Exploration of Efficient Neural Network for Path Optimization Method arXiv:2109.11710, 220X.XXXXX

Yusuke Namekawa
(Kyoto Univ. $\rightarrow$ Hiroshima Univ.)

in collaboration with

K.Kashiwa, H.Matsuda, A.Ohnishi, H.Takase

Contents

[In memory of Taniguchi-san]

- I have 36 papers with Taniguchi-san
- Taniguchi-san and I are core members of an open source code for lattice QCD "Bridge++" : now I maintain Taniguchi-san's codes

Bridge++ party at Irish pub
(Feb. 2019)

From left:
(Namekawa's hand), Aoyama-san, Matsufuru-san, Taniguchi-san
[photo by Kanaya-san]

1 Main message

- For path optimization in a gauge theory,

  "it is efficient to employ a neural network which respects the gauge symmetry"

  ex. gauge invariant input / gauge covariant neural network

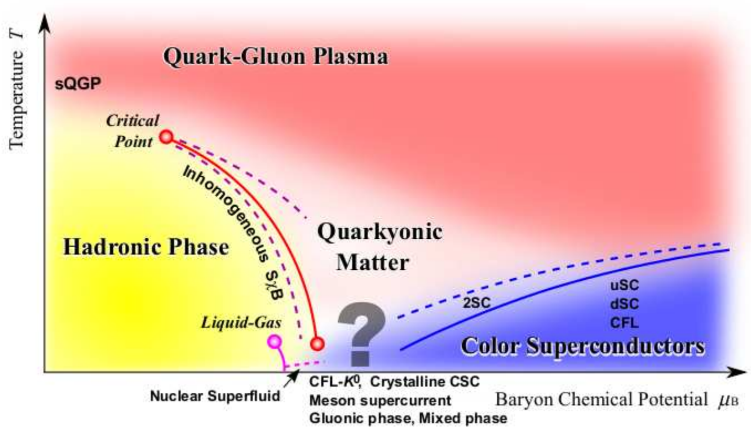
  cf. similar idea is used as a part of gauge equivariant convolutional neural network T.Cohen et al.(2019); Favoni et al.(2020)

  ◊ Gauge variant neural network works but costs a lot

2 Motivation

- QCD at high density has been investigated only by Complex Langevin Method(CLM) Sexty(2013),...,YN et al.(2021) due to sign problem
  $\leftarrow$ Action becomes complex at $\mu \neq 0$, which prohibits Monte Carlo simulation using probability weight $P = e^{-S}$
- CLM can not cover the whole phase diagram of QCD due to validity condition of CLM
  $\rightarrow$ Alternative approach is needed

  ◇ See next page



Fukushima,Hatsuda(2011)

[Several methods to overcome sign problem]

- Complex Langevin method Parisi(1983),Klauder(1984),...

  [Low cost, generally work only in limited parameter region

  Revived by clarifying correctness criterion Aarts et al.(2009), Nishimura,Shimasaki(2015)

  Calculated equation of state of QCD at finite density Sexty(2019), Attanasio et al.(2022) ⟶ cf. Itou-san's talk on 9/22]

- Lefschetz thimble method Witten(2010),Cristoforetti et al.(2012),Fujii et al.(2013),Alexandru et al.(2015),...

  [Middle cost, difficult to resolve sign and ergodicity problems simultaneously

  ◇Tempered Lefschetz thimble method Fukuma,Umeda(2017),...

  [Middle cost, resolve sign and ergodicity problems simultaneously]

  ◇Worldvolume Hybrid Monte Carlo(WV-HMC) method Fukuma,Matsumoto(2020),...

  [Low cost, resolve sign and ergodicity problems simultaneously]

  ⟶ See Fukuma-san's talk@Lattice2022(Aug. 8th, 2022)]

- Path optimization method Mori,Kashiwa,Ohnishi(2017),...

  [Middle cost, find optimized path by machine learning, deal with ergodicity problem by tempering]

- Tensor network Levin,Nave(2007),...

  [High cost → See talks on 9/21]

[Path optimization method(POM)] Mori et al.(2017),Alexandru et al.(2018),Bursa,Kroyter(2018),...

- POM is a method which complexifies dynamical variables and deforms the integration path using machine learning to minimize sign problem
- POM has been successful in models with small redundant degrees of freedom, but is not efficient with large gauge degrees of freedom
  - ◇ One solution is gauge fixing but costs a lot Mori et al.(2019),...
  - ◇ We found gauge invariant input / gauge covariant neural network works well
    $\to$ This talk

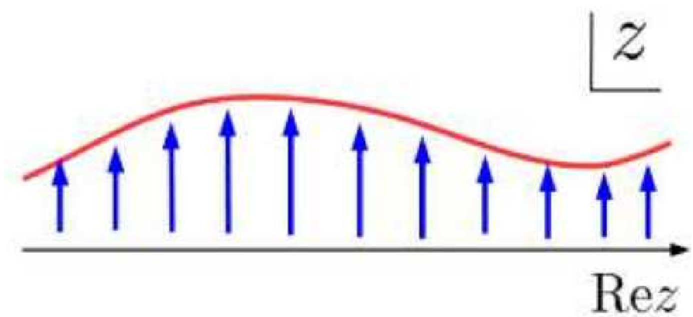
NB. Cauchy's integral theorem ensures the following equality

$$\langle \mathcal{O} \rangle := \frac{1}{Z}\int_R DU\,\mathcal{O}e^{-S[U]} = \frac{1}{Z}\int_C \mathcal{D}\mathcal{U}\,\mathcal{O}e^{-S[\mathcal{U}]}$$

$\mathcal{O}$: observable, Z : partition func, S : action, U : link variable

$$U_{x,\mu} := e^{igA_\mu(x)} \to \mathcal{U}_{x,\mu} = e^{-g\mathrm{Im}\mathcal{A}_\mu(x)}e^{ig\mathrm{Re}\mathcal{A}_\mu(x)}$$

$A_\mu(x) \in \mathbb{R} \to \mathcal{A}_\mu(x) \in \mathbb{C}$ determined by machine learning



Path optimization

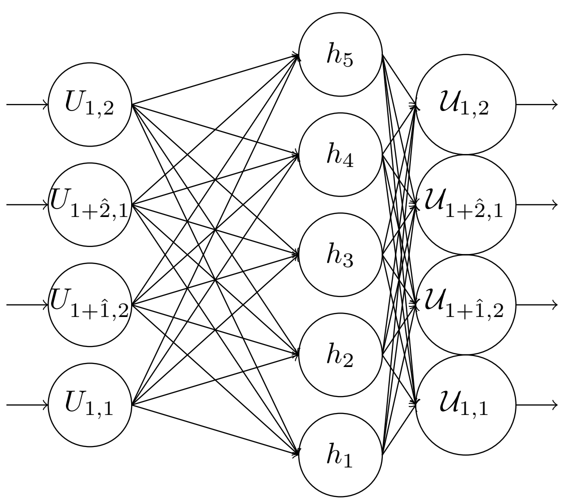
Ohnishi(2017)

[Gauge variant neural network]

$$\underbrace{U_{x,\mu}}_{\text{input layer}} \to \text{hidden layer} \to \underbrace{\mathcal{U}_{x,\mu}}_{\text{output layer}}$$

- Machine learning chooses best path which enhances phase factor
  $e^{i\theta} := Je^{-S}/|Je^{-S}|, \qquad J := \det(\partial\mathcal{U}/\partial U)$

  ◇ Averaged phase factor $|\langle\exp(i\theta)\rangle|$ is an indicator of sign problem:
  $|\langle\exp(i\theta)\rangle| = 1$ for mild, $|\langle\exp(i\theta)\rangle| = 0$ for severe



Input layer : $t_i = U_{x,\mu}$

Hidden layer : $h_j = F(w^{(1)}_{ji} t_i + b_j)$

Output layer : $z_n = \omega_n F(w^{(2)}_{nj} h_j + b_n)$

$w, b, \omega :=$ parameters of neural network

$F(x) := \tanh(x)$, activation func

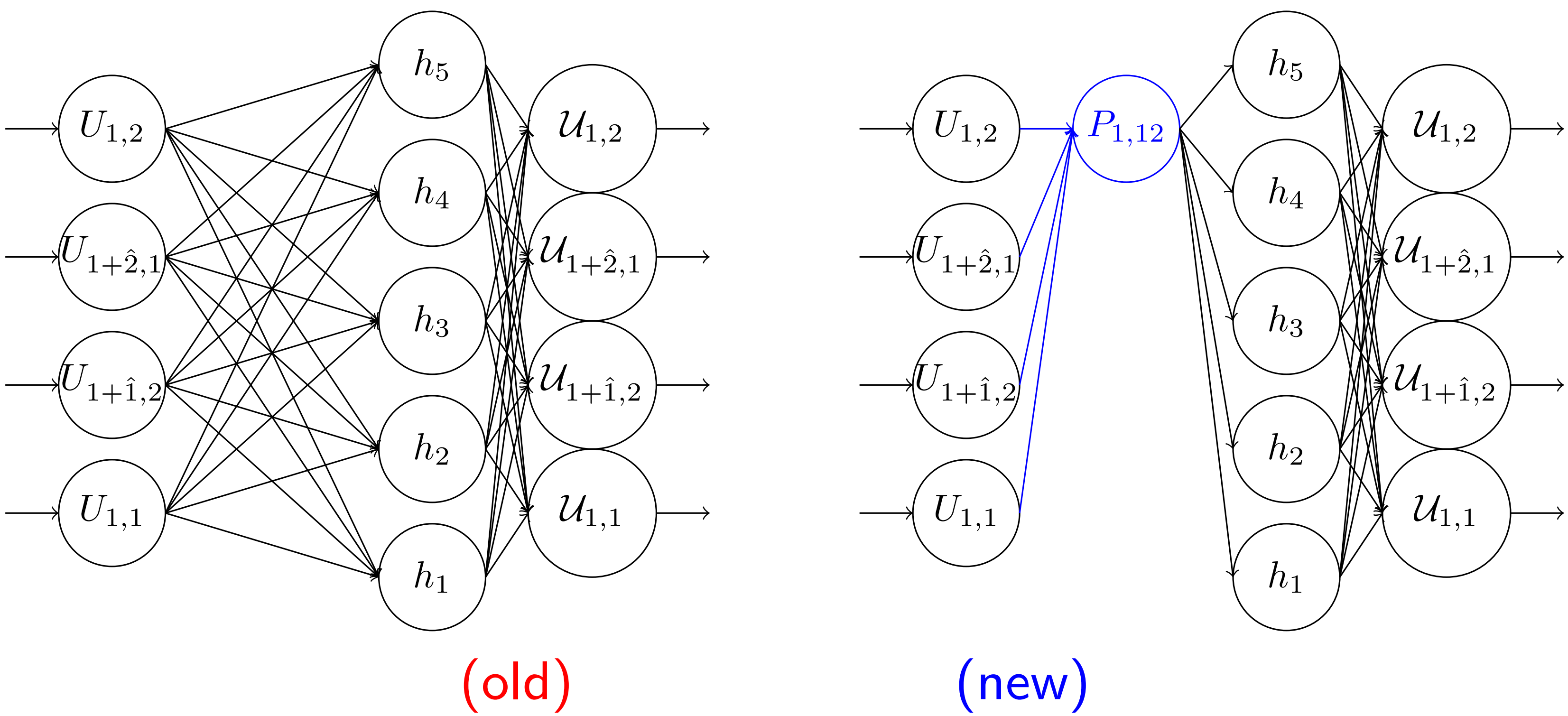
$\mathcal{F}_{\text{cost}}(t) := |Z| \left(|\langle e^{i\theta(t)} \rangle_{\text{pq}}|^{-1} - 1 \right)$, pq : phase quenched

$$\langle \mathcal{O} \rangle_{\text{pq}} := \frac{1}{Z} \int \mathcal{D}U \left[\mathcal{O} \,\middle|\, J\, e^{-S} \right|\Big]_{\mathcal{U}\in\mathcal{C}}$$

[Gauge invariant neural network] YN et al.(2021)

$$\underbrace{U_{x,\mu} \to P_{x,\mu\nu}}_{\text{input layer}} \to \text{hidden layer} \to \underbrace{\mathcal{U}_{x,\mu}}_{\text{output layer}}$$

- We adopt gauge invariant plaquette in the input layer
  $P_{x,12} := U_{x,1}\, U_{x+\hat{1},2}\, U^{-1}_{x+\hat{2},1}\, U^{-1}_{x,2}$
  - ◇ Similar idea is used as a part of gauge equivariant convolutional neural network
    T.Cohen et al.(2019); Favoni et al.(2020)



[Gauge covariant neural network] Tomiya,Nagai(2021)

$$\underbrace{U_{x,\mu}}_{\text{input layer}} \to \underbrace{\tilde{\mathcal{U}}^{(1)}_{x,\mu} \to \cdots \to}_{\text{hidden layer}} \underbrace{\mathcal{U}_{x,\mu}}_{\text{output layer}}$$

$$\tilde{\mathcal{U}}^{(l+1)}_{x,\mu} = \exp[iW^{(l)}_{x,\mu}]\,\tilde{\mathcal{U}}^{(l)}_{x,\mu}, \quad W^{(l)}_{x,\mu} := \sum_{\nu\neq\mu}\left(\rho^{(l)}_{+}\mathcal{P}^{(l)}_{x,\mu\nu} + \rho^{(l)}_{-}\mathcal{P}^{(l)}_{x,\mu\nu}{}^{-1}\right)$$

$\rho^{(l)}_{\pm}$: parameters in neural network, (l) : number of smearing

- The hidden layer is constructed by Stout-like smearing, which is gauge covariant
- We use $N_{\mathrm{stout}} = 2$ in this work

3 Application to 2-dim $U(1)$ gauge theory

- Sign problem is originated from the complex coupling $\beta = 1/(ga)^2 \in \mathbb{R} \to \mathbb{C}$
- Analytic result has been obtained
  $\to$ Good testbed for new approach Kashiwa,Mori(2020),Pawlowski et al.(2021)

cf. 2-dim $U(1)$ + θ-term, another type of sign problem, is investigated by tensor renormalization

Kuramashi and Yoshimura(2019) and complex Langevin Hirasawa et al.(2020)

$$S = -\frac{\beta}{2} \sum_x \left(P_{x,12} + P_{x,12}^{-1} \right)$$

$$\beta = 1/(ga)^2 \in \mathbb{R} \to \mathbb{C}$$

$$P_{x,12} := U_{x,1}\, U_{x+\hat{1},2}\, U_{x+\hat{2},1}^{-1}\, U_{x,2}^{-1}$$

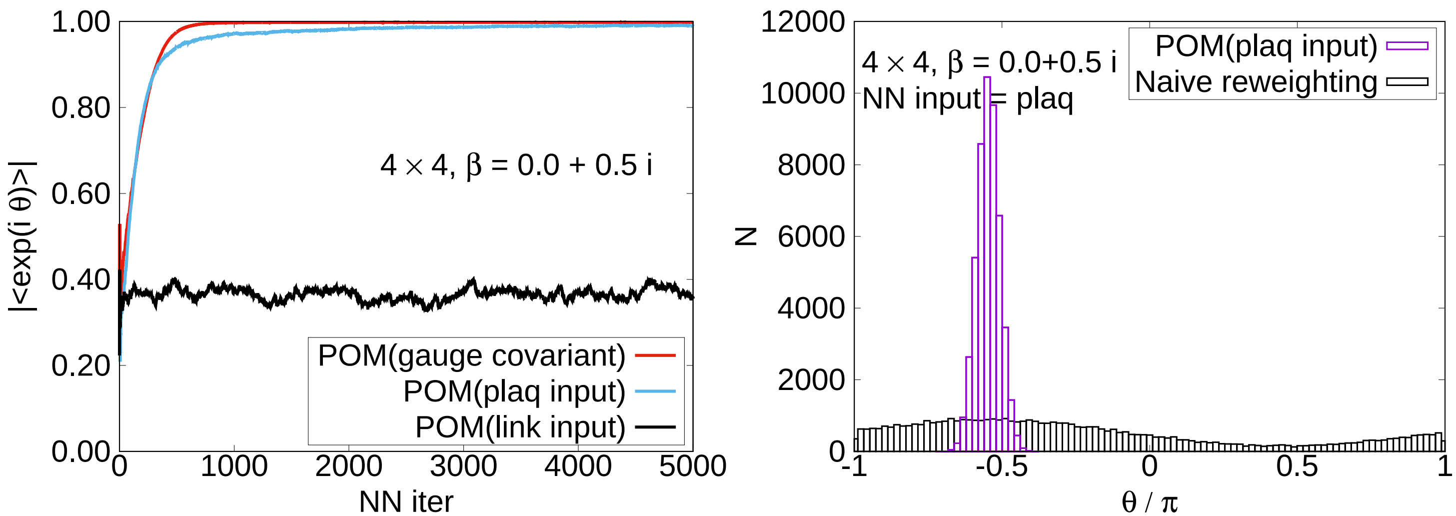
[Analytic result] Wiese(1988),...

$$Z := \int dU e^{-S} = \sum_{n=-\infty}^{+\infty} I_n(\beta)^V$$

$$I_n(\beta) := \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi\, e^{\beta \cos\phi - in\phi}$$

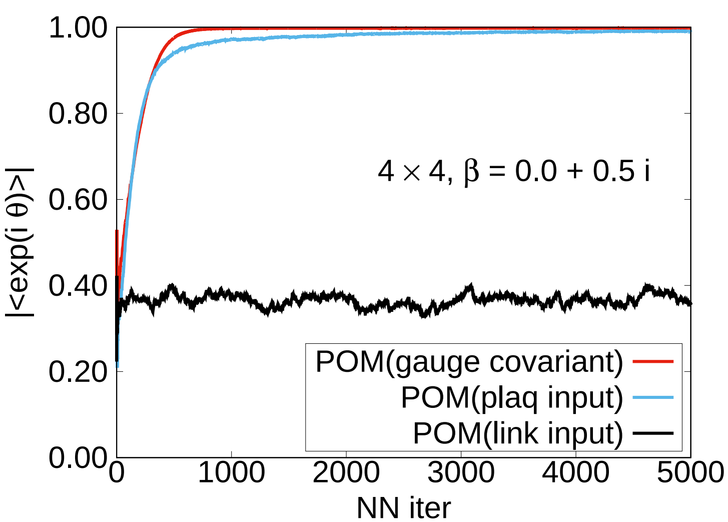
[Neural network iteration dependence of average phase factor]

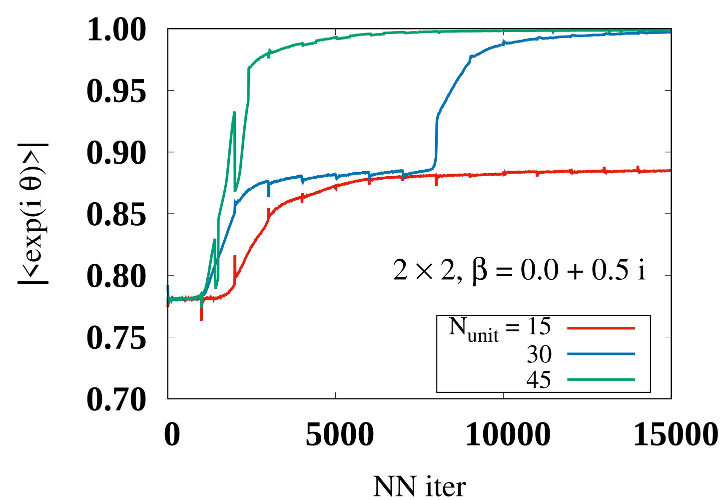
- Gauge invariant input and gauge covariant neural network successfully enhances averaged phase factor $|\langle \exp(i\theta) \rangle|$
  - ♢ The peak structure in histogram of the averaged phase factor is made clear
- (See next page for gauge variant link-variable input)



[Neural network iteration dependence of average phase factor(continued)]

- Naive link-variable input does not enhance the averaged phase factor by 5000 neural network iterations with $N_{\mathrm{unit}} = 16$ hidden layer units
- Naive link-variable input with much larger neural network iterations and larger hidden layer units enhances the averaged phase factor
  - ♢ Neural network can learn the gauge symmetry, but the cost is high $\rightarrow$ It is better to employ gauge invariant input or gauge covariant neural network

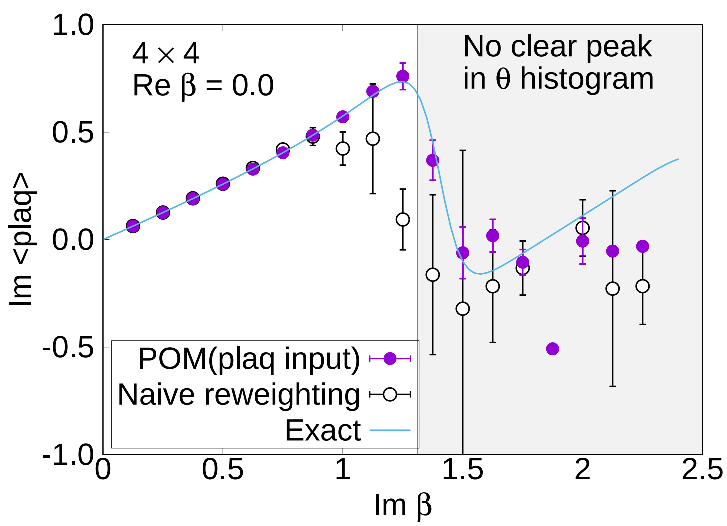


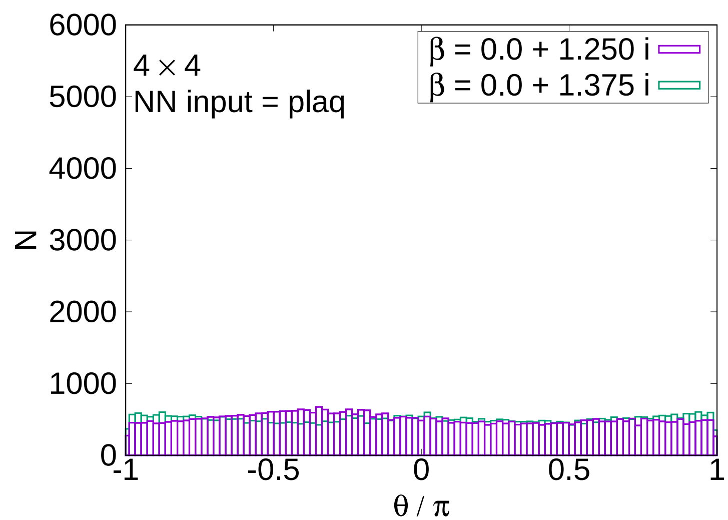


Calculated by H.Matsuda

[Comparison with the analytic solution]

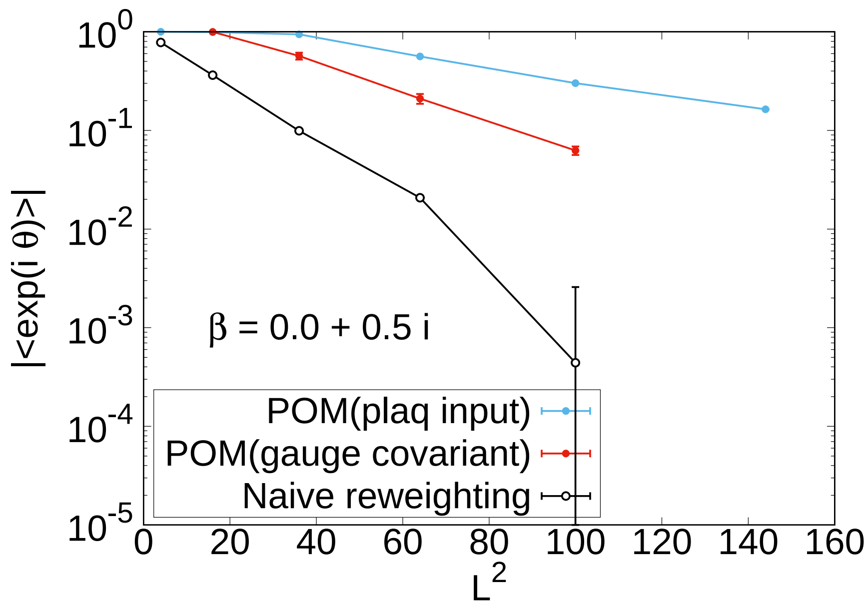
- At small $\mathrm{Im}\beta$, the POM successfully reproduces the exact solution with controlled error
- At large $\mathrm{Im}\beta$, the POM fails due to no enhancement of the peak in histogram of the averaged phase factor
  $\rightarrow$ Further improvement is required, such as tempering





[Volume dependence]

- Enhancement of the averaged phase factor is confirmed
  - ◇ Gauge invariant input / gauge covariant neural network shows milder volume dependence than that of naive reweighting



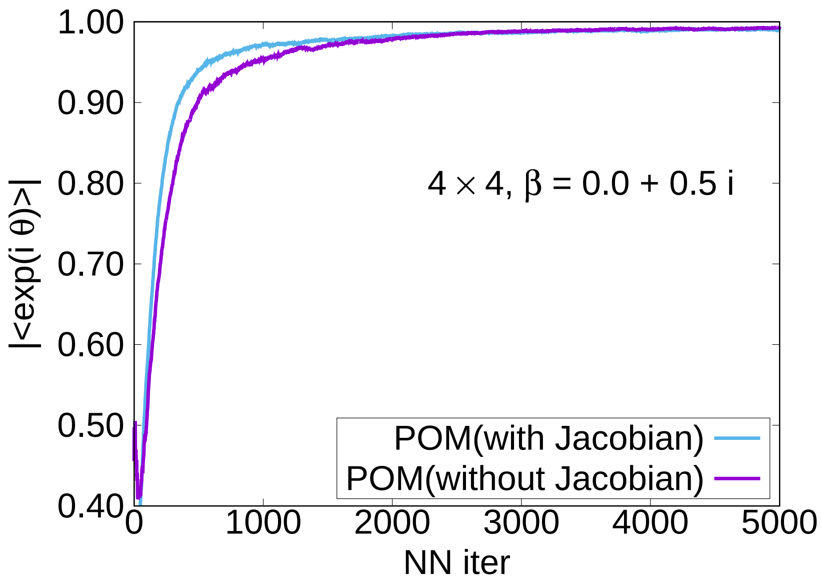
[Test approximated Jacobian in neural network]

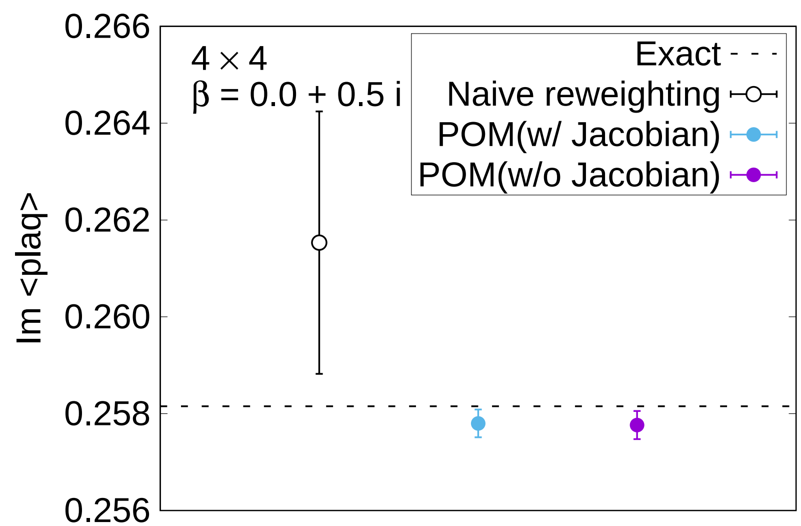
For a large scale simulation, cost reduction is necessary

- The main bottleneck is cost of Jacobian $O(N_{\rm dof}^3)$
- We test $J = 1$ approximation in the learning process
  $\leftarrow$ We still need the exact Jacobian for final output and measurement

cf. WV-HMC needs no explicit form of Jacobian in Monte-Carlo update Fukuma,Matsumoto(2020)

$\diamondsuit$ POM using $J = 1$ approximated neural network can enhance the averaged phase factor with a slightly larger error by 1%





4 Summary

We explored efficient ways for the path optimization method, which reduces sign problem by complexification of path using machine learning

- Gauge invariant input / gauge covariant neural network successfully enhances the average phase factor, i.e., reduces sign problem
  $\leftarrow$ Gauge variant neural network can also enhance the average phase factor with much larger cost
- $J = 1$ approximated neural network still leads to large enhancement of the average phase factor (at least in our setup)

[Future direction]

- Try non-Abelian case, such as $SU(2)$, $SU(3)$
- Test another type of sign problem, such as system with θ-term