Path optimization for the sign problem in low-dimensional QCD and QCD effective models at finite density

Akira Ohnishi¹, Yuto Mori², Kouji Kashiwa³

Yukawa Inst. for Theoretical Physics, Kyoto U.,
 Dept. Phys., Kyoto U., 3. Fukuoka Inst. Tech.

YITP Molecule-type Workshop on Frontiers in Lattice QCD and related topics April 15-26, 2019, Kyoto, Japan.

International Molecule-type Workshop **Frontiers in Lattice QCD and related topics** April 15 - April 26 2019 Yukawa Institute for Theoretical Physics, Kyoto University







Collaborators

Akira Ohnishi¹, Yuto Mori², Kouji Kashiwa³

Yukawa Inst. for Theoretical Physics, Kyoto U.,
 Dept. Phys., Kyoto U., 3. Fukuoka Inst. Tech.







Y. Mori (PhD stu.) K. Kashiwa AO (11 yrs ago)

- 1D integral: Y. Mori, K. Kashiwa, AO, PRD 96 ('17), 111501(R) [arXiv:1705.05605]
- φ⁴ w/ NN: Y. Mori, K. Kashiwa, AO, PTEP 2018 ('18), 023B04 [arXiv:1709.03208]
- Lat 2017: AO, Y. Mori, K. Kashiwa, EPJ Web Conf. 175 ('18), 07043 [arXiv:1712.01088]
- NJL thimble: Y. Mori, K. Kashiwa, AO, PLB 781('18),698 [arXiv:1705.03646]
- PNJL w/ NN: K. Kashiwa, Y. Mori, AO, PRD 99 ('19), 014033 [arXiv:1805.08940 [hep-ph]
- PNJL with vector int. using NN: K. Kashiwa, Y. Mori, AO, arXiv:1903.03679 [hep-lat]
- 0+1D QCD: Y. Mori, K. Kashiwa, AO, in prep; AO, Y. Mori, K. Kashiwa, arXiv:1812.11506 (Lat2018 proc.)



The Sign Problem

When the action is complex, strong cancellation occurs in the Boltzmann weight at large volume. = The Sign Problem

$$\mathcal{Z} = \int \mathcal{D}x \, e^{-S(x)}, \ |\mathcal{Z}| \ll \mathcal{Z}_{pq} = \int \mathcal{D}x \, \left| e^{-S(x)} \right| \ (\text{at large } V)$$

Fermion det. is complex at finite density

 $\det D(\mu) = (\det D(-\mu^*))^* \to S_{\text{eff}} = S_{\text{boson}} - \log \det D \in \mathbb{C}$ RHIC, LHC, **Early Universe** Difficulty in studying T Lattice QCD finite density in LQCD Heavy-Ion Collisions (BES, FAIR \rightarrow Heavy-Ion Collisions, NICA, J-PARC) Neutron Star, Sym. Nucl. CSC Matter **Binary Neutron** 0 Star Mergers, **Ouark Matter** Pure Neut Nuclei, ... Matter Sym. E **Neutron Star** 1

 $\delta = (N-Z)/A \ (or Y_Q(hadron) = Q_h/B \sim (1-\delta)/2)$



Approaches to the Sign Problem

- Standard approaches
 - Taylor exp., Imag. μ (Analytic cont. / Canonical), Strong coupling
- Integral in Complexified variable space
 - Lefschetz thimble method Witten ('10), Cristoforetti+ (Aurora) ('12), Fujii+ ('13), Alexandru+ ('16).
 - Complex Langevin method Parisi ('83), Klauder ('83), Aarts+ ('11), Nagata+ ('16); Seiler+ ('13), Ito+ ('16).

Path optimization method

Mori, Kashiwa, AO ('17,'18,'19); Kashiwa, Mori, AO ('18,19); AO, Mori, Kashiwa ('18,'19); Alexandru+('18), Bursa, Kroyter ('18)



Integral in Complexified Variable Space

Simple Example: Gaussian integral (bosonized repulsive int.) Mori, Kashiwa, AO ('18b)



Complexified variable methods = Extension of the saddle point integral



Lefschetz thimble & Complex Langevin methods

Lefschetz thimble method

Witten ('10), Cristoforetti+ (Aurora) ('12), Fujii+ ('13), Alexandru+ ('16).

• Flow eq. from a fixed point $\sigma \rightarrow$ thimble (Im(S)=const.)

$$\mathcal{J}_{\sigma}: \frac{dz_i(t)}{dt} = \overline{\left(\frac{\partial S}{\partial z_i}\right)} \to \frac{dS}{dt} = \sum_i \left|\frac{\partial S}{\partial z_i}\right|^2 \in \mathbb{R}, \quad \mathcal{C} = \sum_{\sigma} n_{\sigma} \mathcal{J}_{\sigma}$$

Problems:Phase of Jacobian, Multimodal prb., Stokes phenomena, ...

Complex Langevin method

Parisi ('83), Klauder ('83), Aarts+ ('11), Nagata+('16); Seiler+ ('13), Ito+ ('16).

Complex Langevin eq.→ Configs.

$$\frac{dz_i}{dt} = -\frac{\partial S}{\partial z_i} + \eta_i(t)(\eta_i : \text{White noize}), \ \langle \mathcal{O}(x) \rangle = \langle \mathcal{O}(z) \rangle$$

Problems: Wrong conversion, Boundary terms, ...



Path optimization method

Mori, Kashiwa, AO ('17,'18,'19); Kashiwa, Mori, AO ('18,19); AO, Mori, Kashiwa ('18,'19); Alexandru+('18), Bursa, Kroyter ('18)

Integration path is optimized to evade the sign problem, i.e. to enhance the average phase factor.

$$APF = \langle e^{i\theta} \rangle_{pq} = \int dx e^{-S} / \int dx |e^{-S}| = \mathcal{Z}/\mathcal{Z}_{pq}$$

Sign Problem → Optimization Problem

Cauchy(-Poincare) theorem: the partition fn. is invariant if

- the Boltzmann weight W=exp(-S) is holomorphic (analytic), and the path does not go across the poles and cuts of W.
- S is singular but W is not singular when fermion det.=0.



z

Application of POM to Field Theory

Cost function: a measure of the seriousness of the sign problem. $\mathcal{F}[z(x)] = |\mathcal{Z}| \left(\left| \langle e^{i\theta} \rangle_{pq} \right|^{-1} - 1 \right)$

Optimization: Gradient Descent or Neural Network

• Neural network = Combination of linear and non-linear transf.

 $a_{i} = g(W_{ij}^{(1)}x_{j} + b_{i}^{(1)}) \text{ variational}$ $f_{i} = g(W_{ij}^{(2)}a_{j} + \underline{b_{i}^{(2)}}) \text{ parameters}$ $z_{i} = x_{i} + i(\underline{\alpha_{i}}f_{i}(x) + \underline{\beta_{i}})$ $g(x) = \tanh x \text{ (activation fn.)}$

 Universal approximation theorem Any fn. can be reproduced at (hidden layer unit #) → ∞ G. Cybenko, MCSS 2 ('89) 303 K. Hornik, Neural networks 4('91) 251





Optimization of many parameters

Stochastic Gradient Descent method, E.g. ADADELTA algorithm M. D. Zeiler, arXiv:1212.5701 Grad. Desc. :

 $dc_i/dt = -\partial \mathcal{F}/\partial c_i$ Learning rate par. in (j+1)th step $c_i^{(j+1)} = c_i^{(j)} - \eta v_i^{(j+1)}$ mean sq. ave. of v decay rate $r_{i}^{(j+1)} = \gamma r_{i}^{(j)} + (1 - \gamma) (F_{i}^{(j)})^{2}$ $s_i^{(j+1)} = \gamma s_i^{(j)} + (1-\gamma)(v_i^{(j+1)})^2$ gradient $-F_i = \partial \mathcal{F} / \partial c_i$ Machine learning evaluated ~ Educated algorithm in MC to generic problems Cost fn. (batch training) Y TP -A. Ohnishi, FLQCD, Apr. 19, 2019 9

Hybrid Monte-Carlo with Neural Network

Initial Config. on Real Axis

HMC
$$H(x,p) = \frac{p^2}{2} + \operatorname{Re}S(z(x))$$
 Jaco

Jacobian → via Metropolis judge

Do k = 1, Nepoch Do j = 1, Nconf/Nbatch

Mini-batch training of Neural Network

$$c_i^{(j+1)} = c_i^{(j)} - \eta v_i^{(j+1)}$$

Grad. wrt parameters (Nbatch configs.) $F_i = \frac{1}{N_{\text{batch}}}$

New Nbatch configs. by HMC

$$F_{i} = \frac{1}{N_{\text{batch}}} \sum_{n} \frac{\partial \mathcal{F}(n)}{\partial c_{i}}$$
$$H(x, p) = \frac{p^{2}}{2} + \text{Re}S(z(x))$$

Enddo Enddo

Nbatch ~ 10, Nconfig ~ 10,000, Nepoch ~ (10-20)





VITP KVOID

Benchmark test (2): Complex φ^4 theory at finite μ

- Complex Langevin & Lefschetz thimble work.
 G. Aarts, PRL102('09)131601; H. Fujii, et al., JHEP 1310 (2013) 147
- How about POM ?
 - 1+1D Complex φ⁴ theory
 Y. Mori, K. Kashiwa, AO, PTEP 2018 ('18), 023B04 [arXiv:1709.03208]

$$\mathcal{L} = \partial_{\mu} \phi^* \partial^{\mu} \phi - m^2 \phi^* \phi - \lambda (\phi^* \phi)^2$$

$$S = \sum_{x} \left[\frac{(4+m^2)}{2} \phi_{a,x} \phi_{a,x} + \frac{\lambda}{4} (\phi_{a,x} \phi_{a,x})^2 - \phi_{a,x} \phi_{a,x+\hat{1}} - \cosh \mu \phi_{a,x} \phi_{a,x+\hat{0}} + i\epsilon_{ab} \sinh \mu \phi_{a,x} \phi_{b,x+\hat{0}} \right]$$

$$\left(\phi = \frac{1}{\sqrt{2}} (\underline{\phi_1} + i\underline{\phi_2}) \right) \qquad \text{complex}$$

$$Complexify$$



POM in 1+1D φ^4 theory

POM for 1+1D φ⁴ theory

Y. Mori, K. Kashiwa, AO, PTEP 2018 ('18), 023B04 [arXiv:1709.03208]

- 4^2 , 6^2 , 8^2 lattices, $\lambda = m = 1$
- $\mu_c \sim 0.96$ in the mean field approximation



POM also works !

- Enhancement of the APF after optimization.
- Density is suppressed at $\mu < m$. (Silver Blaze)



Path Optimization Method w/ Neural Network seems to work in 1D integral and simple field theories.

How about gauge theory ? What happens when phase transition occurs ?



Contents

Introduction of Path Optimization Method

Y. Mori, K. Kashiwa, AO, PRD 96 ('17), 111501(R) [1705.05605] Y. Mori, K. Kashiwa, AO, PTEP 2018 ('18), 023B04 [1709.03208] AO, Y. Mori, K. Kashiwa, EPJ Web Conf. 175 ('18), 07043 [1712.01088](Lat 2017)

Application to gauge theory: 1-dimensional QCD

Mori, K Kashiwa, AO, in prep. AO, Y. Mori, K. Kashiwa, PoS LATTICE2018 ('19), 023 (1-15) [1812.11506]

Application to QCD effective models

K. Kashiwa, Y. Mori, AO, PRD99('19)014033 [1805.08940] K. Kashiwa, Y. Mori, AO, arXiv:1903.03679 [hep-lat]

Discussions

Summary



Application to Gauge Theory: 1 dimensional QCD



0+1 dimensional QCD

0+1 dimensional QCD (1 dim. QCD) with one species of staggered fermion on a 1xN_τ lattice

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



Bilic+('88), Ravagli+('07), Aarts+('10, CLM), Bloch+('13, subset), Schmidt+('16, LTM), Di Renzo+('17, LTM)

- A toy model, but includes the actual source of 3+1D QCD sign prob.
- Reduces to a diagonalized one-link problem.

$$\mathcal{Z} = \int dU e^{-S} = \int dx_1 dx_2 J \left[\frac{8}{3\pi^2} \prod_{a < b} \sin^2 \left(\frac{z_a - z_b}{2} \right) \right] \left[\prod_a (X_N + 2\cos(z_a - i\mu)) \right]$$

Haar measure exp(-S)

 \rightarrow Analytic results are known.



Fermion determinant in 1 dim. QCD

Fermion determinant (Temporal gauge) reduces to Nc x Nc det.



det
$$D = \det \left[X \otimes 1_c + (-1)^{N_\tau} e^{\mu/T} U + e^{-\mu/T} U^{-1} \right] \checkmark N_c \quad I_k = 2m_q(\boldsymbol{x}, \tau_k)$$

For constant σ , X is obtained as

Partition Function in 1 dim. QCD

Partition Function

$$D = X + e^{\mu/T}U + e^{-\mu/T}U^{-1}$$

$$\mathcal{Z} = \int dU \det D(U) = \frac{\sinh[(N_c + 1)E/T]}{\sinh(E/T)} + 2\cosh(N_c \,\mu/T)$$

$$= X^3 - 2X + 2\cosh(N_c \,\mu/T) \quad (N_c = 3)$$

$$\det D = X^3 + N_c X(N_c \overline{P}_U P_U - 1) + 2\cosh(N_c \mu/T)$$

$$+ N_c X^2(e^{\mu/T} P_U + e^{-\mu/T} \overline{P}_U) + N_c X(e^{2\mu/T} \overline{P}_U + e^{-2\mu/T} P_U)$$

$$+ N_c e^{\mu/T}(N_c \overline{P}_U^2 - 2P_U) + N_c e^{-\mu/T}(N_c P_U^2 - 2\overline{P}_U)$$

$$(N_c = 3, P_U = \operatorname{tr} U/N_c, \int dU \overline{P}_U P_U = 1/N_c^2)$$

Chiral condensate, Quark number density, Polyakov loop $\sigma = \frac{T}{V} \frac{\partial}{\partial m} \log Z, \quad n_q = -\frac{T}{V} \frac{\partial}{\partial \mu} \log Z$ $P = \left\langle \frac{1}{N_c} \operatorname{Tr} U \right\rangle = \frac{1}{N_c} \frac{(X^2 - 1)e^{-\mu/T} + Xe^{2\mu/T}}{X^3 - 2X + 2\cosh(3\mu/T)} \qquad Faldt, Petersson ('86)$ Bilic, Demeterfi ('88)

1 dim. QCD in diagonalized gauge (1)

■ 2 variable problem \rightarrow 2D mesh point integral \rightarrow y_{1,2}(x₁,x₂) are variational parameters by themselves.





- Average phase factor > 0.997
 - (Normal) gradient descent
 - Good enough for small lattice in 3+1D.

$$(APF)_{0+1}^{L^3} \simeq 0.21(L=8)$$





1 dim. QCD in diagonalized gauge (2)

Jacobian is also important !

 $|\mathrm{Im}(JW)| \ll |\mathrm{Im}W| \quad (W = H\exp(-S))$

There are six regions with large stat. weight | JW |.
 Symmetry : S(-z) = (S(z*))*, z_i ↔ z_j(i, j = 1, 2, 3)
 → Problematic in sampling in Hybrid MC





1 dim. QCD w/o diagonalized gauge fixing (1)

Complexification of link variable

$$U \to \mathcal{U}(U) = U \prod_{a=1}^{N_c^2 - 1} e^{-y_i \lambda_i}$$
$$= U e^{-y_1 \lambda_1} e^{-y_2 \lambda_2} \cdots e^{-y_8 \lambda_8} \in \mathrm{SL}(3)$$

Derivative wrt y's is easy. Parametrization deps. is taken care by J.

Hybrid Monte-Carlo in 1 dim. QCD

$$H = \frac{P^2}{2} + \operatorname{Re}(S(\mathcal{U}(U)))$$

■ 8 variables → path optimization using Neural Network



1 dim. QCD w/o diagonalized gauge fixing (2)





1 dim. QCD w/o diagonalized gauge fixing (3)

Statistical weight distribution in diagonalized gauge
 ~ Config. dist. in Hybrid MC w/o diag. gauge fixing









Polyakov-loop-extended NJL (PNJL) model

- Sign problem is more severe around the phase boundary.
 e.g. S. Tsutsui et al., 1811.07647; Y. Ito et al., 1811.12688.
 → Let us discuss QCD effecitve models !
- Polyakov-loop-extended Nambu-Jona-Lasinio (PNJL) model with vector coupling

$$\mathcal{L}_E = \bar{q}(\not\!\!D + m_0)q - G\left[(\bar{q}q)^2 + (\bar{q}i\gamma_5\boldsymbol{\tau}q)^2\right] + \mathcal{V}_g(\Phi,\bar{\Phi}) + G_v(\bar{q}\gamma_\mu q)^2$$

Polyakov Vector

Bosonizaiton & Truncation to homogeneous aux. field

$$\mathcal{V} = -2N_f \int^{\Lambda} \frac{d^3p}{(2\pi)^3} \left[N_c E_p + T \log(f^- f^+) \right]$$

$$+ G(\sigma^2 + \pi^2) + G_v \omega^2 + \mathcal{V}_g$$

$$f^- = 1 + 3\Phi e^{-\beta E_p^-} + 3\overline{\Phi} e^{-2\beta E_p^-} + e^{-3\beta E_p^-}$$

$$E_p = \sqrt{p^2 + (m_0 - 2G\sigma)^2}, \ E_p^{\mp} = E_p \mp \tilde{\mu}, \ \tilde{\mu} = \mu - 2iG_v \omega$$
A. Ohnishi, FLQCD, Apr. 19, 2019 26

Repulsive interaction causes the sign problem

Hubbard-Stratonovich transformation of repulsive interaction

$$\exp[-\alpha(\bar{q}\Gamma q)^2] = \int d\omega \exp[-\alpha\omega^2 + i\alpha\omega\bar{q}\Gamma q]$$

This "model" sign problem causes trouble in Shell Model Monte-Carlo, Strong-coupling LQCD, ...



POM for PNJL (w/o vector coupling)

Partition function

$$\mathcal{Z} = \int dX \exp(-\beta \mathcal{V}/V) = \int dX \exp(-k\mathcal{V}/T^4) \ (dX = dA_3 dA_8 d\sigma d\omega \dots)$$

Homogenous field ansatz & const. k approx. → Results converge to mean field results at large k Cristoforetti, Hell, Klein, Weise ('10) (MC-NJL)

POM works in PNJL !

Average phase factor ~ 1, Pol. loop converges to MF results.





POM for PNJL with vector coupling

- POM works in PNJLv
 - Pol. loop converges to MF results.
 - Average phase factor is enhanced significantly, but we still find the region, APF < 1 and we need special care for the initial config. distribution. (Optimization is not automatic in this case.)





Do we describe multi thimbles ?

- It seems yes.
- Statistical weight in NJLv
 on (σ, ω₄) plain after optimization
 on the 2D mesh.
 Three peaks (or should arg)
 - → Three peaks (or shoulders)
- Transition from the Nambu-Goldstone phase to the Wigner phase occurs at μ=(340-350) MeV
- In the V → ∞ limit, this corresponds to the phase transition. (We may need exchange MC or different tempering.)



Summary

- Complexified variable methods (LTM, CLM, POM) are promising tools to tackle the sign problem. [c.f. Talk by Fukuma]
- Path optimization method has been demonstrated to work in 1 dim. integral, 1+1 dim. scalar theory at finite density, 0+1 dim. QCD, and PNJL model w/ and w/o vector coupling.
 - POM does not suffer from zero point of fermion det., since it is not a singular point of the Boltzmann weight.
 - Complex phase from Jacobian and the Boltzmann weight cancels with each other, and the residual sign problem is evaded.
 - In 1 dim. QCD, an apparent multimodal problem in the diag. gauge can be avoided by calc. w/o diag. gauge fixing.
 - Sometimes, the average phase factor does not easily grow during the optimization. Improving the opt. method and/or knowledge of preferred path would be necessary (not yet *ab initio*).
- To do: 1+1 D QCD, Reducing cost O(V³), ...



Prospect

- Path optimization in 3+1 D field theories would require reduction of numerical cost.
 - **Imaginary part**

= f (real parts of same point and nearest neighbor points) may be a good guess.

Deep learning (# of hidden layers > 3) may be helpful to explore complex paths, which human beings (~ 7 layers) cannot imagine, while "Understanding" the results of machine learning need to be done by human beings (at present).

Defelipe 2011a (Review). The evolution of the brain, the human nature of cortical circuits, and intellectual creativity. Front Neuroanat 5, 29.





Thank you for your attention !

