Gauge invariant input to neural network for path optimization method arXiv:2109.11710

Yusuke Namekawa
(Kyoto Univ.)

in collaboration with

K.Kashiwa, A.Ohnishi, H.Takase

Contents

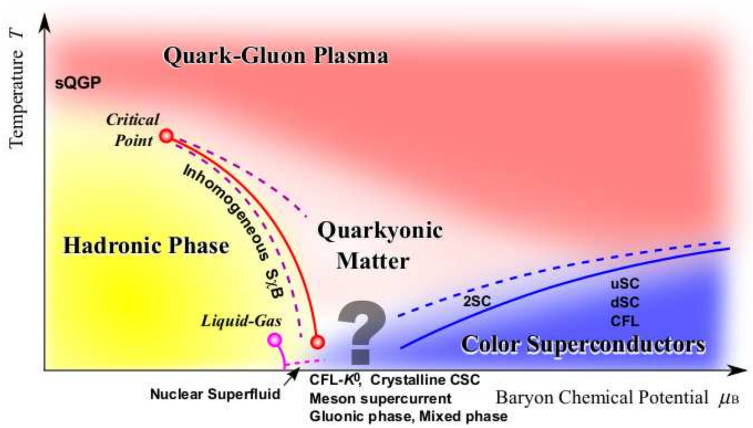
1 Motivation

QCD at high density has been investigated by Complex Langevin Method(CLM)

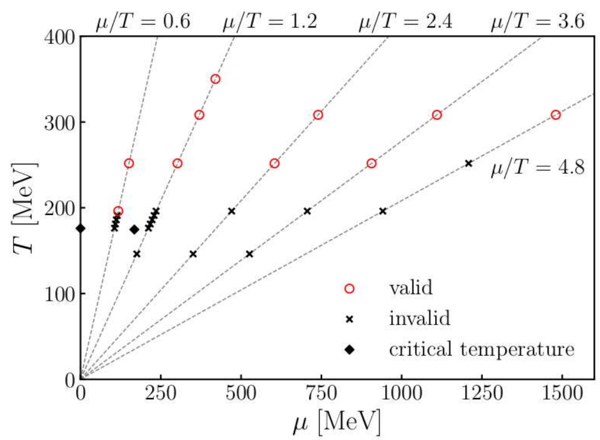
Sexty(2013), ..., Ito et al. including YN(2020), ... ; See Jaeger's talk(3rd day) and Ziegler's poster

- CLM is sign problem free and low cost
- But, CLM can not cover the whole phase diagram of QCD due to validity condition of CLM (two conditions must be satisfied)
  - ◇ Excursion problem must be under control Aarts et al.(2011)
  - ◇ Singular drift must be under control Nishimura and Shimasaki(2015)
    ← (Consistency with the boundary term has been confirmed Scherzer et al.(2019))

→ Alternative approach is needed to cover the whole phase diagram



Fukushima,Hatsuda(2011)



CLM on $24^3 \times 12$

Tsutsui et al. including YN(coming soon)

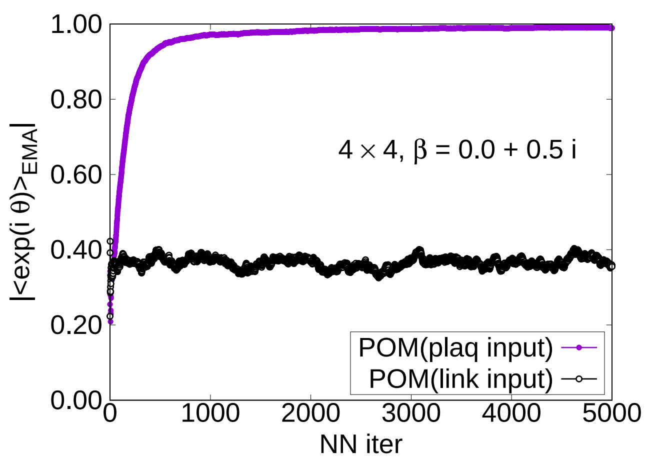
[List of alternative approaches to QCD at high density]

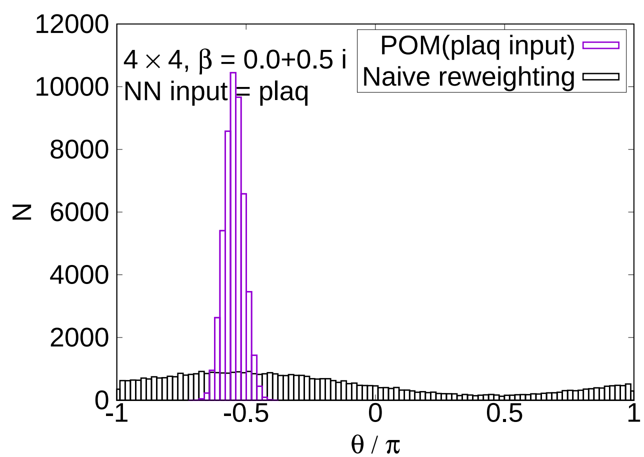
- Tensor renormalization group method
  Levin and Nave(2006), ... ; See Akiyama's talk(1st day)
- Lefschetz thimble method
  Witten(2010), ... , Fukuma et al. including YN(2021), ... ; See Fukuma's talk(2nd day)
- Sign reweighting method de Forcrand et al.(2003), ... ; See Borsányi's talk(5th day)
- Path optimization method(this talk) Mori et al.(2017), ...

2 Short summary

Path optimization method(POM) Mori et al.(2017),... significantly reduces sign problem in $U(1)$ gauge theory with complex coupling, if gauge invariant input is employed for neural network(NN)

- POM is a method which complexifies dynamical variables and deforms the integration path using machine learning to minimize sign problem
  - ◇ Naive link-variable input to neural network does not work
  - ◇ We found gauge invariant input to neural network successfully reduces the sign problem, as indicated by enhancement of the average phase factor(left panel) and the histogram(right panel)





$e^{i\theta} := Je^{-S}/|Je^{-S}|$

$J := \det(\partial \mathcal{U}/\partial U)$

$\mathcal{U}$: complexified link variable

[Path Optimization Method] Mori et al.(2017),...

A Monte Carlo scheme that modifies the integration path to minimize the sign problem by machine learning via neural network

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

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

NB. Cauchy's integral theorem ensures this equality

[Neural network]

- Input(old): link variable $U_{x,\mu} := e^{igA_\mu(x+\hat{\mu}/2)}$
  g : gauge coupling, $A_\mu(x)$: gauge field $\in \mathbb{R} \to \mathcal{A}_\mu(x) \in \mathbb{C}$
- Input(new): gauge invariant plaquette $P_{x,12} := U_{x,1}\,U_{x+\hat{1},2}\,U^{-1}_{x+\hat{2},1}\,U^{-1}_{x,2}$
- Output: $y_x \in \mathbb{R}$, relating to $\mathcal{U}_{x,\mu} = U_{x,\mu}\,e^{-y_x}$
  $\leftarrow$ Machine learning chooses best y_x which enhances phase factor $e^{i\theta} := Je^{-S}/|Je^{-S}|$,
  $J := \det(\partial\mathcal{U}/\partial U)$

$$y_n = \omega_n F(w^{(2)}_{nj}h_j + b_j), \quad i,j = 1,\cdots,2\times n_{\text{dof}}$$

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

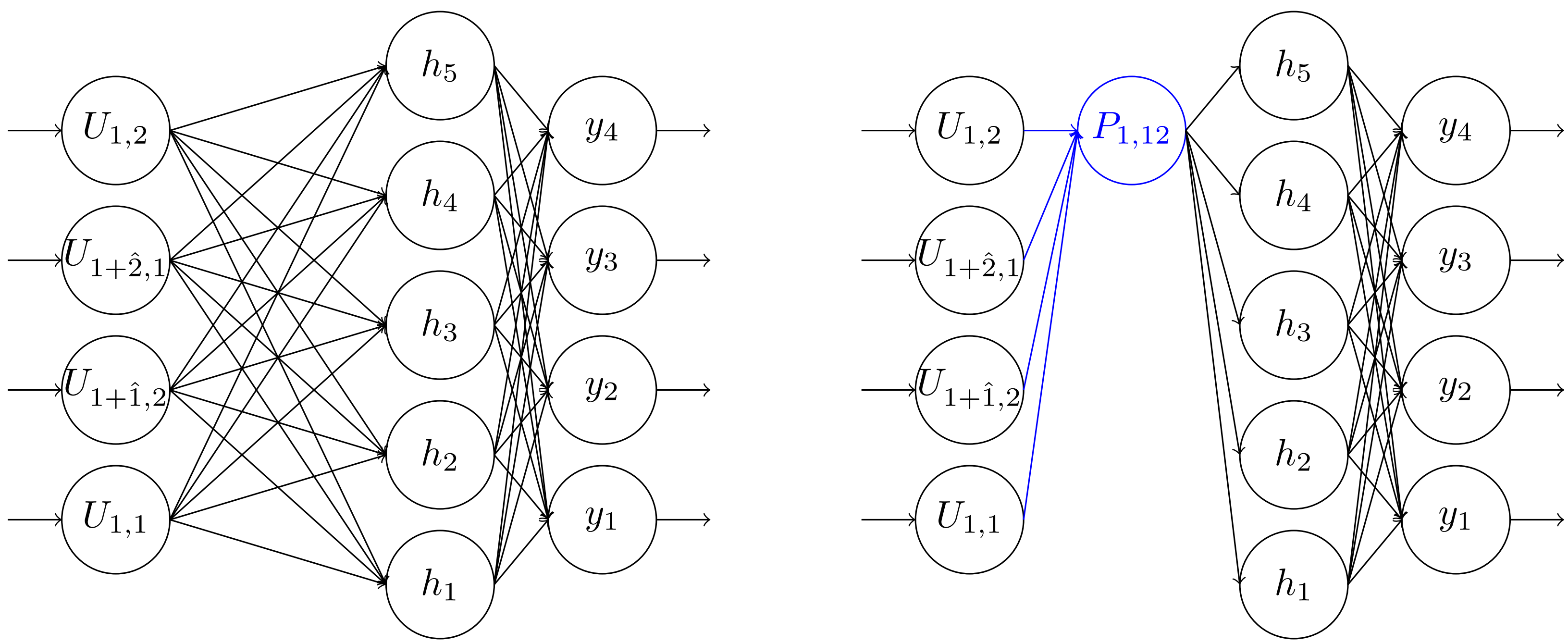
$t :=$ input, $w, b, \omega :=$ parameters of neural network

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

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

[Path Optimization Method(continued)] Mori et al.(2017),...

- Input(old): link variable $U_{x,\mu} := e^{igA_\mu(x+\hat{\mu}/2)}$
  g : gauge coupling, $A_\mu(x)$: gauge field $\in \mathbb{R} \to \mathcal{A}_\mu(x) \in \mathbb{C}$
- Input(new): gauge invariant plaquette $P_{x,12} := U_{x,1}\, U_{x+\hat{1},2}\, U^{-1}_{x+\hat{2},1}\, U^{-1}_{x,2}$
- Output: $y_x \in \mathbb{R}$, related to $\mathcal{U}_{x,\mu} = U_{x,\mu}\, e^{-y_x}$
  $\leftarrow$ Machine learning chooses best y_x which enhances phase factor $e^{i\theta}$



(old) (new)

[Application of POM to 2-dim $U(1)$ gauge theory with complex coupling]

- 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)

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

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

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

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

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

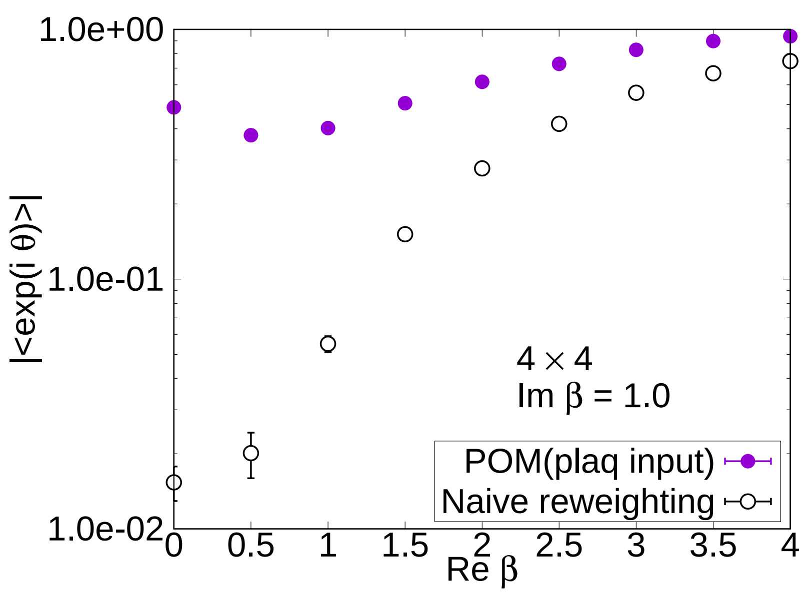
$$I_n(\beta) := \frac{1}{2\pi} \int_{-\pi}^{\pi} d\phi\, e^{\beta \cos\phi - in\phi}$$

[$U(1)$ gauge theory with complex coupling]

- Pure imaginary coupling $\beta = 1/(ga)^2 \in \mathbb{R} \to \mathbb{C}$ mimics the case of real-time action $\exp(iS)$, instead of $\exp(-S_E)$

  ◇ Pure imaginary $\exp(iS)$ leads to the severest sign problem
  $\leftarrow$ We challenge this case by path optimization method(POM)

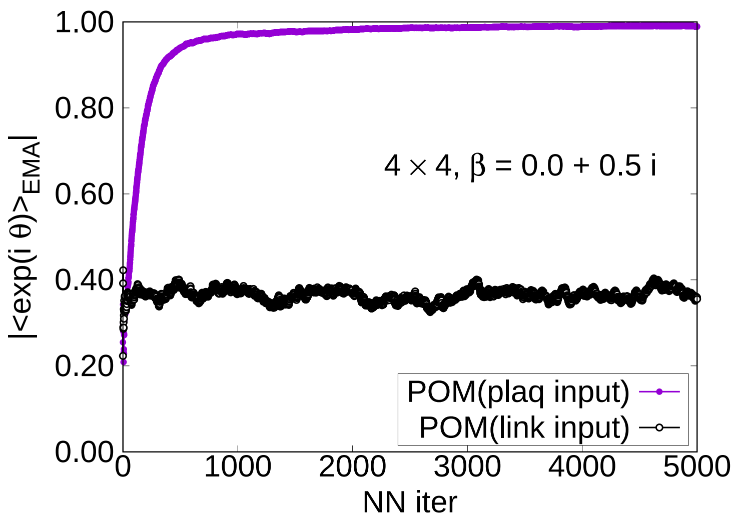
- ($|\left\langle e^{i\theta}\right\rangle|$ is an indicator of sign problem: $|\left\langle e^{i\theta}\right\rangle| = 1$ for mild, $|\left\langle e^{i\theta}\right\rangle| = 0$ for severe)

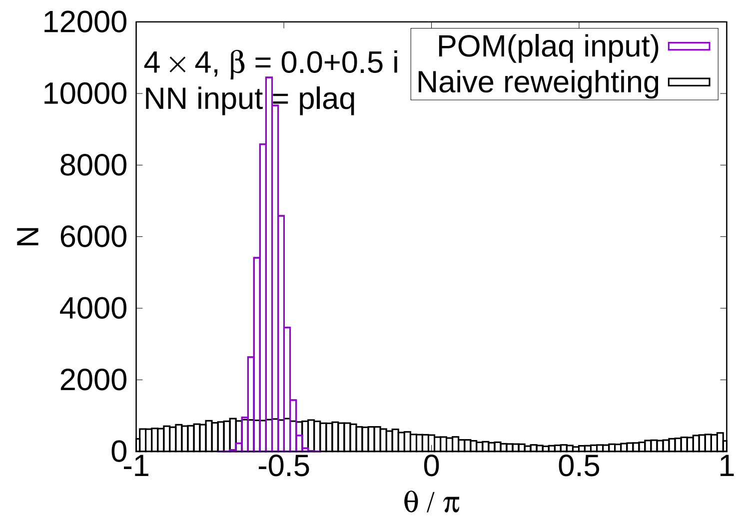


3 Simulation result

[Neural network iteration dependence of average phase factor]

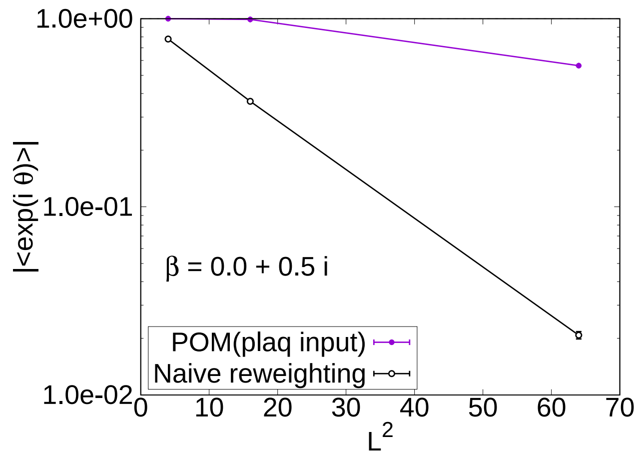
- Neural network with gauge invariant input successfully enhances average phase factor $\langle \exp(i\theta) \rangle$ (an indicator of sign problem), though naive link-variable input fails
  - ◇ Similar enhanced result can be obtained by fixing gauge completely, in expense of additional cost for gauge fixing
- The enhancement is clearly observed in histogram of the phase data

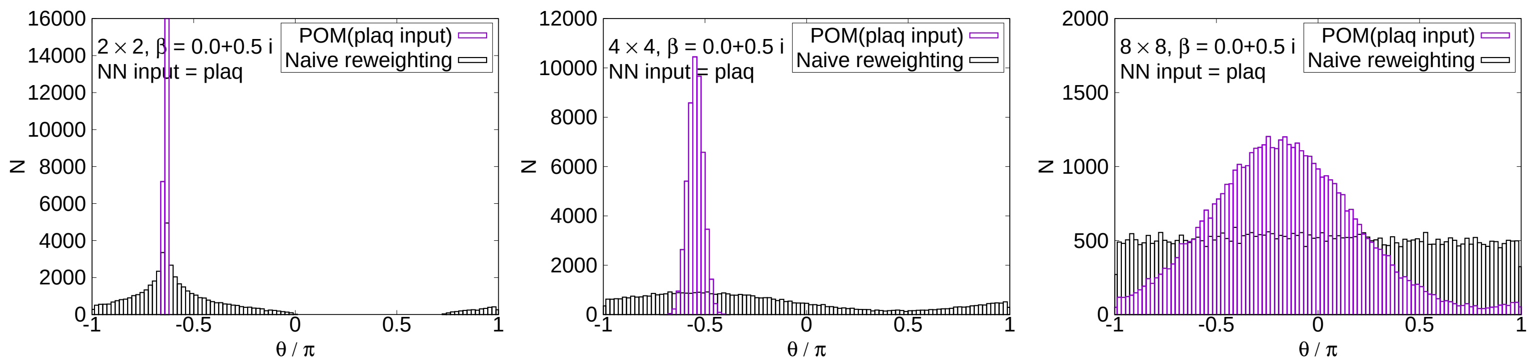




[Volume dependence]

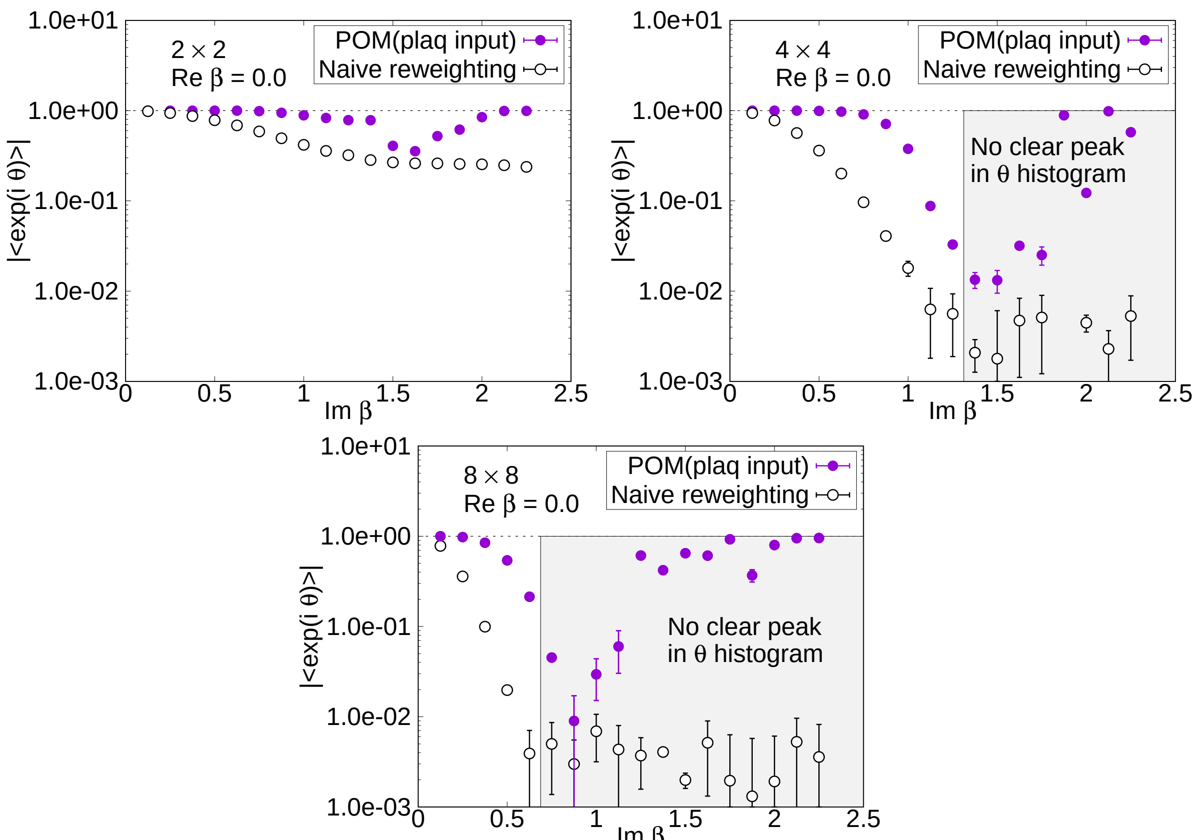
- Enhancement by POM is confirmed in 2×2, 4×4, and 8×8 lattices
  - ◇ The enhancement is maximally 550% in the average phase factor
- Enhancement decreases on larger volume
  - ◇ Volume dependence is milder than that of naive reweighting





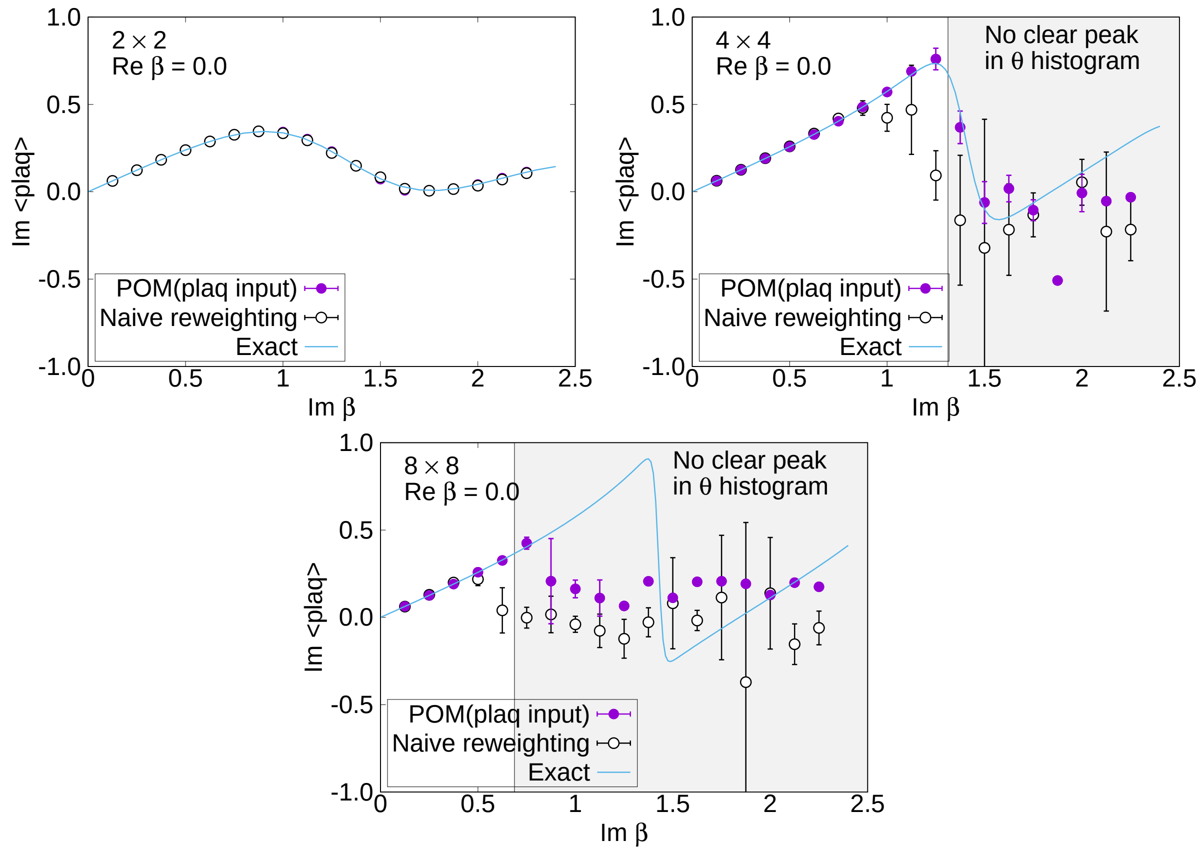
[β-dependence of $|\langle \exp(i\theta)\rangle|$]

- Clear improvement over the naive reweighting method is observed
- Enhancement of the average phase factor becomes less clear, if we have no clear peak in the histogram of the phases
    - ◇ At high β_i, POM becomes unstable probably due to limitation of statistics and/or multimodality effect



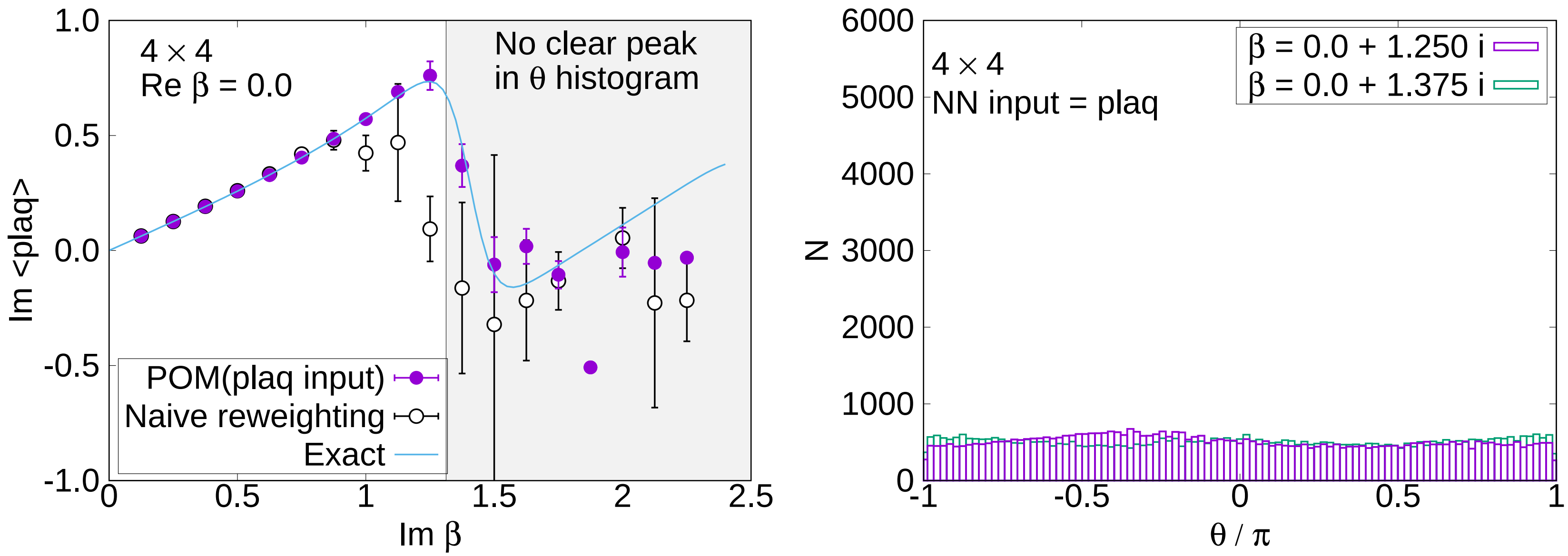
[β-dependence of Im $\langle\text{plaq}\rangle$]

- Results by POM reproduce the exact solution, as long as we find a peak in the histogram of the phases
- Deviations from the exact solution are also observed, if we have no clear peak in the histogram of the phases
  $\rightarrow$ Further improvement is required especially on larger volumes



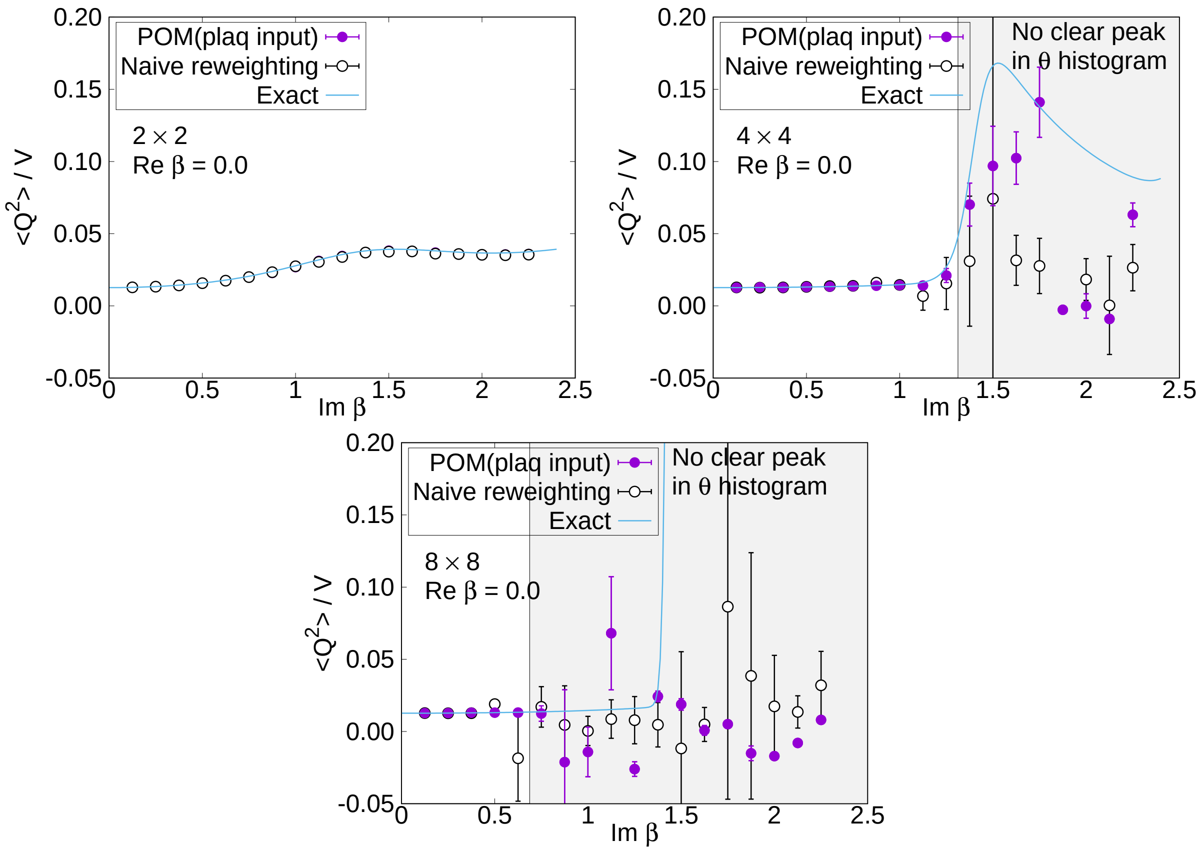
[Histogram of phase around β_c]

- Results by POM reproduce the exact solution, as long as we find a peak in the histogram of the phases
- Deviations from the exact solution are also observed, if we have no clear peak in the histogram of the phases
  $\rightarrow$ Further improvement is required especially on larger volumes



[β-dependence of topological charge]

- Results by POM reproduce the exact solution, as long as we find a peak in the histogram of the phases
- Deviations from the exact solution are also observed, if we have no clear peak in the histogram of the phases
  $\rightarrow$ Further improvement is required especially on larger volumes



4 Summary

Path optimization method significantly reduces sign problem in $U(1)$ gauge theory with complex coupling, chosen to be pure imaginary(the severest sign problem region)

- Gauge invariant plaq input successfully reduces sign problem, leading to maximally 550% enhancement of the average phase factor
- For large β_i on large volume, machine learning fails to find a peak in the histogram of phase data $\rightarrow$ Further improvement is needed

[Future direction]

- Use larger Wilson and Polyakov loops as inputs to neural network
- Test other gauge-symmetry respecting neural networks Favoni et al.(2020), Tomiya and Nagai(2021), combination of path optimization with action optimization Tsutsui and Doi(2015,2017),Lawrence(2020)
    - ♢ Try larger volumes(severer sign problem), finite density and θ-term (another type of sign problem) with improved POM

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
  → Required #data blows up exponentially as the system size with the degrees of freedom N_{dof} increases

$$\langle \mathcal{O} \rangle = \frac{\left\langle O e^{-i\mathrm{Im}S} \right\rangle_{\mathrm{pq}}}{\left\langle e^{-i\mathrm{Im}S} \right\rangle_{\mathrm{pq}}}, \quad \langle f(z) \rangle_{\mathrm{pq}} := (1/Z_R) \int d\mathcal{U} f(z) e^{-\mathrm{Re}S}$$

$$\approx \frac{e^{-O(N_{\mathrm{dof}})} \pm O(1/\sqrt{N_{\mathrm{data}}})}{e^{-O(N_{\mathrm{dof}})} \pm O(1/\sqrt{N_{\mathrm{data}}})}$$

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