Numerical sign problem and the tempered Lefschetz thimble method

Masafumi Fukuma (Dept Phys, Kyoto Univ)

QCD phase diagram and lattice QCD

Oct 26, 2021 @YITP (online)

Based on work with

Nobuyuki Matsumoto, Yusuke Namekawa and Naoya Umeda
(RIKEN/BNL) (Dept Phys, Kyoto U) (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**, "Worldvolume approach to the tempered Lefschetz thimble method" [PTEP2021(2021)023B08, arXiv:2012.08468]
-- **MF**, **Matsumoto** and **Namekawa**, "Statistical analysis method for the Worldvolume Monte Carlo algorithm" [to appear in PTEP, arXiv:2107.06858]

1. Introduction

Overview

The **numerical sign problem** is one of the major obstacles
when performing first-principles 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 dynamics of quantum fields

Today, I would like to show that **[MF-Umeda, 1703.00861]**
a new algorithm "Tempered Lefschetz Thimble Method" (TLTM)
and its extension "Worldvolume-TLTM" (WV-TLTM) **[MF-Matsumoto, 2012.08468]**
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, 1703.00861]**
- 1-dim and 2-dim Hubbard model **[MF-Matsumoto-Umeda, 1906.04243]**
- chiral random matrix model (Stephanov model) **[MF-Matsumoto, 2012.08468]**

I also would like to discuss the computational scaling of WV-TLTM
[MF-Matsumoto-Namekawa, work in progress]

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

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

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 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 $p_{\mathrm{eq}}(x) \equiv e^{-S(x)} / \int dx\, e^{-S(x)}$ 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 $p_{\mathrm{eq}}(x)$ (N_{conf} : sample size)

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

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 $p_{\mathrm{eq}}(x) \equiv e^{-S(x)} / \int dx\, e^{-S(x)}$ 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 $p_{\mathrm{eq}}(x)$ (N_{conf} : sample size)

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

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

Sign problem

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

$$= \frac{\langle e^{-i S_I(x)} \mathcal{O}(x) \rangle_{S_R}}{\langle e^{-i S_I(x)} \rangle_{S_R}}$$

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

Markov chain Monte Carlo (MCMC) simulation:

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: **probability distribution function**

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

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

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

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

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

Sign problem

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

$= \dfrac{\langle e^{-i S_I(x)} \mathcal{O}(x) \rangle_{S_R}}{\langle e^{-i S_I(x)} \rangle_{S_R}}$

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

Markov chain Monte Carlo (MCMC) simulation:

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: (probability distribution function)

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

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

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

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

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

⇒ Require $O(1/\sqrt{N_{\text{conf}}}) \lesssim e^{-O(N)}$ ⇒ $N_{\text{conf}} \gtrsim e^{O(N)}$ 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 ($\beta \sim N$)

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

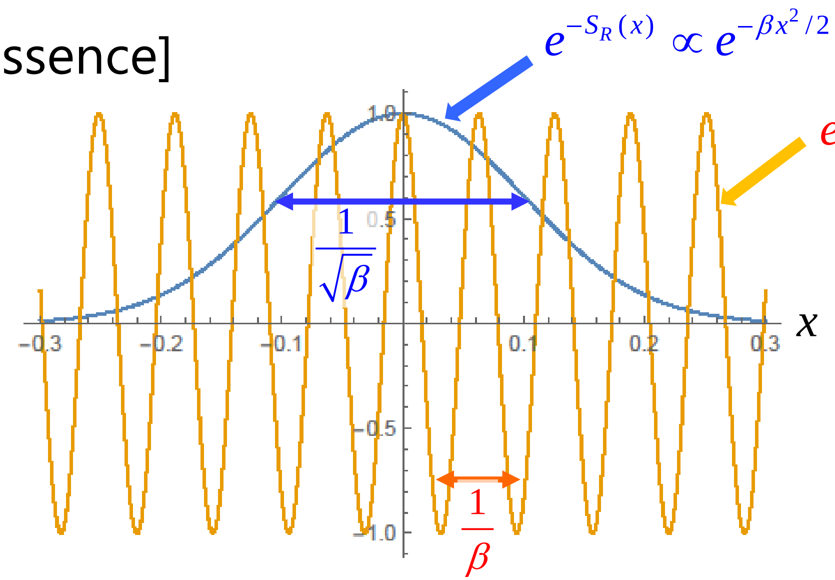
numerically $$\approx \frac{\left(\beta^{-1}-1\right)e^{-\beta/2} \pm O(1/\sqrt{N_{\rm conf}})}{e^{-\beta/2} \pm O(1/\sqrt{N_{\rm conf}})}$$

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

$\Rightarrow$ Necessary sample size:

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

[Essence]



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

Various approaches

■ **method 1: no use of reweighting**

▼ complex Langevin method **[Parisi 1983, Klauder 1983, Aarts et al. 2009]**

(may show a wrong convergence problem) ($\Leftarrow$ wrong results w/ small stat errors)

■ **method 2: deformation of the integration surface**

▼ Lefschetz thimble method **[Witten 2010, Cristoforetti et al. 2012, Fujii et al. 2013, Alexandru et al. 2015]**

⇩

Tempered Lefschetz thimble method (TLTM) **[MF-Umeda 2017] [MF-Umeda-Matsumoto 2019]**

⇩

worldvolume TLTM (WV-TLTM) **[MF-Matsumoto 2020]**

▼ path optimization method (POM) **[Mori-Kashiwa-Ohnishi 2017, Alexandru et al. 2018]**

■ **method 3: no use of MC in the first place**

▼ tensor network **[Levin-Nave 2007, ...]**

- good at calculating the free energy
- but not so much for correl fcns
- complementary to MC approach?

Various approaches

■ **method 1: no use of reweighting**

▼ complex Langevin method **[Parisi 1983, Klauder 1983, Aarts et al. 2009]**

(may show a wrong convergence problem) (⇐ wrong results w/ small stat errors)

■ **method 2: deformation of the integration surface**

▼ Lefschetz thimble method **[Witten 2010, Cristoforetti et al. 2012, Fujii et al. 2013, Alexandru et al. 2015]**

⇓

Tempered Lefschetz thimble method (TLTM) **[MF-Umeda 2017] [MF-Umeda-Matsumoto 2019]**

⇓

worldvolume TLTM (WV-TLTM) **[MF-Matsumoto 2020]**

▼ path optimization method (POM) **[Mori-Kashiwa-Ohnishi 2017, Alexandru et al. 2018]**

■ **method 3: no use of MC in the first place**

▼ tensor network **[Levin-Nave 2007, ...]**

- good at calculating the free energy
- but not so much for correl fcns
- complementary to MC approach?

Plan

1. Introduction (done)
2. Lefschetz thimble method
3. Tempered Lefschetz thimble method (TLTM)
   and its worldvolume extension (WV-TLTM)
4. Application to various models
   - Hubbard model
   - Chiral random matrix model (Stephanov model)
   - Computational scaling of WV-TLTM
5. Summary and outlook

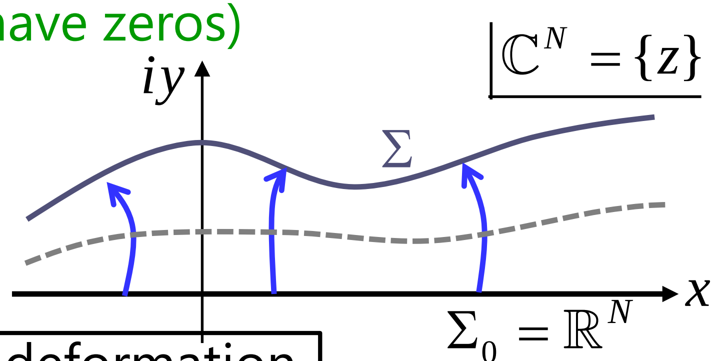
2. Lefschetz thimble method

Basic idea of the thimble method (1/2)

■ complexification of dyn variable: $x=(x^i)\in\mathbb{R}^N \;\Rightarrow\; z=(z^i=x^i+iy^i)\in\mathbb{C}^N$

assumption (satisfied for most cases) $\big(S(x)$: action, $\mathcal{O}(x)$: observable$\big)$

$e^{-S(z)}$, $e^{-S(z)}\mathcal{O}(z)$: entire fcns over $\mathbb{C}^N$ (can have zeros)



Cauchy's theorem

Integrals do not change under continuous deformation of integration surface : $\Sigma_0=\mathbb{R}^N \;\rightarrow\; \Sigma\;(\subset\mathbb{C}^N)$

(boundary at $|x|\to\infty$ 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

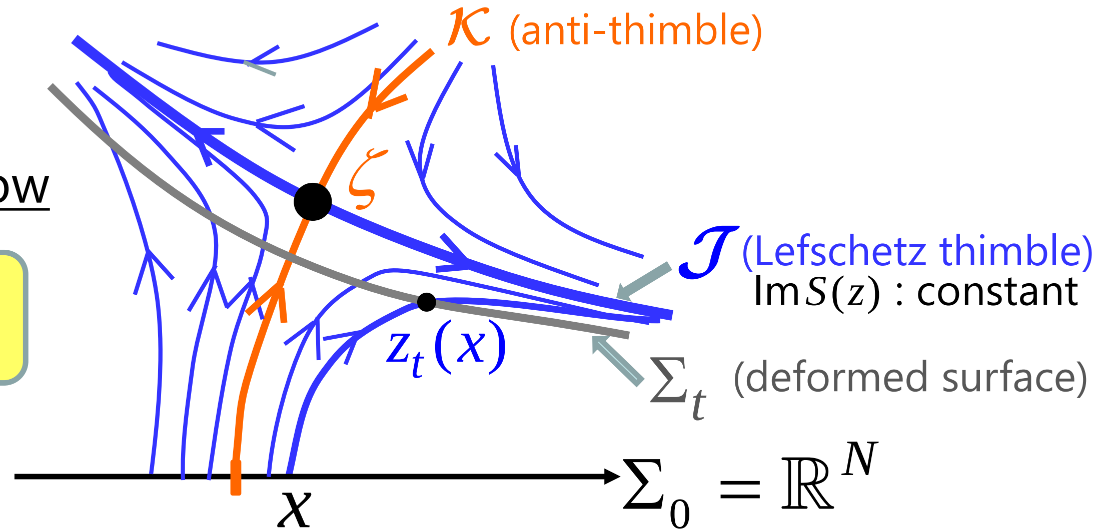
sign problem will be significantly reduced if $\mathrm{Im}\,S(z)$ is almost constant on Σ

Basic idea of the thimble method (2/2)

■ prescription for deformation

anti-holomorphic gradient flow

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



property

$$[S(z_t)]^{\cdot} = \partial S(z_t) \cdot \dot{z}_t = \left|\partial S(z_t)\right|^2 \geq 0$$

⇨ $[\mathrm{Re}\, S(z_t)]^{\cdot} \geq 0$: always increases except at crit pt ζ (ζ : crit pt $\Leftrightarrow \partial S(\zeta) = 0$)

$[\mathrm{Im}\, S(z_t)]^{\cdot} = 0$: always constant

⇨ $\Sigma_t \xrightarrow{t\to\infty} \mathcal{J}$ (Lefschetz thimble) $\equiv$ set of orbits starting from ζ

$\mathrm{Im}\, S(z)$: constant on $\mathcal{J}$ $(= \mathrm{Im}\, S(\zeta))$

⇨ **Sign problem is expected to disappear on Σ_t at a sufficiently large t**

How does the sign problem disappear?

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

$$\langle \mathcal{O}(x)\rangle \equiv \frac{\langle e^{-i\,\mathrm{Im}\,S(x)}\mathcal{O}(x)\rangle_{\Sigma_0(\mathrm{rewt})}}{\langle e^{-i\,\mathrm{Im}\,S(x)}\rangle_{\Sigma_0(\mathrm{rewt})}} \approx \frac{e^{-O(N)} \pm O(1/\sqrt{N_{\mathrm{conf}}})}{e^{-O(N)} \pm O(1/\sqrt{N_{\mathrm{conf}}})}$$

(N : DOF, N_{conf} : sample size)

⇒ need a huge size of sample : $N_{\mathrm{conf}} \simeq e^{O(N)}$ **sign problem**

flow

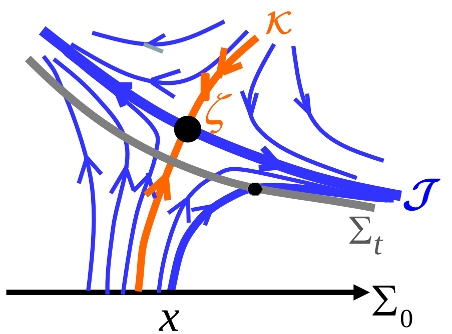
- Integration on a deformed surface Σ_t (flow time t)

$$\langle \mathcal{O}(x)\rangle = \frac{\langle e^{i\theta(z)}\mathcal{O}(z)\rangle_{\Sigma_t}}{\langle e^{i\theta(z)}\rangle_{\Sigma_t}} \approx \frac{e^{-e^{-\lambda t}O(N)} \pm O(1/\sqrt{N_{\mathrm{conf}}})}{e^{-e^{-\lambda t}O(N)} \pm O(1/\sqrt{N_{\mathrm{conf}}})}$$

$\left(e^{i\theta(z)} \equiv e^{-i\,\mathrm{Im}\,S(z)}\dfrac{dz}{|dz|} \right)$

$\left[e^{\lambda t} = O(N) \Leftrightarrow t = O(\log N) \right]$

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



[λ : (typical) singular value of Hessian $\partial_i\partial_j S(\zeta)$]

⇒ **Sign problem disappears at flow time $t = O(\log N)$**

Example: Gaussian revisited

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

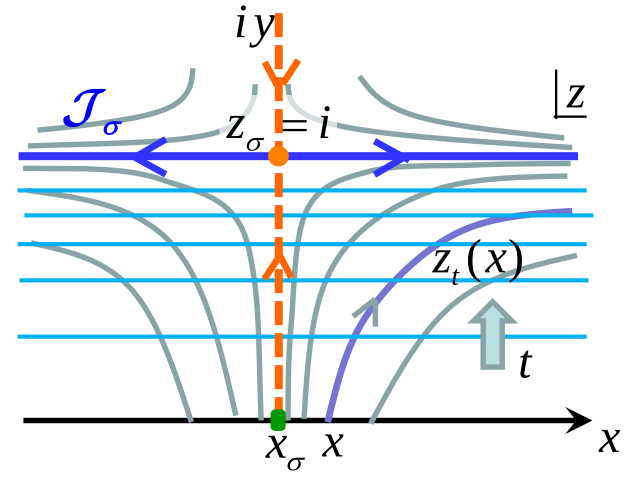
$\dot{z}_t = \overline{S'(z_t)} = \beta(\bar{z}+i)$ with $z_{t=0}=x$

⇨ $z_t(x) = x e^{\beta t} + i(1-e^{-\beta t}) \quad \therefore |dz| = e^{\beta t} dx$

exponential growth of coefficient

⇨ $\begin{cases} \mathrm{Re}\, S(z_t(x)) = \dfrac{1}{2}\beta e^{2\beta t}(x^2 - e^{-4\beta t}) - \beta t \\ \mathrm{Im}\, S(z_t(x)) = -\beta x \quad \therefore e^{i\theta(z_t(x))} = e^{i\beta x} \end{cases}$

⇨ $$\langle f(z)\rangle_{\Sigma_t} = \frac{\int_{\Sigma_t} |dz|\, e^{-\mathrm{Re}\, S(z)} f(z)}{\int_{\Sigma_t} |dz|\, e^{-\mathrm{Re}\, S(z)}} = \frac{\int dx\, e^{-\beta e^{2\beta t} x^2/2} f(z_t(x))}{\int dx\, e^{-\beta e^{2\beta t} x^2/2}}$$

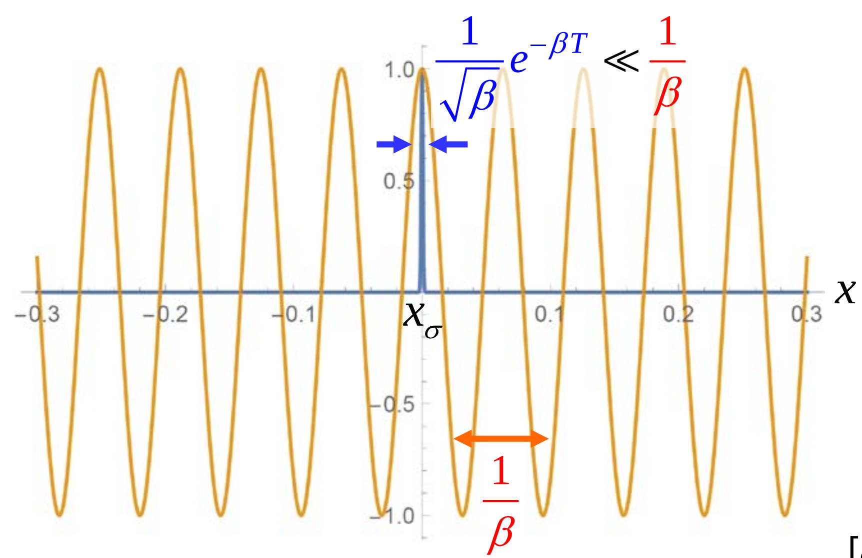


No small numbers appear

if we take a large t $(\equiv T)$ s.t. $e^{-\beta T} \ll \dfrac{1}{\sqrt{\beta}}$

$$\langle x^2 \rangle = \frac{\langle e^{i\theta(z)} z^2 \rangle_{\Sigma_T}}{\langle e^{i\theta(z)} \rangle_{\Sigma_T}}$$

$$= \frac{e^{-(\beta/2)e^{-2\beta T}}\left(\beta^{-1}-1\right)}{e^{-(\beta/2)e^{-2\beta T}}} = \frac{O(1)}{O(1)}$$



<u>NB</u>. Logarithmic increase is enough:

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

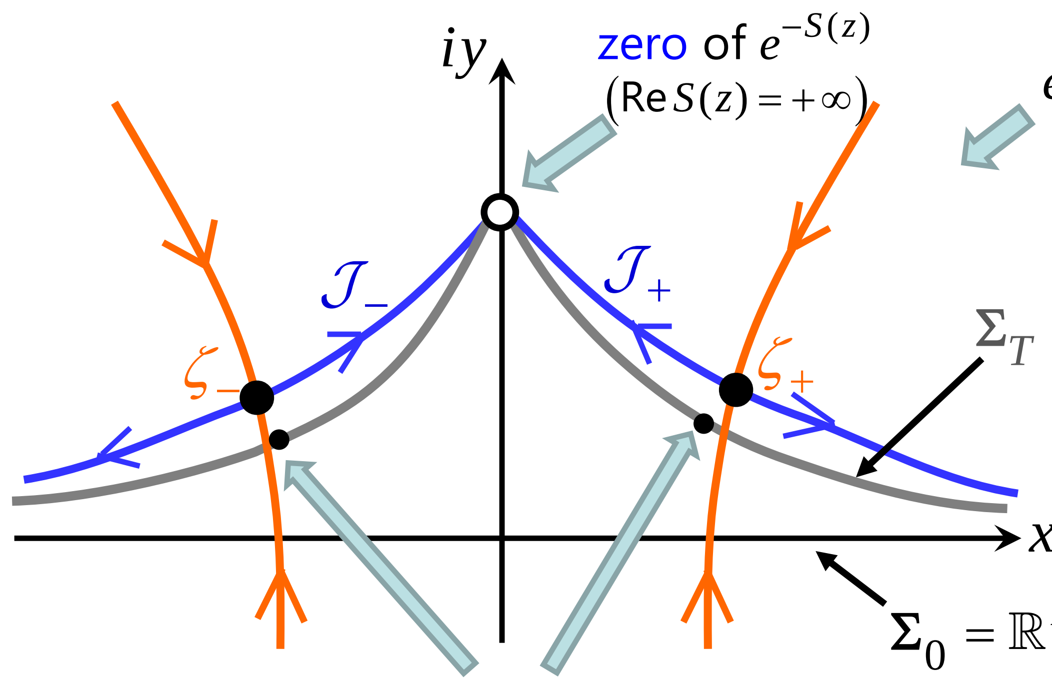
3. Tempered Lefschetz thimble method and its worldvolume extension (TLTM & WV-TLTM)

Ergodicity problem

[Fukuma-Umeda 1703.00861]

Sign problem resolved? **NO!**

Actually, there comes out another problem at large t : **Ergodicity problem**



[Marinari-Parisi 1992]
[Swendsen-Wang 1986, Geyer 1991
Hukushima-Nemoto 1996]

solution : **Implement the <u>tempering</u> to the thimble method**

[MF-Umeda 2017]

using the flow time as a temp param

Tempered Lefschetz thimble method (TLTM) (1/2)

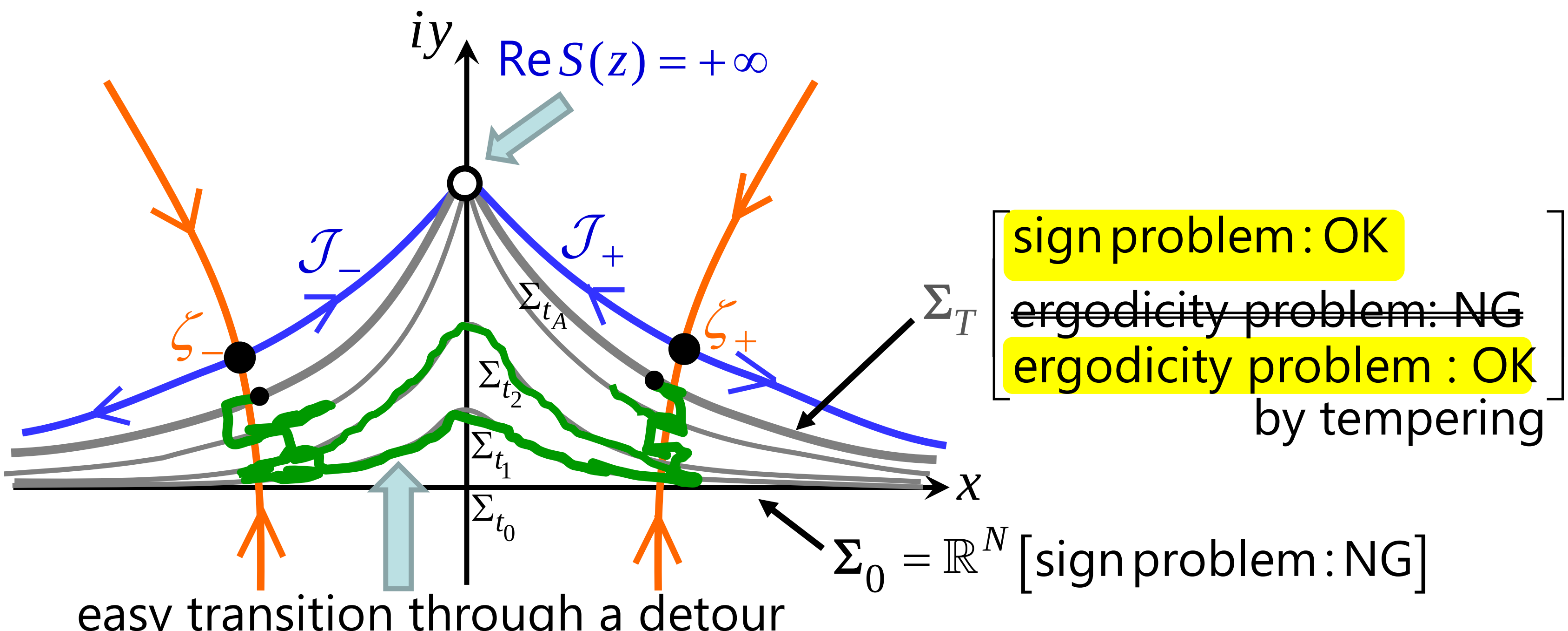
[Fukuma-Umeda 1703.00861]

■ TLTM

(1) Introduce replicas in between the initial integ surface $\Sigma_0 = \mathbb{R}^N$ and the target deformed surface Σ_T as $\left\{\Sigma_{t_0=0},\ \Sigma_{t_1},\ \Sigma_{t_2},\ \ldots,\ \Sigma_{t_A=T}\right\}$

(2) Setup a Markov chain for the extended config space $\{(x, t_a)\}$

(3) After equilibration, estimate observables with a subsample on Σ_T



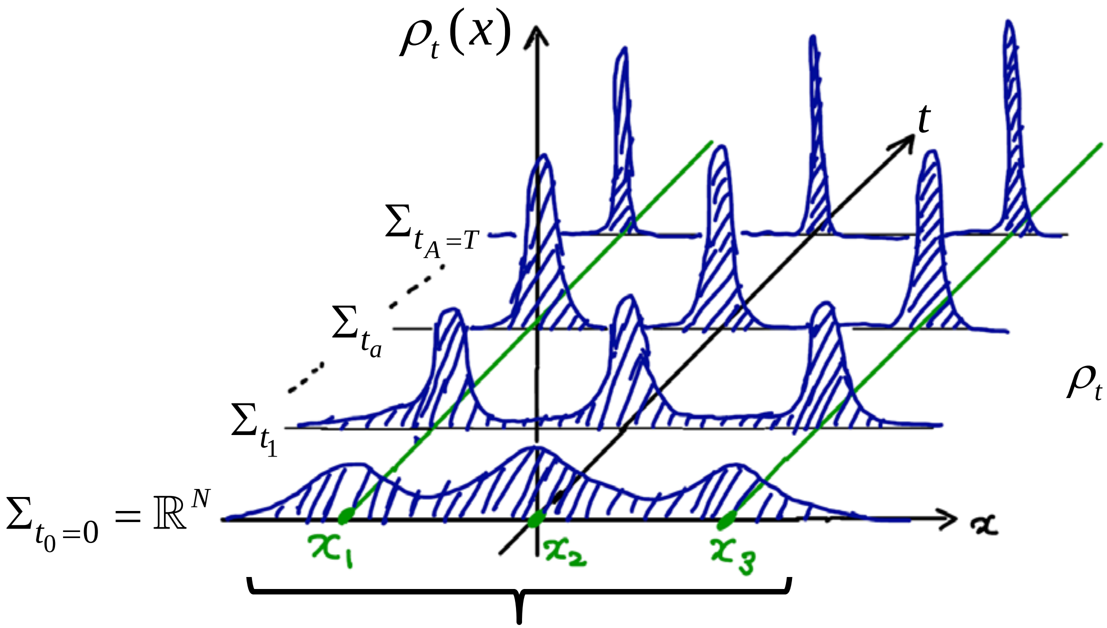
Sign and ergodicity problems solved simultaneously !

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

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

Important point in TLTM:

NO "tiny overlap problem" in TLTM



$$\rho_t(x) \propto e^{-\mathrm{Re}\, S(z_t(x))} \left| \det \frac{\partial z_t(x)}{\partial x} \right|$$

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

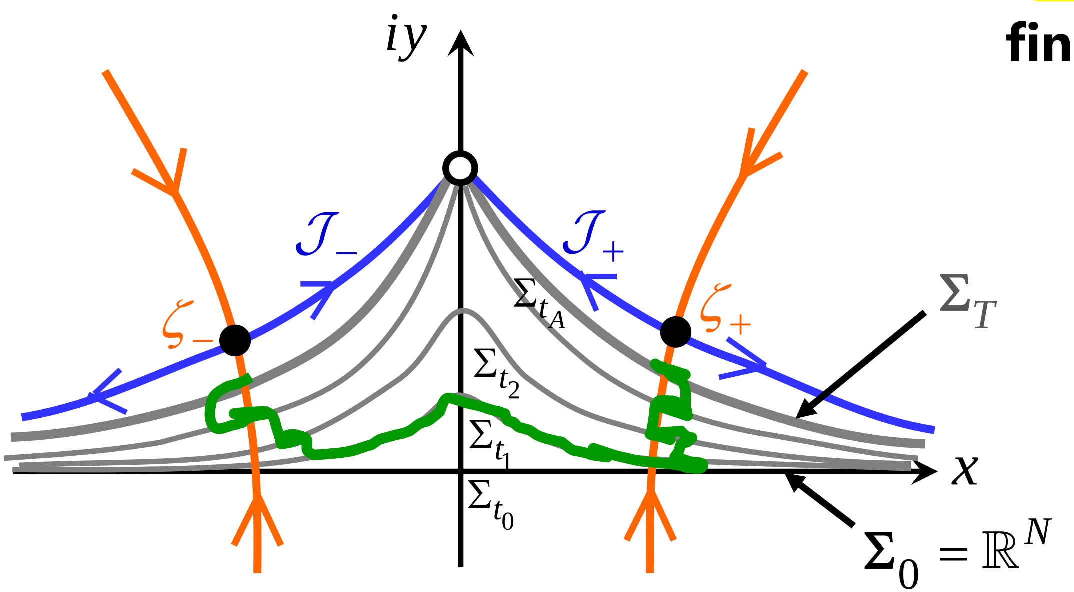
⇨ We can expect significant overlap between adjacent replicas!

Pros and cons of original TLTM

■ TLTM **[MF-Umeda 2017]**

Replicas introduced in between Σ_0 and Σ_T : $\left\{\Sigma_{t_0=0},\ \Sigma_{t_1},\ \Sigma_{t_2},\ \ldots,\ \Sigma_{t_A=T}\right\}$

finite discrete set of replicas



Pros : can be applied to any systems in principle
once formulated by path integrals with continuous variables

Cons : large comput cost at large DOF

- necessary # of replicas $\propto O(N^{0-1})$
- need to calculate Jacobian $J_t(x) = \partial z_t(x) / \partial x \propto O(N^3)$
  everytime we exchange configs between adjacent replicas

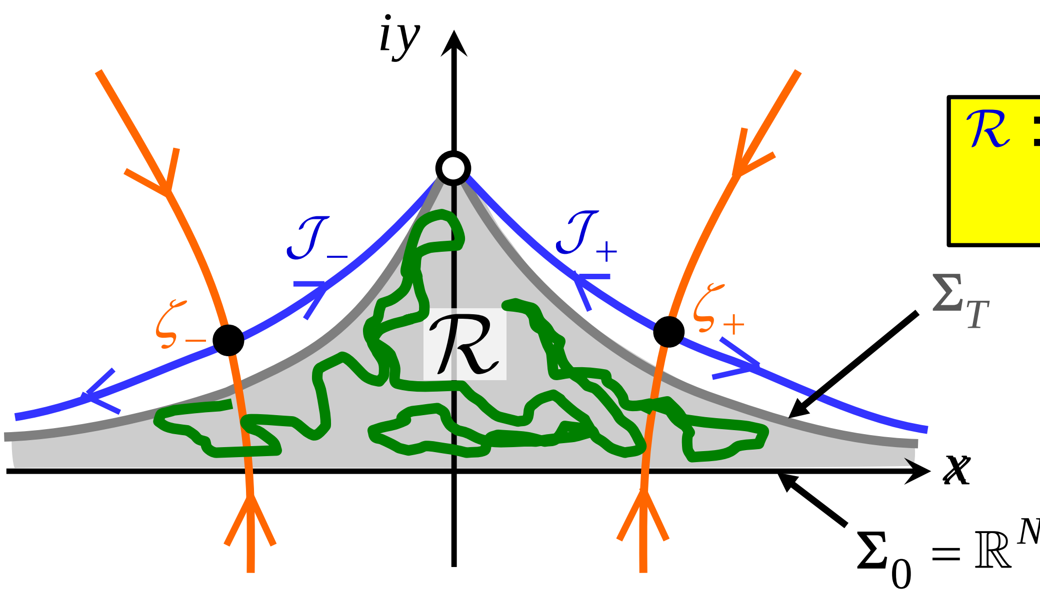
Worldvolume TLTM (WV-TLTM) (1/2)

[MF-Matsumoto 2012.08468]

■ Worldvolume TLTM (WV-TLTM)

HMC on a continuous accumulation of integ surfaces, $\mathcal{R} \equiv \bigcup_{0 \le t \le T} \Sigma_t$

"**worldvolume**"



$\mathcal{R}$: orbit of integration surface in the "target space" $\mathbb{C}^N = \mathbb{R}^{2N}$

(orbit of particle $\to$ worldline
orbit of string $\to$ worldsheet
orbit of surface (membrane) $\to$ worldvolume)

Pros : can be applied to any systems once formulated by path integrals with continuous variables

⊕ major reduction of comput cost at large DOF

- No need to introduce replicas explicitly
- No need to calculate Jacobian $J_t(x) = \partial z_t(x) / \partial x$ in MD process
- Configs can move largely due to the use of HMC

Worldvolume TLTM (WV-TLTM) (2/2)

[MF-Matsumoto 2012.08468]

■ Basic Idea

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

⇦ t-independent (Cauchy's theorem)

⇦ t-independent

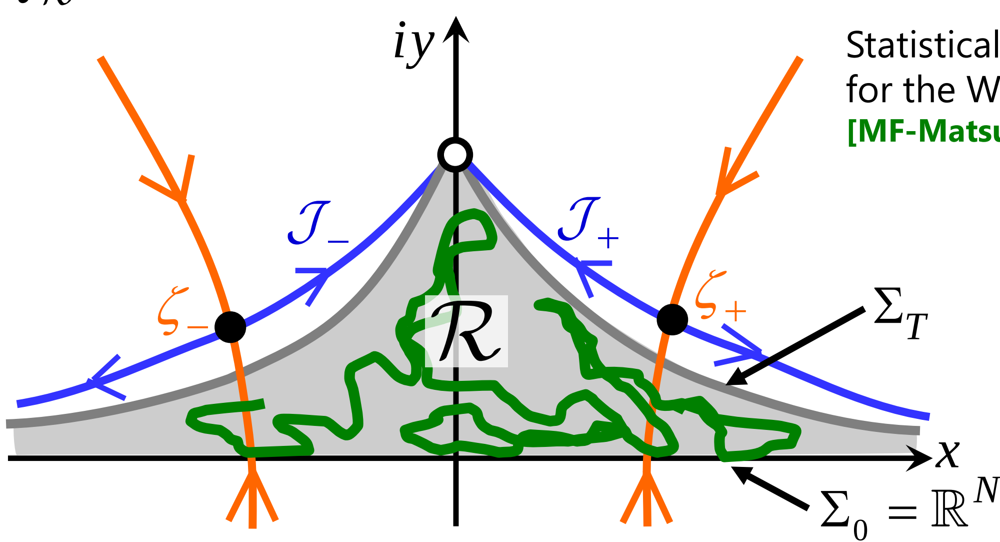
$$= \frac{\int_0^T dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\, e^{-S(z)} \mathcal{O}(z)}{\int_0^T dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\, e^{-S(z)}}$$

($W(t)$: arbitrary function)

(chosen s.t. the appearance prob at different t are almost the same)

$$= \frac{\int_{\mathcal{R}} dt\, dz_t\, e^{-W(t)} e^{-S(z)} \mathcal{O}(z)}{\int_{\mathcal{R}} dt\, dz_t\, e^{-W(t)} e^{-S(z)}}$$

⇐ Path integrals over the worldvolume $\mathcal{R}$

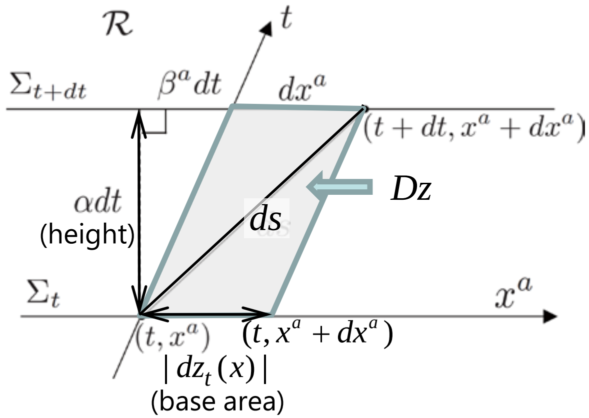


Statistical analysis method for the WV-TLTM is established in **[MF-Matsumoto-Namekawa 2107.06858]**

[MF-Matsumoto 2012.08468]

■ Preparation

$$\langle \mathcal{O}(x)\rangle = \frac{\int_{\mathcal{R}} dt\, dz_t\; e^{-W(t)} e^{-S(z)} \mathcal{O}(z)}{\int_{\mathcal{R}} dt\, dz_t\; e^{-W(t)} e^{-S(z)}}$$



natural measure to appear in HMC on $\mathcal{R}$

= vol element Dz of the induced metric

ADM decomposition

$$ds^2 = \alpha^2 dt^2 + \gamma_{ab}(dx^a + \beta^a dt)(dx^b + \beta^b dt) \quad (\alpha : \text{lapse})$$

⇨ $$Dz = \alpha\, dt\, |dz_t(x)| = \alpha\, |\det J|\, dt\, dx \qquad \left(J = \frac{\partial z_t(x)}{\partial x} \right)$$

⇨ $$dt\, dz_t(x) = Dz \frac{dt\, dz_t(x)}{Dz} = Dz \frac{dt\, dx \det J}{dt\, dx\, \alpha\, |\det J|} = Dz\, \alpha^{-1}(z) e^{i\varphi(z)} \qquad \left(e^{i\varphi(z)} \equiv \frac{\det J}{|\det J|} \right)$$

⇨ $$\langle \mathcal{O}(x)\rangle = \frac{\int_{\mathcal{R}} dt\, dz_t\; e^{-W(t)} e^{-S(z)} \mathcal{O}(z)}{\int_{\mathcal{R}} dt\, dz_t\; e^{-W(t)} e^{-S(z)}} = \frac{\int_{\mathcal{R}} Dz\, \alpha^{-1}(z) e^{i\varphi(z)}\, e^{-W(t)} e^{-\mathrm{Re}\, S(z) - i\, \mathrm{Im}\, S(z)} \mathcal{O}(z)}{\int_{\mathcal{R}} Dz\, \alpha^{-1}(z) e^{i\varphi(z)}\, e^{-W(t)} e^{-\mathrm{Re}\, S(z) - i\, \mathrm{Im}\, S(z)}}$$

$$= \frac{\int_{\mathcal{R}} Dz\; e^{-\mathrm{Re}\, S(z)} e^{-W(t)}\; \alpha^{-1}(z) e^{i\varphi(z)} e^{-i\, \mathrm{Im}\, S(z)}\, \mathcal{O}(z)}{\int_{\mathcal{R}} Dz\; e^{-\mathrm{Re}\, S(z)} e^{-W(t)}\; \alpha^{-1}(z) e^{i\varphi(z)} e^{-i\, \mathrm{Im}\, S(z)}}$$

$$= \frac{\int_{\mathcal{R}} Dz\; e^{-V(z)} A(z)\, \mathcal{O}(z)}{\int_{\mathcal{R}} Dz\; e^{-V(z)} A(z)} \equiv \frac{\langle A(z)\mathcal{O}(z)\rangle_{\mathcal{R}}}{\langle A(z)\rangle_{\mathcal{R}}}$$

$$\left(\langle f(z)\rangle_{\mathcal{R}} \equiv \frac{\int_{\mathcal{R}} Dz\; e^{-V(z)} f(z)}{\int_{\mathcal{R}} Dz\; e^{-V(z)}} \right)$$

Appendix: Details on WV-TLTM (2/2)

[MF-Matsumoto 2012.08468]

■ Algorithm

$$\langle \mathcal{O}(x) \rangle = \frac{\langle A(z)\mathcal{O}(z)\rangle_{\mathcal{R}}}{\langle A(z)\rangle_{\mathcal{R}}}$$

$$\left(\langle f(z) \rangle_{\mathcal{R}} \equiv \frac{\int_{\mathcal{R}} Dz\, e^{-V(z)} f(z)}{\int_{\mathcal{R}} Dz\, e^{-V(z)}} \right)$$

$$\begin{cases} V(z) = \mathrm{Re}\, S(z) + W(t(z)) & \text{: potential} \\ A(z) = \alpha^{-1}(z) e^{i\varphi(z)} e^{-i\,\mathrm{Im}\, S(z)} & \text{: reweighting factor} \end{cases}$$

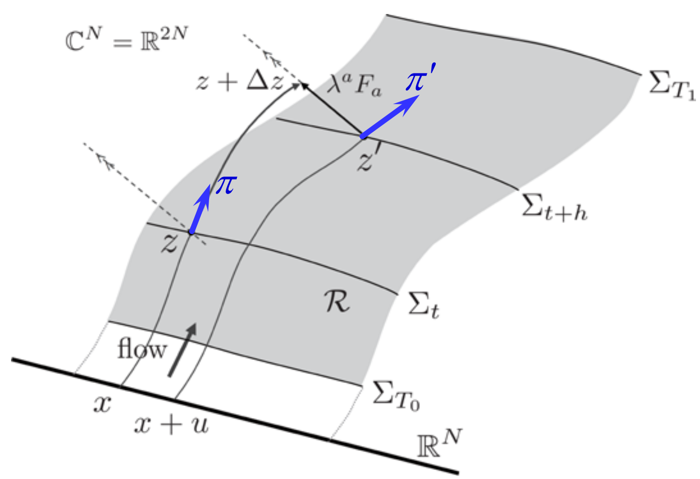
HMC on a constrained space **[Andersen 1983, Leimkuhler-Skeel 1994]**

$\langle f(z) \rangle_{\mathcal{R}}$ is estimated with RATTLE

$$\begin{cases} \pi_{1/2} = \pi - \Delta s\, \bar{\partial} V(z) - \lambda^a F_a(z) \\ z' = z + \Delta s\, \pi_{1/2} \\ \pi' = \pi - \Delta s\, \bar{\partial} V(z') - \lambda'^a F_a(z') \end{cases}$$

λ^a and λ'^a are determined s.t.

$$\begin{cases} z' \in \mathcal{R} \ \text{ and } \ \lambda^a\, \mathrm{Im}[J_a^\dagger(z) E_0(z)] = 0 \\ \pi' \in T_{z'}\mathcal{R} \ \text{ and } \ \lambda^a\, \mathrm{Im}[J_a^\dagger(z') E_0(z')] = 0 \end{cases}$$



cf) RATTLE on $\mathcal{J} = \Sigma_\infty$ [Fujii et al. 2013]
RATTLE on Σ_t [Alexandru@Lattice2019, MF-Matsumoto-Umeda 2019]

4. Application to various models

(WV-)TLTM has been successfully applied to ...

— (0+1)dim massive Thirring model **[MF-Umeda 1703.00861]**

— 2dim Hubbard model **[MF-Matsumoto-Umeda 1906.04243, 1912.13303]**

— chiral random matrix model (a toy model of finite density QCD) **[MF-Matsumoto 2012.08468]**

— anti-ferro Ising on triangular lattice **[MF-Matsumoto 2020, JPS meeting]**

So far successful for all the models when applied,
though the system sizes are yet small (DOF $N \leq 200$)

TLTM has been successfully applied to …

— (0+1)dim massive Thirring model **[MF-Umeda 1703.00861]**

— 2dim Hubbard model **[MF-Matsumoto-Umeda 1906.04243, 1912.13303]**

— chiral random matrix model (a toy model of finite density QCD)
[MF-Matsumoto 2012.08468]

— anti-ferro Ising on triangular lattice **[MF-Matsumoto 2020, JPS meeting]**

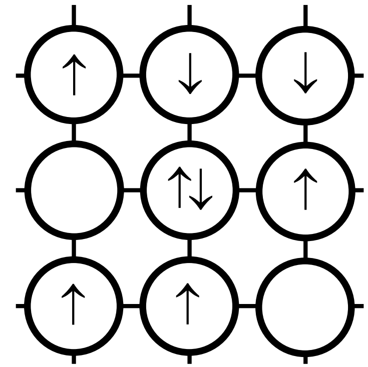
So far successful for all the models when applied,
though the system sizes are yet small (DOF $N \leq 200$)

4-1. Hubbard model (using original TLTM)

Hubbard model (1/3)

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

(fermion bilinear) (four fermion)



$(N_s$: # of sites$)$

- $n_{\mathbf{x},\sigma} \equiv c^{\dagger}_{\mathbf{x},\sigma} c_{\mathbf{x},\sigma}$ ($\mathbf{x}$: site, $\sigma = \uparrow, \downarrow$: spin)
- $\kappa (>0)$: hopping parameter, $K = (K_{\mathbf{xy}})$: adjacent matrix
- μ : chemical potential
- $U (>0)$: strength of on-site replusive potential

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

HS ⇨

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

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

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

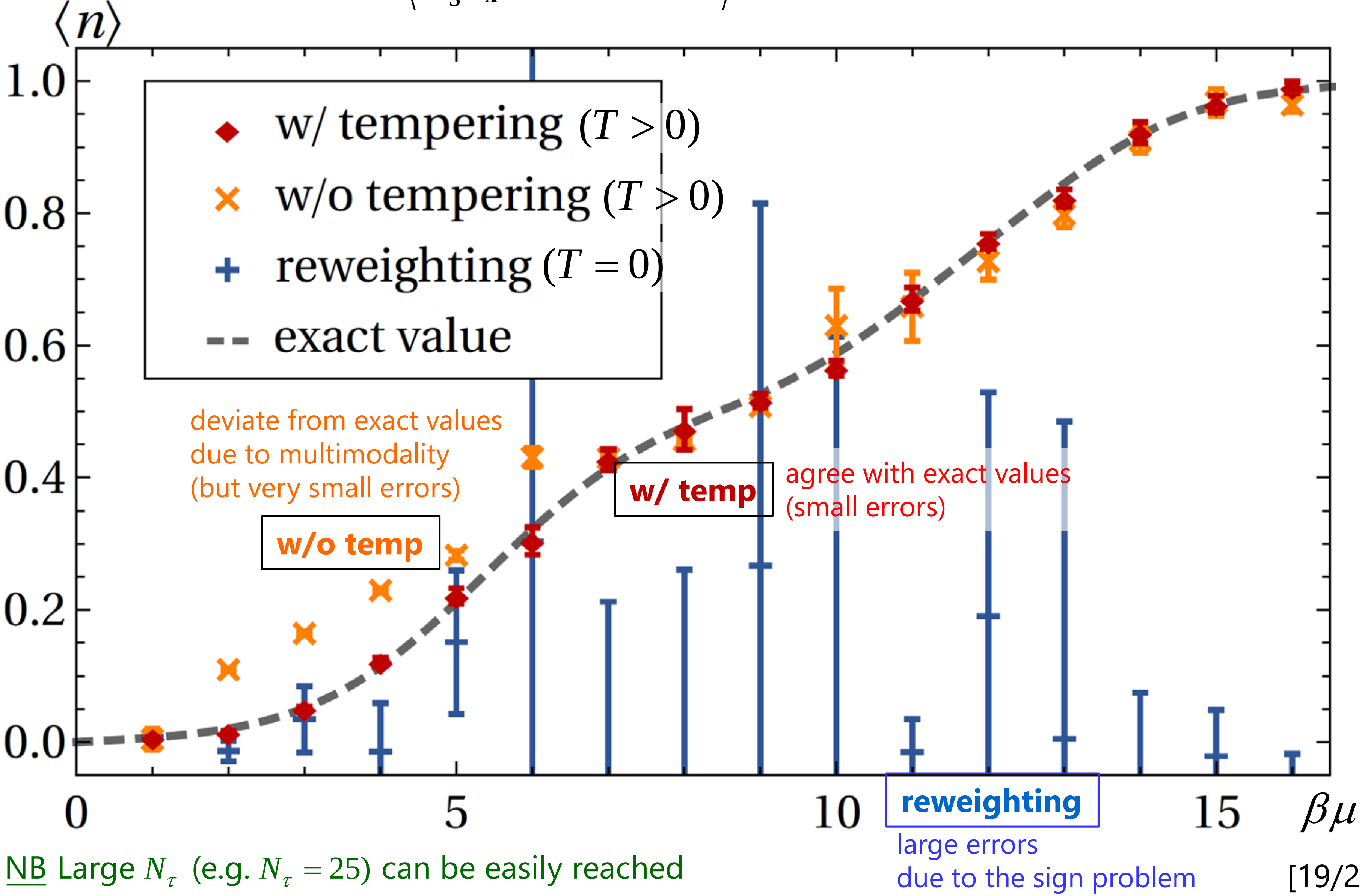
(For half filling $(\mu = 0)$: $\det M_a[\phi] \det M_b[\phi] = \left| \det M_a[\phi] \right|^2 \geq 0 \;\Rightarrow$ No sign problem)

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

[MF-Matsumoto-Umeda 1906.04243]

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

[MF-Matsumoto-Umeda 1906.04243]

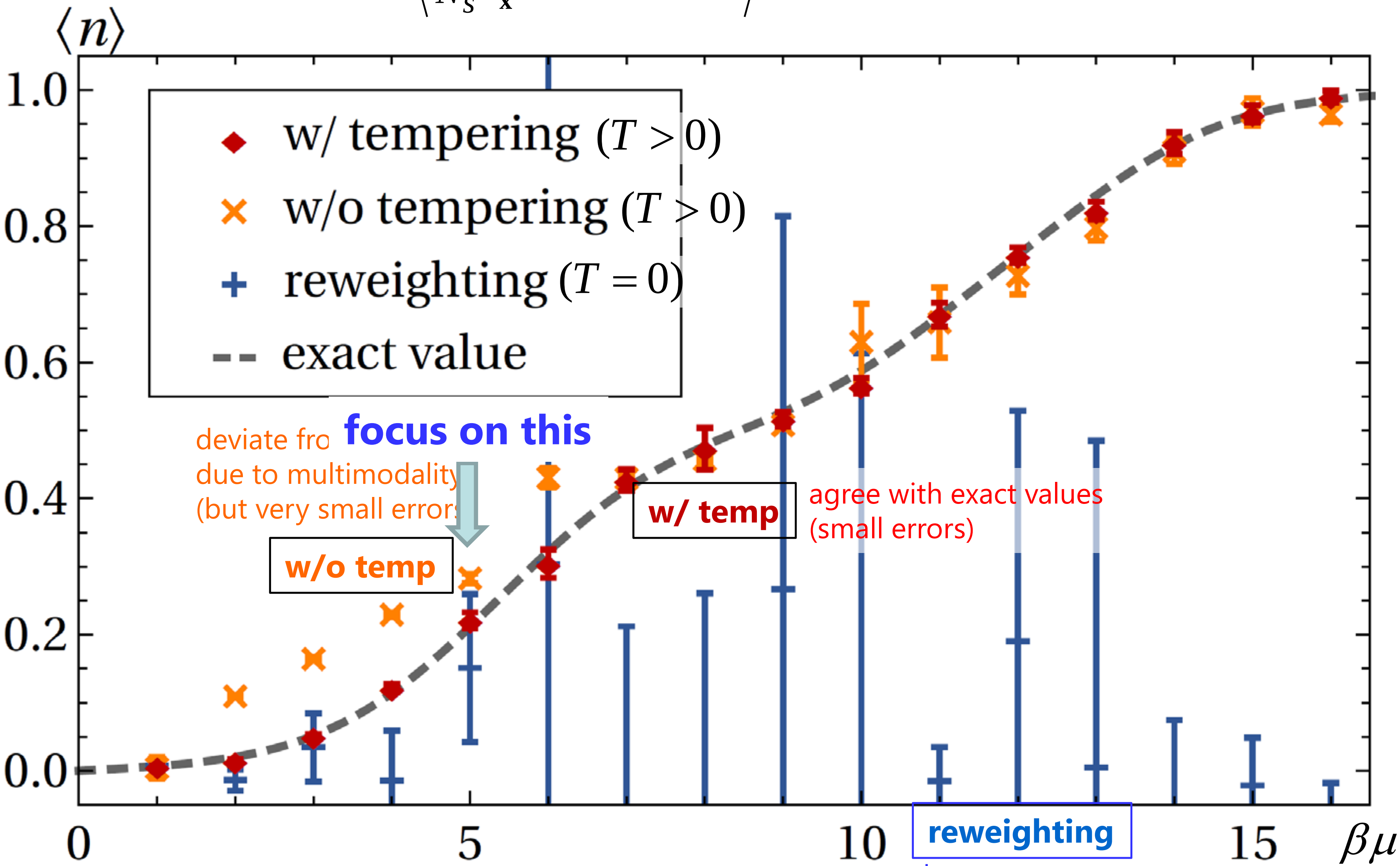


NB Large N_τ (e.g. $N_\tau = 25$) can be easily reached

Hubbard model (2/3)

[MF-Matsumoto-Umeda 1906.04243]

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



large errors
due to the sign problem

NB Large N_τ (e.g. $N_\tau = 25$) can be easily reached

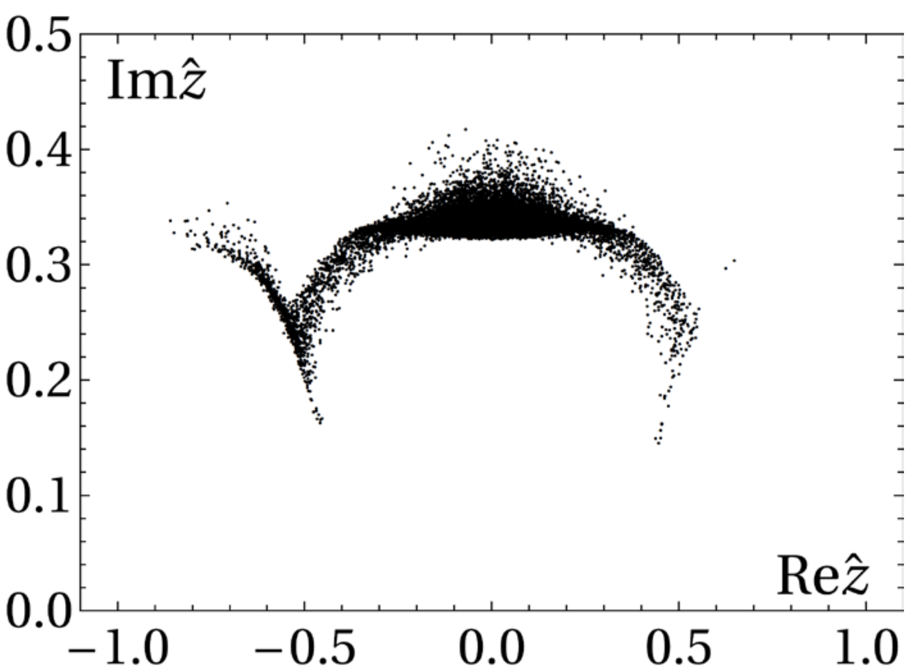
Hubbard model (3/3)

[MF-Matsumoto-Umeda 1906.04243]

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

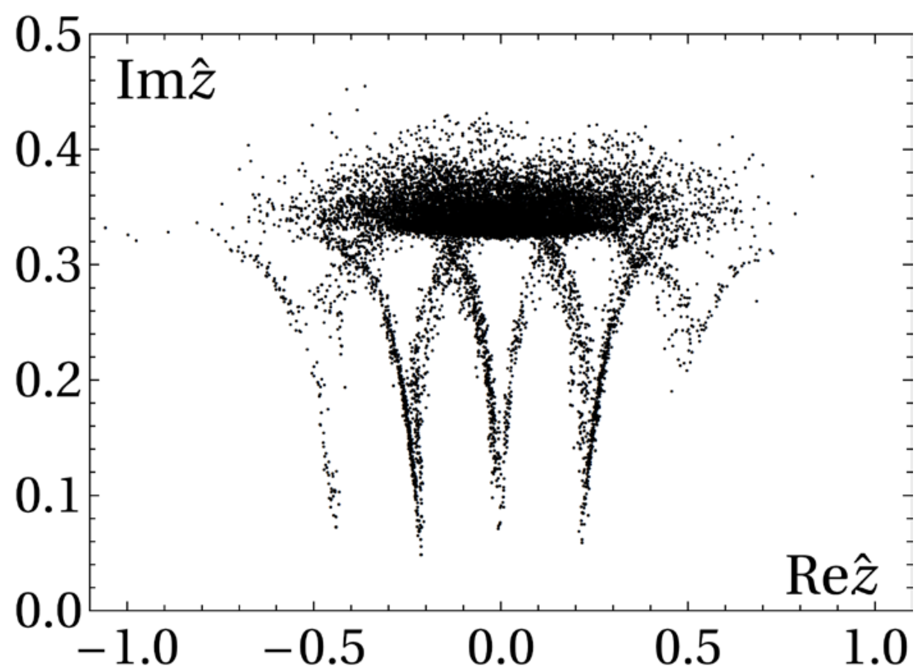
$\left(\text{projected on a plane : } \hat{z} = (N_\tau N_s)^{-1} \sum_{\ell,\mathbf{x}} z_{\ell,\mathbf{x}}\right)$

w/o temp



stuck to a small # of thimbles

w/ temp



distributed widely
over many thimbles

Comment on the Generalized LTM

[MF-Matsumoto-Umeda 1906.04243]

Alexandru et al. (2015) made a very interesting proposal
for reconciling the sign and ergodicity problems:

Choose a flow time that is sufficiently large so as to resolve the sign problem
but at the same time is not too large so as to avoid the ergodicity problem.

Our experience says it is NOT possible in many cases.

In fact, in most cases, the sign problem gets relaxed only after Σ_t reaches a zero

We confirm this for various values of $\beta\mu$

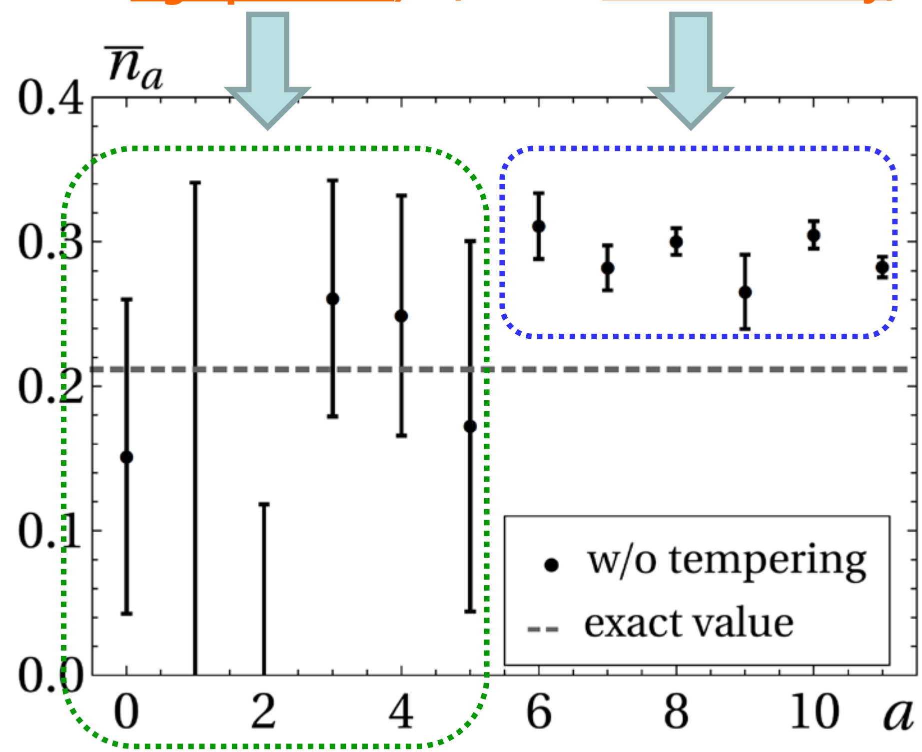
$$\begin{bmatrix} N_\tau = 5,\ \ N_s = 2\times 2 \\ \beta\kappa = 3,\ \ \beta U = 13,\ \ 0 \le T \le 0.4 \left(\Leftrightarrow 0 \le a \le 10\right) \\ N_{\rm conf} = 5{,}000 \sim 25{,}000 \text{ depending on } \beta\mu \\ \langle n \rangle = \dfrac{\langle e^{i\theta(z)} n(z) \rangle_{\Sigma_{t_a}}}{\langle e^{i\theta(z)} \rangle_{\Sigma_{t_a}}} \approx \overline{n}_a \end{bmatrix}$$

It is a hard task to find an intermediate flow time that solves both the sign and ergodicity problems simultaneously

Example: $\beta\mu = 5$

large stat errors (due to sign problem)

wrong value (due to multimodality)



4-2. Chiral random matrix model and the computational scaling (using WV-TLTM)

Chiral random matrix model (1/2)

[MF-Matsumoto 2012.08468]

■ finite density QCD

$$\left(\{\gamma_\mu,\gamma_\nu\}=2\delta_{\mu\nu},\ \gamma_\mu=\gamma_\mu^\dagger=\begin{pmatrix}0 & \sigma_\mu\\ \sigma_\mu^\dagger & 0\end{pmatrix}\right)$$

$$\begin{aligned} Z_{\text{QCD}} &= \operatorname{tr} e^{-\beta(H-\mu N)} \\ &= \int [dA_\mu][d\psi d\bar\psi]\, e^{(1/2g^2)\int \operatorname{tr} F_{\mu\nu}^2 + \int[\bar\psi(\gamma_\mu D_\mu + m)\psi + \mu\psi^\dagger\psi]} \\ &= \int [dA_\mu]\, e^{(1/2g^2)\int \operatorname{tr} F_{\mu\nu}^2}\ \mathrm{Det}\begin{pmatrix} m & \sigma_\mu(\partial_\mu + A_\mu)+\mu \\ \sigma_\mu^\dagger(\partial_\mu + A_\mu)+\mu & m\end{pmatrix}\end{aligned}$$

⇓ toy model

■ chiral random matrix model **[Stephanov 1996, Halasz et al. 1998]**

$$Z_{\text{Steph}} = \int d^2W\ e^{-n\,\operatorname{tr}W^\dagger W}\ \det\begin{pmatrix} m & iW+\mu \\ iW^\dagger+\mu & m\end{pmatrix}$$

(quantum field replaced by a matrix incl spacetime DOF)

$(T=0,\ N_f=1)$

$W=(W_{ij})=(X_{ij}+iY_{ij}):\ n\times n$ complex matrix

(DOF : $N=2n^2 \ \Leftrightarrow\ 4L^4(N_c^2-1)$)

■ role as an important benchmark model

- well approximates the qualitative behaviour of QCD at large n
- complex Langevin suffers from wrong convergence **[Bloch et al. 2018]**

Chiral random matrix model (2/2)

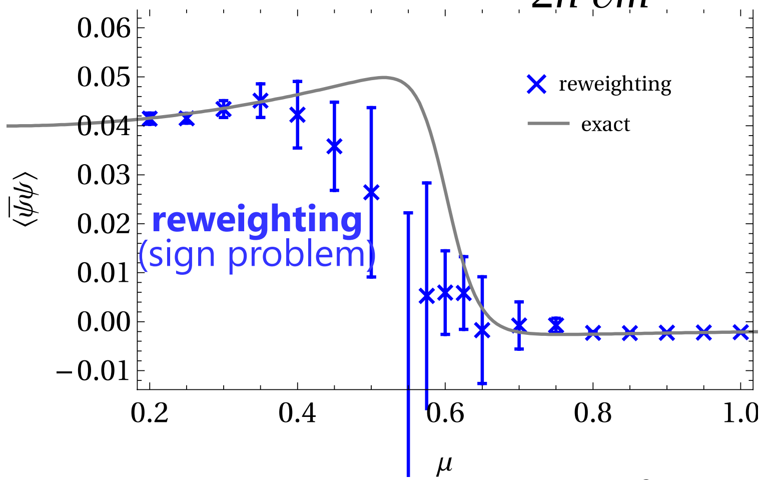
[MF-Matsumoto 2012.08468]

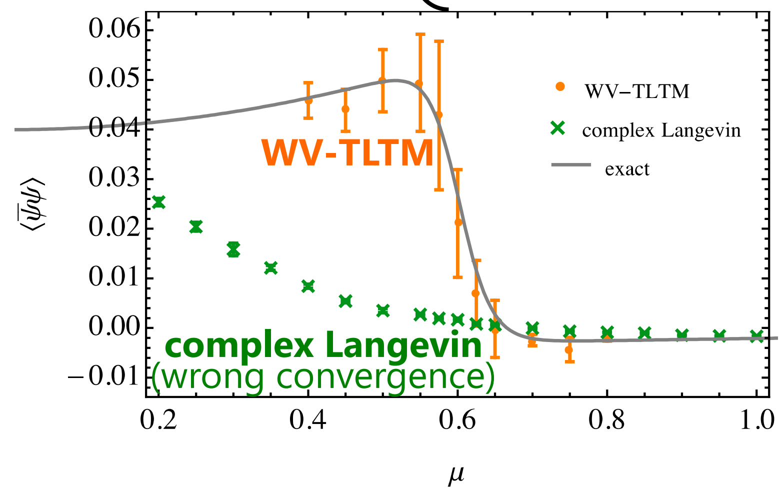
matrix size : $n = 10$ (DOF : $N = 200$)

(now easy at large DOF compared to the original TLTM)

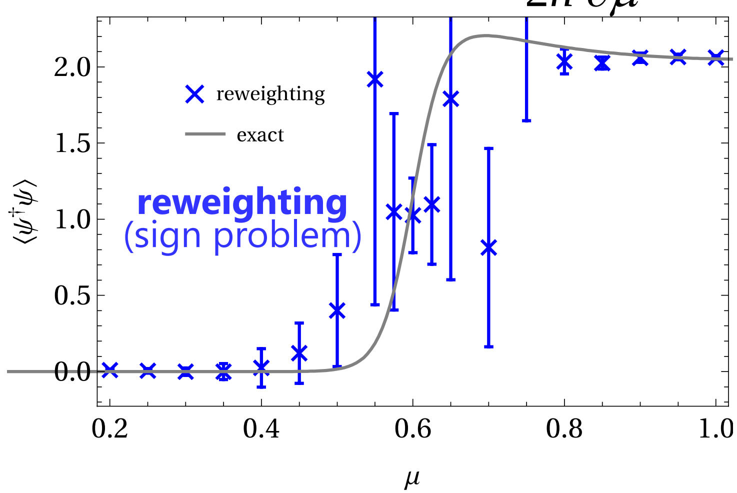
<u>sample size</u>
reweighting : 10k
complex Langevin :10k
WV-TLTM : 4k-17k

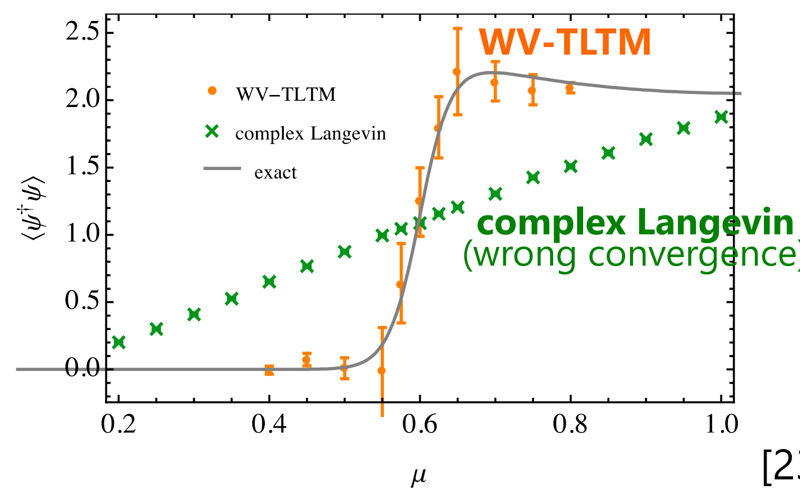
<u>**chiral condensate**</u> $\langle \bar{\psi}\psi \rangle \equiv \dfrac{1}{2n}\dfrac{\partial}{\partial m}\ln Z_{\text{Steph}}$ $[m = 0.004,\ T = 0]$





<u>**baryon # density**</u> $\langle \psi^{\dagger}\psi \rangle \equiv \dfrac{1}{2n}\dfrac{\partial}{\partial \mu}\ln Z_{\text{Steph}}$





Scaling of computational cost

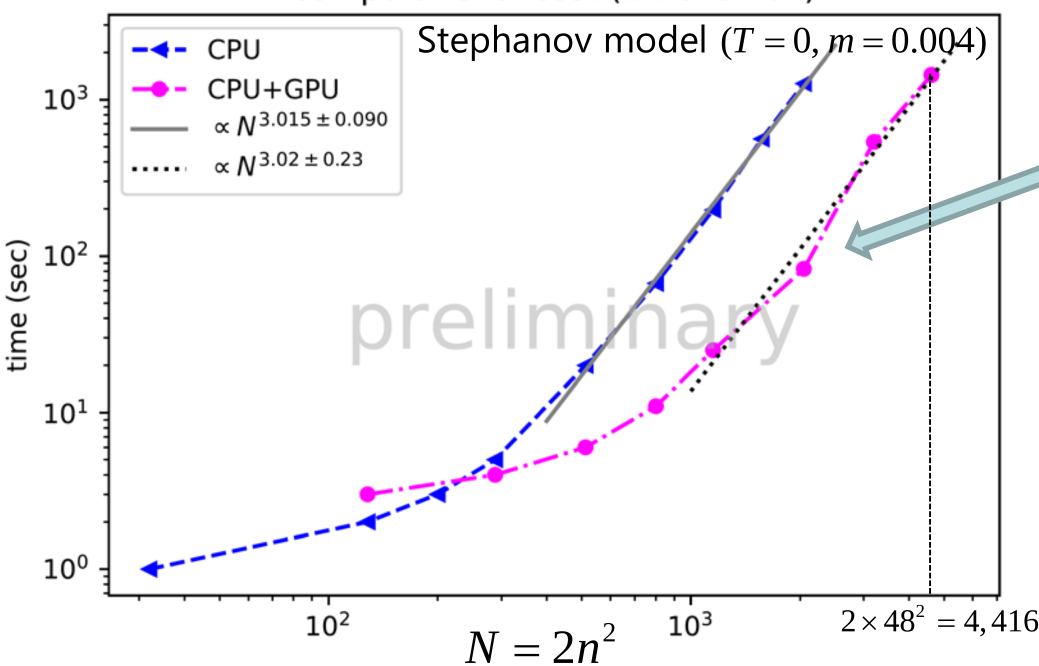
[MF-Matsumoto-Namekawa, work in progress]

We need to make a linear inversion of $J_t x = b$ several times in generating a config.
The code on this part is currently being updated : **Phase 1** ⇒ **Phase 2**

Phase 1 - Use a direct method (LU decomp) in linear inversion $\Leftarrow O(N^3)$

- J_t is calculated explicitly by solving a diff eq $\dot{J}_t = \overline{\partial^2 S(z_t)} \cdot J_t$ $\Leftarrow O(N^3)$

computational cost (1 transition)



$O(N^3)$ (not exponential)

$[O(N^{\gtrsim 3.5})$ with original TLTM$]$

(Ongoing: comput scaling with fixed statistical errors)

computed on Yukawa-21@YITP

(CPU: Xeon Platinum 8280
GPU: NVIDIA Tesla V100)

Phase 2 - Use an iterative method (BiCGStab) in linear inversion $\Leftarrow O(N^{2\sim 3})$

- Only need to solve a flow eq for a vector v_t tangent to Σ_t, $\dot{v}_t = \overline{\partial^2 S(z_t)} \cdot v_t$ $\Leftarrow O(N^2)$

⇒ We expect $O(N^{2\sim 3})$ for generating a config (work in progress)

5. Summary and outlook

Summary and outlook

■ Summary

▼ TLTM has a potential to be a solution to the sign problem

- Sign and ergodicity problems are solved simultaneously

▼ TLTM has been successfully applied to various models

(yet only to toy models w/ small DOF at this stage)

- finite density QCD ⟸ chiral random matrix model [MF-Matsumoto]
- QMC :
  - strongly correl electron systems ⟸ 1D/2D Hubbard model [MF-Matsumoto-Umeda]
  - frustrated classical / quatum spin systems ⟸ antiferro Ising on trianglular lattice [MF-Matsumoto]

■ Outlook [MF-Matsumoto-Namekawa, work in progress]

▼ Large-scale computation for large-size systems w/ WV-TLTM

▼ Further improvements of algorithm

▼ Combining various algorithms

(e.g.) TRG (non-MC) : good at calculating free energy — cf) TRG for 2D YM: [MF-Kadoh-Matsumoto 2107.14149]

▼ Particularly important: MC calc for time-dependent systems

⟹ first-principles calc of nonequilibrium processes such as early universe, heavy ion collision experiments, …

Thank you.