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

Yusuke Namekawa
(Hiroshima Univ.)

in collaboration with

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

Contents

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

  Favoni et al.(2020)

  ◊ Gauge variant neural network works but costs a lot

2 Motivation

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
      $\rightarrow$ This talk

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

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

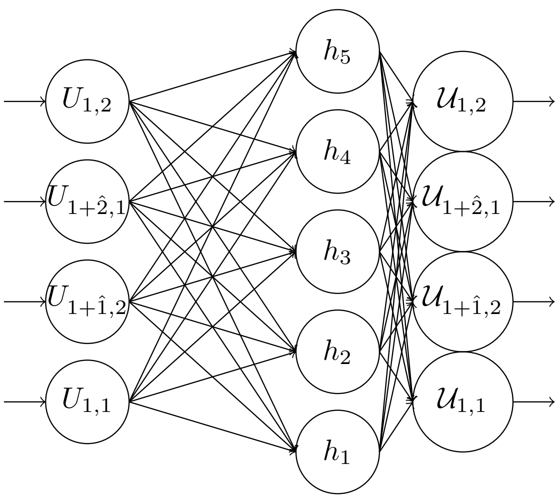
$U_{x,\mu} := e^{igA_\mu(x+\hat{\mu}/2)} \rightarrow \mathcal{U}$, $A_\mu(x) \in \mathbb{R} \rightarrow \mathcal{A}_\mu(x) \in \mathbb{C}$

NB. Cauchy's integral theorem ensures this equality

[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 $e^{i\theta} := Je^{-S}/|Je^{-S}|$, $J := \det(\partial\mathcal{U}/\partial U)$
  - ◇ 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



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

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

$t :=$ input, $w, b, \omega :=$ 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

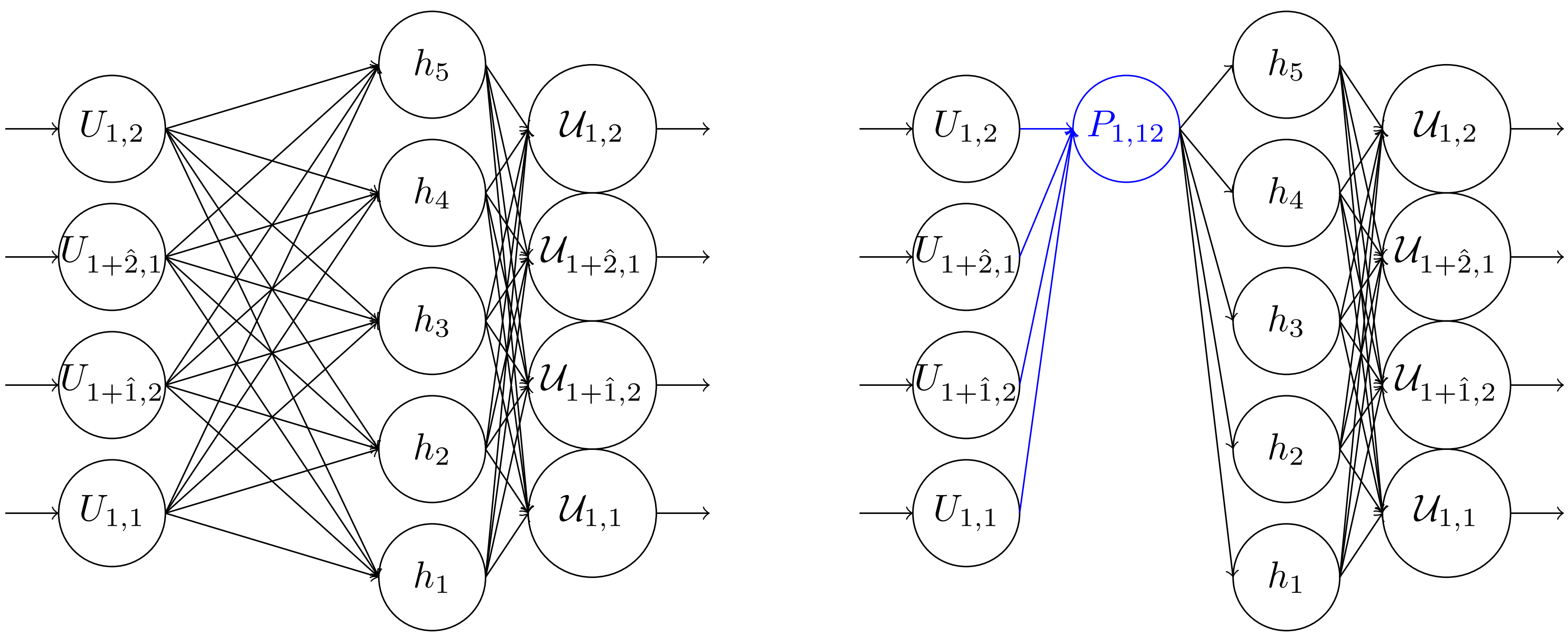
$$\langle \mathcal{O} \rangle_{\text{pq}} := \frac{1}{Z} \int \mathcal{D}U \left[\mathcal{O} \,\middle|\, J\, e^{-S} \right| \Big]_{\mathcal{U} \in \mathcal{C}}$$

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

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

- 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 Favoni et al.(2020)



(old) (new)

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

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

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

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

- The hidden layer is constructed by Stout-like smearing, which is gauge covariant
- We use $N_{\text{stout}} = 2$ 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 $U(1)+\theta$-term, another type of sign problem, is investigated by tensor renormalization Kuramashi and Yoshimura(2019) and complex Langevin Hirasawa et al.(2020)

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

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

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

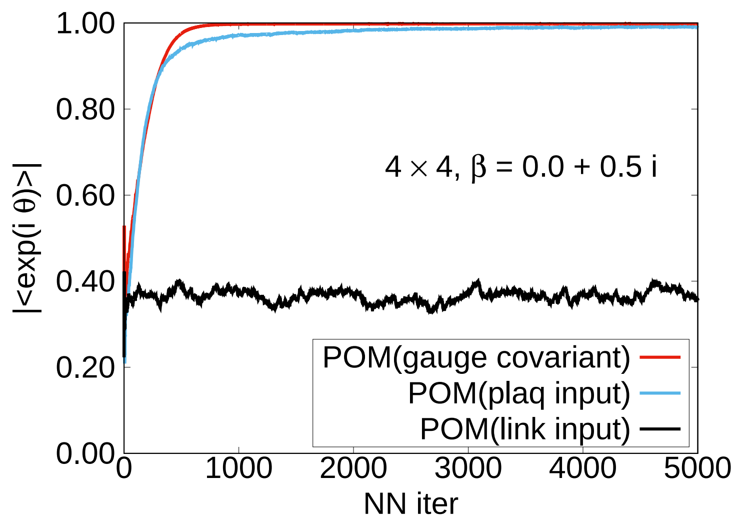
[Analytic result] Wiese(1988),...

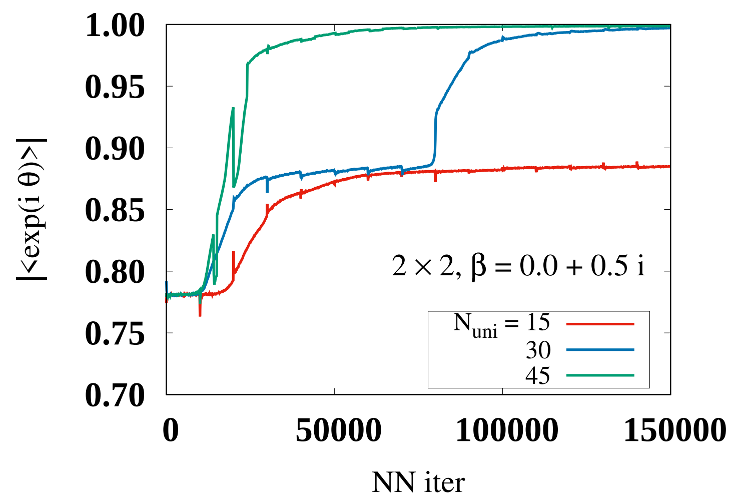
$$Z := \int dU e^{-S} = \sum_{n=-\infty}^{+\infty} I_n(\beta)^V$$

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

[Neural network iteration dependence of average phase factor]

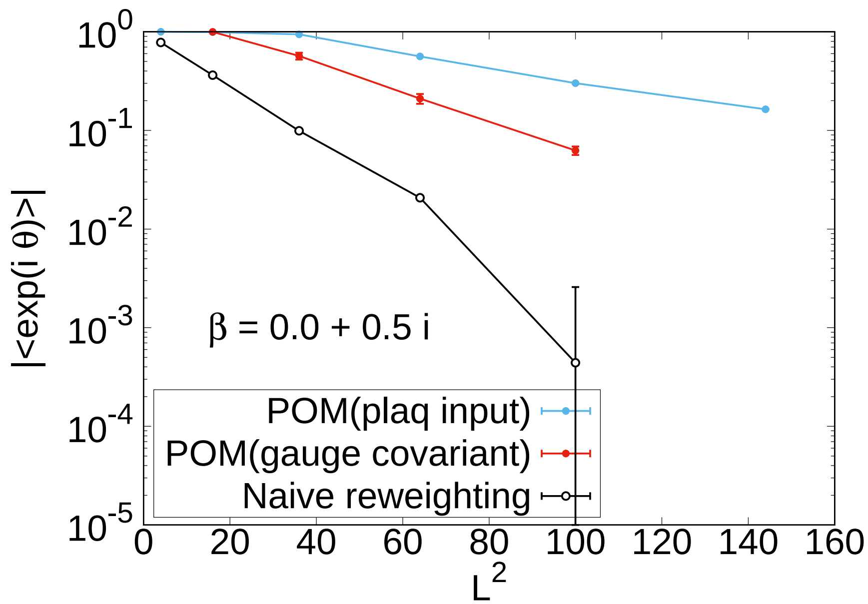
- Neural network with gauge invariant/covariant input successfully enhances averaged phase factor $|\langle \exp(i\theta) \rangle|$
- Naive link-variable input does not enhance the averaged phase factor by 5000 neural network iteration with $N_{\rm uni} = 16$ hidden layer units
  $\rightarrow$ Naive link-variable input with much larger neural network iterations and larger hidden layer units enhances the averaged phase factor





[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



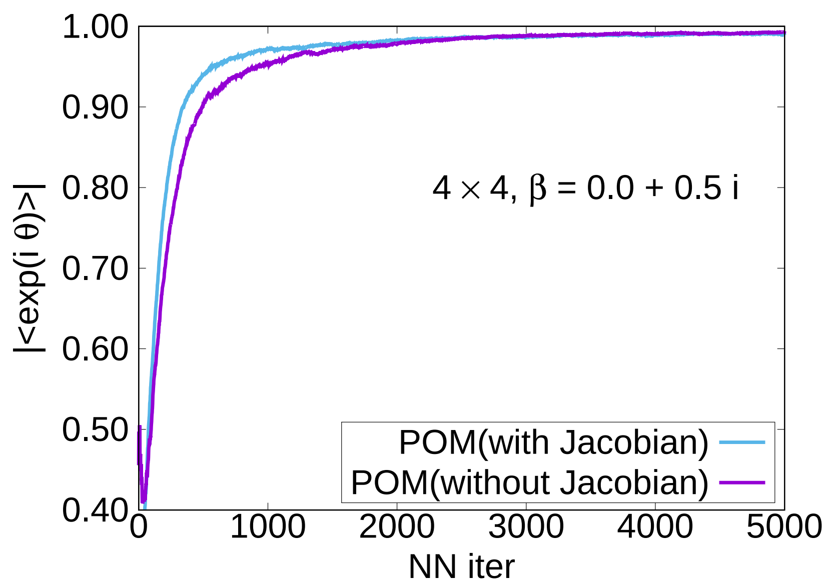
[Test approximated Jacobian in neural network]

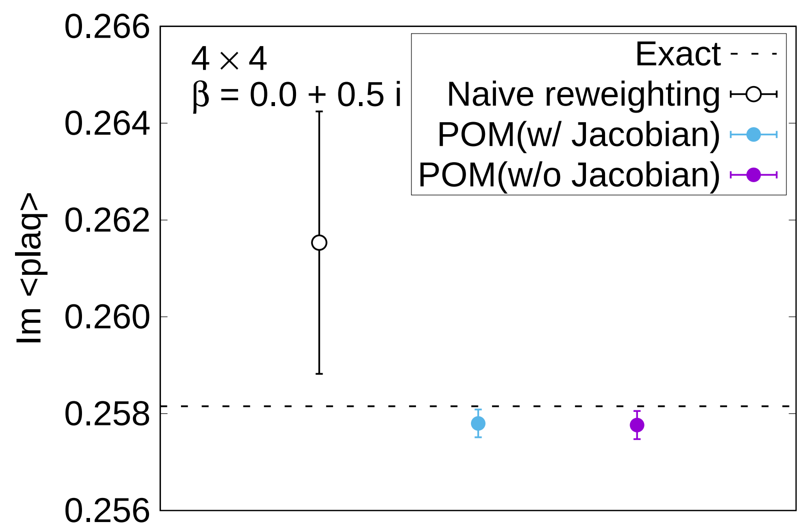
- Calculation of Jacobian is $O(N_{\rm dof}^3)$, which is the main bottleneck
- We test $J = 1$ approximation in the neural network
  $\leftarrow$ We still need the exact Jacobian for final output and measurement

  cf. worldvolume Lefschetz thimble method removes explicit Jacobian in Monte-Carlo update

  Fukuma,Matsumoto(2020); Fukuma's talk

  ♢ 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
  $\leftarrow$ Gauge variant neural network can also enhance the average phase factor with much larger cost
- $J = 1$ approximated neural network still leads to enhancement of the average phase factor at least in our setup

[Future direction]

- Test other types of sign problem, such as finite density QCD and θ-term

Appendix

[Sign problem (overlap problem)]

- Direct Monte Carlo is not possible, because complex part cannot be regarded as probability
- Naive reweighting suffers from sever cancellation between denominator and numerator
  $\rightarrow$ Required #data blows up exponentially as the system size with the degrees of freedom $N_{\rm dof}$ increases

$$\begin{aligned}\langle \mathcal{O} \rangle &= \frac{\left\langle O e^{-i{\rm Im}S} \right\rangle_{\rm pq}}{\left\langle e^{-i{\rm Im}S} \right\rangle_{\rm pq}}, \quad \langle f(z) \rangle_{\rm pq} := (1/Z_R) \int d\mathcal{U} f(z) e^{-{\rm Re}S} \\ &\approx \frac{e^{-O(N_{\rm dof})} \pm O(1/\sqrt{N_{\rm data}})}{e^{-O(N_{\rm dof})} \pm O(1/\sqrt{N_{\rm data}})}\end{aligned}$$

$$\therefore e^{-O(N_{\rm dof})} \gg O(1/\sqrt{N_{\rm data}}) \quad \text{i.e.,} \quad N_{\rm data} \gg e^{O(N_{\rm dof})}$$