Sign problem and the tempered Lefschetz thimble method

Masafumi Fukuma (Dept Phys, Kyoto Univ)

August 5, 2020 @ APLAT 2020

Based on work with

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

-- **MF** and **Umeda**, "Parallel tempering algorithm for integration over Lefschetz thimbles" [PTEP2017(2017)073B01, arXiv:1703.00861]
-- **MF**, **Matsumoto** and **Umeda**, "Applying the tempered Lefschetz thimble method to the Hubbard model away from half filling", [PRD100(2019)114510, arXiv:1906.04243]
-- **MF**, **Matsumoto** and **Umeda**, "Implementation of the HMC algorithm on the tempered Lefschetz thimble method", [arXiv:1912.13303]
-- **MF** and **Matsumoto**, some on-going work

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

-- **MF**, **Matsumoto** and **Umeda**,
[JHEP1712(2017)001, arXiv:1705.06097], [JHEP1811(2018)060, arXiv:1806.10915], [2004.00975]

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
(such as the Hubbard model)
③ θ vacuum with finite θ
④ Real time QM/QFT

Today, I would like to show that [MF-Umeda, PTEP2017(2017)073B01, arXiv:1703.00861]
a new algorithm "Tempered Lefschetz Thimble Method" (TLTM)
may be a promising method towards solving the sign problem,
by exemplifying its effectiveness for various models

- (0+1)-dim massive Thirring model [MF-Umeda, arXiv:1703.00861]
- 1-dim and 2-dim Hubbard model [MF-Matsumoto-Umeda, arXiv:1906.04243]
- chiral random matrix model (Stephanov model) [MF-Matsumoto-Umeda, in preparation]

The last part (application to Stephanov model) will be discussed
in Matsumoto's talk (next talk) with a refinement of the algorithm
[MF-Matsumoto-Umeda, arXiv:1912.13303]

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 $\{x^{(k)}\}_{k=1,\dots,N_{\text{conf}}}$ from $p_{\text{eq}}(x)$ $\left(N_{\text{conf}}\ \text{: sample size}\right)$

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

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

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

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 $p_{\mathrm{eq}}(x)$ $\big(N_{\mathrm{conf}}$: sample size$\big)$

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

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 :

Sign problem

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

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

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 $p_{\text{eq}}(x)$ (N_{conf} : sample size)

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

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

Sign problem

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

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

Markov chain Monte Carlo (MCMC) simulation:

probability distribution function

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

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

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

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

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

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

Sign problem

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

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

Markov chain Monte Carlo (MCMC) simulation:

probability distribution function

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

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

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

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

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

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

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

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, ...] [Fujii et al. 2013, ...] [Alexandru et al. 2015, ...]

(3) Others (tensor network, path-optimization, quantum computation, ...)
[Kuramashi, Takeda, Kadoh, ...] [Kashiwa-Mori-Ohnishi, Alexandru et al, ...] [Chakraborty-Honda-Izubuchi-Kikuchi-Tomiya, Kharzeev-Kikuchi, ...]

Advantages/disadvantages:

(1) CLM
Pros: fast $\propto O(N)$ $(N: \mathrm{DOF})$
Cons: "wrong convergence problem" [Ambjørn-Yang 1985, Aarts et al. 2011, Nagata-Nishimura-Shimasaki 2016]
(giving incorrect values with small errors)

(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, ...]

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

Pros: Works well even when multiple thimbles are relevant
Cons: Expensive $\propto O(N^{3\sim 4})$ ⇐ Jacobian determinant + tempering

Plan

1. Introduction (done)
2. Tempered Lefschetz thimble method (TLTM)
3. Applying TLTM to various models
4. Conclusion and outlook

2. Tempered Lefschetz thimble method (TLTM)

[MF-Umeda PTEP2017(2017)073B01, 1703.00861]
[MF-Matsumoto-Umeda PRD100(2019)114510, 1906.04243]
[MF-Matsumoto-Umeda 1912.13303]

Basic idea in Lefschetz thimble methods

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

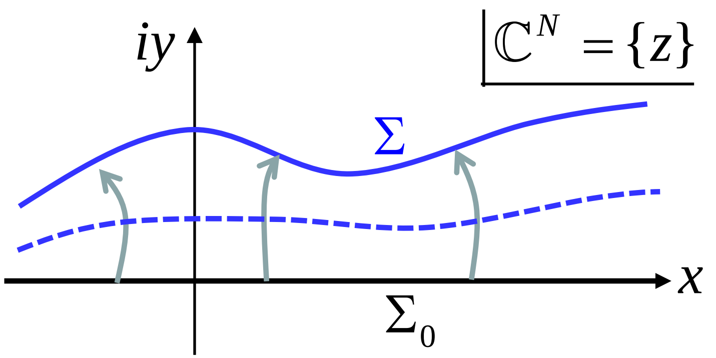
Complexify the variable: $x=(x^i)\in\mathbb{R}^N \;\Rightarrow\; z=(z^i=x^i+iy^i)\in\mathbb{C}^N$

Assumption: $e^{-S(z)}$, $e^{-S(z)}\mathcal{O}(z)$: 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) :

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



⇧ severe sign problem

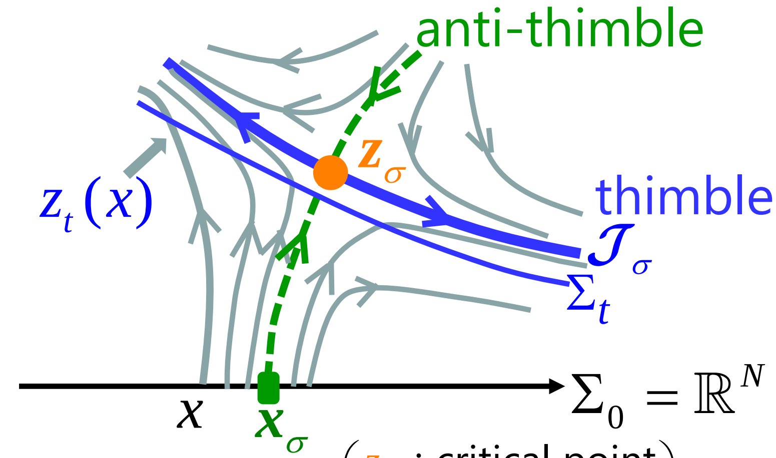
⇧ sign problem will get much reduced if $\mathrm{Im}\,S(z)$ is almost constant on Σ

Construction of Σ

<u>Antiholomorphic gradient flow</u>:

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

⇨ $\Sigma_t \equiv z_t(\mathbb{R}^N)$



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

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

⇨ $[\mathrm{Re}\, S(z_t)]^\bullet \geq 0$: real part always increases along the flow
$[\mathrm{Im}\, S(z_t)]^\bullet = 0$: 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)

Comments on Lefschetz thimble method

Common misunderstanding on Lefschetz thimble methods:

"The method eventually will encounter the sign problem for large DOF because it is based on the reweighting..."

But this is NOT true !

- On the original surface $\Sigma_0 = \mathbb{R}^N$ (flow time $t=0$)

$$\langle \mathcal{O}(x) \rangle \equiv \frac{\langle e^{-iS_I(x)} \mathcal{O}(x) \rangle_{\Sigma_0}}{\langle e^{-iS_I(x)} \rangle_{\Sigma_0}} \approx \frac{e^{-O(N)} \pm O(1/\sqrt{N_{\text{conf}}})}{e^{-O(N)} \pm O(1/\sqrt{N_{\text{conf}}})}$$

One needs a very large N_{conf}

$N_{\text{conf}} \simeq e^{O(N)}$

(sign problem)

- On a flowed surface Σ_T (flow time $t = T$)

$$\langle \mathcal{O}(x) \rangle \equiv \frac{\langle e^{i\theta(z)} \mathcal{O}(z) \rangle_{\Sigma_T}}{\langle e^{i\theta(z)} \rangle_{\Sigma_T}} \approx \frac{e^{-e^{-\lambda T} O(N)} \pm O(1/\sqrt{N_{\text{conf}}})}{e^{-e^{-\lambda T} O(N)} \pm O(1/\sqrt{N_{\text{conf}}})}$$

[λ : eigenvalue of $\partial_i \partial_j S(z_\sigma)$]

[Set $T = O(\log N)$]

$$= \frac{O(1) \pm O(1/\sqrt{N_{\text{conf}}})}{O(1) \pm O(1/\sqrt{N_{\text{conf}}})}$$

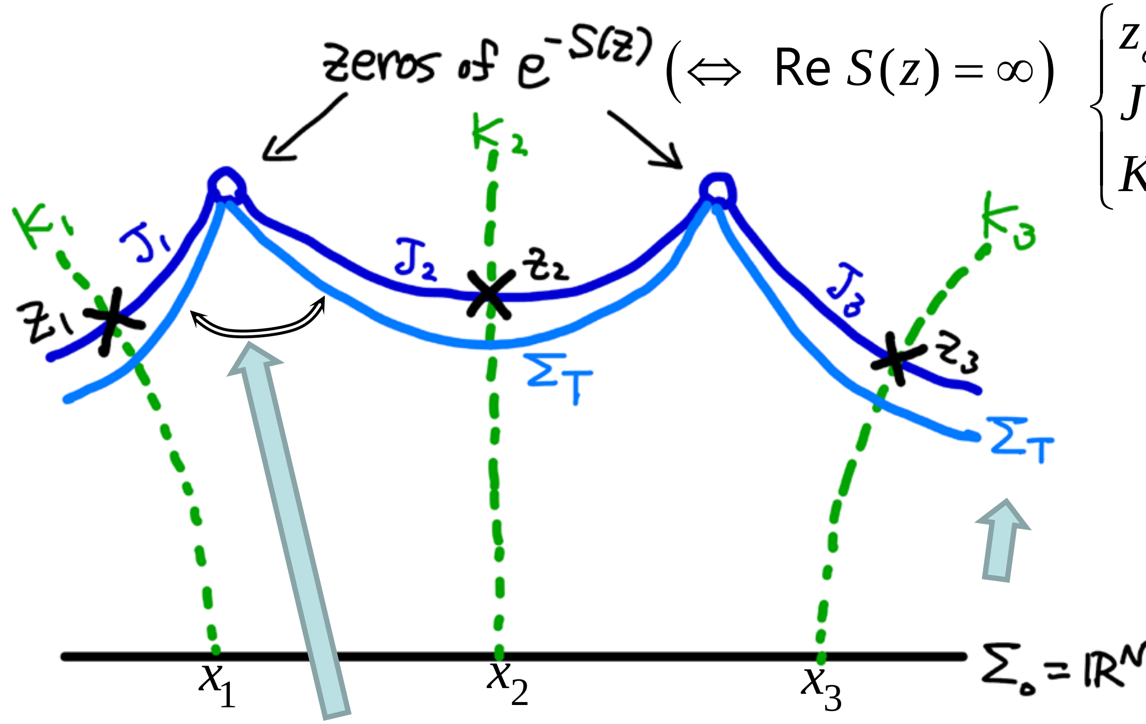
No longer needs that large N_{conf}

Sign problem is expected to disappear at flow time $T = O(\log N)$

Ergodicity problem in Lefschetz thimble methods

Flow time T needs to be large enough to solve the sign problem $\left(T = O(\ln N)\right)$.
However, this introduces a new problem, "ergodicity (multimodal) problem".

$\begin{cases} z_\sigma : \text{ critical pt} \\ J_\sigma : \text{ thimble} \\ K_\sigma : \text{ anti-thimble} \end{cases}$



transitions among regions separated by zeros
become indefinitely difficult as $t = T$ increases

Dilemma between the sign problem and the ergodicity problem

(for small T) (for large T)

Tempered Lefschetz thimble method (TLTM) (1/3)

[MF-Umeda 1703.00861]
[MF-Matsumoto-Umeda 1906.04243, 1912.13303]

In order to solve the dilemma between the sign problem (for small T) and the ergodicity problem (for large T), we implement the parallel tempering (= replica exchange MCMC) method.

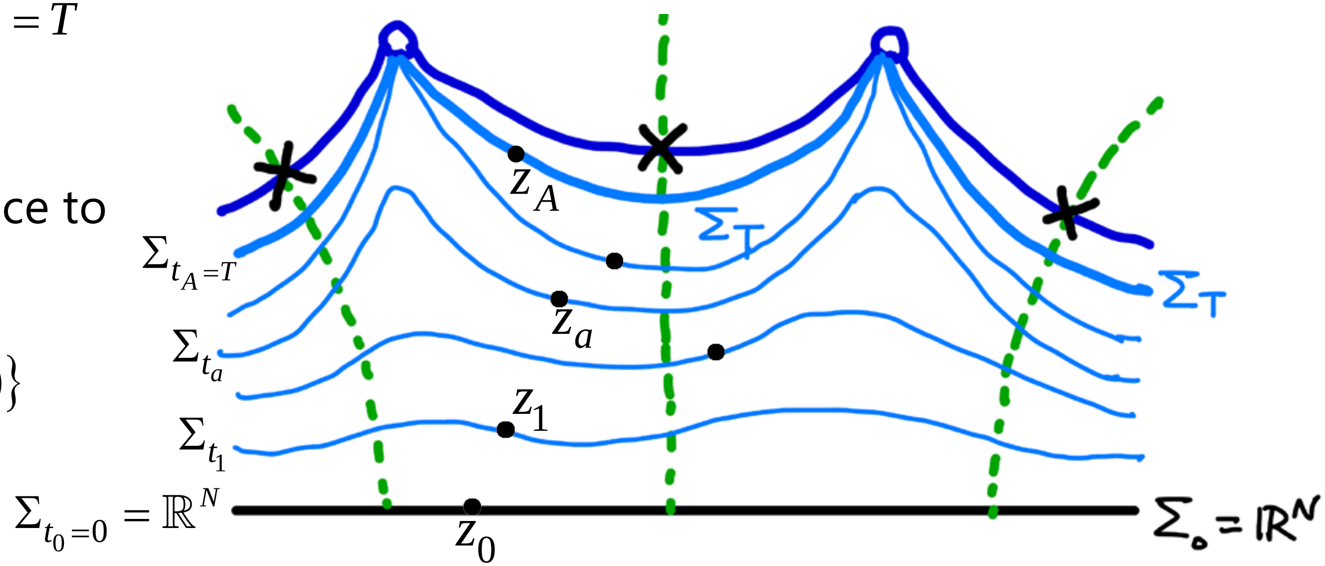
[Swendsen-Wang 1986, Geyer 1991, Hukushima-Nemoto 1996]

⇓

(1) Introduce a tempering parameter set $\{t_a\}$ $(a = 0, 1, \ldots, A)$
with $t_0 = 0 < t_1 < \cdots < t_A = T$

(2) Extend the config space to

$$\Sigma_{\text{tot}} = \Sigma_{t_0} \times \Sigma_{t_1} \times \cdots \times \Sigma_{t_A} = \left\{ \vec{z} = (z_0, z_1, \ldots, z_A) \right\}$$



(3) Construct a Markov chain $\vec{z}^{(k)} \to \vec{z}^{(k+1)}$ s.t. it gives the equilib distribution:

$$p_{\text{eq}}(\vec{z}) \prod_a |dz_a| \propto \prod_a e^{-\text{Re}\, S(z_a)} |dz_a|$$

Tempered Lefschetz thimble method (TLTM) (1/3)

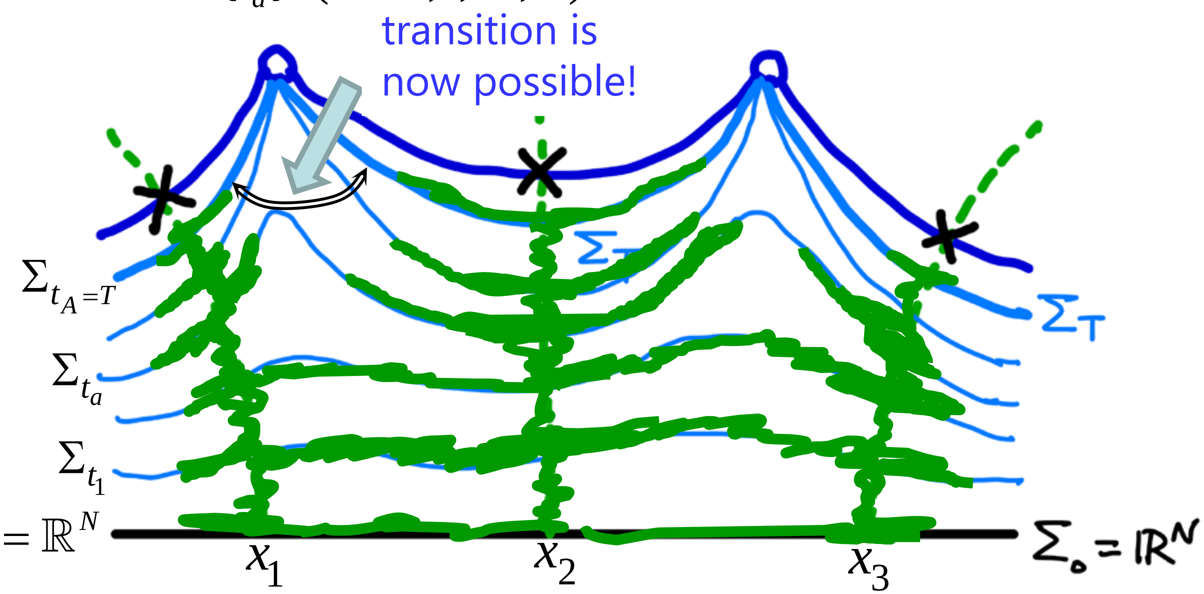
[MF-Umeda 1703.00861]
[MF-Matsumoto-Umeda 1906.04243, 1912.13303]

In order to solve the dilemma between the sign problem (for small t) and the ergodicity problem (for large t), we implement the parallel tempering (= replica exchange MCMC) method.
[Swendsen-Wang 1986, Geyer 1991, Hukushima-Nemoto 1996]

(1) Introduce a tempering parameter set $\{t_a\}$ $(a=0,1,\ldots,A)$
with $t_0 = 0 < t_1 < \cdots < t_A = T$

(2) Extend the config space to

$$\Sigma_{\text{tot}} = \Sigma_{t_0} \times \Sigma_{t_1} \times \cdots \times \Sigma_{t_A} = \left\{ \vec{z} = (z_0, z_1, \ldots, z_A) \right\}$$



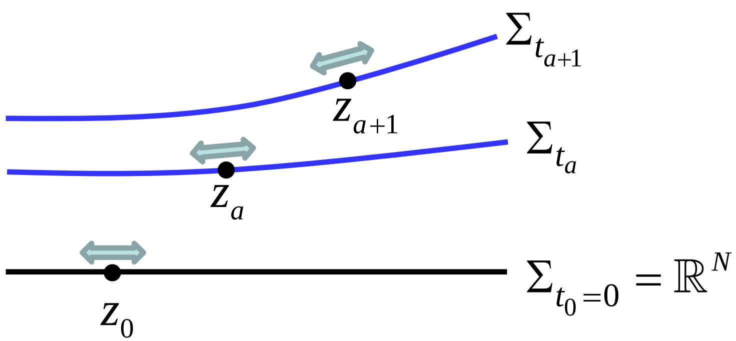
(3) Construct a Markov chain $\vec{z}^{(k)} \to \vec{z}^{(k+1)}$ s.t. it gives the equilib distribution:

$$p_{\text{eq}}(\vec{z}) \prod_a |dz_a| \propto \prod_a e^{-\text{Re}\, S(z_a)} |dz_a|$$

Tempered Lefschetz thimble method (TLTM) (2/3)

The Markov chain consists of the following two processes:

(A) Transitions on Σ_{t_a} $(a = 0, 1, \ldots, A)$

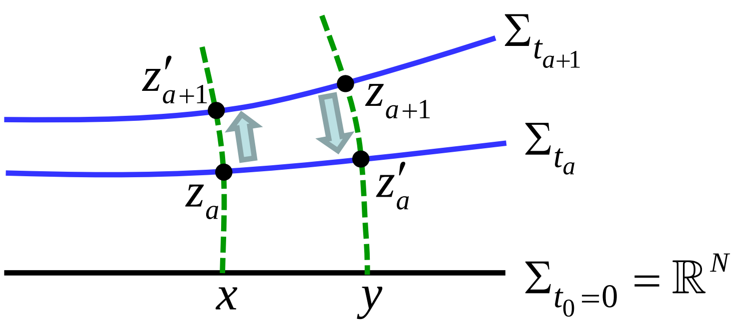


$$(z_0, \cdots, z_a, z_{a+1}, \cdots, z_A)$$
$$\Downarrow \quad \Downarrow \quad \Downarrow \quad \Downarrow$$
$$(z_0', \cdots, z_a', z_{a+1}', \cdots, z_A')$$

This can be realized by
- Metropolis **[MF-Umeda 1703.00861, MF-Matsumoto-Umeda 1906.04243]**
- HMC **[MF-Matsumoto-Umeda 1912.13303]**

⇨ Matsumoto's talk (next talk)

(B) Swapping of configs between adjacent replicas



$$(z_0, \cdots, z_a, z_{a+1}, \cdots, z_A)$$
$$\times$$
$$(z_0', \cdots, z_a', z_{a+1}', \cdots, z_A')$$

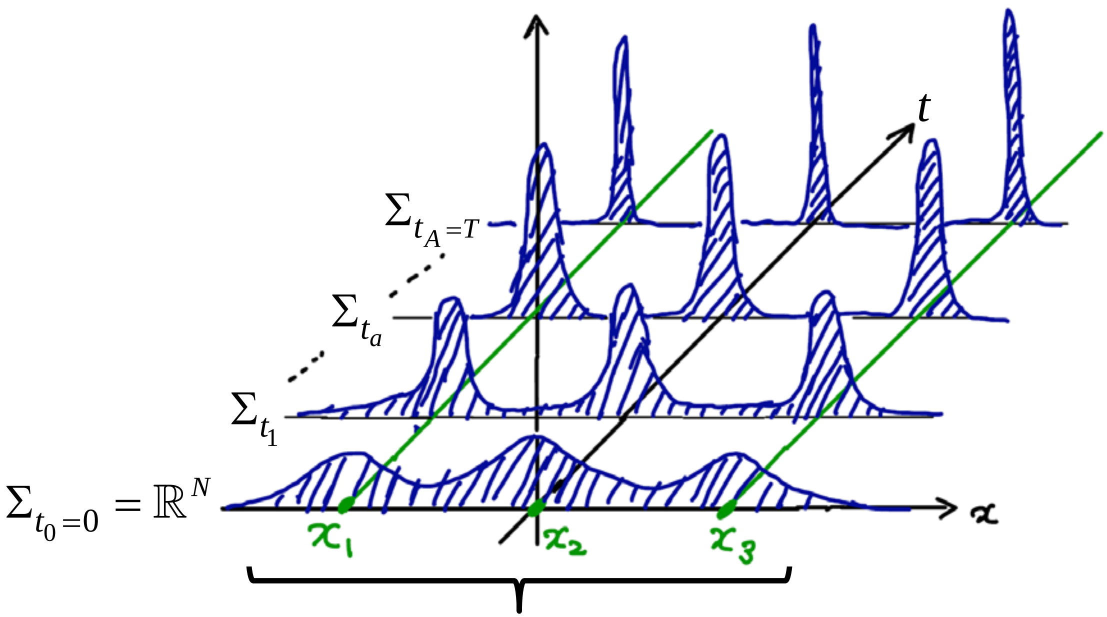
(This is actually an exchange of initial configs $x \leftrightarrow y$ for $z_a = z_{t_a}(x)$ and $z_{a+1} = z_{t_{a+1}}(y)$)

Tempered Lefschetz thimble method (TLTM) (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**

Analysis

[MF-Matsumoto-Umeda 1906.04243]

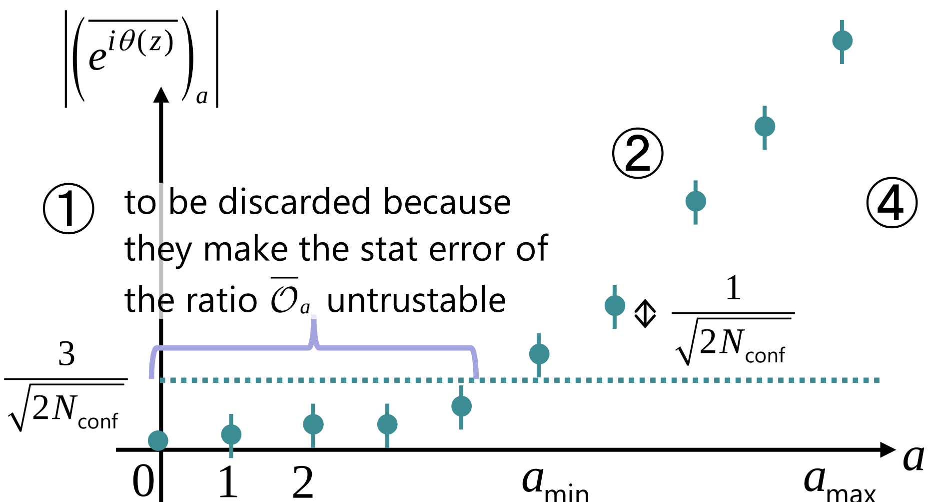
Consider the estimates of $\langle \mathcal{O} \rangle$ at various flow times t_a:

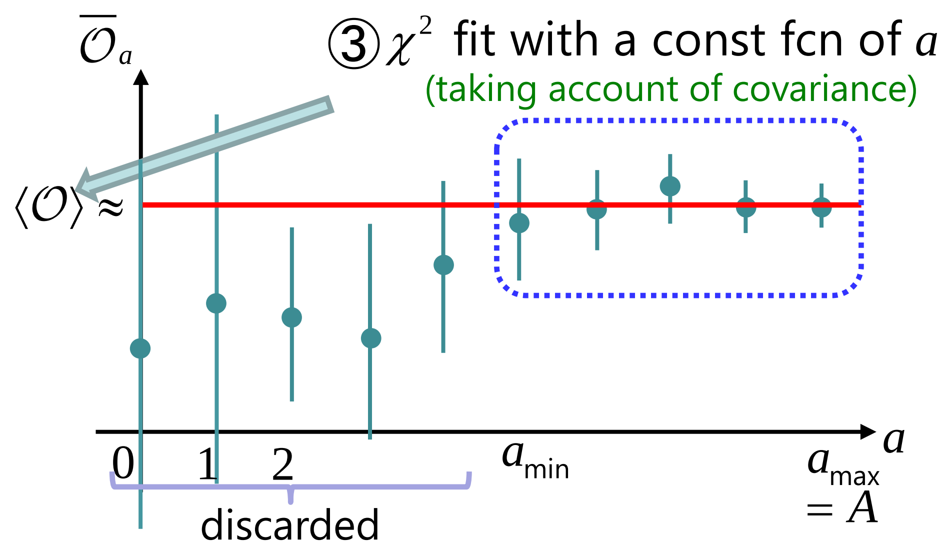
$$\langle \mathcal{O} \rangle = \frac{\langle e^{i\theta(z)} \mathcal{O}(z) \rangle_{\Sigma_{t_a}}}{\langle e^{i\theta(z)} \rangle_{\Sigma_{t_a}}} \approx \frac{(1/N_{\text{conf}}) \sum_{k=1}^{N_{\text{conf}}} e^{i\theta(z_a^{(k)})} \mathcal{O}(z_a^{(k)})}{(1/N_{\text{conf}}) \sum_{k=1}^{N_{\text{conf}}} e^{i\theta(z_a^{(k)})}} = \frac{\left(\overline{e^{i\theta(z)} \mathcal{O}(z)}\right)_a}{\left(\overline{e^{i\theta(z)}}\right)_a} \equiv \overline{\mathcal{O}}_a \quad (a = 0, 1, \dots, A)$$

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!





3. Applying TLTM to various models

[MF-Matsumoto-Umeda, work in progress]

The models to which TLTM has been applied

- (0+1)-dim massive Thirring model
  [MF-Umeda, PTEP2017(2017)073B01, arXiv:1703.00861]
- 1-dim & 2-dim Hubbard model
  [MF-Matsumoto-Umeda, PRD100(2019)114510, arXiv:1906.04243]
  [MF-Matsumoto-Umeda, arXiv:1912.13303]
- chiral random matrix model (Stephanov model)
  [MF-Matsumoto-Umeda, in preparation]

The following projects are also in progress:

- 2-dim frustrated spin systems (classical & quantum)
- θ-vacuum of 2-dim & 4-dim pure Yang-Mills with finite θ

(For the geometrical optimization of tempering algorithms,
see: **[MF-Matsumoto, JHEP1712(2017)001, arXiv:1705.06097,
JHEP1811(2018)060, arXiv:1806.10915]**)

The models to which TLTM has been applied

- (0+1)-dim massive Thirring model
  [MF-Umeda, PTEP2017(2017)073B01, arXiv:1703.00861]
- 1-dim & 2-dim Hubbard model
  [MF-Matsumoto-Umeda, PRD100(2019)114510, arXiv:1906.04243]
  [MF-Matsumoto-Umeda, arXiv:1912.13303]
- chiral random matrix model (Stephanov model)
  [MF-Matsumoto-Umeda, in preparation]
  ⇒ Matsumoto's talk (next talk)

The following projects are also in progress:

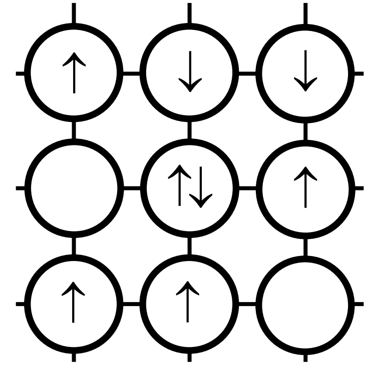
- 2-dim frustrated spin systems (classical & quantum)
- θ-vacuum of 2-dim & 4-dim pure Yang-Mills with finite θ

(For the geometrical optimization of tempering algorithms,
see: **[MF-Matsumoto, JHEP1712(2017)001, arXiv:1705.06097,
JHEP1811(2018)060, arXiv:1806.10915]**)

Hubbard model

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

(fermion bilinear) (four fermion)



$(N_s$: # of sites$)$

$$\left\{\begin{array}{l} n_{\mathbf{x},\sigma}\equiv c^{\dagger}_{\mathbf{x},\sigma}c_{\mathbf{x},\sigma} \\ \kappa(>0)\ :\ \text{hopping parameter} \\ \mu\ :\ \text{chemical potential} \\ U(>0)\ :\ \text{strength of on-site replusive potential} \end{array}\right\}$$

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

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

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

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

(For half filling $(\mu=0)$: $\det M_a[\phi]\det M_b[\phi]=\left|\det M_a[\phi]\right|^2\geq 0\ \Rightarrow$ No sign problem)

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

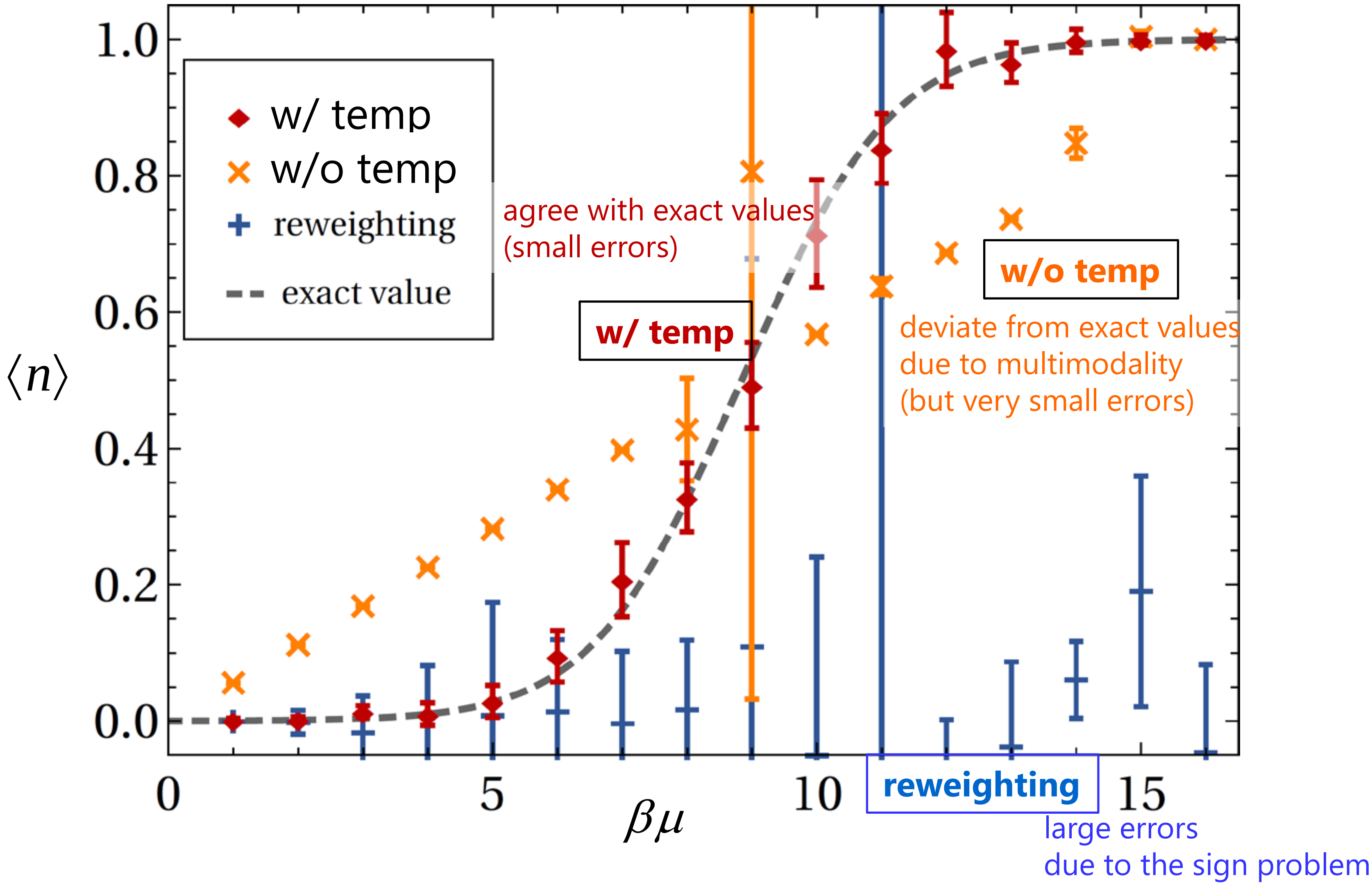
[MF-Matsumoto-Umeda 1906.04243]

Results for 1D lattice (1/2)

[MF-Matsumoto-Umeda 2019]

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

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

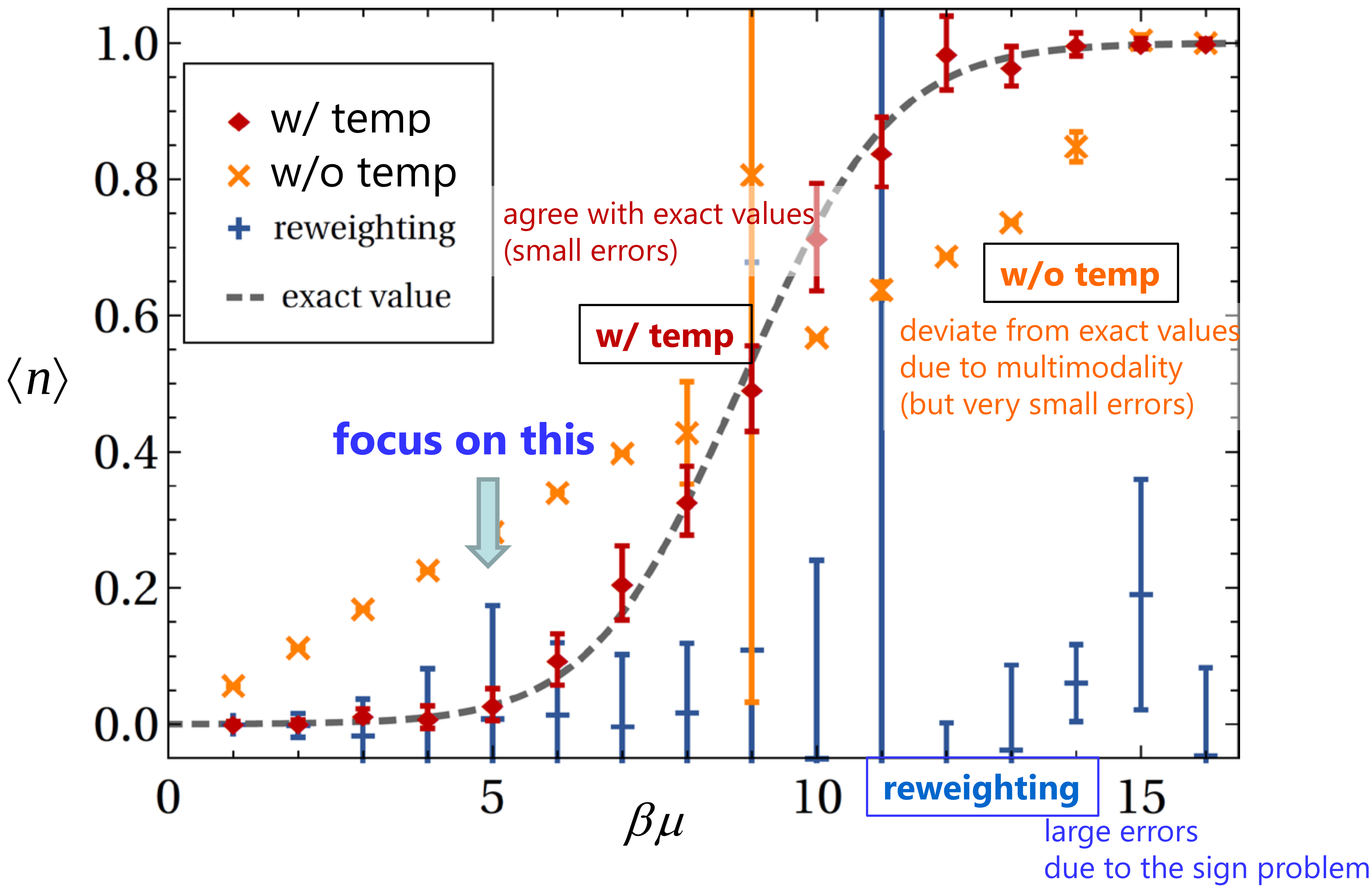


Results for 1D lattice (1/2)

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

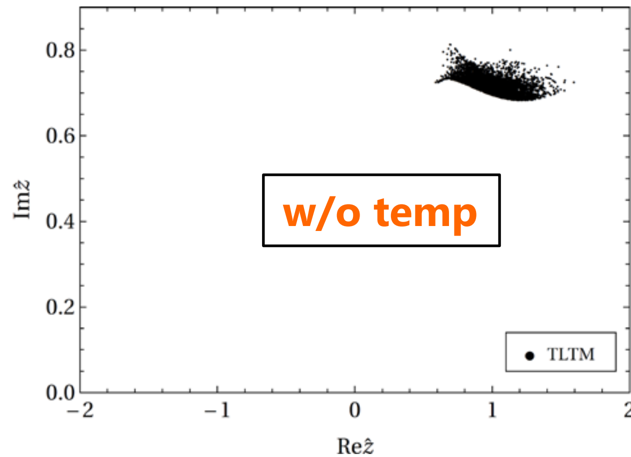


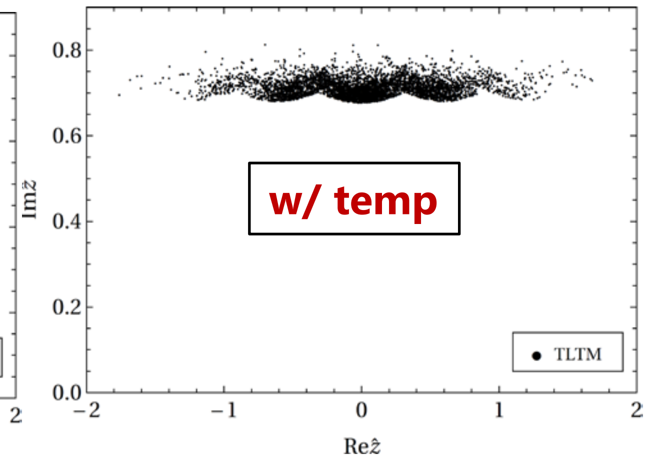
Results for 1D lattice (2/2)

[MF-Matsumoto-Umeda 2019]

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

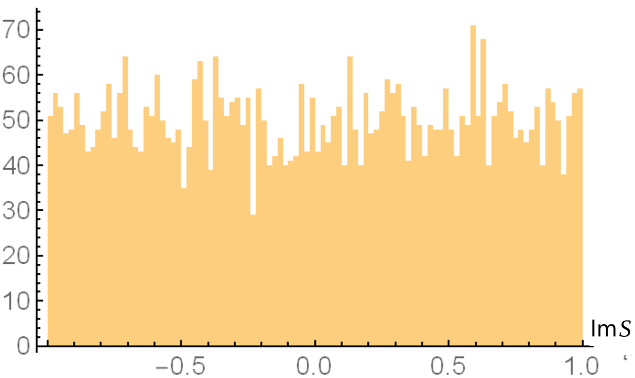
(projected on a plane)





Histogram of ImS(z)/π

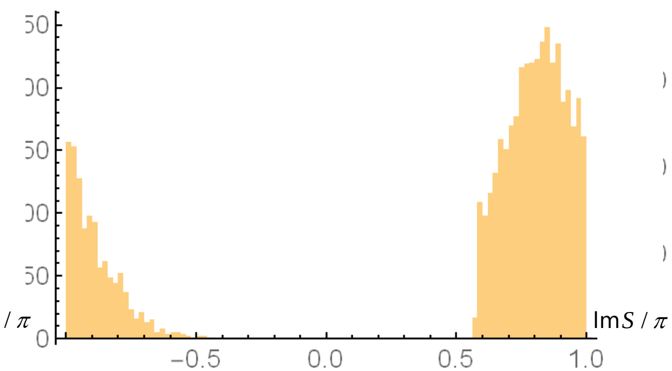
reweighting



distributing uniformly from −π to +π

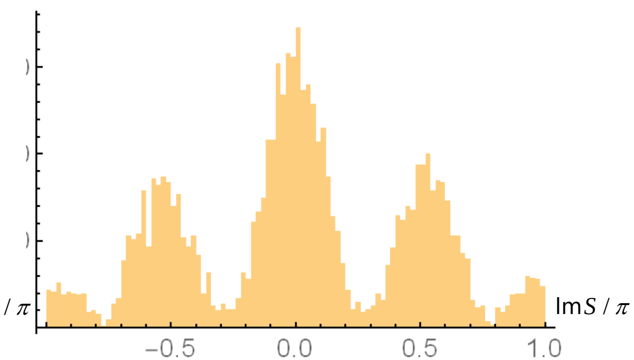
⇨ severe sign problem

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)

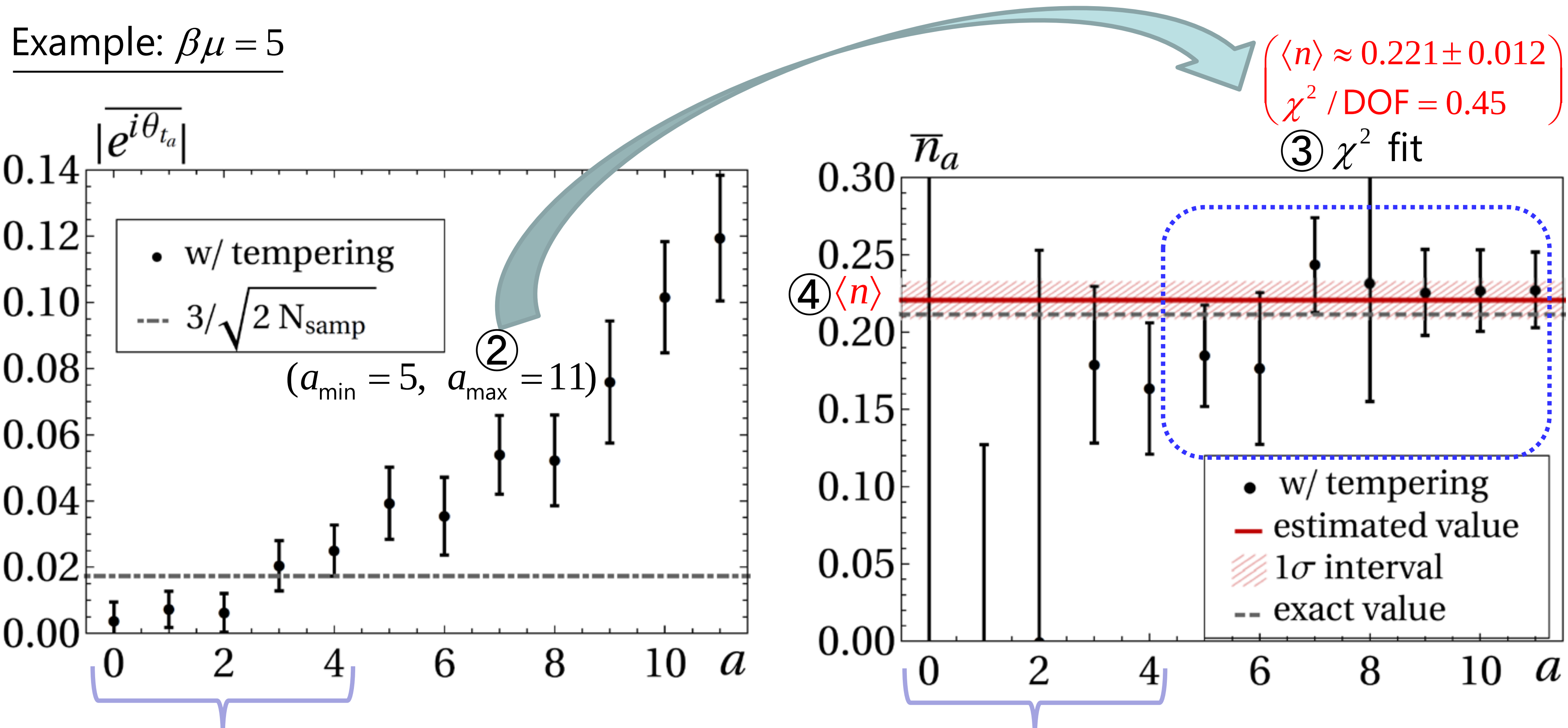
Results for 2D lattice (0/2)

[MF-Matsumoto-Umeda 1906.04243]

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

$$\left(\langle n \rangle = \frac{\langle e^{i\theta(z)} n(z) \rangle_{\Sigma_{t_a}}}{\langle e^{i\theta(z)} \rangle_{\Sigma_{t_a}}} \approx \overline{n}_a \right)$$

Example: $\beta\mu = 5$



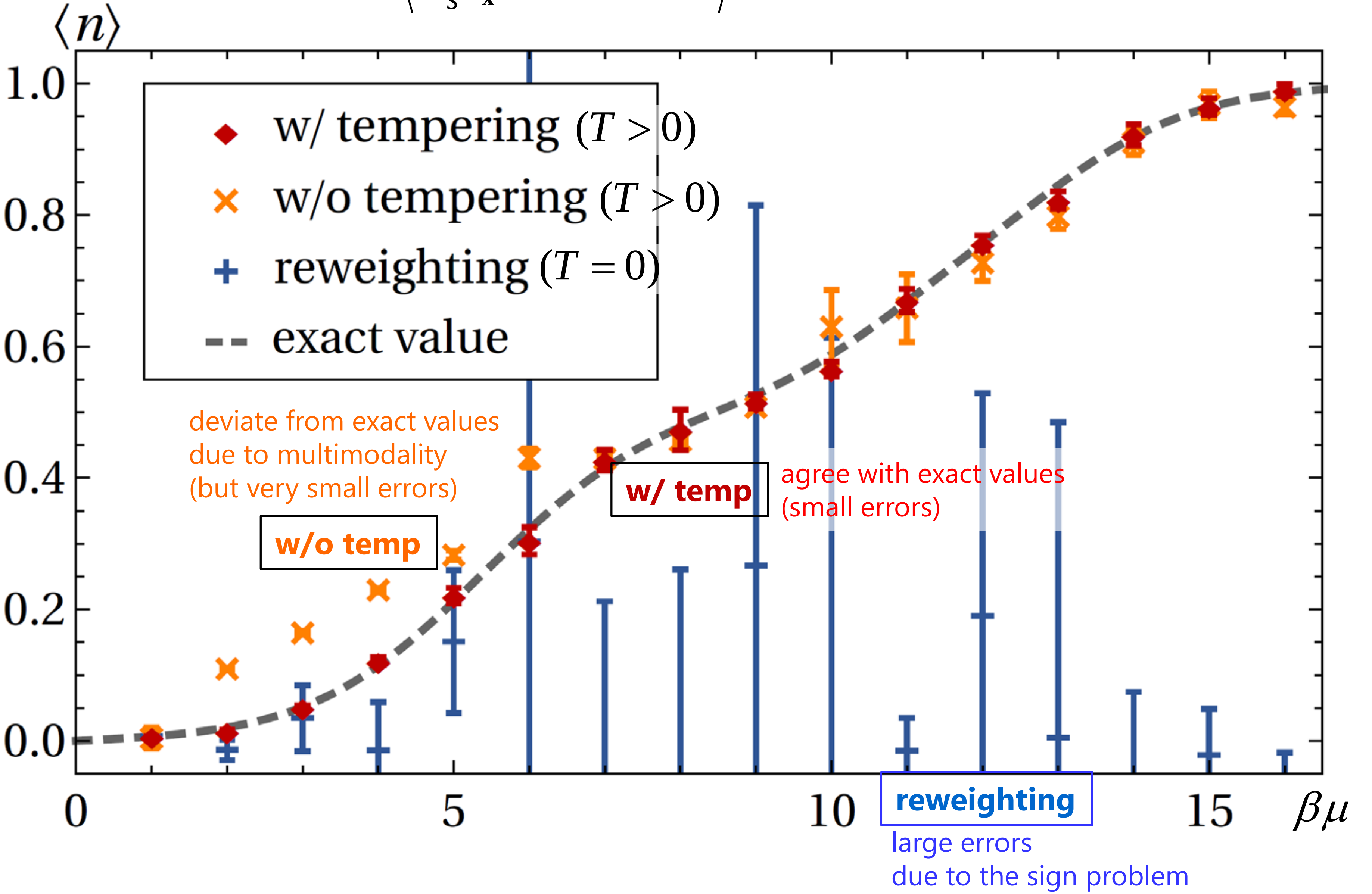
① discarded

discarded

Results for 2D lattice (1/2)

[MF-Matsumoto-Umeda 1906.04243]

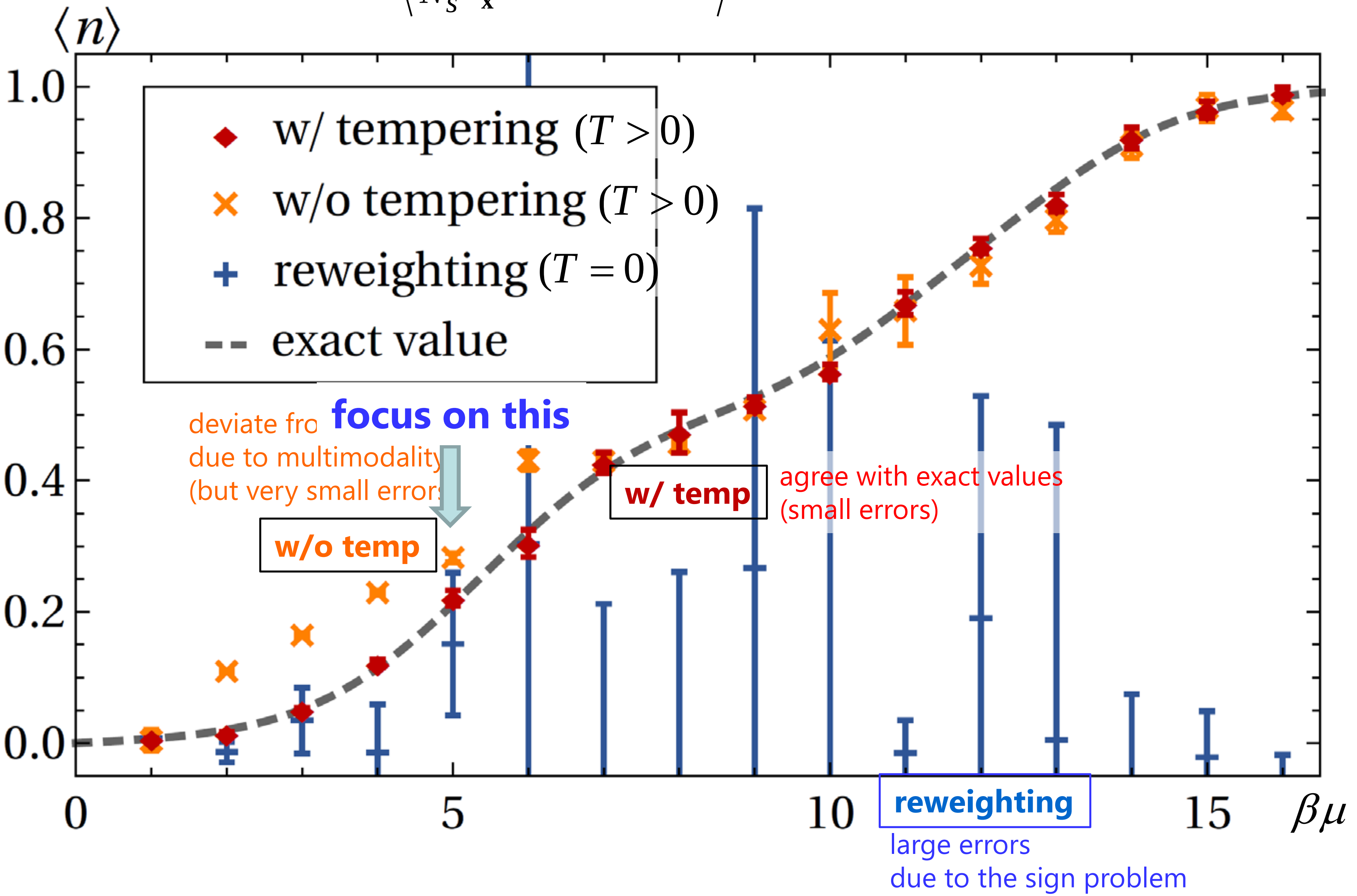
$$\begin{bmatrix} N_\tau = 5, N_S = 2\times 2 \\ \beta\kappa = 3, \quad \beta U = 13 \end{bmatrix} \qquad \langle n \rangle = \left\langle \frac{1}{N_S} \sum_{\mathbf{x}} \left(n_{\mathbf{x},\uparrow} + n_{\mathbf{x},\downarrow} - 1 \right) \right\rangle$$



Results for 2D lattice (1/2)

[MF-Matsumoto-Umeda 1906.04243]

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

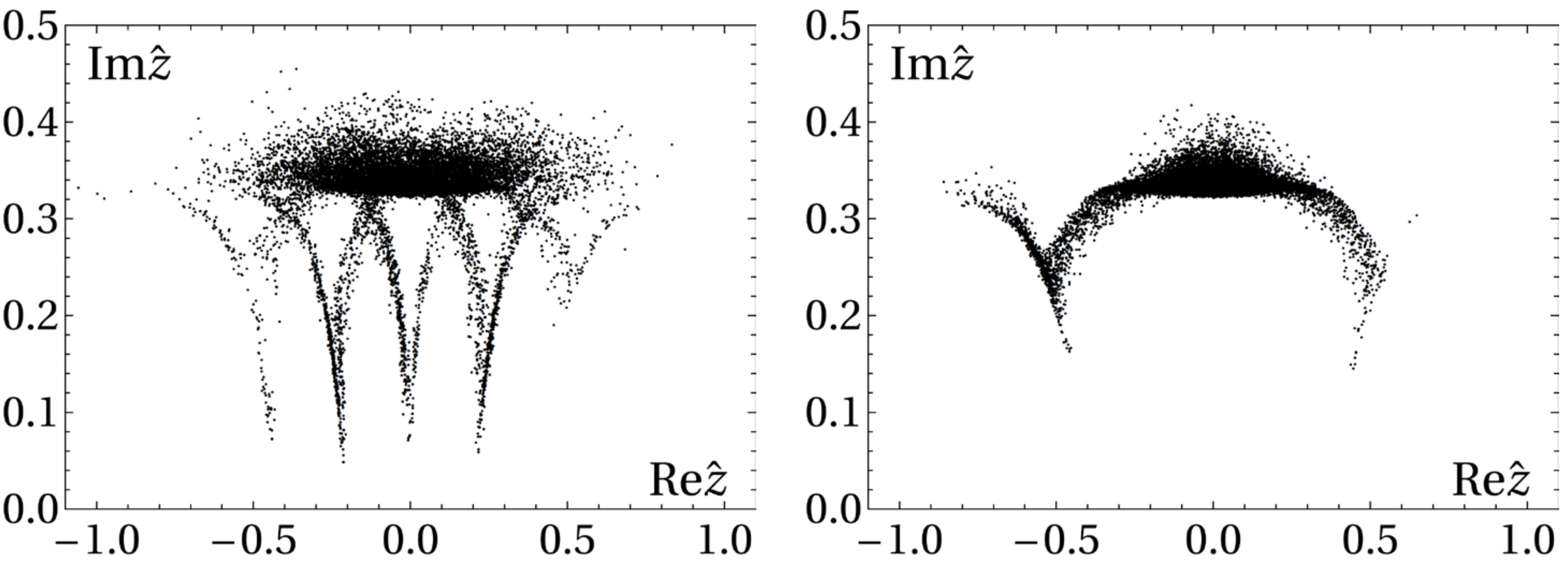


Results for 2D lattice (2/2)

[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

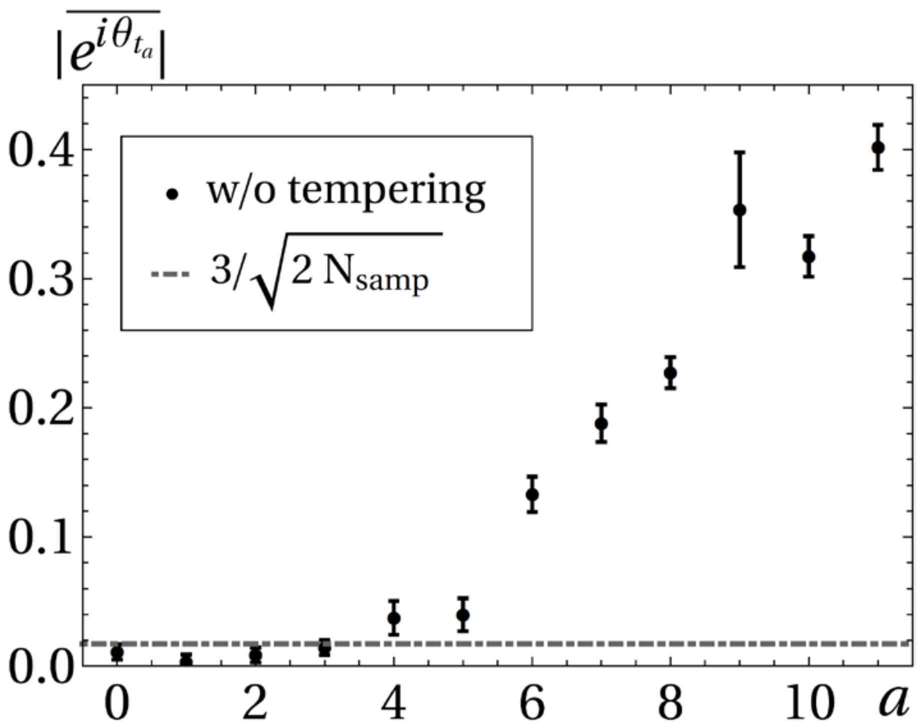
Comment on the Generalized LTM

[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 $\beta\mu$

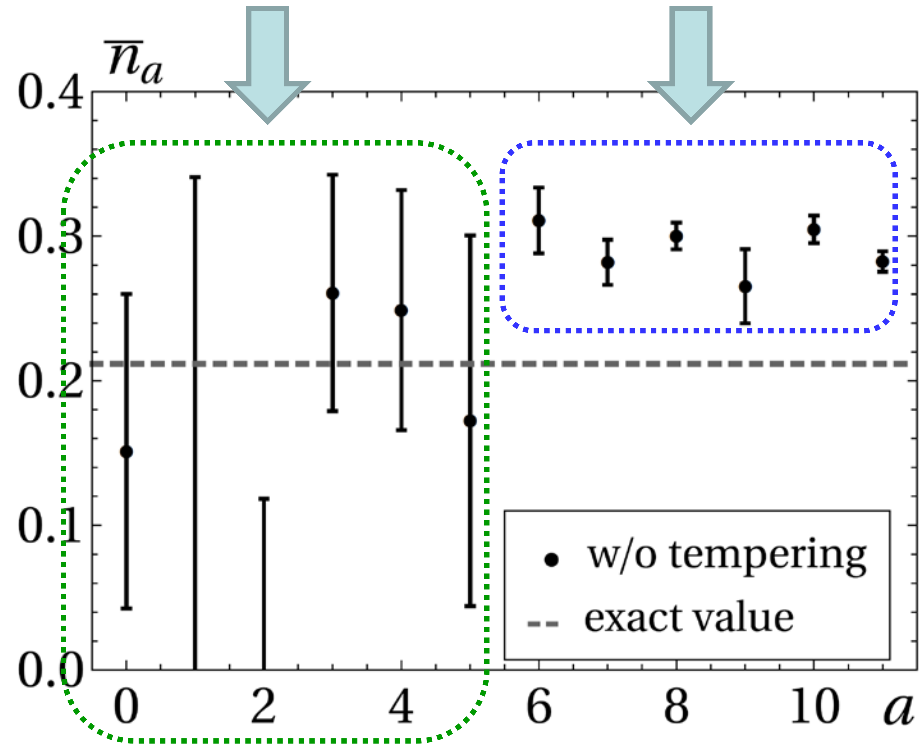
$$\left(\langle n \rangle = \frac{\langle e^{i\theta(z)} n(z) \rangle_{\Sigma_{t_a}}}{\langle e^{i\theta(z)} \rangle_{\Sigma_{t_a}}} \approx \overline{n}_a \right)$$

Example: $\beta\mu = 5$



large stat errors (due to sign problem)

wrong value (due to multimodality)



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

4. 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)
- TLTM works nicely in various models avoiding both the sign and ergodicity problems simultaneously

Outlook:

- Investigate the Stephanov model of larger sizes to understand the computational scaling [expected to be $O(\mathrm{DOF}^{3\sim 4})$]
- Apply TLTM to the following four typical subjects:
  ① Finite density QCD
  ② Quantum Monte Carlo
  ③ θ vacuum
  ④ Real time QFT
- Keep developing more efficient algorithms with less computational cost

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)
- TLTM works nicely in various models avoiding both the sign and ergodicity problems simultaneously

Outlook:

- Investigate the Stephanov model of larger sizes to understand the computational scaling [expected to be $O(\mathrm{DOF}^{3\sim 4})$]
- Apply TLTM to the following four typical subjects:
  ① Finite density QCD (Stephanov model ⇒ …) **[MF-Matsumoto, work in progress]**
  ② Quantum Monte Carlo (Hubbard model, frustrated spin systems ⇒ …)
  ③ θ vacuum (2 dim ⇒ 4 dim)
  ④ Real time QFT (QM ⇒ QFT)
- Keep developing more efficient algorithms with less computational cost

Thank you.