2004).
- [17] K. Fukushima and Y. Hidaka, *Phys. Rev. D* **75**, 036002 (2007).
- [18] Y. Tanizaki, H. Nishimura, and K. Kashiwa, *Phys. Rev. D* **91**, 101701 (2015).
- [19] Y. Mori, K. Kashiwa, and A. Ohnishi, *Phys. Lett. B* **781**, 688 (2018).
- [20] E. Witten, *AMS/IP Stud. Adv. Math.* **50**, 347 (2011).
- [21] M. Cristoforetti, F. Di Renzo, and L. Scorzato (AuroraScience Collaboration), *Phys. Rev. D* **86**, 074506 (2012).
- [22] H. Fujii, D. Honda, M. Kato, Y. Kikukawa, S. Komatsu, and T. Sano, *J. High Energy Phys.* **10** (2013) 147.
- [23] Y. Mori, K. Kashiwa, and A. Ohnishi, *Phys. Rev. D* **96**, 111501 (2017).
- [24] A. Ohnishi, Y. Mori, and K. Kashiwa, *EPJ Web Conf.* **175**, 07043 (2018).
- [25] Y. Mori, K. Kashiwa, and A. Ohnishi, *Prog. Theor. Exp. Phys.* **2018**, 023B04 (2018).
- [26] K. Kashiwa, Y. Mori, and A. Ohnishi, *Phys. Rev. D* **99**, 014033 (2019).
- [27] K. Kashiwa, T. Hell, and W. Weise, *Phys. Rev. D* **84**, 056010 (2011).
- [28] M. Cristoforetti, T. Hell, B. Klein, and W. Weise, *Phys. Rev. D* **81**, 114017 (2010).
- [29] S. Roessner, C. Ratti, and W. Weise, *Phys. Rev. D* **75**, 034007 (2007).
- [30] C. Ratti, M. A. Thaler, and W. Weise, *Phys. Rev. D* **73**, 014019 (2006).
- [31] G. Parisi and Y.-s. Wu, *Sci. Sin.* **24**, 483 (1981).
- [32] G. Parisi, *Phys. Lett.* **131B**, 393 (1983).
- [33] A. Alexandru, G. Basar, P. F. Bedaque, G. W. Ridgway, and N. C. Warrington, *J. High Energy Phys.* **05** (2016) 053.
- [34] J. Nishimura and S. Shimasaki, *EPJ Web Conf.* **175**, 07018 (2018).
- [35] Y. Ito and J. Nishimura, *J. High Energy Phys.* **12** (2016) 009.
- [36] F. Di Renzo and G. Eruzzi, *Proc. Sci.*, LATTICE2016 (2016) 047 [arXiv:1611.08223].
- [37] F. Di Renzo and G. Eruzzi, *Phys. Rev. D* **97**, 014503 (2018).
- [38] F. Attanasio and B. Jger, *Eur. Phys. J. C* **79**, 16 (2019).
- [39] F. Bursa and M. Kroyter, *J. High Energy Phys.* **12** (2018) 054.
- [40] A. Alexandru, P. F. Bedaque, H. Lamm, and S. Lawrence, *Phys. Rev. D* **97**, 094510 (2018).
- [41] A. Alexandru, P. F. Bedaque, H. Lamm, and S. Lawrence, *Phys. Rev. D* **96**, 094505 (2017).
- [42] G. Cybenko, *Math. Control Signals Syst.* **2**, 303 (1989).
- [43] K. Hornik, *Neural Netw.* **4**, 251 (1991).
- [44] Y. Mori, K. Kashiwa, and A. Ohnishi, arXiv:1904.11140.
- [45] H. Nishimura, M. C. Ogilvie, and K. Pangeni, *Phys. Rev. D* **90**, 045039 (2014).
- [46] H. Nishimura, M. C. Ogilvie, and K. Pangeni, *Phys. Rev. D* **91**, 054004 (2015).