Sign problem and the tempered Lefschetz thimble method

Masafumi Fukuma (Dept Phys, Kyoto Univ)

Dec 4, 2019
KEK Theory workshop 2019

Based on work with

Nobuyuki Matsumoto (Kyoto Univ) & **Naoya Umeda** (PwC)

-- **MF** and **Umeda**, "Parallel tempering algorithm for integration over Lefschetz thimbles" [arXiv:1703.00861, PTEP2017(2017)073B01]
-- **MF**, **Matsumoto** and **Umeda**, "Applying the tempered Lefschetz thimble method to the Hubbard model away from half-filling" [arXiv:1906.04243, to appear in PRD]

Also, for the geometrical optimization of tempering algorithms and its application to QG:

-- **MF**, **Matsumoto** and **Umeda**
[arXiv:1705.06097, JHEP1712(2017)001], [arXiv:1806.10915, JHEP1811(2018)060]

and for the geometry of tempered stochastic matrix models (= AdS BH) :

-- **MF** and **Matsumoto** [arXiv:1912.*****] ⇨ **Matsumoto's poster**

1. Introduction

Overview

The **numerical sign problem** is one of the major obstacles
when performing numerical calculations in various fields of physics

Typical examples:

① Finite density QCD
② Quantum Monte Carlo simulations of quantum statistical systems
③ Real time QM/QFT

Today, I would like to

-- explain what the sign problem is

-- argue that **[MF-Umeda 1703.00861, MF-Matsumoto-Umeda 2019]**
a new algorithm “Tempered Lefschetz thimble method” (TLTM)
is a promising method towards solving the sign problem,
by exemplifying its effectiveness for:

② Quantum Monte Carlo simulations
of strongly correlated electron systems,
especially the Hubbard model away from half-filling

Sign problem

Our main concern is to estimate: $\langle \mathcal{O}(x) \rangle_S \equiv \dfrac{\int dx\, e^{-S(x)} \mathcal{O}(x)}{\int dx\, e^{-S(x)}}$

$\begin{cases} x = (x^i) \in \mathbb{R}^N\text{: dynamical variable (real-valued)} \\ S(x)\text{: action, } \mathcal{O}(x)\text{: observable} \end{cases}$

Markov chain Monte Carlo (MCMC) simulation:

probability distribution function

When $S(x) \in \mathbb{R}$, one can regard $p_{\text{eq}}(x) \equiv e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF:

$$0 \le p_{\text{eq}}(x) \le 1, \quad \int dx\, p_{\text{eq}}(x) = 1$$

⇨ Generate a sample $\{x^{(k)}\}_{k=1,\ldots,N_{\text{conf}}}$ from $p_{\text{eq}}(x)$

⇨ $\langle \mathcal{O}(x) \rangle \approx \dfrac{1}{N_{\text{conf}}} \sum_{k=1}^{N_{\text{conf}}} \mathcal{O}(x^{(k)})$

Sign problem:

When $S(x) = S_R(x) + i\, S_I(x) \in \mathbb{C}$, one cannot regard $e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF

⇨ Reweighting method :

Sign problem

Our main concern is to estimate: $\left\langle \mathcal{O}(x) \right\rangle_S \equiv \dfrac{\int dx\, e^{-S(x)} \mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \dfrac{\int dx\, e^{-S_R(x)} e^{-i S_I(x)} \mathcal{O}(x)}{\int dx\, e^{-S_R(x)} e^{-i S_I(x)}}$

$$\begin{cases} x = (x^i) \in \mathbb{R}^N\text{: dynamical variable (real-valued)} \\ S(x)\text{: action, } \mathcal{O}(x)\text{: observable} \end{cases}$$

Markov chain Monte Carlo (MCMC) simulation:

probability distribution function

When $S(x) \in \mathbb{R}$, one can regard $p_{\text{eq}}(x) \equiv e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF:

$$0 \le p_{\text{eq}}(x) \le 1, \quad \int dx\, p_{\text{eq}}(x) = 1$$

⇨ Generate a sample $\{x^{(k)}\}_{k=1,\ldots,N_{\text{conf}}}$ from $p_{\text{eq}}(x)$

⇨ $\left\langle \mathcal{O}(x) \right\rangle \approx \dfrac{1}{N_{\text{conf}}} \displaystyle\sum_{k=1}^{N_{\text{conf}}} \mathcal{O}(x^{(k)})$

Sign problem:

When $S(x) = S_R(x) + i\, S_I(x) \in \mathbb{C}$, one cannot regard $e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF

⇨ Reweighting method :

Sign problem

Our main concern is to calculate: $\langle \mathcal{O}(x) \rangle_S \equiv \dfrac{\int dx\, e^{-S(x)} \mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \dfrac{\int dx\, e^{-S_R(x)} e^{-i S_I(x)} \mathcal{O}(x)}{\int dx\, e^{-S_R(x)} e^{-i S_I(x)}}$

$\begin{cases} x = (x^i) \in \mathbb{R}^N \text{: dynamical variable (real-valued)} \\ S(x) \text{: action, } \mathcal{O}(x) \text{: observable} \end{cases}$

Markov chain Monte Carlo (MCMC) simulation:

probability distribution function

When $S(x) \in \mathbb{R}$, one can regard $p_{\mathrm{eq}}(x) \equiv e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF:

$$0 \le p_{\mathrm{eq}}(x) \le 1, \quad \int dx\, p_{\mathrm{eq}}(x) = 1$$

⇒ Generate a sample $\{x^{(k)}\}_{k=1,\ldots,N_{\mathrm{conf}}}$ from $p_{\mathrm{eq}}(x)$

⇒ $\langle \mathcal{O}(x) \rangle \approx \dfrac{1}{N_{\mathrm{conf}}} \displaystyle\sum_{k=1}^{N_{\mathrm{conf}}} \mathcal{O}(x^{(k)})$

Sign problem:

When $S(x) = S_R(x) + i S_I(x) \in \mathbb{C}$, one cannot regard $e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF

⇒ Reweighting method : treat $e^{-S_R(x)} / \int dx\, e^{-S_R(x)}$ as a PDF

Sign problem

Our main concern is to calculate: $\langle \mathcal{O}(x) \rangle_S \equiv \dfrac{\int dx\, e^{-S(x)} \mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \dfrac{\int dx\, e^{-S_R(x)} e^{-i S_I(x)} \mathcal{O}(x)}{\int dx\, e^{-S_R(x)} e^{-i S_I(x)}}$

$\begin{cases} x = (x^i) \in \mathbb{R}^N \text{: dynamical variable (real-valued)} \\ S(x)\text{: action, } \mathcal{O}(x)\text{: observable} \end{cases}$

Markov chain Monte Carlo (MCMC) simulation:

probability distribution function

When $S(x) \in \mathbb{R}$, one can regard $p_{\mathrm{eq}}(x) \equiv e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF:

$$0 \le p_{\mathrm{eq}}(x) \le 1, \quad \int dx\, p_{\mathrm{eq}}(x) = 1$$

⇨ Generate a sample $\{x^{(k)}\}_{k=1,\dots,N_{\mathrm{conf}}}$ from $p_{\mathrm{eq}}(x)$

⇨ $\langle \mathcal{O}(x) \rangle \approx \dfrac{1}{N_{\mathrm{conf}}} \displaystyle\sum_{k=1}^{N_{\mathrm{conf}}} \mathcal{O}(x^{(k)})$

Sign problem:

When $S(x) = S_R(x) + i S_I(x) \in \mathbb{C}$, one cannot regard $e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF

⇨ Reweighting method : treat $e^{-S_R(x)} / \int dx\, e^{-S_R(x)}$ as a PDF

⇨ $\langle \mathcal{O}(x) \rangle_S \equiv \dfrac{\langle e^{-i S_I(x)} \mathcal{O}(x) \rangle_{S_R}}{\langle e^{-i S_I(x)} \rangle_{S_R}} = \dfrac{e^{-O(N)}}{e^{-O(N)}} = O(1) \quad (N\text{ : DOF})$

Sign problem

Our main concern is to calculate: $\langle \mathcal{O}(x) \rangle_S \equiv \dfrac{\int dx\, e^{-S(x)} \mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \dfrac{\int dx\, e^{-S_R(x)} e^{-i S_I(x)} \mathcal{O}(x)}{\int dx\, e^{-S_R(x)} e^{-i S_I(x)}}$

$\begin{cases} x = (x^i) \in \mathbb{R}^N\text{: dynamical variable (real-valued)} \\ S(x)\text{: action, } \mathcal{O}(x)\text{: observable} \end{cases}$

Markov chain Monte Carlo (MCMC) simulation:

probability distribution function

When $S(x) \in \mathbb{R}$, one can regard $p_{\text{eq}}(x) \equiv e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF:

$$0 \le p_{\text{eq}}(x) \le 1, \quad \int dx\, p_{\text{eq}}(x) = 1$$

⇒ Generate a sample $\{x^{(k)}\}_{k=1,\dots,N_{\text{conf}}}$ from $p_{\text{eq}}(x)$

⇒ $\langle \mathcal{O}(x) \rangle \approx \dfrac{1}{N_{\text{conf}}} \sum_{k=1}^{N_{\text{conf}}} \mathcal{O}(x^{(k)})$

Sign problem:

When $S(x) = S_R(x) + i\, S_I(x) \in \mathbb{C}$, one cannot regard $e^{-S(x)} / \int dx\, e^{-S(x)}$ as a PDF

⇒ Reweighting method : treat $e^{-S_R(x)} / \int dx\, e^{-S_R(x)}$ as a PDF

⇒ $\langle \mathcal{O}(x) \rangle_S \equiv \dfrac{\langle e^{-i S_I(x)} \mathcal{O}(x) \rangle_{S_R}}{\langle e^{-i S_I(x)} \rangle_{S_R}} \approx \dfrac{e^{-O(N)} \pm O(1/\sqrt{N_{\text{conf}}})}{e^{-O(N)} \pm O(1/\sqrt{N_{\text{conf}}})}$ (N : DOF, N_{conf} : sample size)

⇒ Require $O(1/\sqrt{N_{\text{conf}}}) < e^{-O(N)}$ ⇒ $N_{\text{conf}} \simeq e^{O(N)}$ sign problem!

Example: Gaussian

Let us consider
$$\begin{cases} S(x) = \frac{\beta}{2}(x-i)^2 \equiv S_R(x) + i\,S_I(x) \\ \mathcal{O}(x) = x^2 \qquad \boxed{\beta \gg 1} \end{cases} \qquad \begin{pmatrix} S_R(x) = \frac{\beta}{2}(x^2-1) \\ S_I(x) = -\beta x \end{pmatrix}$$

large β mimics large DOF ($\beta \sim N$)

$$\Rightarrow \quad \langle x^2 \rangle_S = \frac{\left\langle e^{-i S_I(x)} x^2 \right\rangle_{S_R}}{\left\langle e^{-i S_I(x)} \right\rangle_{S_R}} = \frac{\left(\beta^{-1} - 1\right) e^{-\beta/2}}{e^{-\beta/2}}$$

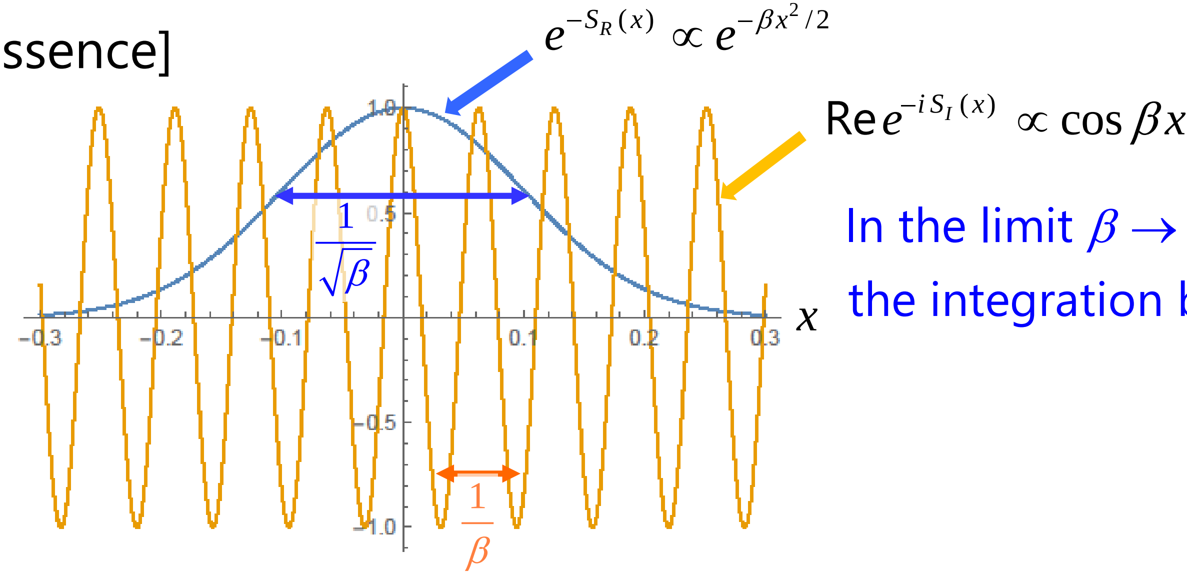
numerically $$\approx \frac{\left(\beta^{-1}-1\right)e^{-\beta/2} \pm O(1/\sqrt{N_{\text{conf}}})}{e^{-\beta/2} \pm O(1/\sqrt{N_{\text{conf}}})}$$

(NB : The num and the denom are estimated separately.)

⇒ Necessary sample size:

$$1/\sqrt{N_{\text{conf}}} \lesssim O(e^{-\beta/2}) \quad \Leftrightarrow \quad \boxed{N_{\text{conf}} \gtrsim O(e^{\beta})}$$

[Essence]



In the limit $\beta \to \infty$ $\left(\because 1/\beta \ll 1/\sqrt{\beta}\right)$, the integration becomes highly oscillatory

Approaches to the sign problem

Various approaches:

(1) Complex Langevin method (CLM) **[Parisi 1983]**

(2) (Generalized) Lefschetz thimble method ((G)LTM) **[Cristoforetti et al. 2012, ...]** **[Alexandru et al. 2015, ...]**

(3) ...

Advantages/disadvantages:

(1) CLM

Pros: fast $\propto O(N)$ (N : DOF)

Cons: "wrong convergence problem" **[Ambjørn-Yang 1985, Aarts et al. 2011, Nagata-Nishimura-Shimasaki 2016]**

(2) LTM

Pros: No wrong convergence problem *iff* only a single thimble is relevant

Cons: Expensive $\propto O(N^3)$ ⇐ Jacobian determinant

Ergodicity problem if more than one thimble are relevant (wrong convergence de facto)

⇓

(2′) **TLTM** (Tempered Lefschetz thimble method) **[MF-Umeda 1703.00861, MF-Matsumoto-Umeda 1906.04243]**

"facilitate transitions among thimbles by tempering the system with the flow time"

Pros: Works well even when multi thimbles are relevant

Cons: Expensive $\propto O(N^{3\sim 4})$ ⇐ Jacobian determinant + tempering

Plan

1. Introduction (done)
2. (Generalized) LTM (GLTM)
3. Tempered LTM (TLTM)
4. Applying the TLTM to the Hubbard model
   - 1D case
   - 2D case
5. Conclusion and outlook

2. (Generalized) Lefschetz thimble method (GLTM)

[Cristoforetti et al. 1205.3996, 1303.7204, 1308.0233]
[Fujii-Honda-Kato-Kikukawa-Komatsu-Sano 1309.4371]
[Alexandru et al. 1512.08764]

Lefschetz thimble method (1/2)

[cf. Prof. Dunne's talk]

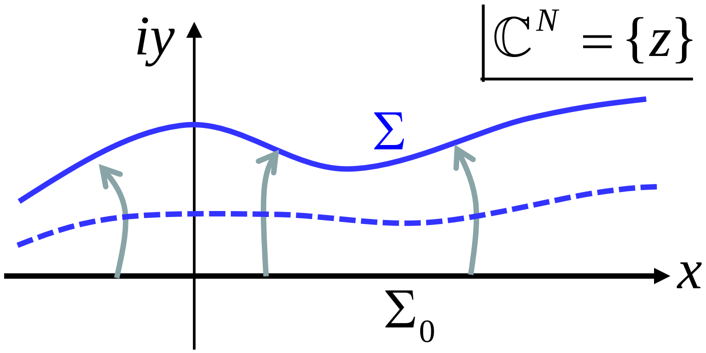
Complexify the variable: $x=(x^i)\in\mathbb{R}^N \;\Rightarrow\; z=(z^i=x^i+iy^i)\in\mathbb{C}^N$

<u>Assumption</u>: $e^{-S(z)}$, $e^{-S(z)}\mathcal{O}(z)$: entire functions over $\mathbb{C}^N$

⇩ Cauchy's theorem

Integral does not change under continuous deformations of the integration region from $\Sigma_0=\mathbb{R}^N$ to $\Sigma\subset\mathbb{C}^N$ (with the boundary at infinity $|x|\to\infty$ kept fixed) :

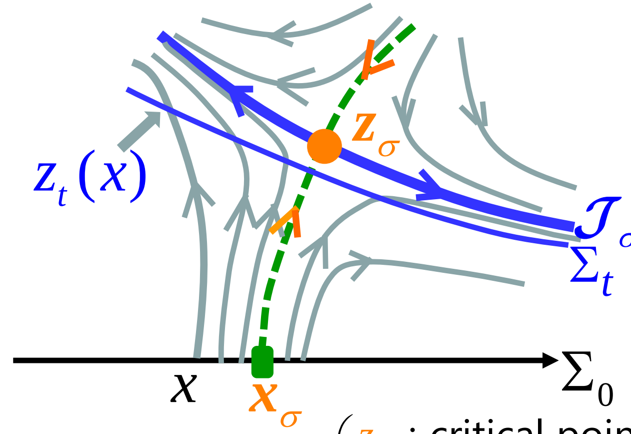
$$\langle\mathcal{O}(x)\rangle_S \equiv \frac{\int_{\Sigma_0} dx\, e^{-S(x)}\,\mathcal{O}(x)}{\int_{\Sigma_0} dx\, e^{-S(x)}} = \frac{\int_{\Sigma} dz\, e^{-S(z)}\,\mathcal{O}(z)}{\int_{\Sigma} dz\, e^{-S(z)}}$$



severe sign problem (⇧ left-hand side)

sign problem will get much reduced if $\mathrm{Im}\,S(z)$ is almost constant on Σ (⇧ right-hand side)

Lefschetz thimble method (2/2)



Prescription:

antiholomorphic gradient flow $\quad \dot{z}_t^i = \overline{\partial_i S(z_t)}$ with $z_{t=0}^i = x^i$

Property: $[S(z_t)]^\cdot = \partial_i S(z_t)\,\dot{z}_t^i = |\partial_i S(z_t)|^2 \geq 0$

(z_σ : critical point $(\partial_i S(z_\sigma) = 0)$)

⇒ $[\mathrm{Re}\,S(z_t)]^\cdot \geq 0$: real part always increases along the flow
$[\mathrm{Im}\,S(z_t)]^\cdot = 0$: imaginary part is kept fixed ″

⇒ In $t \to \infty$, Σ_t approaches a union of Lefschetz thimbles: $\Sigma_t \to \bigcup_\sigma \mathcal{J}_\sigma$

(on each of which $\mathrm{Im}\,S(z)$ is constant)

Expectation value:

$$\langle \mathcal{O}(x)\rangle_S \equiv \frac{\int_{\Sigma_0} dx\, e^{-S(x)}\mathcal{O}(x)}{\int_{\Sigma_0} dx\, e^{-S(x)}} = \frac{\int_{\Sigma_t} dz_t\, e^{-S(z_t)}\mathcal{O}(z_t)}{\int_{\Sigma_t} dz_t\, e^{-S(z_t)}} = \frac{\int_{\Sigma_0} dx\, \det(\partial z_t^i(x)/\partial x^j)\, e^{-S(z_t(x))}\,\mathcal{O}(z_t(x))}{\int_{\Sigma_0} dx\, \det(\partial z_t^i(x)/\partial x^j)\, e^{-S(z_t(x))}}$$

$$\det(\partial z_t^i(x)/\partial x^j)\, e^{-S(z_t(x))} \equiv e^{-S_t^{\mathrm{eff}}(x)}\, e^{i\theta_t(x)}$$

$$= \frac{\left\langle e^{i\theta_t(x)}\mathcal{O}(z_t(x))\right\rangle_{S_t^{\mathrm{eff}}}}{\left\langle e^{i\theta_t(x)}\right\rangle_{S_t^{\mathrm{eff}}}}$$

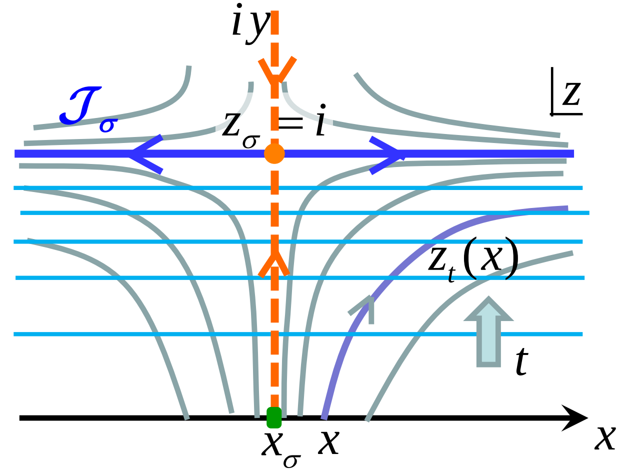
$$e^{-S_t^{\mathrm{eff}}(x)} \equiv e^{-\mathrm{Re}\,S(z_t(x))}\left|\det(\partial z_t^i(x)/\partial x^j)\right|$$

$$e^{i\theta_t(x)} \equiv e^{-i\,\mathrm{Im}\,S(z_t(x)) + i\arg\det(\partial z_t^i(x)/\partial x^j)}$$

Example: Gaussian

Gradient flow: $\left[S(z) = (\beta/2)(z-i)^2 \right]$

$$\dot z_t = \dot x_t + i\,\dot y_t = \overline{S'(z_t)} \Leftrightarrow \begin{cases} \dot x_t = \beta x \\ \dot y = -\beta(y_t - 1) \end{cases} \text{ with } \begin{cases} x_{t=0} = x \\ y_{t=0} = 0 \end{cases}$$



$$\Rightarrow \begin{cases} z_t(x) = x e^{\beta t} + i(1 - e^{-\beta t}) \\ J_t(x) = \dfrac{dz_t(x)}{dx} = e^{\beta t} \end{cases}$$

exponential growth of coefficient

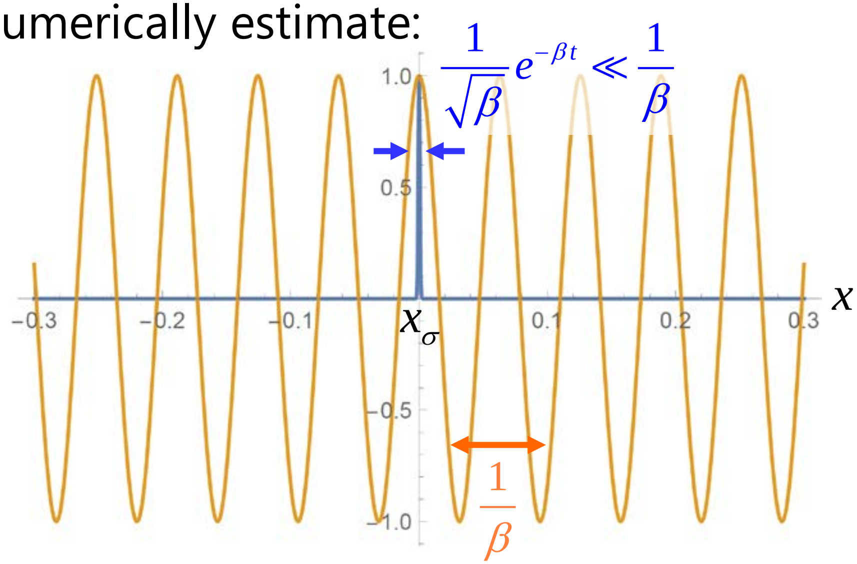
$$\Rightarrow \begin{cases} S_t^{\text{eff}}(x) = \dfrac{1}{2}\beta e^{2\beta t}(x^2 - e^{-4\beta t}) - \beta t \\ \theta_t(x) = \beta x \end{cases}$$

$$\left(J_t(x)\, e^{-S(z_t(x))} = e^{-S_t^{\text{eff}}(x)}\, e^{i\theta_t(x)} \right)$$

Taking a large T s.t. $e^{-\beta T} \ll \dfrac{1}{\sqrt{\beta}}$, we can numerically estimate:

$$\left\langle x^2 \right\rangle_S = \frac{\left\langle e^{i\theta_T(x)} z_T^2(x) \right\rangle_{S_T^{\text{eff}}}}{\left\langle e^{i\theta_T(x)} \right\rangle_{S_T^{\text{eff}}}} = \frac{e^{-(\beta/2)e^{-2\beta T}}\left(\beta^{-1} - 1\right)}{e^{-(\beta/2)e^{-2\beta T}}} = \frac{O(1)}{O(1)}$$

(no small numbers appear!)



NB. Logarithmic increase is enough:

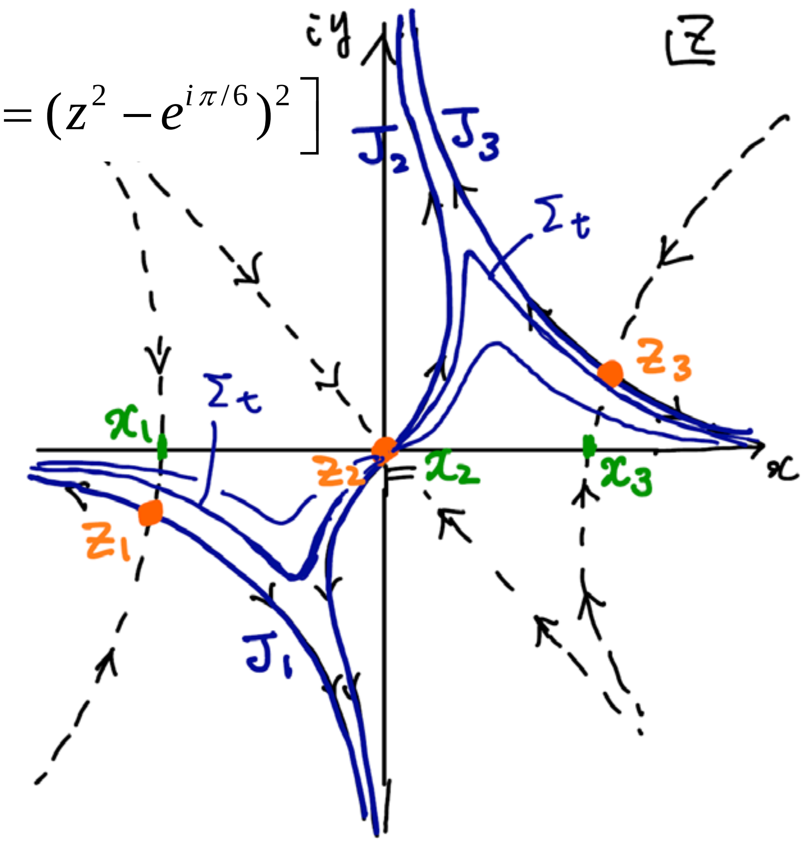
$T \sim O(\log\beta)\ \left(= O(\log N)\right)$

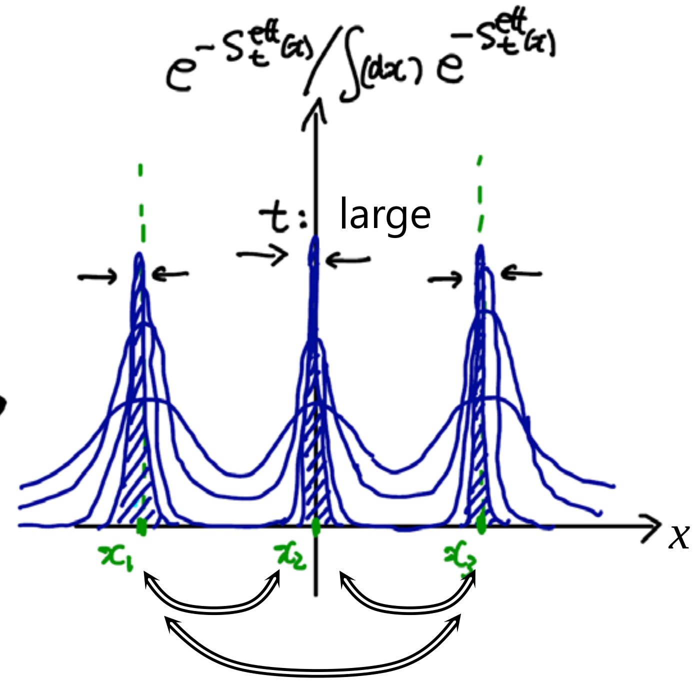
Multimodal problem and Generalized LTM (1/2)

Flow time t needs to be large enough to solve the sign problem

However, this introduces a new problem "ergodicity (multimodal) problem"

$$\left[\, S(z) = (z^2 - e^{i\pi/6})^2 \,\right]$$





transitions among thimbles become indefinitely difficult as t increases

Dilemma between the sign problem and the ergodicity problem

(for small t) (for large t)

Multimodal problem and Generalized LTM (2/2)

Proposal in Generalized LTM: [Alexandru-Basar-Bedaque-Ridgway-Warrington 1512.08764]

Choose an intermediate value of T s.t. it is large enough for the sign problem but at the same time is not too large for the ergodicity (multimodal) problem

flow time ($= T$)	small	medium	large
sign problem	NG	△	**OK**
ergodicity problem	**OK**	△	NG

However, the existence of such T is not obvious a priori

Even when it exists, a very fine tuning will be needed

⇩

Tempered LTM: [MF-Umeda 1703.00861] (cf. [Alexandru-Basar-Bedaque-Warrington 1703.02414])

Implement a tempering method by using the flow time *t* as a dynamical variable

flow time ($= T$)	small	medium	large
sign problem	NG	**OK**	**OK**
ergodicity problem	**OK**	**OK**	**OK**

no fine tuning needed!

3. Tempered Lefschetz thimble method (TLTM)

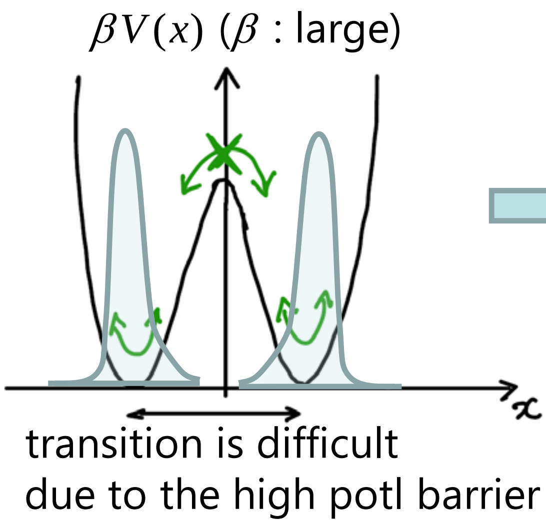
[MF-Umeda 1703.00861]
[MF-Matsumoto-Umeda 1906.04243]

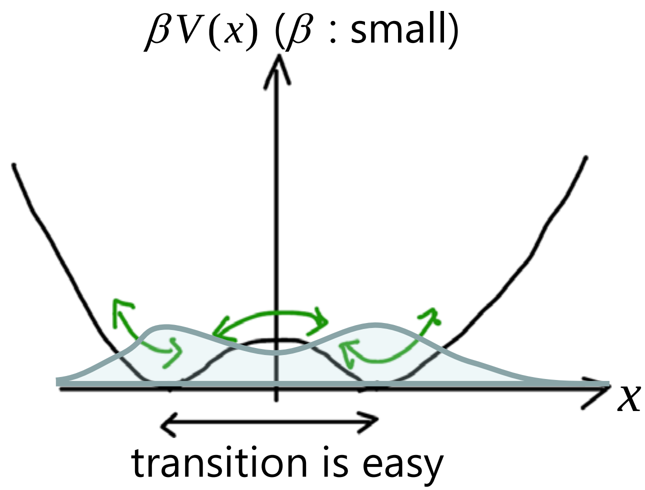
Idea of tempering

[Marinari-Parisi Europhys.Lett.19(1992)451]

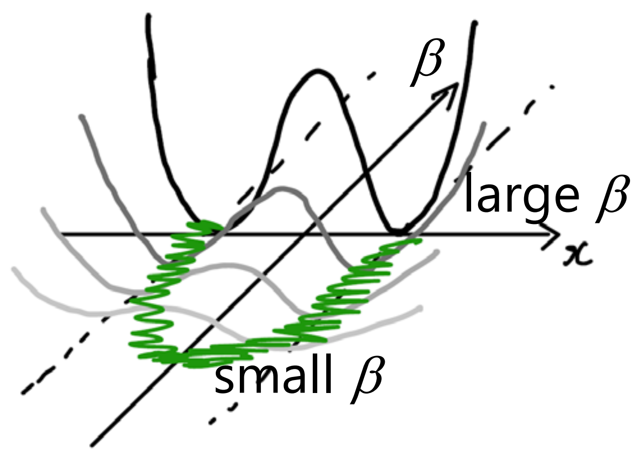
Suppose that the action $S(x;\beta)$ gives a multimodal distribution for the value of β in our main concern (e.g. $S(x;\beta)=\beta V(x)$ with $\beta \gg 1$)

It often happens that multimodality disappears if we take a different value of β (e.g. for $\beta \ll 1$)





In the tempering method,
we extend the config space from $\{x\}$ to $\{(x,\beta)\}$.
Then, transitions between two modes become easy by passing through configs with smaller β



Tempered LTM (1/3)

[MF-Umeda 1703.00861]

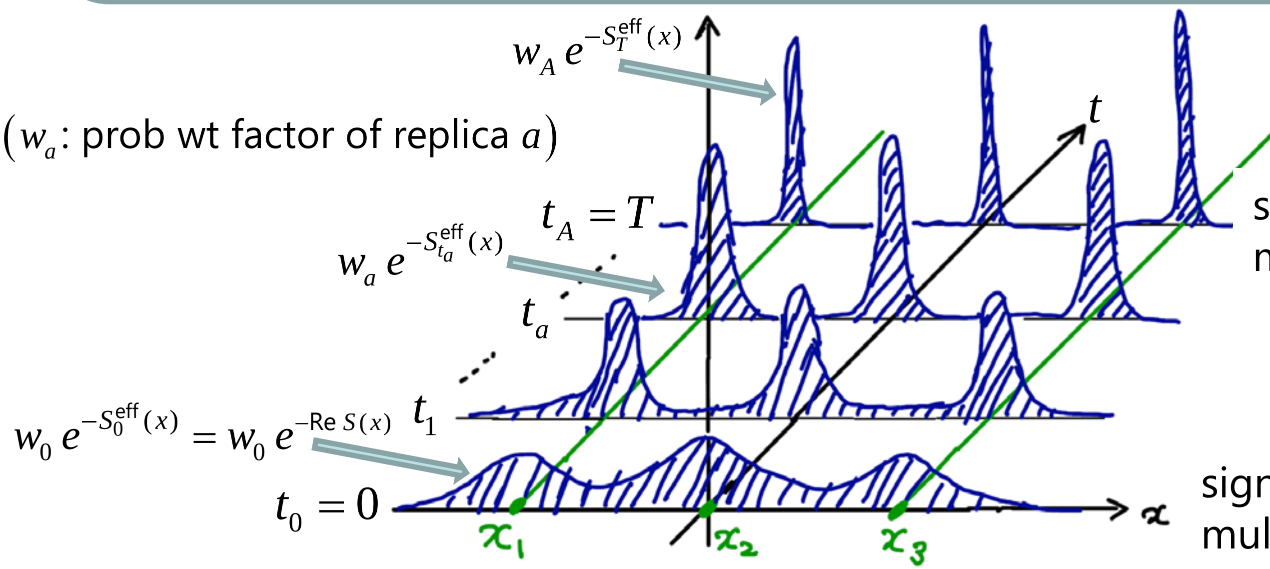
Algorithm of TLTM

(1) Introduce copies of config space labeled by a finite set of flow times

$$\mathcal{A} = \{t_a\}\ (a = 0,1,\dots,A) \quad \left(t_0 = 0 < t_1 < t_2 < \cdots < t_A = T\right),$$

and construct a Markov chain that drives the enlarged system to global equilibrium

$\left(w_a\text{: prob wt factor of replica } a\right)$



sign problem : **OK**
multimodality : NG

sign problem : NG
multimodality : **OK**

Tempered LTM (1/3)

[MF-Umeda 1703.00861]

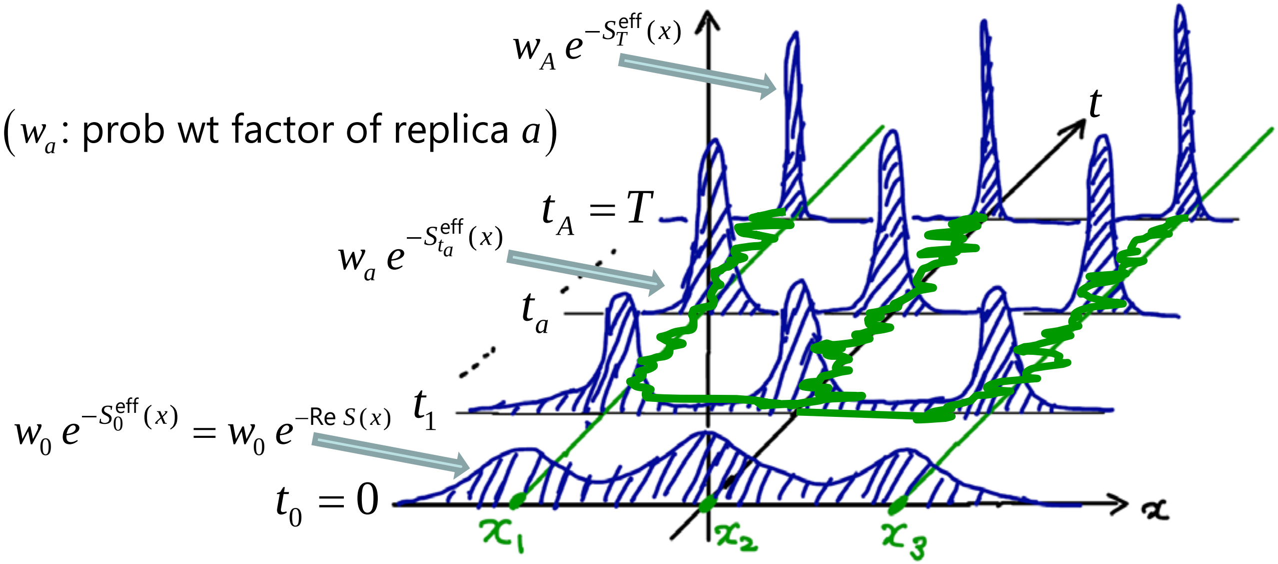
Algorithm of TLTM

(1) Introduce copies of config space labeled by a finite set of flow times

$$\mathcal{A} = \{t_a\}\ (a = 0,1,\ldots,A) \quad \left(t_0 = 0 < t_1 < t_2 < \cdots < t_A = T\right),$$

and construct a Markov chain that drives the enlarged system to global equilibrium

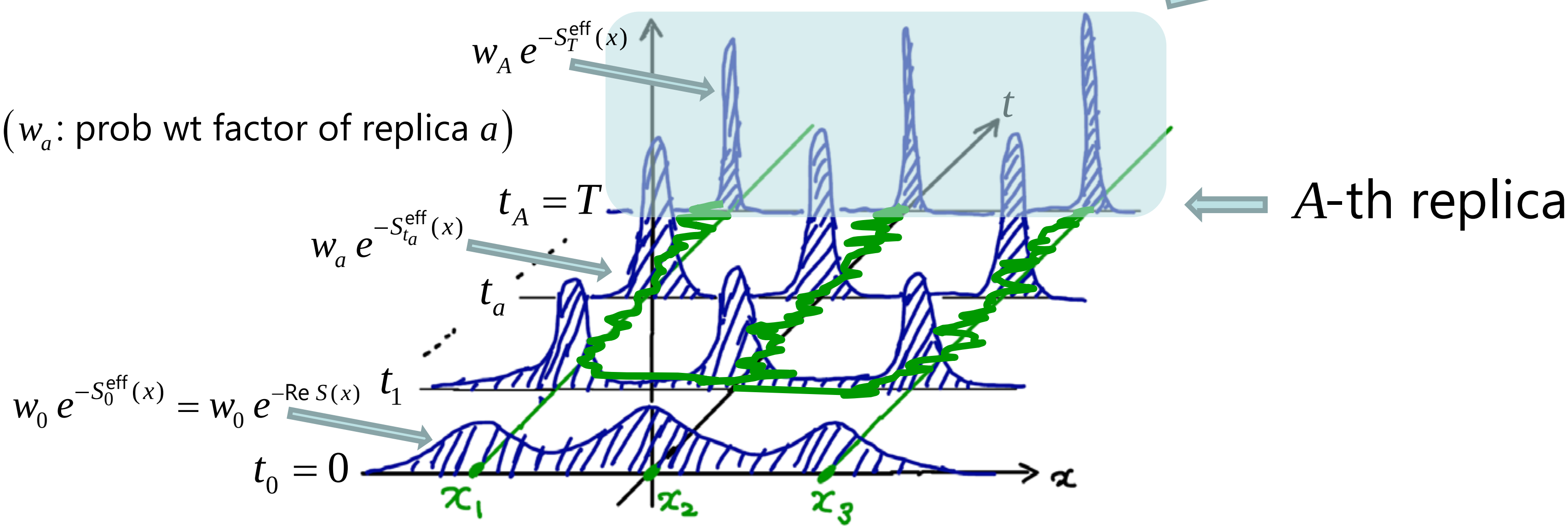
$\left(w_a\text{: prob wt factor of replica } a\right)$



[MF-Umeda 1703.00861]

Algorithm of TLTM

(2) After the enlarged system is relaxed to global equilibrium, evaluate the expectation value by using the subsample at $t_A = T$ $(a = A)$

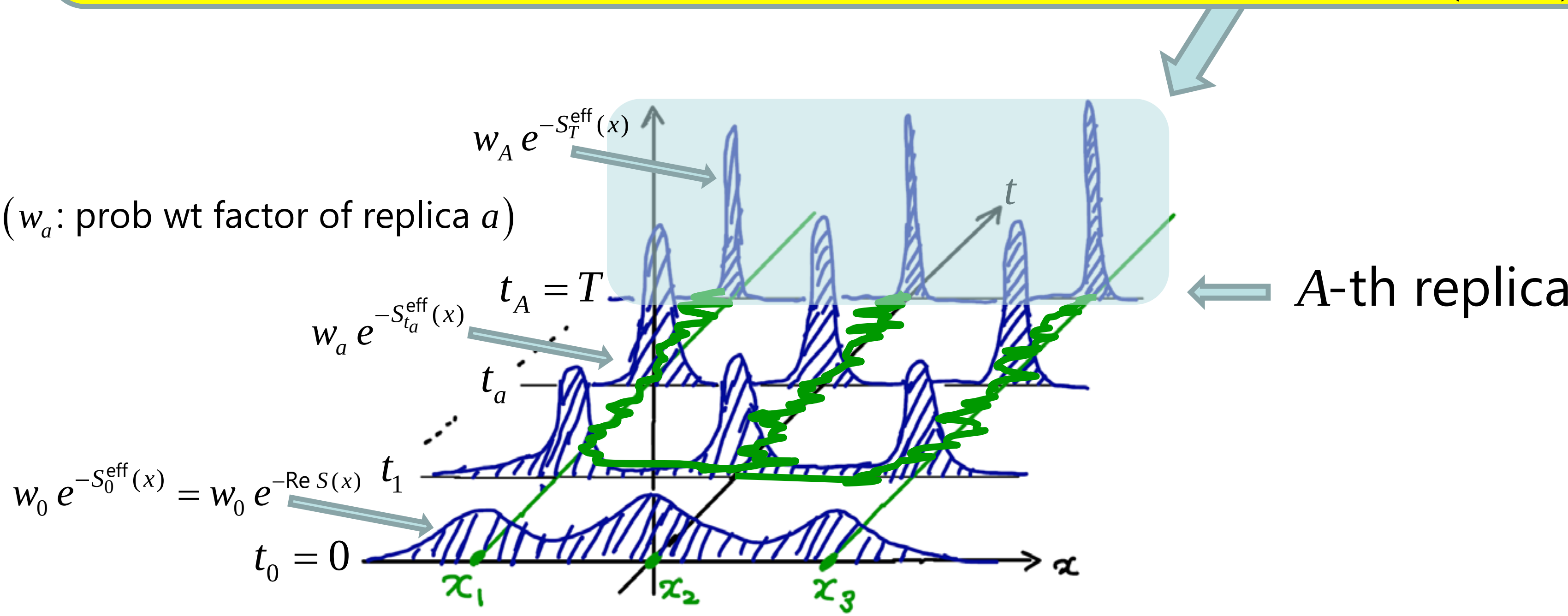


Tempered LTM (2/3)

[MF-Umeda 1703.00861]

Algorithm of TLTM

(2) After the enlarged system is relaxed to global equilibrium, evaluate the expectation value by using the subsample at $t_A = T$ $(a = A)$



$\left(w_a\text{: prob wt factor of replica } a\right)$

NB: various tempering methods $\left(\mathcal{M} \equiv \{x\}\right.$: original config space$\left.\right)$

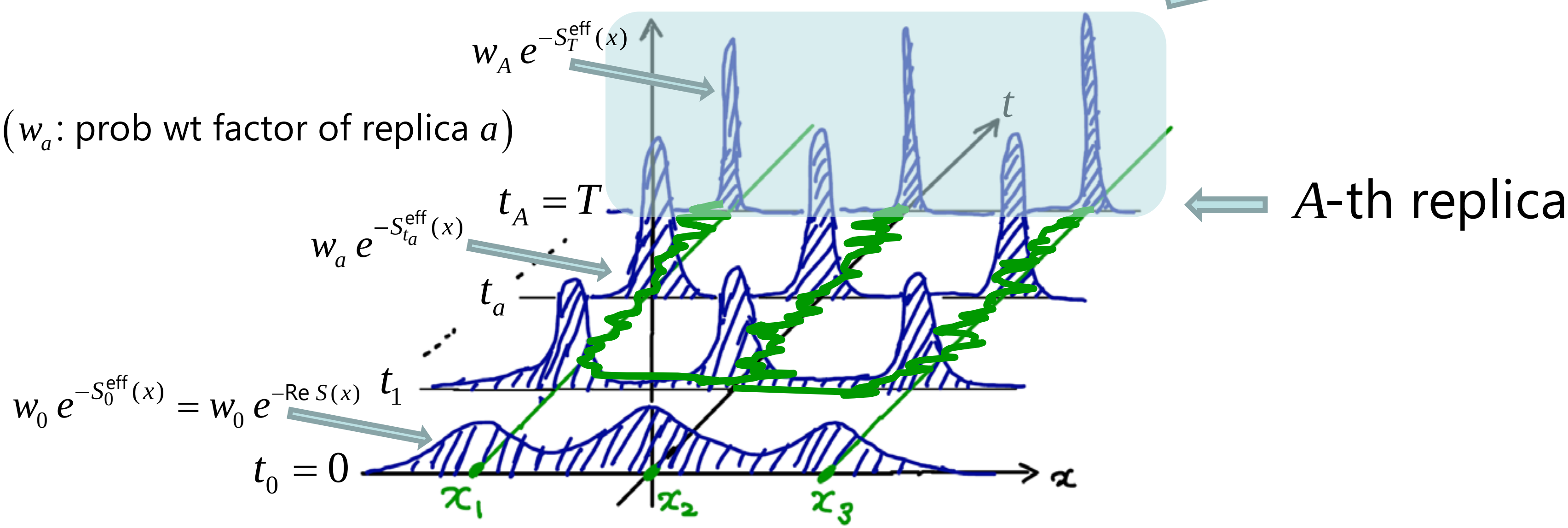
- simulated tempering : enlarged system $\Longleftrightarrow$ $\mathcal{M} \times \mathcal{A} = \{(x, t_a)\}$ $\left(\triangle \left[\text{tedious task to detemine the weights } w_a\right]\right)$
  [Marinari-Parisi 1992]
- parallel tempering (replica exchange MCMC) : enlarged system $\Longleftrightarrow$ $\overbrace{\mathcal{M} \times \mathcal{M} \times \cdots \times \mathcal{M}}^{A+1} = \{(x_0, x_1, \ldots, x_A)\}$ $(\bigcirc)$
  [Swendsen-Wang 1986, Geyer 1991, Nemoto-Hukushima 1996]
  most of relevant steps can be done in parallel processes

Tempered LTM (2/3)

[MF-Umeda 1703.00861]

Algorithm of TLTM

(2) After the enlarged system is relaxed to global equilibrium, evaluate the expectation value by using the subsample at $t_A = T$ $(a = A)$



NB: various tempering methods $\left(\mathcal{M} \equiv \{x\}\right.$: original config space$)$

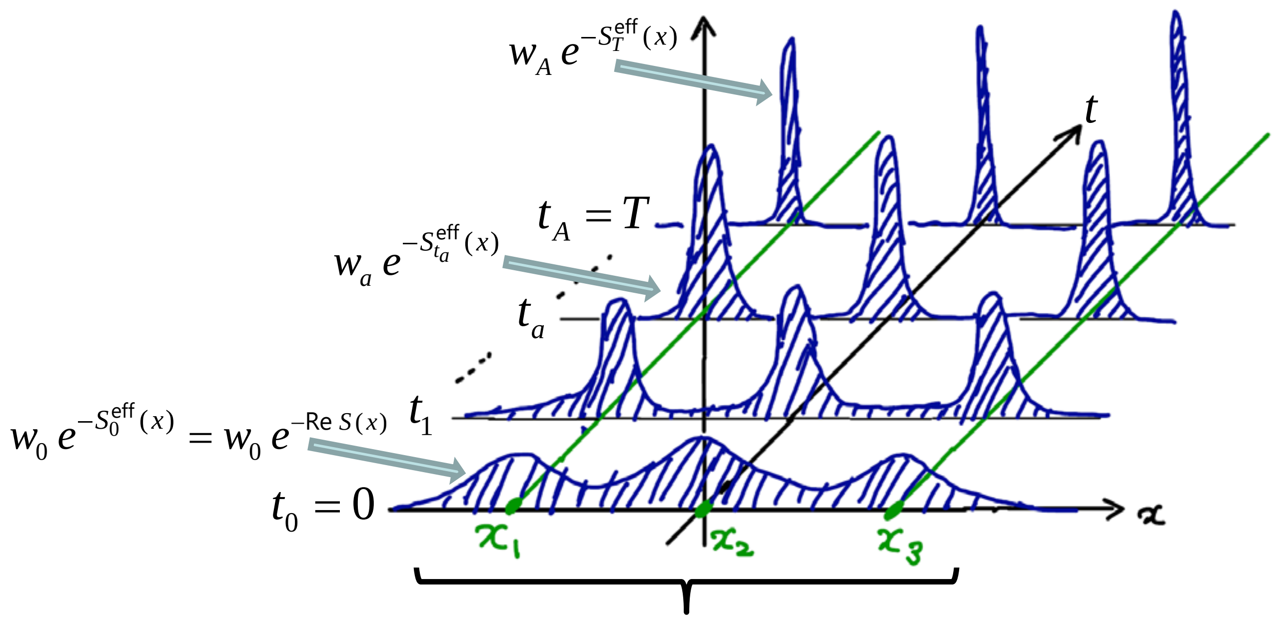
- simulated tempering : enlarged system $\Leftrightarrow$ $\mathcal{M} \times \mathcal{A} = \{(x, t_a)\}$ (△ [tedious task to detemine the weights w_a])
  [Marinari-Parisi 1992]
- parallel tempering (replica exchange MCMC) : enlarged system $\Leftrightarrow$ $\overbrace{\mathcal{M} \times \mathcal{M} \times \cdots \times \mathcal{M}}^{A+1} = \{(x_0, x_1, \ldots, x_A)\}$ (☺)
  most of relevant steps can be done in parallel processes
  [Swendsen-Wang 1986, Geyer 1991, Nemoto-Hukushima 1996]

Tempered LTM (3/3)

[MF-Umeda 1703.00861, MF-Matsumoto-Umeda 1906.04243]

Important points in TLTM:

(1) **NO "tiny overlap problem" in TLTM**



Distribution functions have peaks at the same positions x_σ for varying tempering parameter (which is t in our case)

⇒ We can expect significant overlap between adjacent replicas!

(2) **The growth of computational cost due to the tempering can be compensated by the increase of parallel processes**

Example: (0+1)-dim Massive Thirring model (1/3)

Lorentzian action (dim reduction of (1+1)D model):

[Pawlowski-Zielinski 1302.1622, 1402.6042, Fujii-Kamata-Kikukawa 1509.08176]

$$S_M = \int dt \left[i\bar{\psi}\gamma^0\partial_0\psi - m\bar{\psi}\psi - \frac{g^2}{2}(\bar{\psi}\gamma^0\psi)^2 \right] \quad \left((\gamma^0)^2 = 1_2, \quad \gamma^{0\dagger} = \gamma^0 \right)$$

bosonization + discretization

Grand partition function $Z_{\beta,\mu} = \mathrm{tr}\ \mathrm{e}^{-\beta(H-\mu Q)}$:

$$Z_{\beta,\mu} = \int_{\text{PBC}} (d\phi) e^{-S(\phi)}$$

with

$$(d\phi) = \prod_{n=1}^{N} \frac{d\phi_n}{2\pi}, \quad e^{-S(\phi)} = \det D(\phi) \exp\left[\frac{-1}{2g^2} \sum_{n=1}^{N} (1 - \cos\phi_n) \right]$$

$$D_{nn'}(\phi) = \frac{1}{2}\left(e^{i\phi_n + \mu}\delta_{n+1,n'} - e^{-(i\phi_n+\mu)}\delta_{n-1,n'} - e^{i\phi_N+\mu}\delta_{n,N}\delta_{n',1} + e^{-(i\phi_N+\mu)}\delta_{n,1}\delta_{n',N} \right) + m\delta_{n,n'}$$

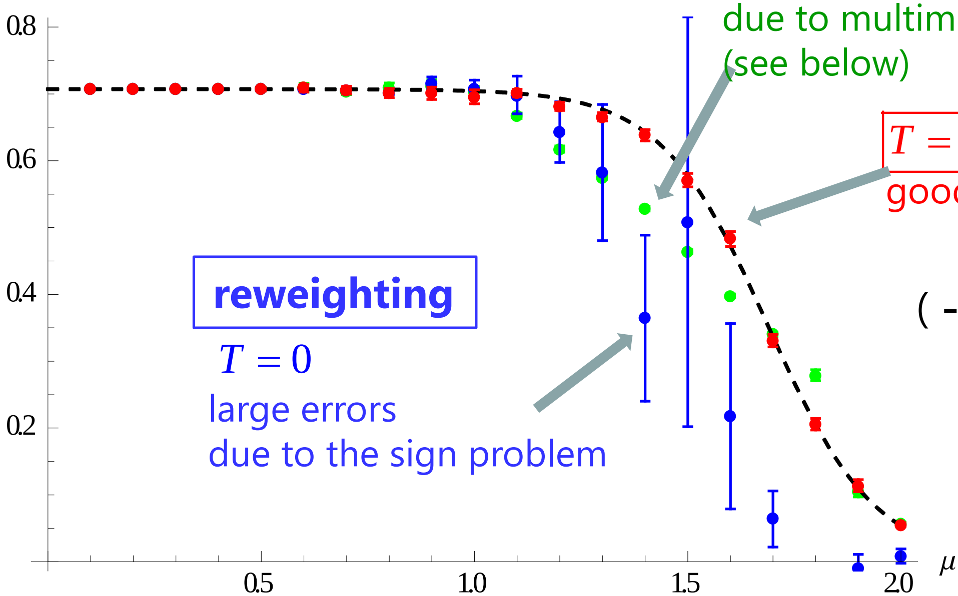
One can show $\boxed{[\det D(\phi;\mu)]^* = \det D(\phi;-\mu)}$ (thus, $\det D \notin \mathbb{R}$ for $\mu \in \mathbb{R}$)

⇨ Sign problem will arise when N is very large

Example: (0+1)-dim Massive Thirring model (2/3)

[MF-Umeda 1703.00861]

Chiral condensate $\langle \bar{\chi}\chi \rangle$



$T = 2$ (w/o temp)
deviate from exact values
due to multimodality
(see below)

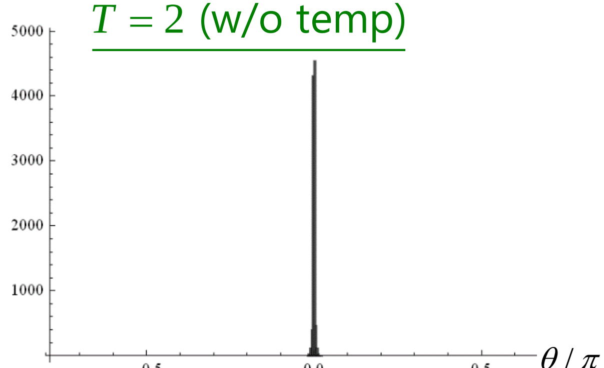
$T = 2$ (w/ temp)
good agreement

reweighting
$T = 0$
large errors
due to the sign problem

(------ : exact values)

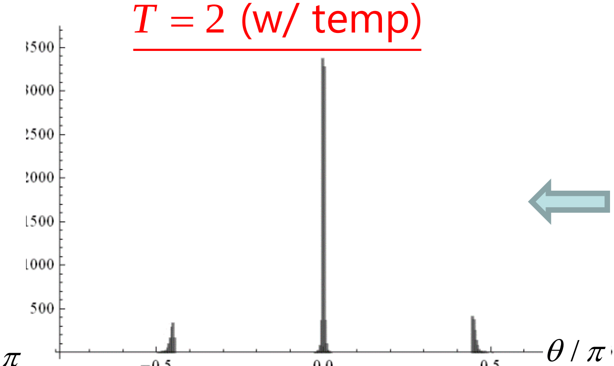
Confirmation of the resolution of multimodality

$T = 2$ (w/o temp)



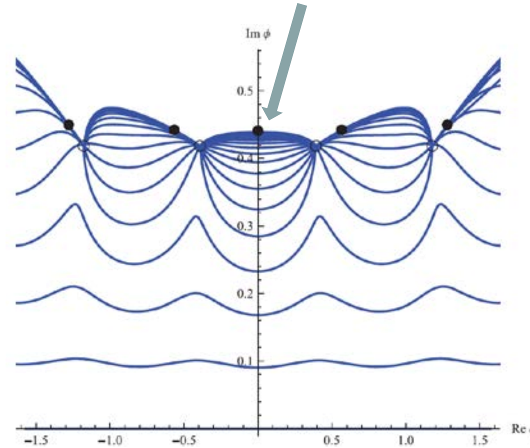
dominated by
a single thimble

$T = 2$ (w/ temp)



contributed by
multiple thimbles (i.e. config space is well explored)

dominant thimble



Example: (0+1)-dim Massive Thirring model (3/3)

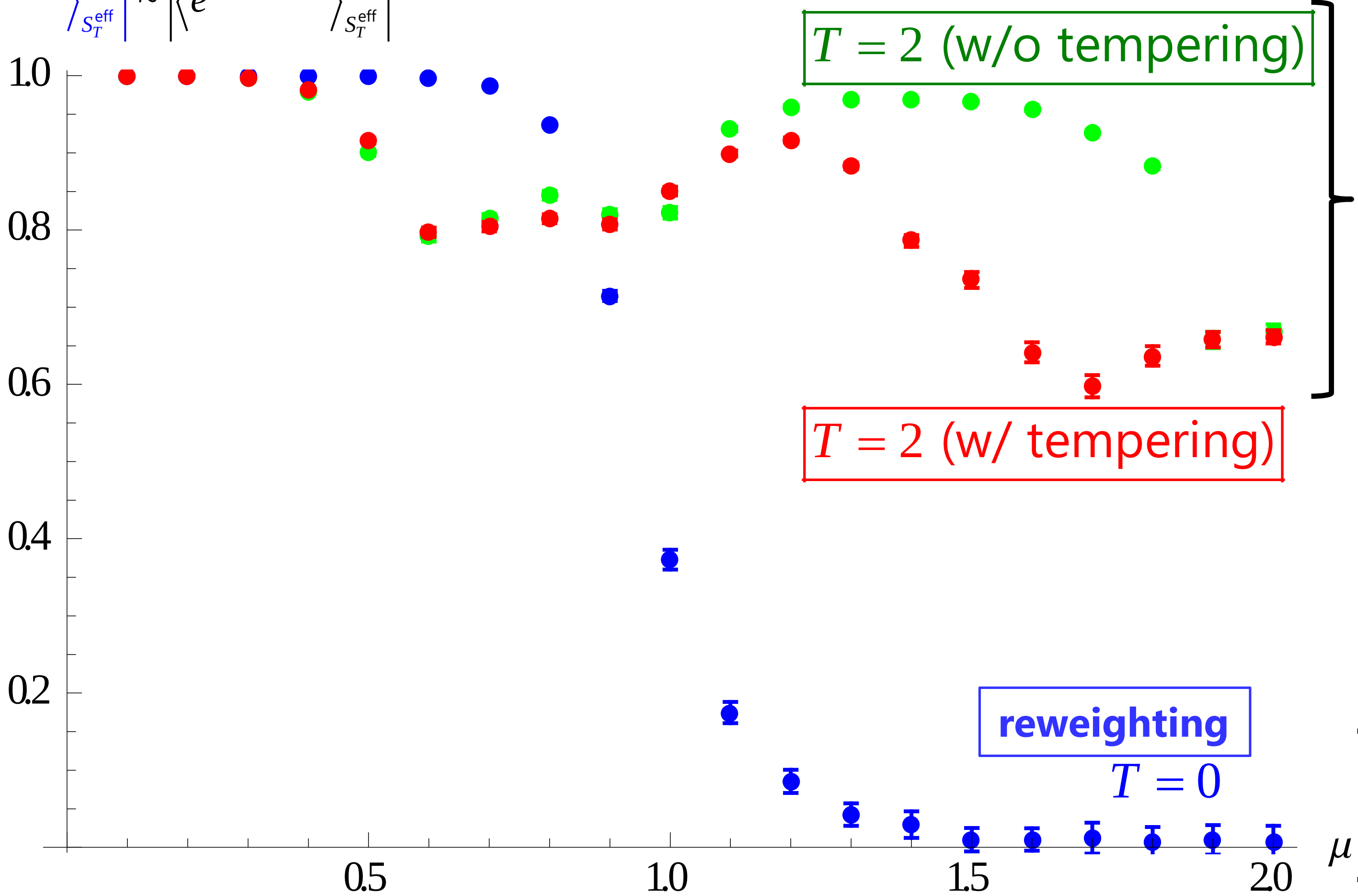
[MF-Umeda 1703.00861]

Confirmation of the resolution of sign problem

$$\left(\langle \mathcal{O}(\phi) \rangle = \frac{\langle e^{i\theta_T(\phi)}\, \mathcal{O}(\phi) \rangle_{S^{\mathrm{eff}}_{t_a}}}{\langle e^{i\theta_T(\phi)} \rangle_{S^{\mathrm{eff}}_T}} \right)$$

sign average

$$\left| \left\langle e^{i\theta_T(\phi)} \right\rangle_{S^{\mathrm{eff}}_T} \right| \sim \left| \left\langle e^{-iS_I(z_T(\phi))} \right\rangle_{S^{\mathrm{eff}}_T} \right|$$



no sign problem at $T = 2$

(NB: sign avarage $\left| \left\langle e^{i\theta_T(\phi)} \right\rangle_{S^{\mathrm{eff}}_T} \right|$ is smaller for the right sampling)

sign problem surely exists for the original action ($T = 0$)

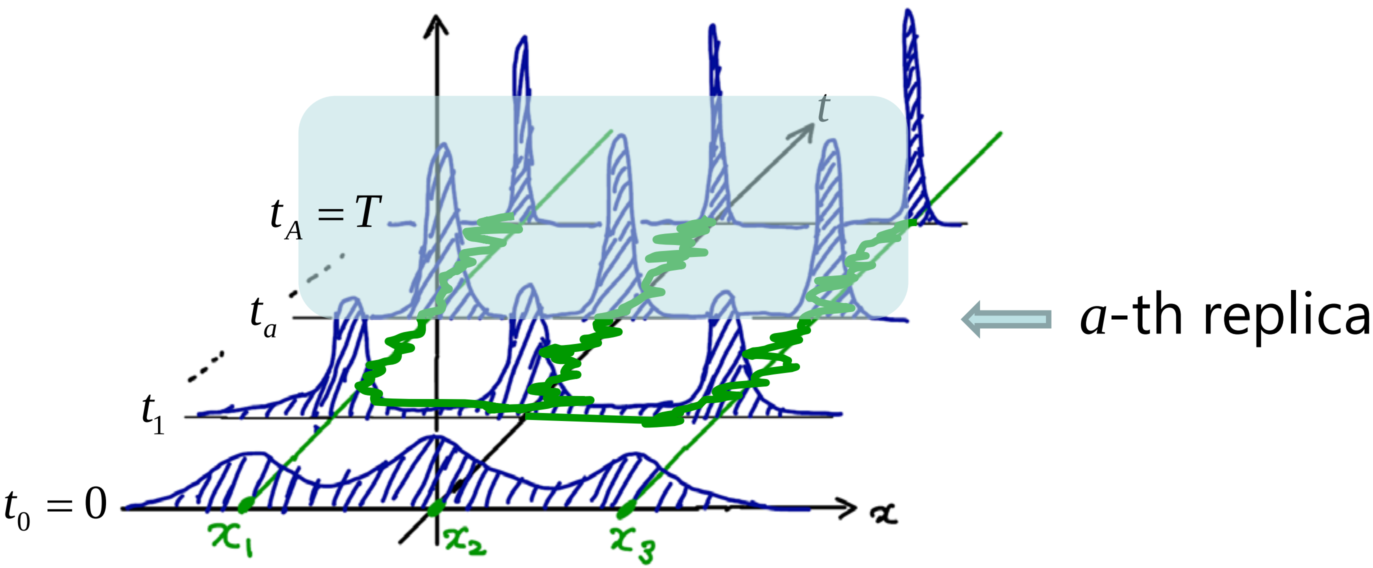
We actually can go further…

[MF-Matsumoto-Umeda 1906.04243]

Consider the estimates of $\langle \mathcal{O} \rangle_S$ at various flow times t_a:

$$\langle \mathcal{O} \rangle_S = \frac{\langle e^{i\theta_{t_a}(x)} \mathcal{O}(z_{t_a}(x)) \rangle_{S^{\mathrm{eff}}_{t_a}}}{\langle e^{i\theta_{t_a}(x)} \rangle_{S^{\mathrm{eff}}_{t_a}}} \approx \frac{\sum_{k=1}^{N_{\mathrm{conf}}} e^{i\theta_{t_a}(x^{(k)})} \mathcal{O}(z_{t_a}(x^{(k)}))}{\sum_{k=1}^{N_{\mathrm{conf}}} e^{i\theta_{t_a}(x^{(k)})}} \equiv \overline{\mathcal{O}}_a \quad (a = 0, 1, \ldots, A)$$

Here the estimation on the RHS is made by using the subsample at t_a:



We actually can go further…

[MF-Matsumoto-Umeda 1906.04243]

Consider the estimates of $\langle \mathcal{O} \rangle_S$ at various flow times t_a:

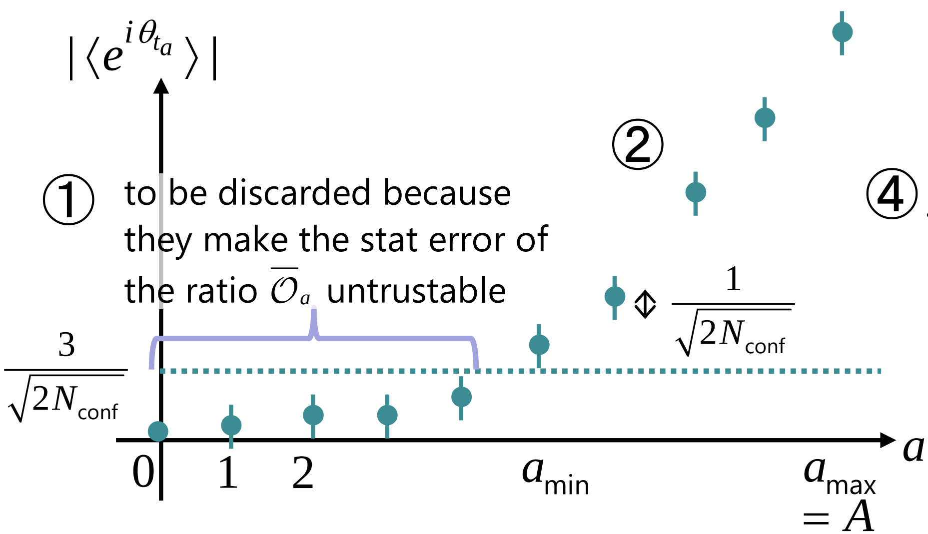
$$\langle \mathcal{O} \rangle_S = \frac{\langle e^{i\theta_{t_a}(x)} \mathcal{O}(z_{t_a}(x)) \rangle_{S^{\rm eff}_{t_a}}}{\langle e^{i\theta_{t_a}(x)} \rangle_{S^{\rm eff}_{t_a}}} \approx \frac{\sum_{k=1}^{N_{\rm conf}} e^{i\theta_{t_a}(x^{(k)})} \mathcal{O}(z_{t_a}(x^{(k)}))}{\sum_{k=1}^{N_{\rm conf}} e^{i\theta_{t_a}(x^{(k)})}} \equiv \overline{\mathcal{O}}_a \quad (a = 0,1,\ldots,A)$$

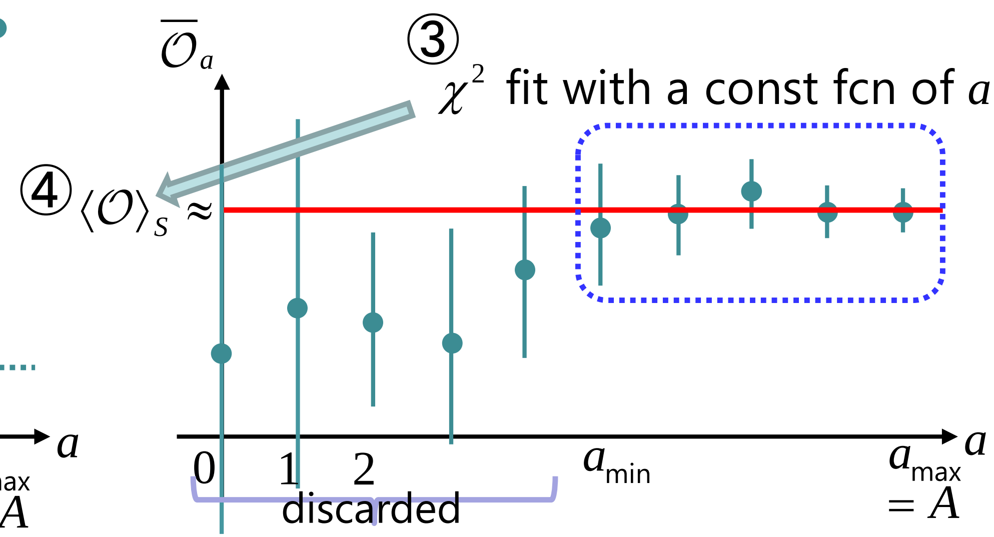
The LHS must be independent of a due to Cauchy's theorem

⇓

The RHS must be the same for all a's within the statistical error margin if the system is in global equilibrium and the sample size is large enough

This gives a method with a criterion for precise estimation in the TLTM!





4. Applying the TLTM to the Hubbard model

[MF-Matsumoto-Umeda 1906.04243]

Hubbard model (1/2)

Hubbard model [Hubbard 1963]

modeling NR electrons in a solid

- $c^\dagger_{\mathbf{x},\sigma},\ c_{\mathbf{x},\sigma}$: creation/anihilation op of an electron on site $\mathbf{x}$ with spin $\sigma(=\uparrow,\downarrow)$

- Hamiltonian

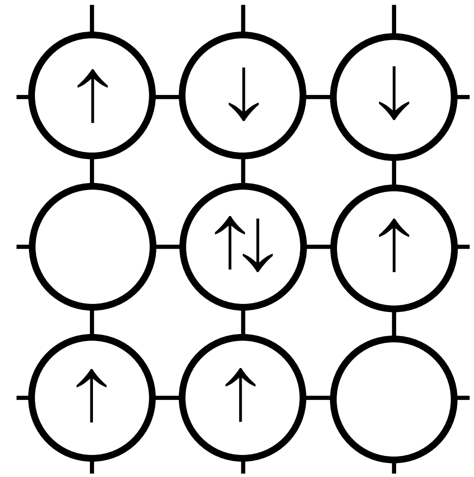
$$H = -\kappa \sum_{\langle \mathbf{x},\mathbf{y}\rangle} \sum_{\sigma} c^\dagger_{\mathbf{x},\sigma} c_{\mathbf{y},\sigma} - \mu \sum_{\mathbf{x}} \left(n_{\mathbf{x},\uparrow} + n_{\mathbf{x},\downarrow} \right) + U \sum_{\mathbf{x}} n_{\mathbf{x},\uparrow} n_{\mathbf{x},\downarrow}$$

$n_{\mathbf{x},\sigma} \equiv c^\dagger_{\mathbf{x},\sigma} c_{\mathbf{x},\sigma}$

$\kappa(>0)$: hopping parameter

μ : chemical potential

$U(>0)$: strength of on-site replusive potential



bipartite lattice

(N_S : # of sites)

$n_{\mathbf{x},\sigma} \to n_{\mathbf{x},\sigma} - 1/2$ s.t. $\mu = 0 \Leftrightarrow$ half-filling $\sum_{\sigma=\uparrow,\downarrow} \left\langle n_{\mathbf{x},\sigma} - 1/2 \right\rangle = 0$

$$H = \underbrace{-\kappa \sum_{\mathbf{x},\mathbf{y}} \sum_{\sigma} K_{\mathbf{xy}}\, c^\dagger_{\mathbf{x},\sigma} c_{\mathbf{y},\sigma} - \mu \sum_{\mathbf{x}} \left(n_{\mathbf{x},\uparrow} + n_{\mathbf{x},\downarrow} - 1 \right)}_{H_1} + \underbrace{U \sum_{\mathbf{x}} \left(n_{\mathbf{x},\uparrow} - \frac{1}{2} \right)\left(n_{\mathbf{x},\downarrow} - \frac{1}{2} \right)}_{H_2}$$

(fermion bilinear) (four fermion)

Hubbard model (2/2)

- Grand partition function (continuous imaginary time) : $Z^{\text{cont}}_{\beta,\mu} = \text{tr}\, e^{-\beta H}$
- Quantum Monte Carlo

$$e^{-\beta H} = e^{-\beta(H_1+H_2)} = \left(e^{-\epsilon(H_1+H_2)}\right)^{N_\tau} \cong \left(e^{-\epsilon H_1} e^{-\epsilon H_2}\right)^{N_\tau} \quad (\beta = N_\tau \epsilon)$$

⇨ Transform $e^{-\epsilon H_2} = \prod_{\mathbf{x}} e^{-\epsilon U \left(n_{\mathbf{x},\uparrow} - 1/2\right)\left(n_{\mathbf{x},\downarrow} - 1/2\right)}$ to a fermion bilinear using a boson ϕ

⇨

$$Z_{\beta,\mu} = \int [d\phi] e^{-S[\phi_{\ell,\mathbf{x}}]} \equiv \int \prod_{\ell=1}^{N_\tau} \prod_{\mathbf{x}} d\phi_{\ell,\mathbf{x}}\, e^{-(1/2)\sum_{\ell,\mathbf{x}} \phi_{\ell,\mathbf{x}}{}^2} \det M_\uparrow[\phi] \det M_\downarrow[\phi]$$

$$M_{\uparrow/\downarrow}[\phi] \equiv 1_{N_s} + e^{\pm\beta\mu} \prod_{\ell} \left(e^{\epsilon\kappa K}\, \text{diag}[e^{\pm i\sqrt{\epsilon U}\phi_{\ell,\mathbf{x}}}] \right) \ : \ N_s \times N_s \text{ matrix}$$

This gives complex actions for non half-filling $(\mu \neq 0)$

(NB: For half-filling $(\mu = 0)$

$$\det M_\uparrow[\phi] \det M_\downarrow[\phi] = \left|\det M_\uparrow[\phi]\right|^2 \geq 0$$

$\Rightarrow$ No sign problem)

⇨ We apply the Tempered LTM to this system $\left(\begin{array}{l} x = (x^i) = (\phi_{\ell,\mathbf{x}}) \in \mathbb{R}^N \\ i = 1,\ldots,N \ \ (N = N_\tau N_s) \end{array} \right)$

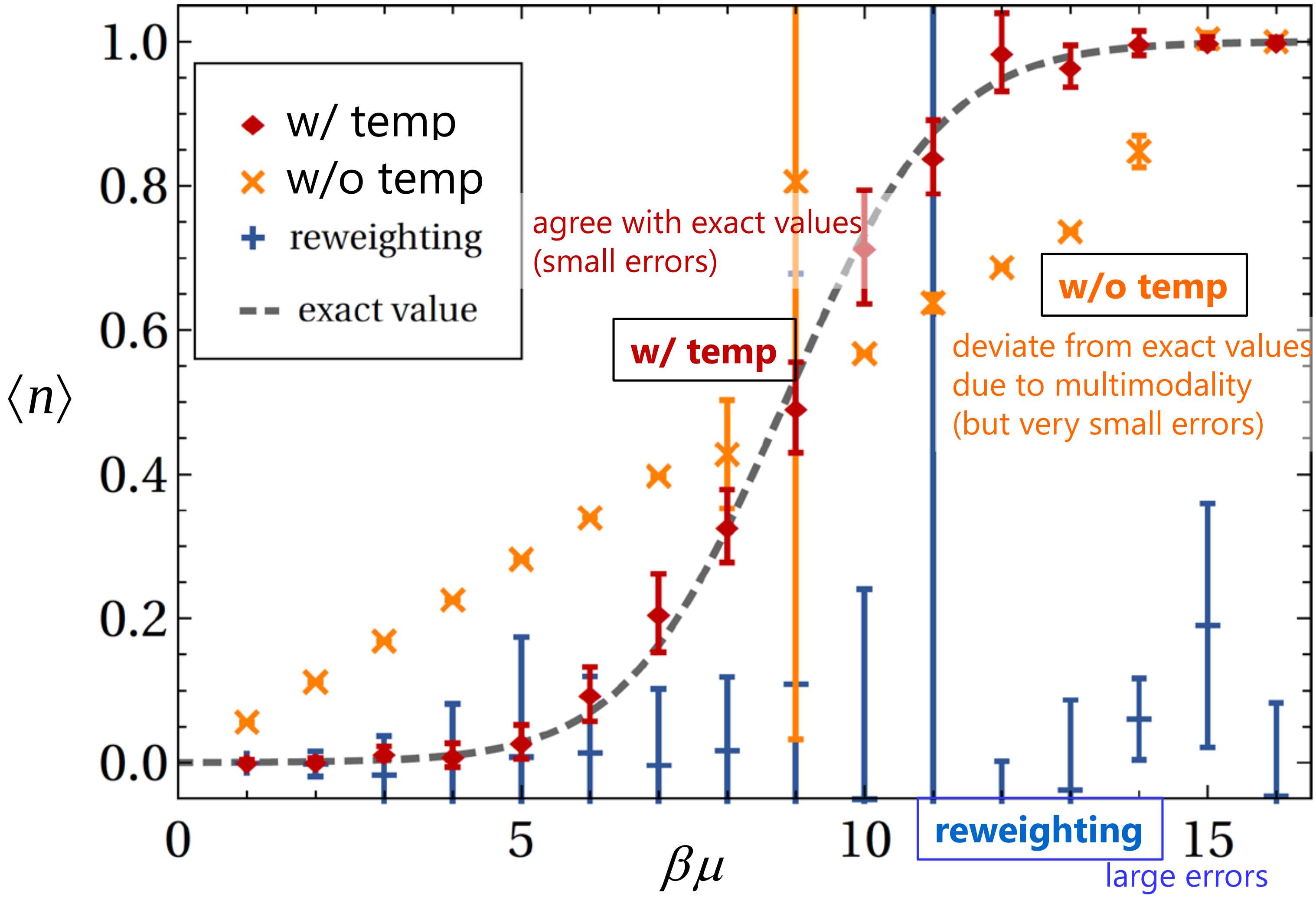
[MF-Matsumoto-Umeda 1906.04243]

Results for 1D lattice (1/3)

[MF-Matsumoto-Umeda 2019]

imaginary time : 2 steps $(N_\tau = 2)$
spatial lattice: 1D periodic lattice with $N_s = 2$
$\beta\kappa = 1, \quad \beta U = 16,$ max flow time $T = 0.4$
sample size: 5,000

number density $n = \dfrac{1}{N_s}\sum_x (n_{x,\uparrow} + n_{x,\downarrow} - 1)$



agree with exact values (small errors)

w/ temp

w/o temp

deviate from exact values due to multimodality (but very small errors)

reweighting

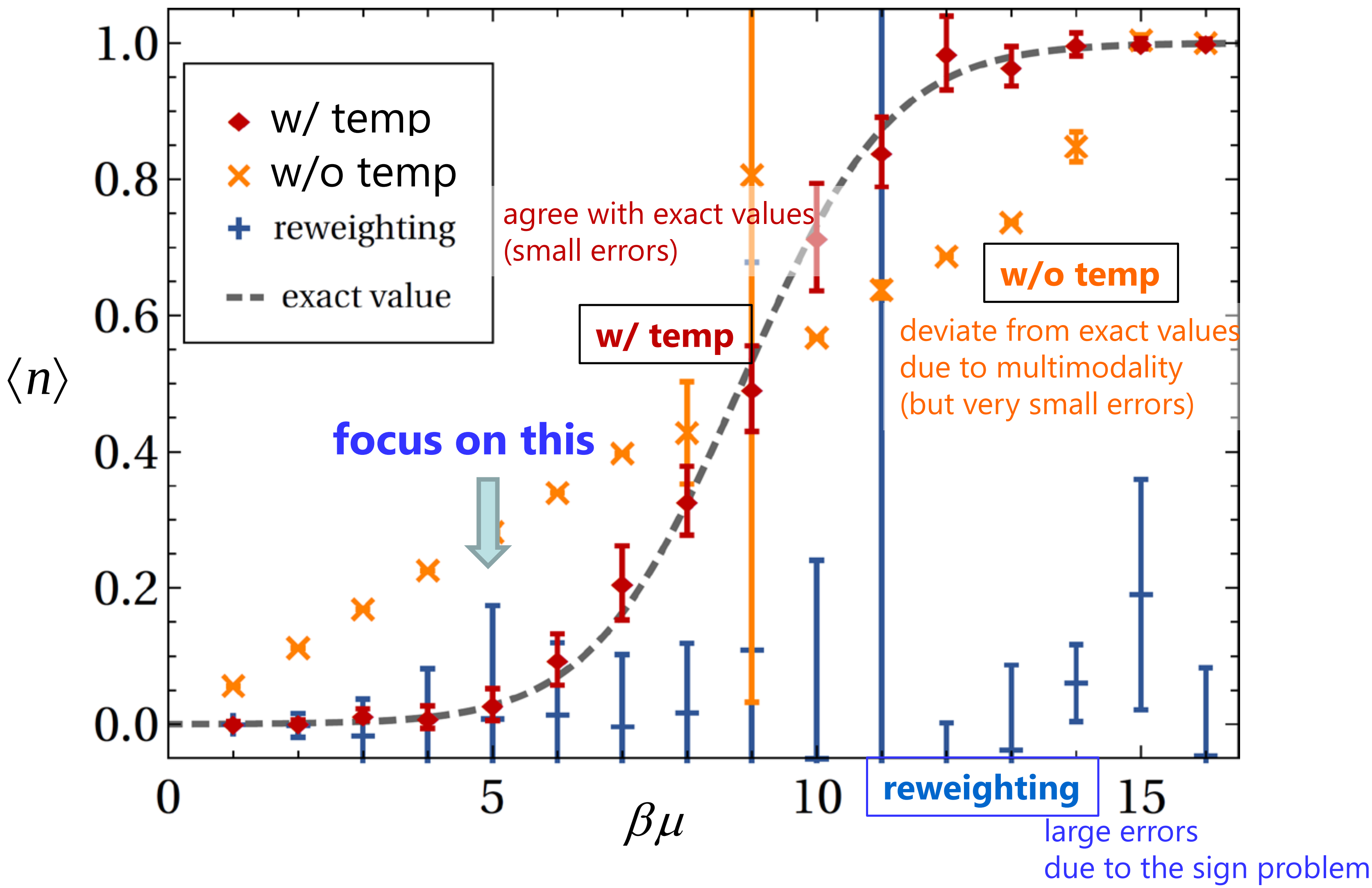
large errors due to the sign problem

Results for 1D lattice (1/3)

[MF-Matsumoto-Umeda 2019]

imaginary time : 2 steps ($N_\tau = 2$)
spatial lattice: 1D periodic lattice with $N_s = 2$
$\beta\kappa = 1, \quad \beta U = 16, \quad$ max flow time $T = 0.4$
sample size: 5,000

number density $n = \frac{1}{N_s} \sum_x \left(n_{x,\uparrow} + n_{x,\downarrow} - 1 \right)$

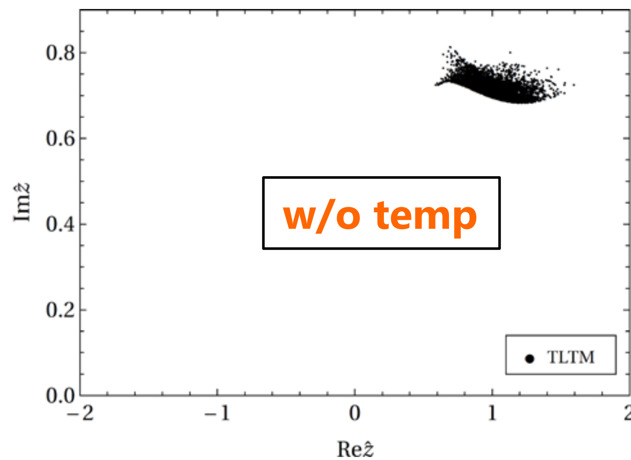


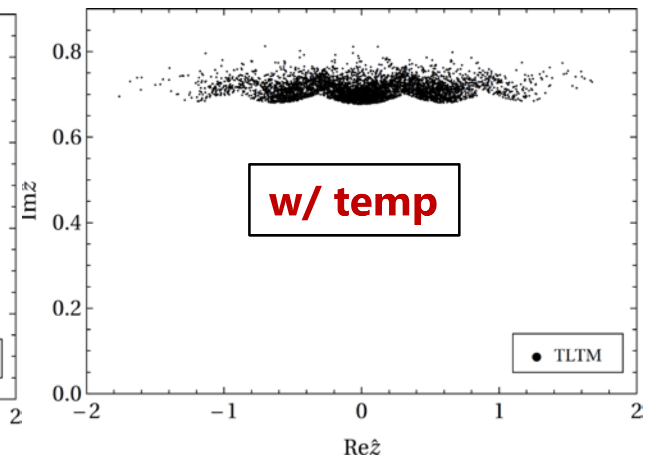
Results for 1D lattice (2/3)

[MF-Matsumoto-Umeda 2019]

Distribution of flowed configs at flow time $T = 0.4$

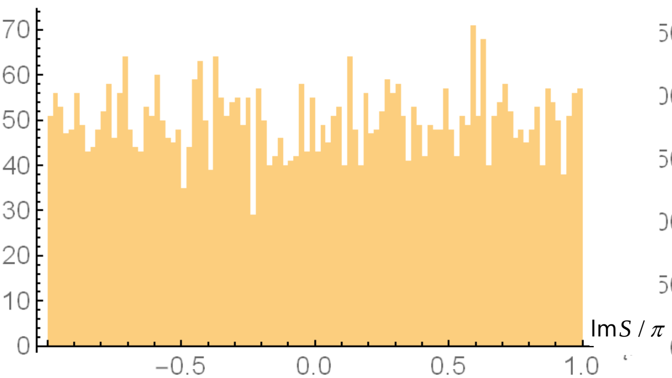
(projected on a plane)





Histogram of ImS(z)/π

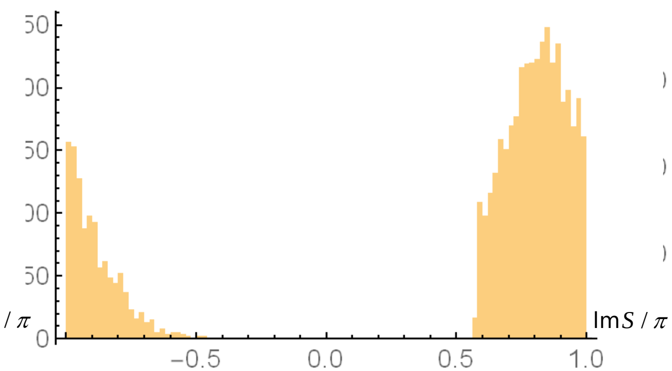
reweighting



distributing uniformly
from −π to +π

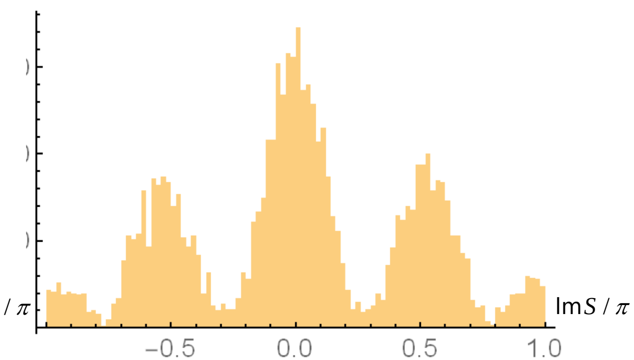
⇨ severe sign problem

w/o temp



peaked at a single angle ~0.8 π
due to the trap to a single thimble
(errors become small
because the thimble is well sampled)

w/ temp



peaked at several angles
because of sufficient transitions
among thimbles
(errors become a bit larger
due to the small size of sampling)

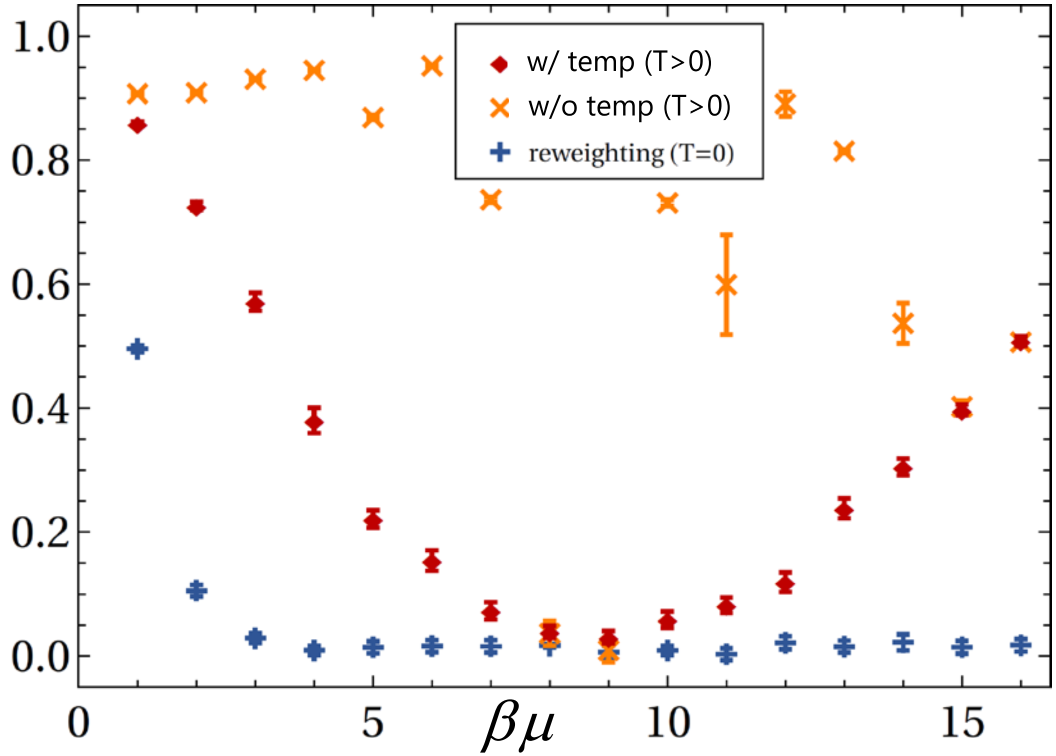
Results for 1D lattice (3/3)

[MF-Matsumoto-Umeda 2019]

sign average

$$\left(\langle \mathcal{O}(x) \rangle = \frac{\langle e^{i\theta_T(x)}\, \mathcal{O}(z_T(x)) \rangle_{S_T^{\text{eff}}}}{\langle e^{i\theta_T(x)} \rangle_{S_T^{\text{eff}}}} \right)$$

$$\left| \langle e^{i\theta_T(x)} \rangle_{S_T^{\text{eff}}} \right|$$



When only a single (or very few) thimble(s) is sampled, the sign average can become larger than the correct sampling due to the absence of phase mixtures among thimbles

⇒ It is generally dangerous to regard the sign average as an index of the "resolution of the sign problem"

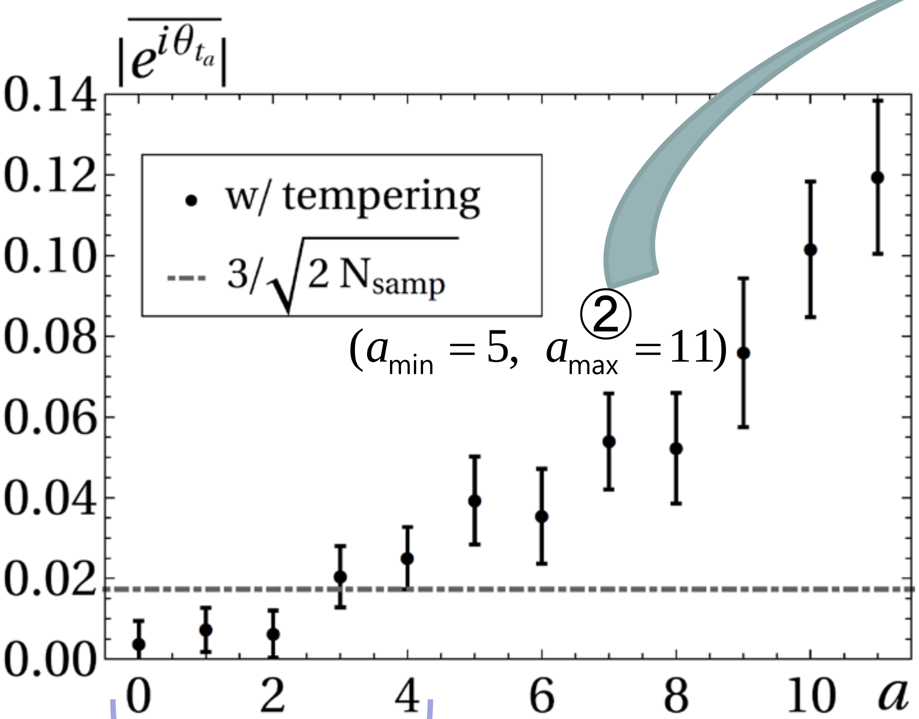
Results for 2D lattice (1/5)

[MF-Matsumoto-Umeda 1906.04243]

imaginary time : 5 steps ($N_\tau = 5$)
spatial lattice: 2D periodic lattice with $N_s = 2\times 2$
$\beta\kappa = 3$ $\beta U = 13$, max flow time $T = 0.5$
sample size: 5,000~25,000 depending on $\beta\mu$

$$\left(\langle n \rangle = \frac{\langle e^{i\theta_{t_a}(x)}\, n(z_{t_a}(x))\rangle_{S^{\rm eff}_{t_a}}}{\langle e^{i\theta_{t_a}(x)}\rangle_{S^{\rm eff}_{t_a}}} \approx \overline{n}_a \right)$$

Example: $\beta\mu = 5$

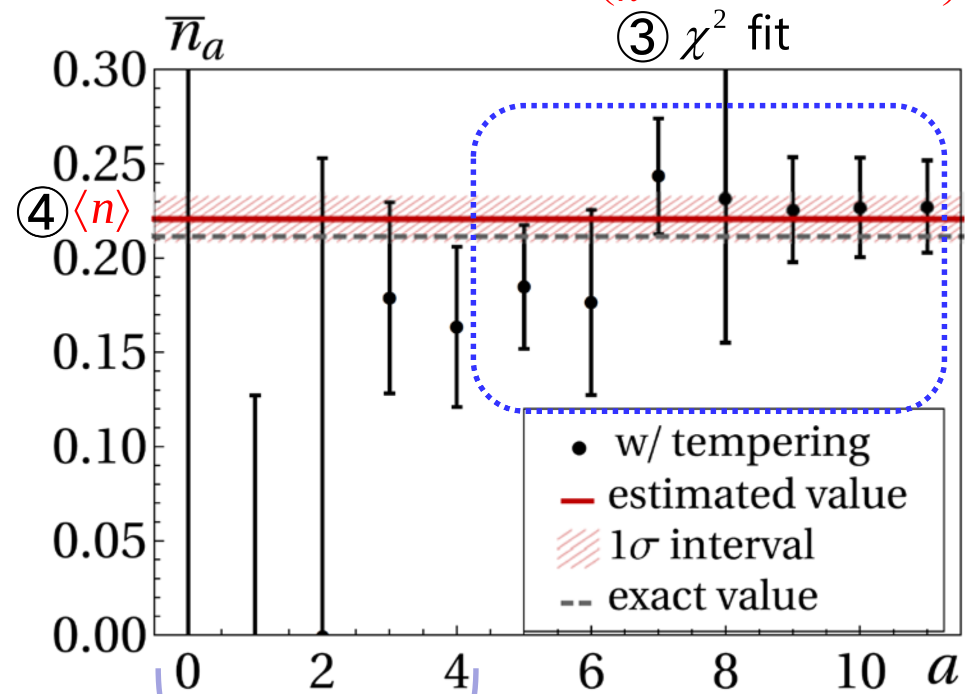


② ($a_{\rm min} = 5$, $a_{\rm max} = 11$)

① discarded

$\left(\langle n \rangle \approx 0.221 \pm 0.012 \quad \chi^2/\mathrm{DOF} = 0.45 \right)$

③ χ^2 fit



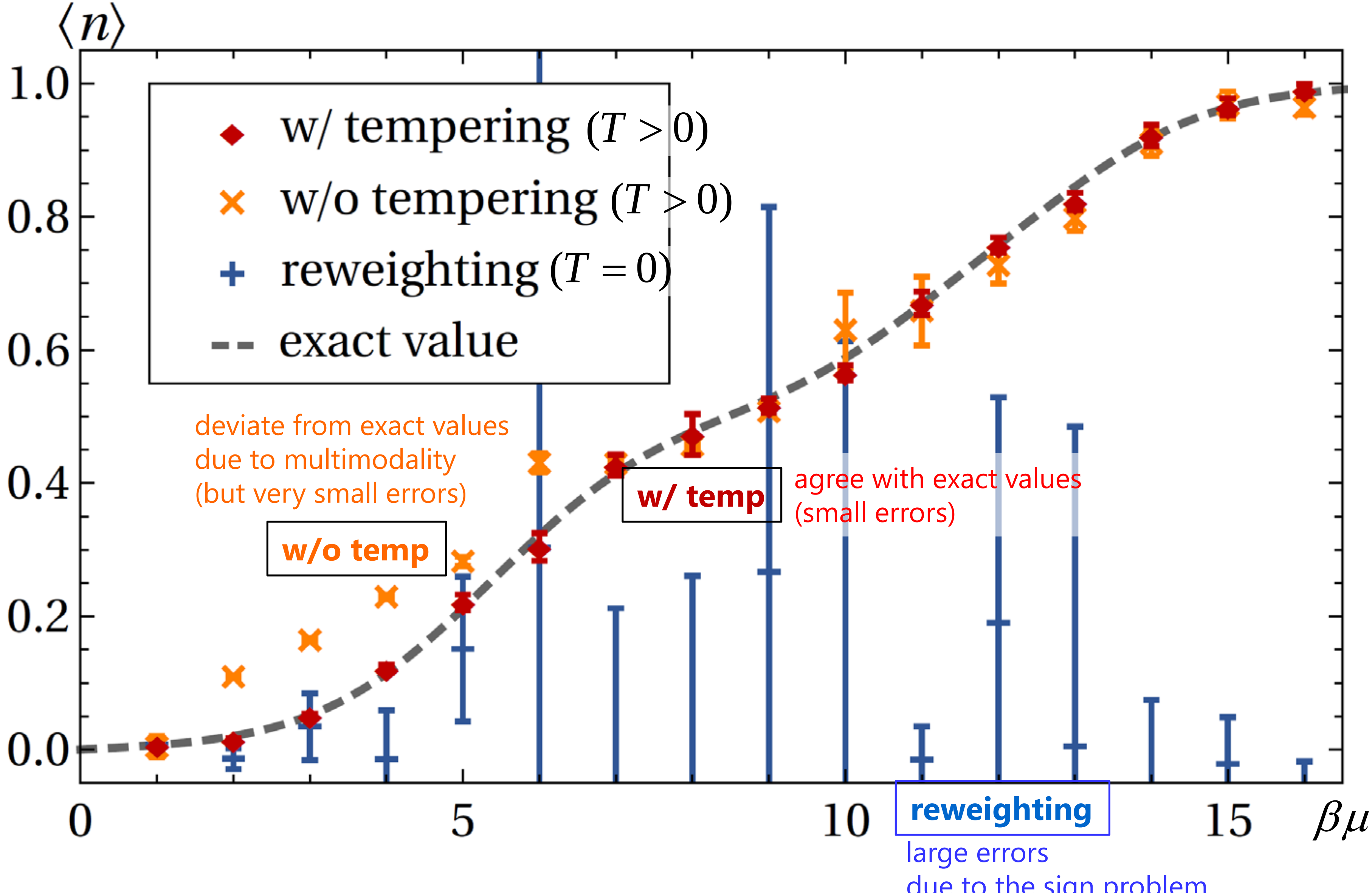
discarded

Results for 2D lattice (2/5)

$$\begin{bmatrix} N_\tau = 5, N_s = 2\times 2 \\ \beta\kappa = 3, \quad \beta U = 13 \end{bmatrix}$$

$$\langle n \rangle = \left\langle \frac{1}{N_s} \sum_{\mathbf{x}} \left(n_{\mathbf{x},\uparrow} + n_{\mathbf{x},\downarrow} - 1 \right) \right\rangle$$

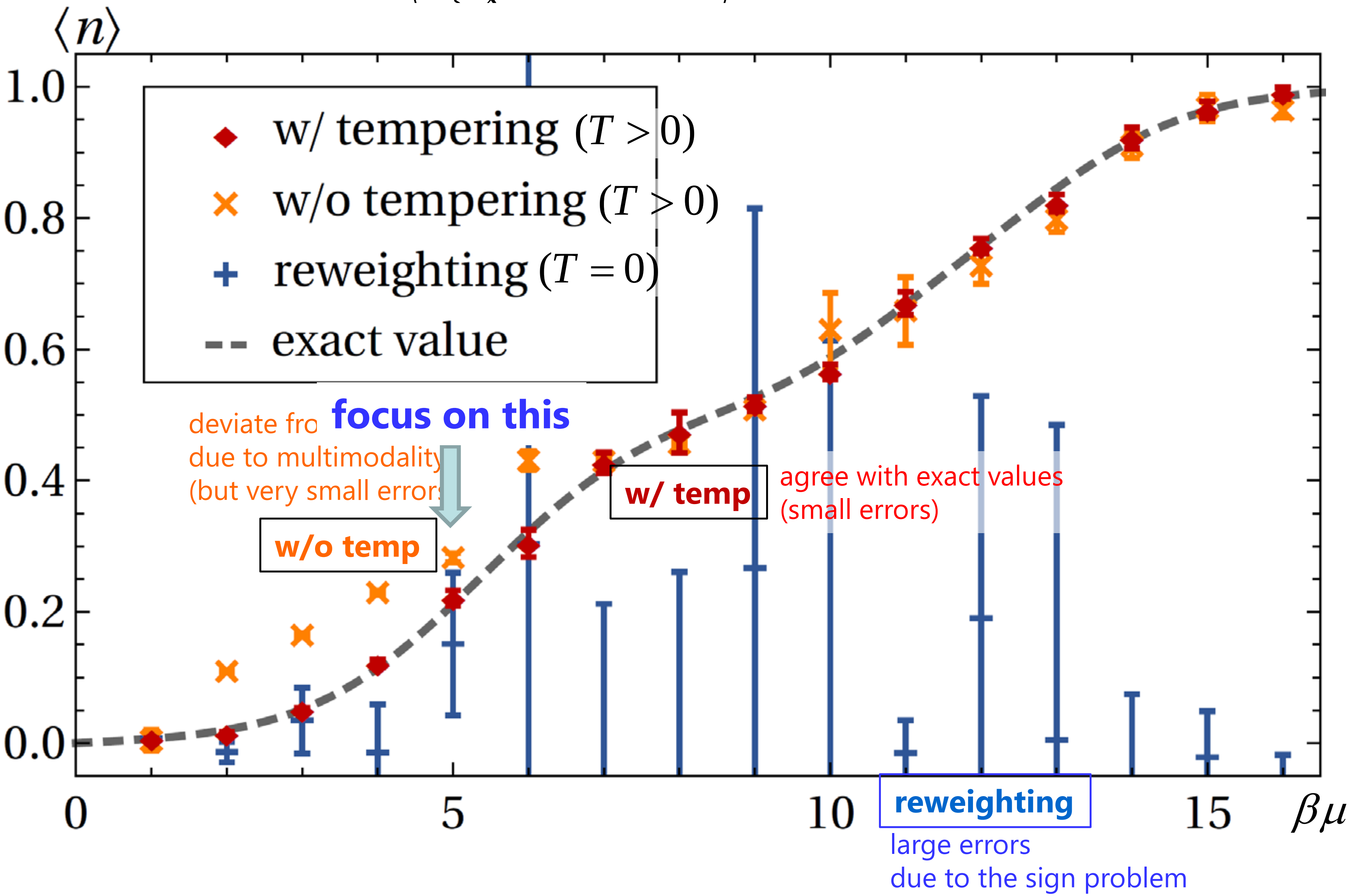
[MF-Matsumoto-Umeda 1906.04243]



Results for 2D lattice (2/5)

$$\begin{bmatrix} N_\tau = 5, N_s = 2\times 2 \\ \beta\kappa = 3, \quad \beta U = 13 \end{bmatrix} \qquad \langle n \rangle = \left\langle \frac{1}{N_s} \sum_{\mathbf{x}} \left(n_{\mathbf{x},\uparrow} + n_{\mathbf{x},\downarrow} - 1 \right) \right\rangle$$

[MF-Matsumoto-Umeda 1906.04243]

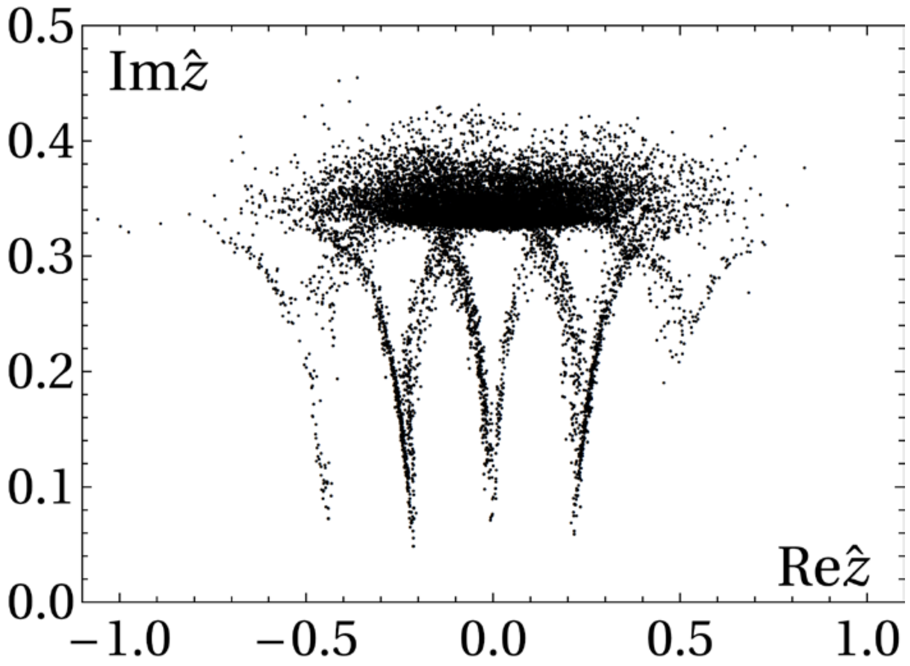


Results for 2D lattice (3/5)

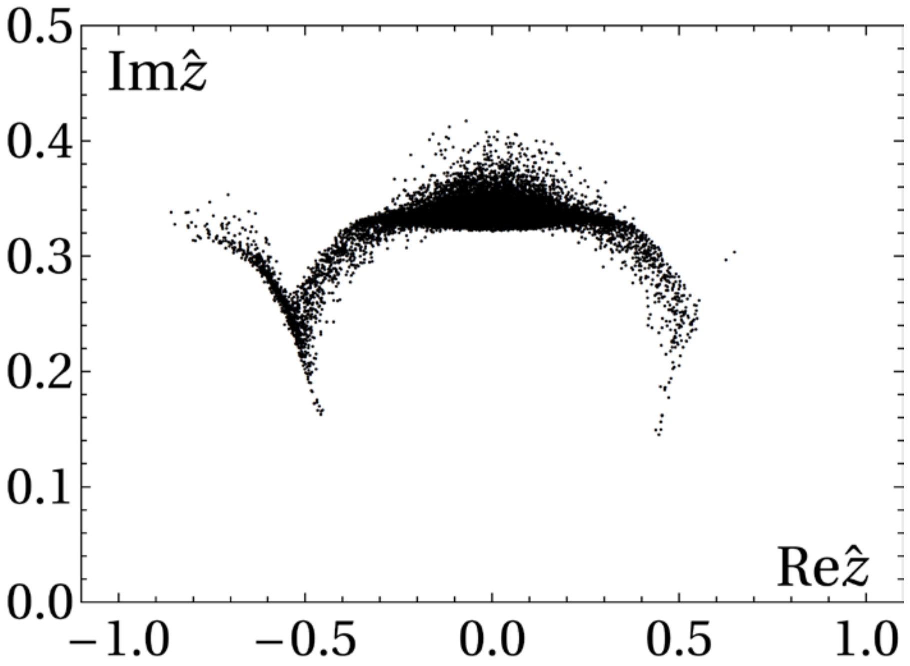
[MF-Matsumoto-Umeda 1906.04243]

Distribution of flowed configs at flow time $T = 0.5$ $(\beta\mu = 5)$
(projected on a plane)

w/ temp



w/o temp



distributed widely
over many thimbles

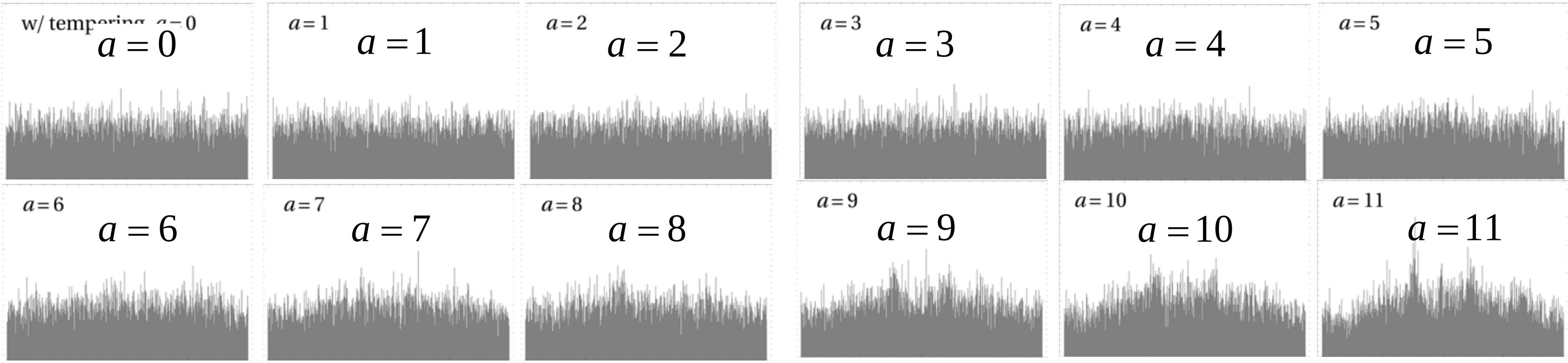
distributed over only
a small number of thimbles

Results for 2D lattice (4/5)

[MF-Matsumoto-Umeda 1906.04243]

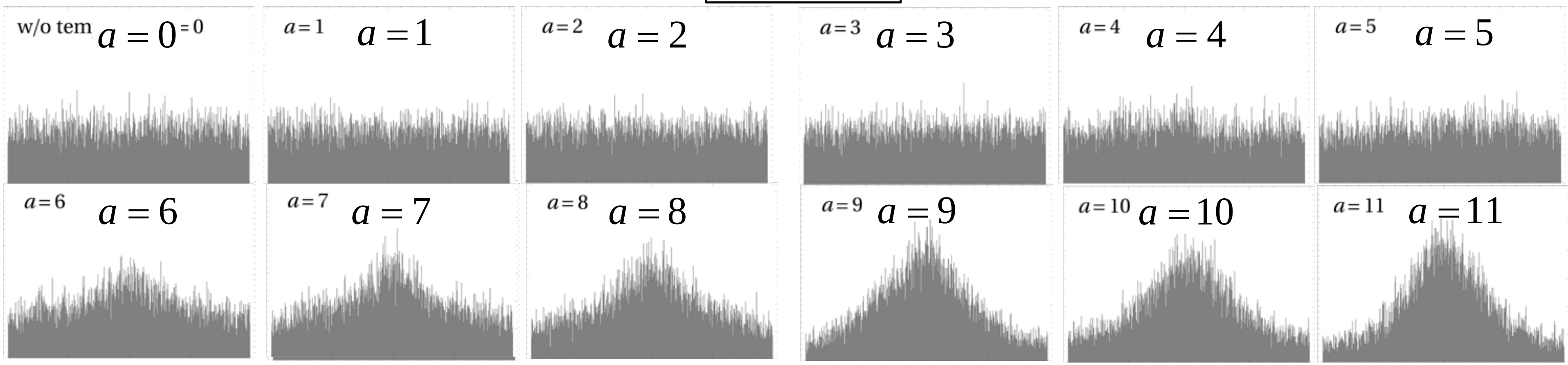
Histogram of $\theta_{t_a} \in [-\pi, \pi)$

w/ temp



many peaks (may not be so obvious because there are so many peaks and the peaks are broadened by Jacobian)

w/o temp



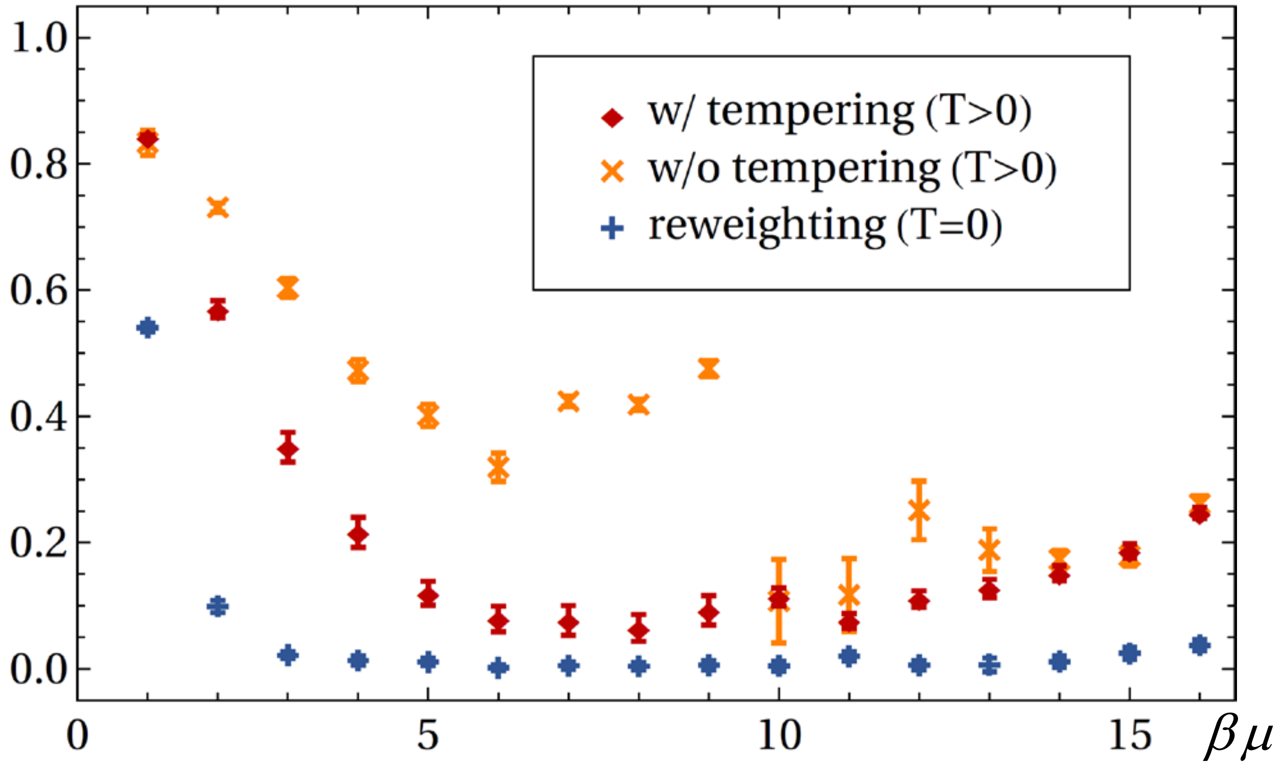
unimodal distribution

Results for 2D lattice (5/5)

[MF-Matsumoto-Umeda 1906.04243]

<u>sign average</u>

$$\left| \langle e^{i\theta_T(x)} \rangle_{S_T^{\text{eff}}} \right| \quad \left(\langle \mathcal{O}(x) \rangle = \frac{\langle e^{i\theta_T(x)}\, \mathcal{O}(z_T(x)) \rangle_{S_T^{\text{eff}}}}{\langle e^{i\theta_T(x)} \rangle_{S_T^{\text{eff}}}} \right)$$



When only a single (or very few) thimble(s) is sampled,
the sign average can become larger than that in the correct sampling
due to the absence of phase mixtures among thimbles

⇒ It is generally dangerous to regard the sign average as an index of the "resolution of the sign problem"

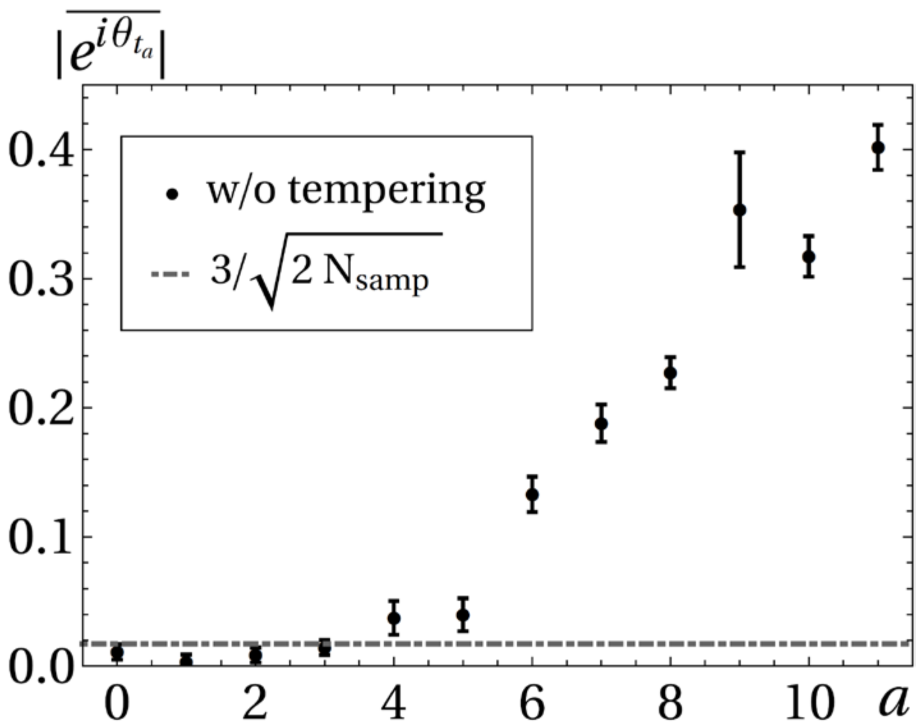
Comment on the Generalized LTM

[MF-Matsumoto-Umeda 1906.04243]

imaginary time : 5 steps $(N_\tau = 5)$
spatial lattice: 2D periodic lattice with $N_s = 2\times 2$
$\beta\kappa = 3,\quad \beta U = 13,\quad 0 \le T \le 0.4 (\Leftrightarrow 0 \le a \le 10)$
sample size: 5,000~25,000 depending on $\beta\mu$

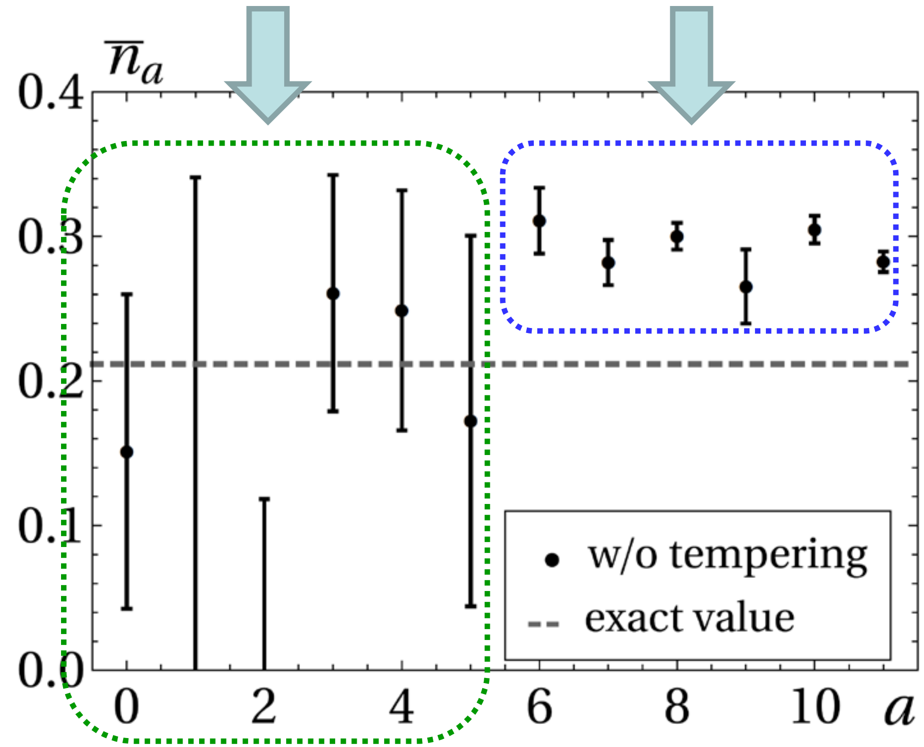
$$\left(\langle n \rangle = \frac{\langle e^{i\theta_{t_a}(x)}\, n(z_{t_a}(x)) \rangle_{S^{\mathrm{eff}}_{t_a}}}{\langle e^{i\theta_{t_a}(x)} \rangle_{S^{\mathrm{eff}}_{t_a}}} \approx \overline{n}_a \right)$$

Example: $\beta\mu = 5$



large stat errors (due to sign problem)

wrong values (due to multimodality)



It is a hard task to find an intermediate flow time that solves both sign problem and multimodality

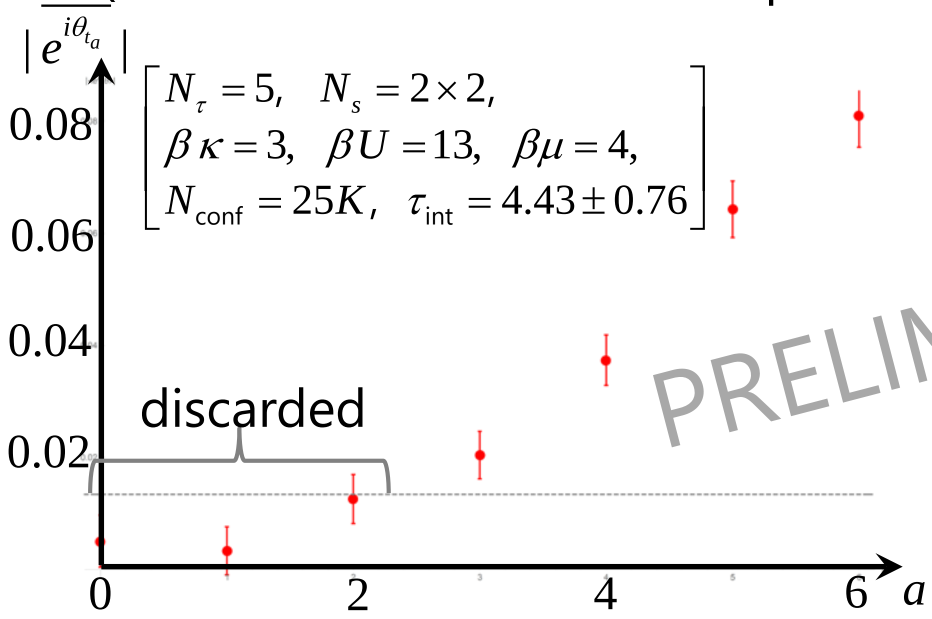
5. Some on-going work

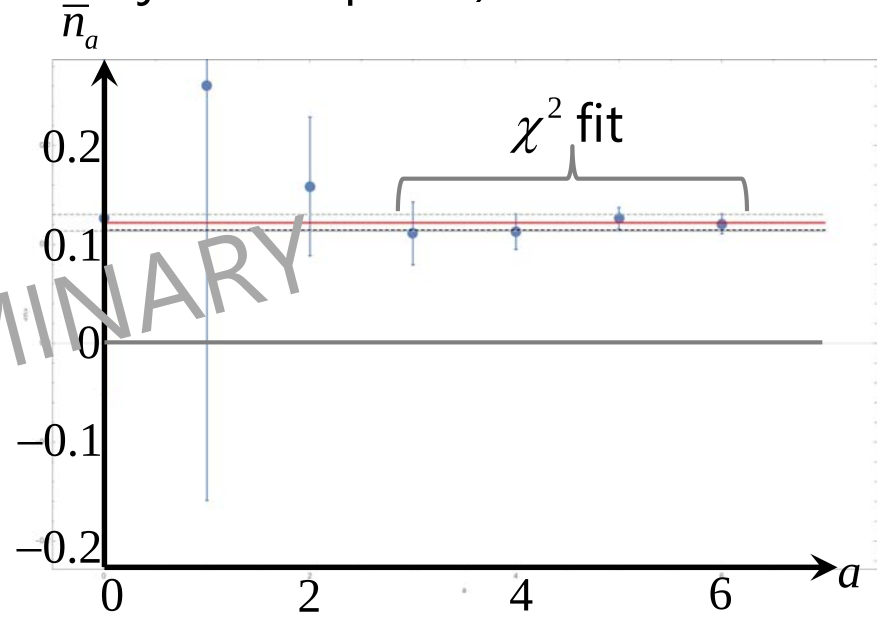
[MF-Matsumoto-Umeda, in preparation]

Some on-going work (1/2)

Implementation of HMC on the TLTM: **[MF-Matsumoto-Umeda, in prep]**

- We implemented the HMC algorithm for transitions at each replica
  [cf. Fujii-Honda-Kato-Kikukawa-Komatsu-Sano 2013, Alexandru in Lattice 2019]
  (our crucial improvement: handling of configs near det zeros + tempering)
- Computational cost gets much reduced with short autocorrelation times
  (at least a few times faster than the Metropolis even for small N)
- We no longer need to tune parameters that required long-term test runs
  (such as determining the variance of the proposal distribution)
- Good features in the TLTM are all preserved
  (such as the sufficient overlaps between adjacent replicas)





Some on-going work (2/2)

Application of TLTM to Stephanov models (chiral matrix models):

Dirac operator $D \Rightarrow 2N \times 2N$ dense complex matrix $D = \begin{pmatrix} m1_N & * \\ * & m1_N \end{pmatrix}$

- It has been known that the CLM does not work for this model even for small N (Gauge cooling is not applicable for this model)
- Multi Lefschetz thimbles again become relevant around critical points
- GLTM gives wrong results or large ambiguities for some parameter region
- TLTM seems to work for all the region of parameters (T, μ, m) , producing numerical results that agree with exact values $(N = 4, 8, 12, \ldots)$

6. Conclusion and outlook

Conclusion and outlook

What we have done:

- We proposed the tempered Lefschetz thimble method (TLTM) as a versatile method towards solving the numerical sign problem
- We further developed it and found an algorithm for a precise estimation with a criterion ensuring global equilibrium and the sample size (the key: $\overline{\mathcal{O}}_a$ should not depend on replica a due to Cauchy's theorem)
- GLTM can easily give incorrect results or large ambiguities
- TLTM works for the Hubbard model and gives correct results, avoiding both the sign and ergodicity problems simultaneously

Outlook: **[MF-Matsumoto, work in progress]**

- Investigate the Hubbard model of larger temporal and spatial sizes to understand the phase structure [computational cost: $O(N^{3\sim4})$]
- More generally, apply the TLTM to the following three typical subjects:
  ① Finite density QCD
  ② Quantum Monte Carlo (incl. the Hubbard model)
  ③ Real time QM/QFT
- Develop a more efficient algorithm with less computational cost (e.g. HMC at each replica **[MF-Matsumoto-Umeda, in prep]**)

Thank you.