Sign problem and the Worldvolume Hybrid Monte Carlo method

Masafumi Fukuma (Dept Phys, Kyoto Univ)

"Challenges and opportunities in Lattice QCD simulations and related fields"

Feb 15-17, 2023 @RIKEN R-CCS

Based on work with

Naoya Umeda (PwC)
Nobuyuki Matsumoto (RIKEN/BNL)
Yusuke Namekawa (Hiroshima Univ)
Issaku Kanamori (RIKEN R-CCS)

Today's talk

- What is the sign problem?
  - Difficulty for complex action
  - Why the naive method (reweighting) fails
- Overview of various approaches
- Argue that
  Worldvolume Hybrid Monte Carlo method
  is a promising method

Introduction

Basics of Monte Carlo methods

■ What we want: expectation values of observables

$$\begin{cases} x = (x^i) \in \mathbb{R}^N : \text{dynamical variable } (N\text{: DOF}) \\ S(x): \text{ action} \\ \mathcal{O}(x): \text{ observable} \\ dx = \prod_{i=1}^{N} dx^i \end{cases}$$

(e.g. scalar field

$$x^i \leftrightarrow \phi(t,\mathbf{x})$$

$$S(x) \leftrightarrow S[\phi] = \int dt\, d^3\mathbf{x} \left[\frac{1}{2}(\partial_t \phi)^2 + \cdots \right]$$

$$dx = \prod_i dx^i \leftrightarrow [d\phi] = \prod_{t,\mathbf{x}} d\phi(t,\mathbf{x})$$

)

$$\langle \mathcal{O} \rangle \equiv \int dx\, \rho(x) \mathcal{O}(x) \qquad \left(\rho(x) \equiv \frac{e^{-S(x)}}{\int dx\, e^{-S(x)}} \right)$$

When $N \gg 1$, numerical multiple integrals are unrealistic

Basics of Monte Carlo methods

■ What we want: expectation values of observables

$$\begin{cases} x = (x^i) \in \mathbb{R}^N : \text{dynamical variable } (N\text{: DOF}) \\ S(x): \text{ action} \\ \mathcal{O}(x): \text{ observable} \\ dx = \prod_{i=1}^{N} dx^i \end{cases}$$

e.g. scalar field

$$x^i \leftrightarrow \phi(t,\mathbf{x})$$

$$S(x) \leftrightarrow S[\phi] = \int dt\, d^3\mathbf{x} \left[\frac{1}{2}(\partial_t \phi)^2 + \cdots \right]$$

$$dx = \prod_i dx^i \leftrightarrow [d\phi] = \prod_{t,\mathbf{x}} d\phi(t,\mathbf{x})$$

$$\langle \mathcal{O} \rangle \equiv \int dx\, \rho(x) \mathcal{O}(x) \qquad \left(\rho(x) \equiv \frac{e^{-S(x)}}{\int dx\, e^{-S(x)}} \right)$$

When $N \gg 1$, numerical multiple integrals are unrealistic

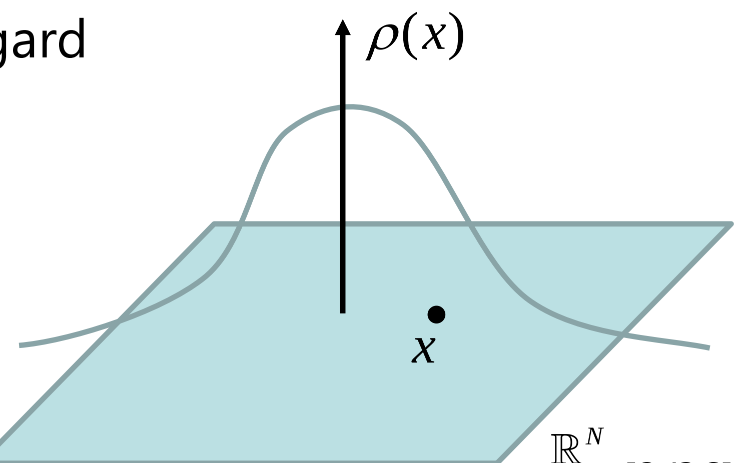
However, when $S(x)$ is real-valued, one can regard

$$\rho(x) = \frac{e^{-S(x)}}{\int dx\, e^{-S(x)}}$$

as a probability distribution $\left(\because \begin{array}{l} \rho(x) \geq 0 \quad (\forall x) \\ \int dx\, \rho(x) = 1 \end{array} \right)$

⇩

Markov chain Monte Carlo method



Basics of Monte Carlo methods

■ What we want: expectation values of observables

$$\begin{cases} x=(x^i)\in\mathbb{R}^N : \text{dynamical variable } (N\text{: DOF}) \\ S(x): \text{ action} \\ \mathcal{O}(x): \text{ observable} \\ dx=\prod_{i=1}^N dx^i \end{cases}$$

$$\left(\begin{array}{l} \text{e.g. scalar field} \\ x^i \leftrightarrow \phi(t,\mathbf{x}) \\ S(x)\leftrightarrow S[\phi]=\int dt\, d^3\mathbf{x}\left[\frac{1}{2}(\partial_t\phi)^2+\cdots\right] \\ dx=\prod_i dx^i \leftrightarrow [d\phi]=\prod_{t,\mathbf{x}} d\phi(t,\mathbf{x}) \end{array}\right)$$

$$\langle\mathcal{O}\rangle\equiv\int dx\,\rho(x)\mathcal{O}(x) \qquad \left(\rho(x)\equiv\frac{e^{-S(x)}}{\int dx\, e^{-S(x)}}\right)$$

When $N\gg 1$, numerical multiple integrals are unrealistic

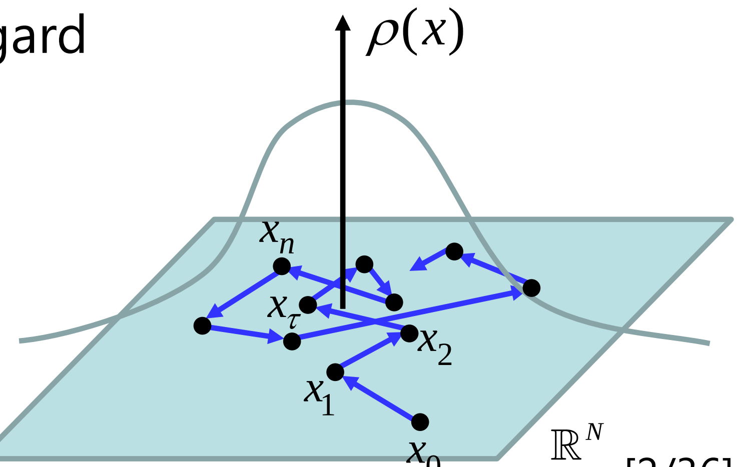
However, <u>when $S(x)$ is real-valued</u>, one can regard

$$\rho(x)=\frac{e^{-S(x)}}{\int dx\, e^{-S(x)}}$$

as a probability distribution $\left(\because\ \begin{array}{l}\rho(x)\geq 0\ \ (\forall x) \\ \int dx\,\rho(x)=1\end{array}\right)$

⇩

Markov chain Monte Carlo method



What if the action is complex-valued?

■ **complex action** $S(x) = \mathrm{Re}\, S(x) + i\, \mathrm{Im}\, S(x) \in \mathbb{C} \quad \big(\mathrm{Im}\, S(x) \neq 0\big)$

$\rho(x)$ cannot be regarded as a probability distribution $\left(\rho(x) = \dfrac{e^{-S(x)}}{\int dx\, e^{-S(x)}} \right)$

■ **Examples**

- finite-density QCD

Q: baryon number
μ: chemical potential

$$Z = \mathrm{tr}\, e^{-\beta(H-\mu Q)}$$

$$= \int [dA_\mu][d\psi d\bar{\psi}]\, e^{(1/2g^2)\int \mathrm{tr}\, F_{\mu\nu}^2 + \int \bar{\psi} D(A_\mu;\mu)\psi}$$

$$\left(\gamma_\mu = \gamma_\mu^\dagger \ (\mu = 1,2,3,0) \quad \{\gamma_\mu, \gamma_\nu\} = 2\delta_{\mu\nu} 1_4 \right)$$

$$\big(D(A_\mu;\mu) = \gamma_\mu D_\mu + m + \mu\gamma_0 \big)$$

$$= \int [dA_\mu]\, e^{(1/2g^2)\int \mathrm{tr}\, F_{\mu\nu}^2} \boxed{\mathrm{Det}\, D(A_\mu;\mu)}$$

$$\in \mathbb{C} \quad \left(\because \left[\mathrm{Det}\, D(A_\mu;\mu)\right]^* = \mathrm{Det}\, D(A_\mu;-\mu^*) \right)$$

- Quantum Monte Carlo (QMC)

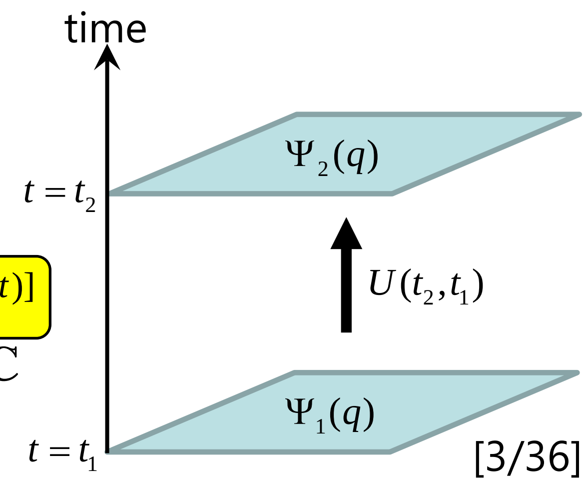
strongly correlated electron systems,
frustrated spin systems, ...

- real-time dynamics of QM/QFT

$$\langle \Psi_2 | U(t_2,t_1) | \Psi_1 \rangle = \int [dq(t)]\, \Psi_2^*\big(q(t_2)\big)\, \Psi_1\big(q(t_1)\big) \boxed{e^{iS[q(t)]}}$$

$$\in \mathbb{C}$$

$$\left(U(t_2,t_1) = T \exp\left[-i \int_{t_1}^{t_2} dt\, H(t) \right] \right)$$



Plan

1. Introduction (done)
2. Sign problem and various approaches
3. Complex Langevin method
4. (Generalized) Lefschetz thimble method
5. Tempered Lefschetz thimble (TLT) method
6. Worldvolume Hybrid Monte Carlo (WV-HMC) method
7. Application to various models
8. Summary and outlook

Plan

1. Introduction
2. Sign problem and various approaches
3. Complex Langevin method
4. (Generalized) Lefschetz thimble method
5. Tempered Lefschetz thimble (TLT) method
6. Worldvolume Hybrid Monte Carlo (WV-HMC) method
7. Application to various models
8. Summary and outlook

Sign problem

■ **reweighting method**: simplest prescription for the complex action

$$e^{-S(x)} = \underbrace{e^{-\mathrm{Re}\,S(x)}}_{\text{used for a wt}} \ \underbrace{e^{-i\,\mathrm{Im}\,S(x)}}_{\text{treated as part of obs}}$$

$$\langle \mathcal{O}(x) \rangle = \frac{\int dx\, e^{-S(x)} \mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \frac{\int dx\, e^{-\mathrm{Re}\,S(x)} e^{-i\,\mathrm{Im}\,S_I(x)} \mathcal{O}(x) \Big/ \int dx\, e^{-\mathrm{Re}\,S(x)}}{\int dx\, e^{-\mathrm{Re}\,S(x)} e^{-i\,\mathrm{Im}\,S_I(x)} \Big/ \int dx\, e^{-\mathrm{Re}\,S(x)}}$$

highly oscillatory for $N \gg 1$

$$= \frac{\langle e^{-i\,\mathrm{Im}\,S_I(x)} \mathcal{O}(x) \rangle_{\text{rewt}}}{\langle e^{-i\,\mathrm{Im}\,S_I(x)} \rangle_{\text{rewt}}}$$

$$= \frac{e^{-O(N)}}{e^{-O(N)}} \quad \big(= O(1) \big)$$

Sign problem

■ **reweighting method**: simplest prescription for the complex action

$$e^{-S(x)} = \underbrace{e^{-\mathrm{Re}\, S(x)}}_{\text{used for a wt}} \; \underbrace{e^{-i\, \mathrm{Im}\, S(x)}}_{\text{treated as part of obs}}$$

$$\langle \mathcal{O}(x) \rangle = \frac{\int dx\, e^{-S(x)} \mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \frac{\int dx\, e^{-\mathrm{Re}\, S(x)} e^{-i\, \mathrm{Im}\, S_I(x)} \mathcal{O}(x) \Big/ \int dx\, e^{-\mathrm{Re}\, S(x)}}{\int dx\, e^{-\mathrm{Re}\, S(x)} e^{-i\, \mathrm{Im}\, S_I(x)} \Big/ \int dx\, e^{-\mathrm{Re}\, S(x)}}$$

highly oscillatory for $N \gg 1$

$$= \frac{\langle e^{-i\, \mathrm{Im}\, S_I(x)} \mathcal{O}(x) \rangle_{\mathrm{rewt}}}{\langle e^{-i\, \mathrm{Im}\, S_I(x)} \rangle_{\mathrm{rewt}}}$$

$$= \frac{e^{-O(N)}}{e^{-O(N)}} \quad \left(= O(1)\right)$$

Sign problem

■ **reweighting method**: simplest prescription for the complex action

$$e^{-S(x)} = \underbrace{e^{-\mathrm{Re}\,S(x)}}_{\text{used for a wt}} \ \underbrace{e^{-i\,\mathrm{Im}\,S(x)}}_{\text{treated as part of obs}}$$

$$\langle \mathcal{O}(x) \rangle = \frac{\int dx\, e^{-S(x)}\, \mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \frac{\int dx\, e^{-\mathrm{Re}\,S(x)}\, e^{-i\,\mathrm{Im}\,S_I(x)}\, \mathcal{O}(x) \Big/ \int dx\, e^{-\mathrm{Re}\,S(x)}}{\int dx\, e^{-\mathrm{Re}\,S(x)}\, e^{-i\,\mathrm{Im}\,S_I(x)} \Big/ \int dx\, e^{-\mathrm{Re}\,S(x)}}$$

highly oscillatory for $N \gg 1$

$$= \frac{\langle e^{-i\,\mathrm{Im}\,S_I(x)}\, \mathcal{O}(x) \rangle_{\mathrm{rewt}}}{\langle e^{-i\,\mathrm{Im}\,S_I(x)} \rangle_{\mathrm{rewt}}}$$

$$= \frac{e^{-O(N)}}{e^{-O(N)}} \quad \left(= O(1)\right)$$

Sign problem

■ **reweighting method**: simplest prescription for the complex action

$$e^{-S(x)} = \underbrace{e^{-\mathrm{Re}\, S(x)}}_{\text{used for a wt}} \; \underbrace{e^{-i\,\mathrm{Im}\, S(x)}}_{\text{treated as part of obs}}$$

$$\langle \mathcal{O}(x) \rangle = \frac{\int dx\, e^{-S(x)}\, \mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \frac{\int dx\, e^{-\mathrm{Re}\, S(x)}\, e^{-i\,\mathrm{Im}\, S_I(x)}\, \mathcal{O}(x) \Big/ \int dx\, e^{-\mathrm{Re}\, S(x)}}{\int dx\, e^{-\mathrm{Re}\, S(x)}\, e^{-i\,\mathrm{Im}\, S_I(x)} \Big/ \int dx\, e^{-\mathrm{Re}\, S(x)}}$$

$$= \frac{\langle e^{-i\,\mathrm{Im}\, S_I(x)}\, \mathcal{O}(x) \rangle_{\mathrm{rewt}}}{\langle e^{-i\,\mathrm{Im}\, S_I(x)} \rangle_{\mathrm{rewt}}} \qquad \left(\langle f(x) \rangle_{\mathrm{rewt}} \equiv \frac{\int dx\, e^{-\mathrm{Re}\, S(x)}\, f(x)}{\int dx\, e^{-\mathrm{Re}\, S(x)}} \right)$$

Sign problem

■ **reweighting method**: simplest prescription for the complex action

$$e^{-S(x)} = \underbrace{e^{-\mathrm{Re}\,S(x)}}_{\text{used for a wt}} \ \underbrace{e^{-i\,\mathrm{Im}\,S(x)}}_{\text{treated as part of obs}}$$

$$\langle \mathcal{O}(x) \rangle = \frac{\int dx\, e^{-S(x)}\, \mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \frac{\int dx\, e^{-\mathrm{Re}\,S(x)}\, e^{-i\,\mathrm{Im}\,S_I(x)}\, \mathcal{O}(x) \Big/ \int dx\, e^{-\mathrm{Re}\,S(x)}}{\int dx\, e^{-\mathrm{Re}\,S(x)}\, e^{-i\,\mathrm{Im}\,S_I(x)} \Big/ \int dx\, e^{-\mathrm{Re}\,S(x)}}$$

highly oscillatory for $N \gg 1$

$$= \frac{\langle e^{-i\,\mathrm{Im}\,S_I(x)}\, \mathcal{O}(x) \rangle_{\mathrm{rewt}}}{\langle e^{-i\,\mathrm{Im}\,S_I(x)} \rangle_{\mathrm{rewt}}} \qquad \left(\langle f(x) \rangle_{\mathrm{rewt}} \equiv \frac{\int dx\, e^{-\mathrm{Re}\,S(x)}\, f(x)}{\int dx\, e^{-\mathrm{Re}\,S(x)}} \right)$$

$$= \frac{e^{-O(N)}}{e^{-O(N)}} \quad \left(= O(1) \right)$$

Sign problem

■ **reweighting method**: simplest prescription for the complex action

$$e^{-S(x)} = \underbrace{e^{-\mathrm{Re}\,S(x)}}_{\text{used for a wt}} \; \underbrace{e^{-i\,\mathrm{Im}\,S(x)}}_{\text{treated as obs}}$$

$$\langle \mathcal{O}(x)\rangle = \frac{\int dx\, e^{-S(x)}\,\mathcal{O}(x)}{\int dx\, e^{-S(x)}} = \frac{\int dx\, e^{-\mathrm{Re}\,S(x)}\, e^{-i\,\mathrm{Im}\,S_I(x)}\,\mathcal{O}(x) \Big/ \int dx\, e^{-\mathrm{Re}\,S(x)}}{\int dx\, e^{-\mathrm{Re}\,S(x)}\, e^{-i\,\mathrm{Im}\,S_I(x)} \Big/ \int dx\, e^{-\mathrm{Re}\,S(x)}}$$

highly oscillatory for $N \gg 1$

$$= \frac{\langle e^{-i\,\mathrm{Im}\,S_I(x)}\,\mathcal{O}(x)\rangle_{\text{rewt}}}{\langle e^{-i\,\mathrm{Im}\,S_I(x)}\rangle_{\text{rewt}}} \qquad \left(\langle f(x)\rangle_{\text{rewt}} \equiv \frac{\int dx\, e^{-\mathrm{Re}\,S(x)}\, f(x)}{\int dx\, e^{-\mathrm{Re}\,S(x)}} \right)$$

$$= \frac{e^{-O(N)}}{e^{-O(N)}} \quad \big(= O(1)\big)$$

But, in numerical calc, the numer and denom are estimated with sample avgs

$$\langle \mathcal{O}(x)\rangle \approx \frac{e^{-O(N)} \pm O(1/\sqrt{N_{\text{conf}}})}{e^{-O(N)} \pm O(1/\sqrt{N_{\text{conf}}})} \qquad \begin{pmatrix} N \;:\; \text{DOF} \\ N_{\text{conf}} \;:\; \text{sample size} \end{pmatrix}$$

$\therefore$ small error $O(1/\sqrt{N_{\text{conf}}}) \le e^{-O(N)} \Leftrightarrow \boxed{N_{\text{conf}} \ge e^{O(N)}}$

SIGN PROBLEM!

Need an exponentially large sample

Example：Gaussian

$$\begin{cases} S(x) = \dfrac{\beta}{2}(x-i)^2 \equiv \mathrm{Re}\,S(x) + i\,\mathrm{Im}\,S(x) \\ \mathcal{O}(x) = x^2 \end{cases} \quad \left(\begin{array}{l} \mathrm{Re}\,S(x) = \dfrac{\beta}{2}(x^2-1) \\ \mathrm{Im}\,S(x) = -\beta x \end{array} \right)$$

$\boxed{\beta \gg 1}$ with $N=1$

large β mimics large DOF

$$\Rightarrow \quad \langle x^2 \rangle = \frac{\langle e^{-i\mathrm{Im}S(x)}\, x^2 \rangle_{\text{rewt}}}{\langle e^{-i\mathrm{Im}S(x)} \rangle_{\text{rewt}}} = \frac{\left(\beta^{-1}-1\right)e^{-\beta/2}}{e^{-\beta/2}}$$

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

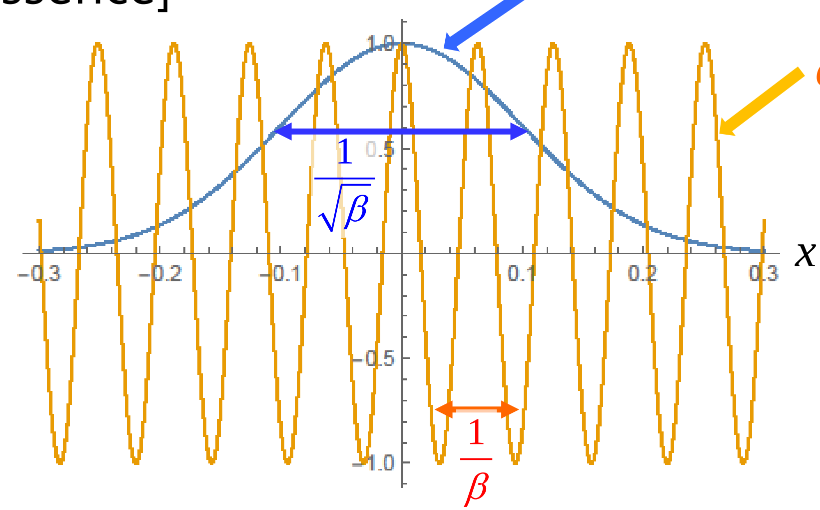
$\Rightarrow$ Necessary sample size:

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

[Essence]

$e^{-\mathrm{Re}\,S(x)} \propto e^{-\beta x^2/2}$

$e^{-i\,\mathrm{Im}\,S(x)} \propto e^{i\beta x}$



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]**

(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]**

⇩

generalized thimble method **[Alexandru et al. 2015]**

⇩

tempered Lefschetz thimble method **[MF-Umeda 2017, Alexandru et al. 2017]**

⇩

Worldvolume Hybrid Monte Carlo method **[MF-Matsumoto 2020, Fujisawa et al. 2021, ...]**

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

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

▼ tensor network (e.g. TRG) **[Levin-Nave 2007, Shimizu-Kuramashi 2014, Kadoh et al. 2020, ...]**

- good at calculating the free energy (but not so much for correl fcns)
- direct treatement of Grassmann variables

⇨ Akiyama-san's talk

Various approaches

■ **method 1 : no use of reweighting**

▼ complex Langevin method **[Parisi 1983, Klauder 1983]**

(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]**

⇩

generalized thimble method **[Alexandru et al. 2015]**

⇩

tempered Lefschetz thimble method **[MF-Umeda 2017, Alexandru et al. 2017]**

⇩

Worldvolume Hybrid Monte Carlo method **[MF-Matsumoto 2020, Fujisawa et al. 2021, ...]**

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

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

▼ tensor network (e.g. TRG) **[Levin-Nave 2007, Shimizu-Kuramashi 2014, Kadoh et al. 2020, ...]**

- good at calculating the free energy
  (but not so much for correl fcns)
- direct treatement of Grassmann variables

⇨ Akiyama-san's talk

Plan

1. Introduction
2. Sign problem and various approaches
3. Complex Langevin method
4. (Generalized) Lefschetz thimble method
5. Tempered Lefschetz thimble (TLT) method
6. Worldvolume Hybrid Monte Carlo (WV-HMC) method
7. Application to various models
8. Summary and outlook

Basic idea

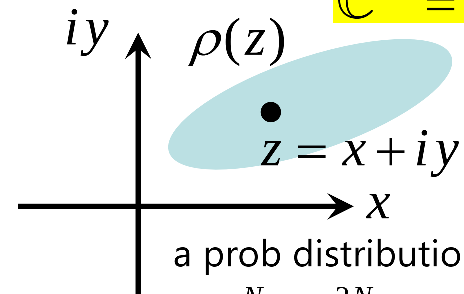
$$\frac{e^{-S(x)}}{\int_{\mathbb{R}^N} dx\, e^{-S(x)}} \in \mathbb{C}$$

does not define
a prob distribution on $\mathbb{R}^N$

$x \in \mathbb{R}^N$

$\to \ z = x + i\,y \in \mathbb{C}^N = \mathbb{R}^{2N}$

$\mathbb{C}^N = \mathbb{R}^{2N}$



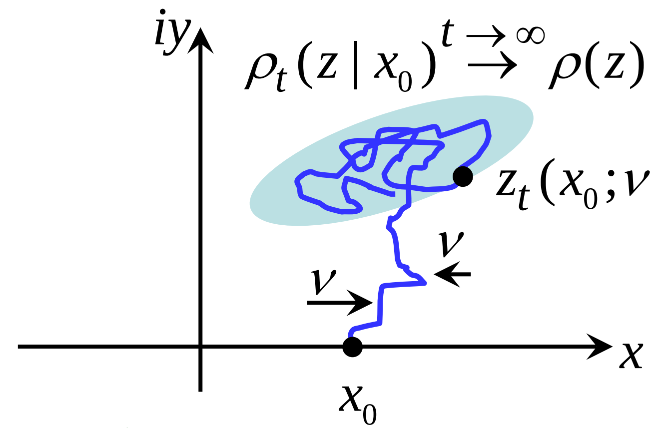
a prob distribution
in $\mathbb{C}^N = \mathbb{R}^{2N}$

■ Algorithm **[Parisi 1983, Klauder 1983]**

$$\begin{cases} \dot{z}_t = -\partial S(z_t) + \nu_t \\ z_{t=0} = x_0 \end{cases} \Rightarrow z_t = z_t(x_0;\nu)$$

ν_t : Gaussian white noise

$$\left(\langle \nu_t \nu_{t'} \rangle_\nu = 2\delta(t-t') \right)$$



$$\left(\rho_t(z\,|\,x_0) \equiv \langle \delta^{2N}(z - z_t(x_0;\nu)) \rangle_\nu \right)$$

The following holds under a certain condition:

$$\int d^{2N}z\, \rho(z)\mathcal{O}(z) \overset{(*)}{=} \frac{\int_{\mathbb{R}^N} dx\, e^{-S(x)}\mathcal{O}(x)}{\int_{\mathbb{R}^N} dx\, e^{-S(x)}} \left(= \langle \mathcal{O}(x) \rangle \right)$$

Then, we have

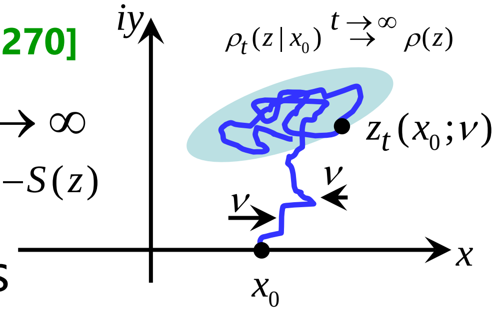
$$\langle \mathcal{O}(x) \rangle \approx \frac{1}{N_{\text{conf}}} \sum_{k=1}^{N_{\text{conf}}} \mathcal{O}(z_{\tau+k})$$

Wrong convergence problem

Condition for $(*)$ to hold : **[Aarts-James-Seiler-Stamatescu 1101.3270]**

$\rho_t(z\,|\,x_o)$: is not spread much in the direction $|\mathrm{Im}\, z| \to \infty$
(t : large) does not have large values near zeros of $e^{-S(z)}$

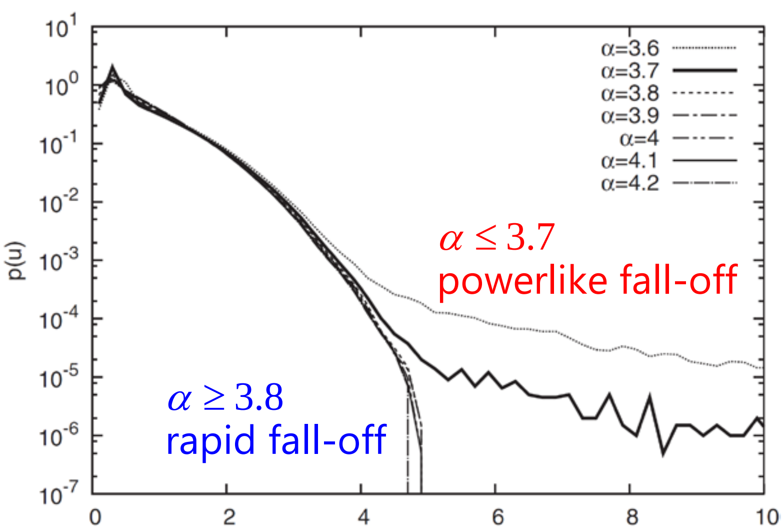
Otherwise, it gives wrong estimates with small stat errors



Explicit form **[Nagata-Nishimura-Shimasaki 1606.07627]**

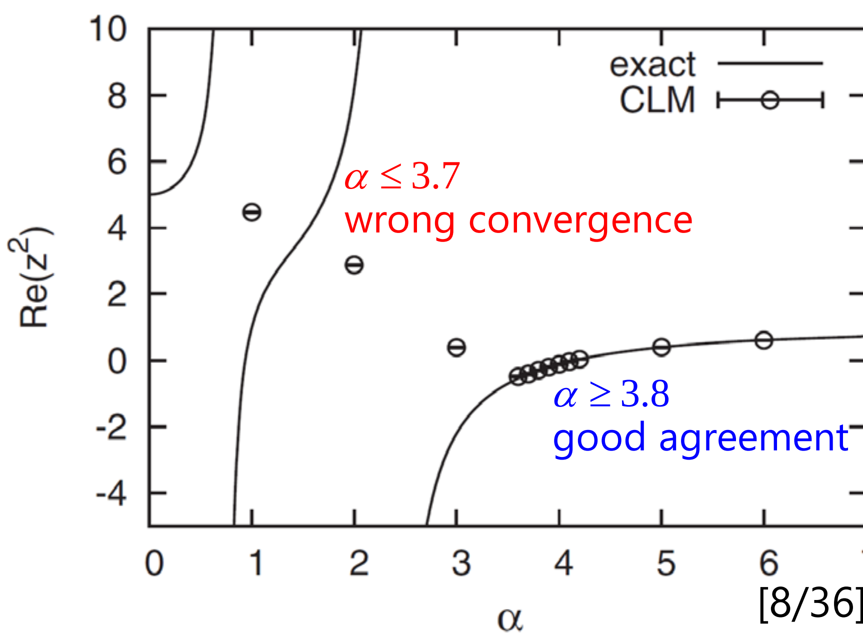
Histogram of $|\partial S(x+iy)|$ decreases rapidly (at least exponentially)

Example : $e^{-S(x)} = (x+i\alpha)^4 e^{-x^2/2}$ **[Nagata-Nishimura-Shimasaki 1606.07627]**



$u \equiv |S'(z)| \quad (z = x+iy)$

$p(u)$: normalized histogram



Plan

1. Introduction
2. Sign problem and various approaches
3. Complex Langevin method
4. (Generalized) Lefschetz thimble method
5. Tempered Lefschetz thimble (TLT) method
6. Worldvolume Hybrid Monte Carlo (WV-HMC) method
7. Application to various models
8. Summary and outlook

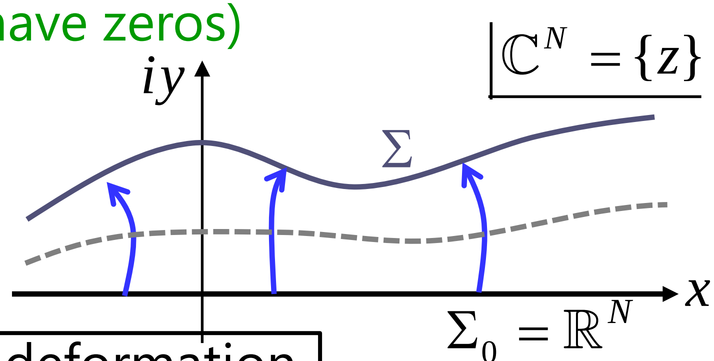
Basic idea of the thimble method (1/2)

■ 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 : $\Sigma_0=\mathbb{R}^N \;\to\; \Sigma\,(\subset\mathbb{C}^N)$

(boundary at $|x|\to\infty$ kept fixed)

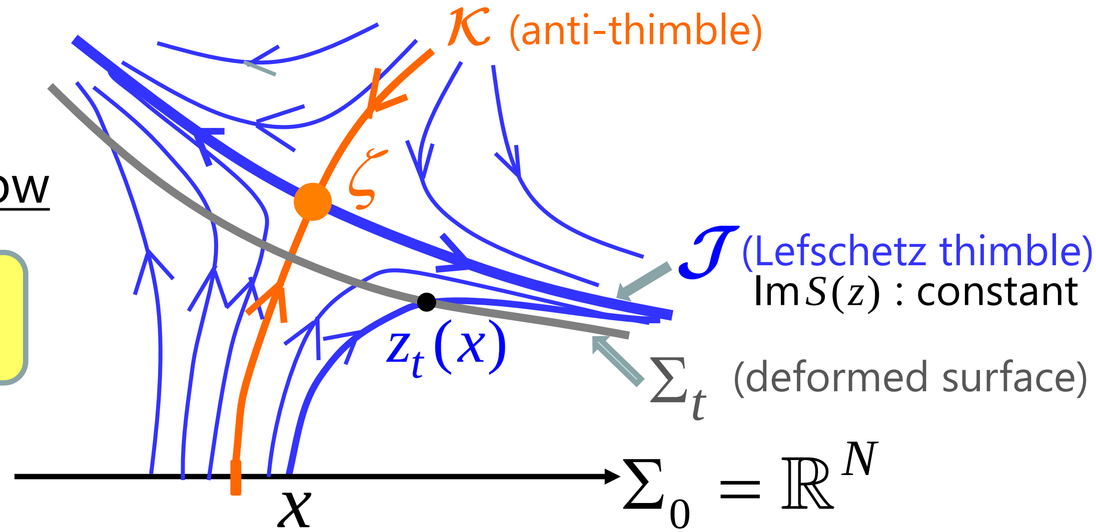
$$\langle\mathcal{O}(x)\rangle \equiv \frac{\int_{\Sigma_0} dx\, e^{-S(x)}\,\mathcal{O}(x)}{\int_{\Sigma_0} dx\, e^{-S(x)}} = \frac{\int_{\Sigma} dz\, e^{-S(z)}\,\mathcal{O}(z)}{\int_{\Sigma} dz\, e^{-S(z)}}$$

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

$$[S(z_t)]^{\cdot} = \partial S(z_t) \cdot \dot{z}_t = |\partial S(z_t)|^2 \geq 0 \quad \Rightarrow \quad \begin{cases} [\mathrm{Re}\, S(z_t)]^{\cdot} \geq 0 \\ [\mathrm{Im}\, S(z_t)]^{\cdot} = 0 \end{cases}$$

⇒ $\begin{cases} \mathrm{Re}\, S(z_t) \text{ : always increases except at crit pt } \zeta \\ \mathrm{Im}\, S(z_t) \text{ : always constant} \end{cases}$ $\left(\begin{array}{l} \zeta \text{ : crit pt} \\ \Leftrightarrow \ \partial S(\zeta) = 0 \end{array} \right)$

⇒ $\Sigma_t \xrightarrow{t \to \infty} \mathcal{J}$ (Lefschetz thimble) ≡ union 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

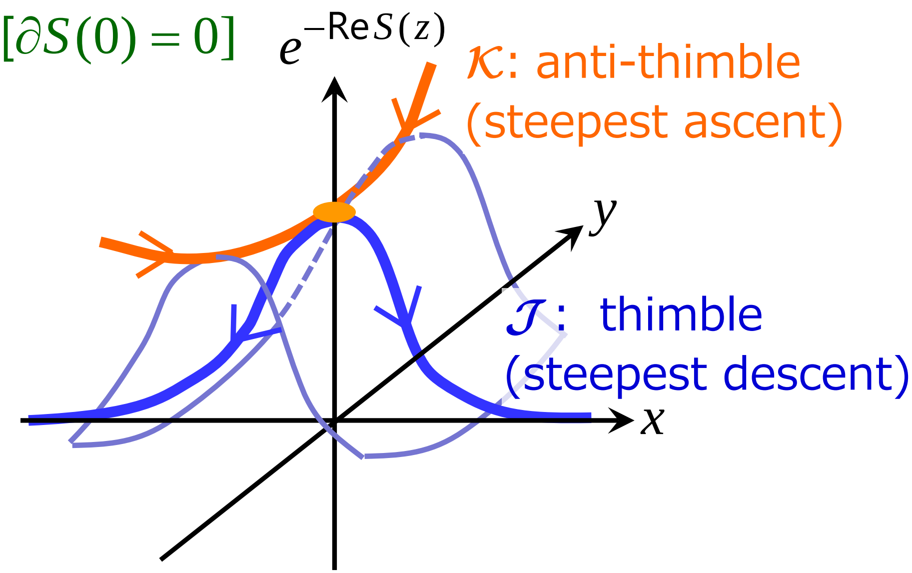
Digression：thimble＝YUBINUKI (指ぬき)

canonical form around a crit pt $\zeta = 0$ $[\partial S(0) = 0]$

■ complex dim 1 ($N = 1$)

$$S(z) = S(0) + \frac{\lambda}{2} z^2 + \cdots \quad (\lambda > 0)$$

⇨ $e^{-\mathrm{Re}\,S(z)} \propto e^{-\lambda(x^2 - y^2)/2}$ ⇨

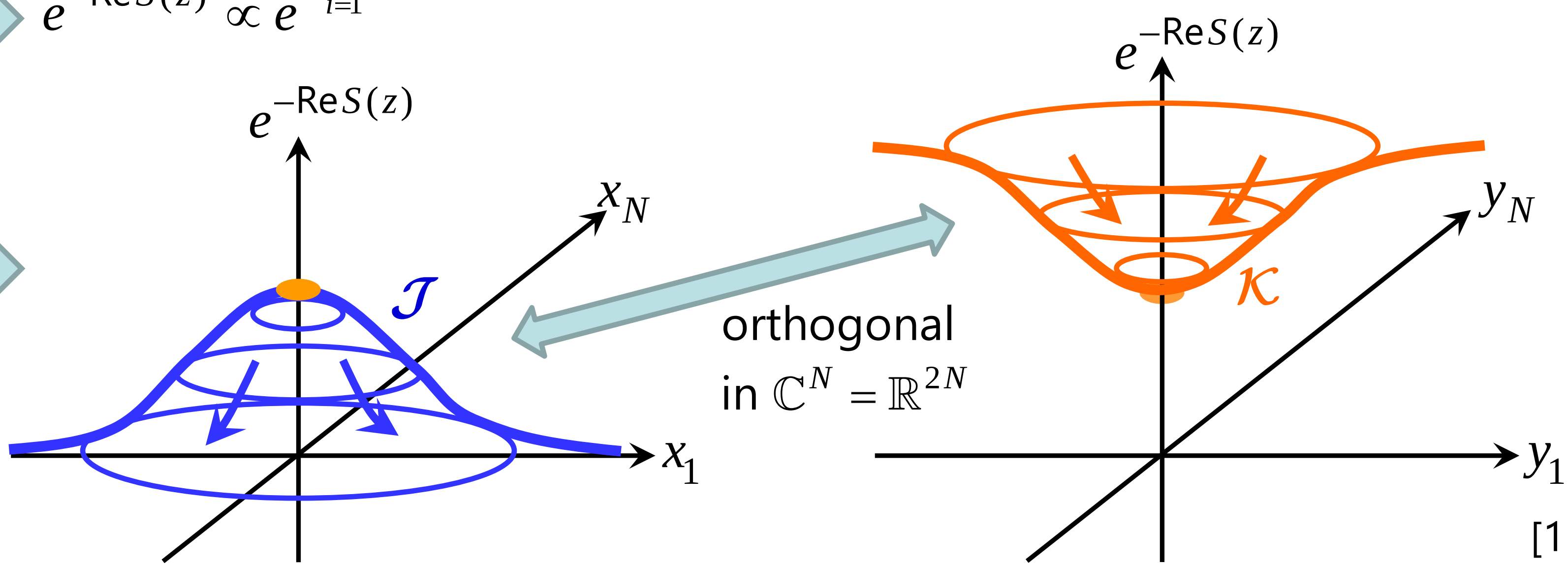


■ complex dim N

$$S(z) = S(0) + \sum_{i=1}^{N} \frac{\lambda_i}{2} z_i^2 + \cdots \quad (\lambda_i > 0)$$

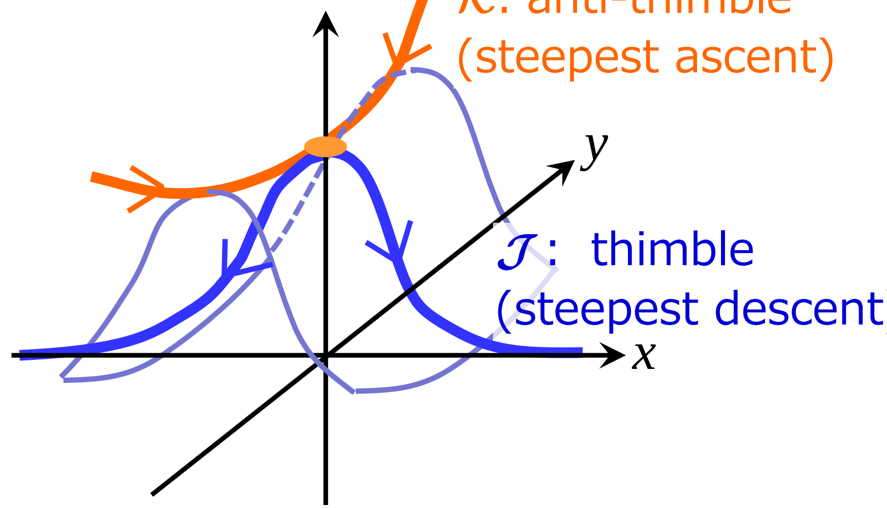
⇨ $e^{-\mathrm{Re}\,S(z)} \propto e^{-\sum_{i=1}^{N} \lambda_i (x_i^2 - y_i^2)/2}$

⇨



Digression：thimble＝YUBINUKI (指ぬき)

canonical form around a crit pt $\zeta = 0$ $[\partial S(0) = 0]$



■ complex dim 1 $(N = 1)$

$$S(z) = S(0) + \frac{\lambda}{2} z^2 + \cdots \quad (\lambda > 0)$$

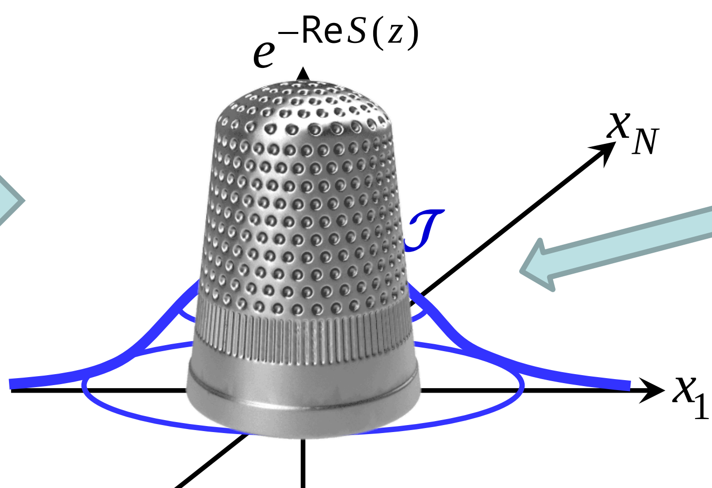
⇨ $e^{-\mathrm{Re}\,S(z)} \propto e^{-\lambda(x^2 - y^2)/2}$ ⇨

■ complex dim N

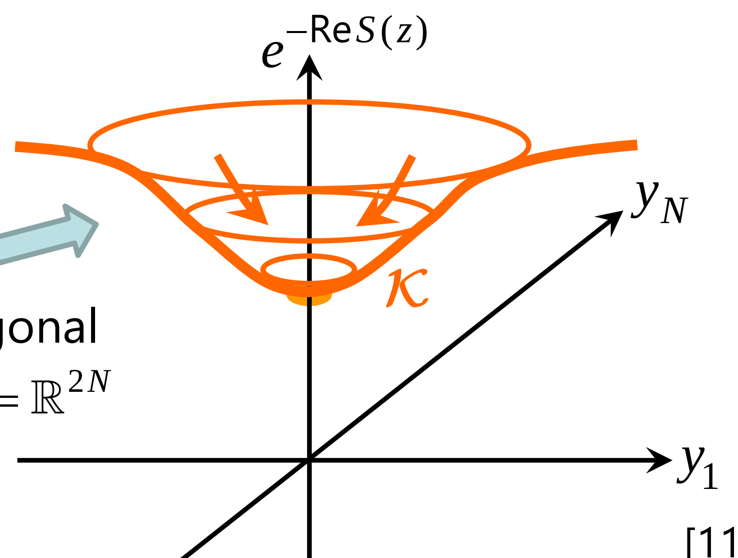
$$S(z) = S(0) + \sum_{i=1}^{N} \frac{\lambda_i}{2} z_i^2 + \cdots \quad (\lambda_i > 0)$$

⇨ $e^{-\mathrm{Re}\,S(z)} \propto e^{-\sum_{i=1}^{N} \lambda_i (x_i^2 - y_i^2)/2}$

⇨



orthogonal
in $\mathbb{C}^N = \mathbb{R}^{2N}$



How does the sign problem disappear?

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

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

$\begin{pmatrix} N \text{ : DOF} \\ N_{\mathrm{conf}} \text{ : sample size} \end{pmatrix}$

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

flow

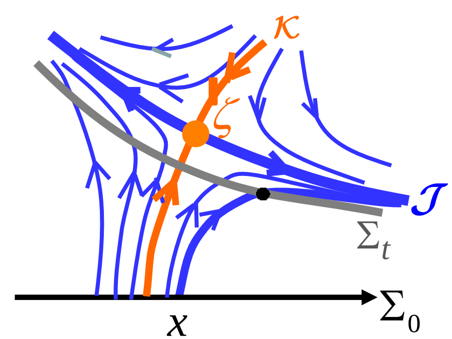
- Integration on a deformed surface Σ_t (flow time t)

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

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

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

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

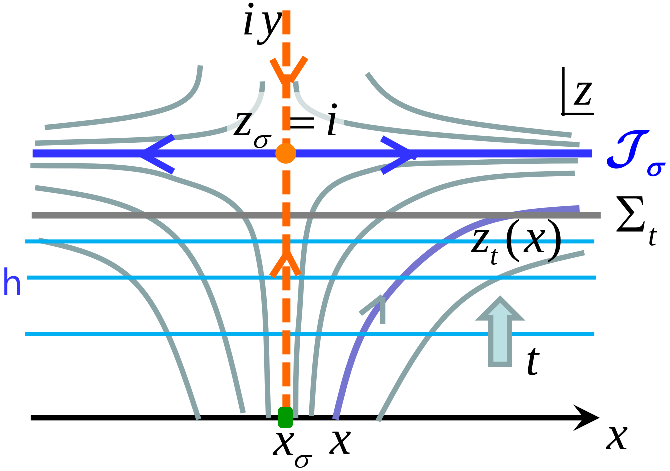


$\left[\lambda : \text{(typical) singular value of Hessian } \partial_i \partial_j S(\zeta) \right]$

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

Example: Gaussian revisited

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



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

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

exponential growth of coefficient

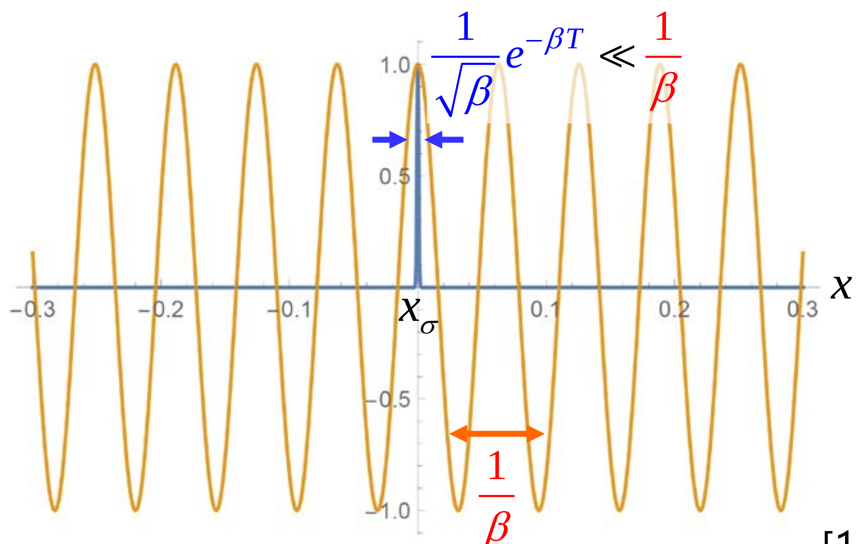
$$\Rightarrow \begin{cases} \mathrm{Re}\, S(z_t(x)) = \dfrac{1}{2} \beta e^{2\beta t} (x^2 - e^{-4\beta t}) \\ \mathrm{Im}\, S(z_t(x)) = -\beta x \ \therefore e^{i\theta(z_t(x))} = e^{i\beta x} \end{cases}$$

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

No small numbers appear

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

$$\langle x^2 \rangle = \frac{\langle e^{i\theta(z)} z^2 \rangle_{\Sigma_T}}{\langle e^{i\theta(z)} \rangle_{\Sigma_T}} = \frac{e^{-(\beta/2) e^{-2\beta T}} (\beta^{-1} - 1)}{e^{-(\beta/2) e^{-2\beta T}}} = \frac{O(1)}{O(1)}$$



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

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

Appendix: Anti-holomorphic gradient flow

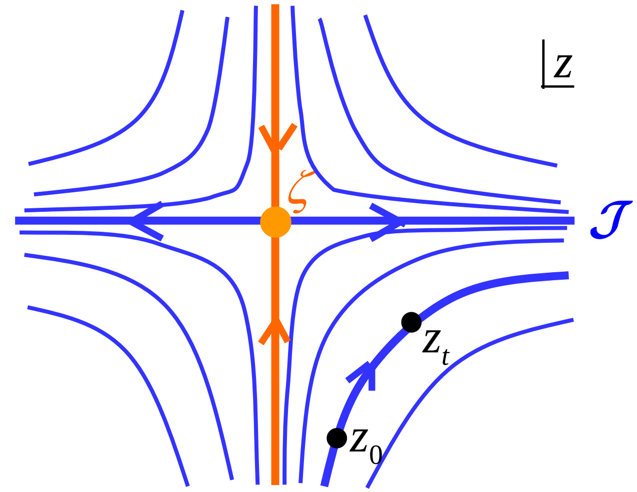
Behaviors around special points

(1) around a critical point $\zeta = \xi + i\eta$:

$$S(z) = \frac{\lambda}{2}(z-\zeta)^2 + \cdots \quad (\lambda > 0)$$

⇨ $\dot{z}_t = \overline{S'(z_t)} \simeq \lambda(\bar{z}_t - \bar{\zeta})$

⇨ $\begin{cases} x_t \simeq \xi + e^{\lambda t}(x_0 - \xi) \\ y_t \simeq \eta + e^{-\lambda t}(y_0 - \eta) \end{cases}$



ζ : critical point $\left(\partial S(\zeta) = 0\right)$

$\mathcal{J} \equiv \{z \mid z \to \zeta \ (t \to -\infty)\}$

: Lefschetz thimble associated with ζ

(2) around a zero z_* of $e^{-S(z)}$:

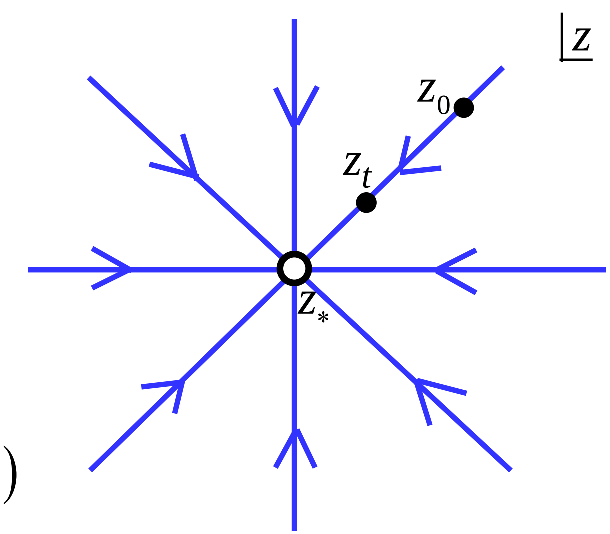
$$e^{-S(z)} = (z - z_*)^{\gamma} f(z) \quad \left(\gamma \in \mathbb{Z}_{>0}\right)$$

⇨ $S(z) = -\gamma \ln(z - z_*) + \cdots$

⇨ $\dot{z}_t = \overline{S'(z_t)} \simeq -\gamma(\bar{z}_t - \bar{z}_*)^{-1}$

⇨ $\begin{cases} |z_t - z_*|^2 \simeq |z_0 - z_*|^2 - 2\gamma t \\ \dfrac{z_t - z_*}{\bar{z}_t - \bar{z}_*} \simeq \text{const} \end{cases}$

(converges to z_* in finite flow time)



Plan

1. Introduction
2. Sign problem and various approaches
3. Complex Langevin method
4. (Generalized) Lefschetz thimble method
5. Tempered Lefschetz thimble (TLT) method
6. Worldvolume Hybrid Monte Carlo (WV-HMC) method
7. Application to various models
8. Summary and outlook

Ergodicity problem in thimble methods

large flow time t ⇨ relaxation of oscillatory integral ⇨ **Sign problem** resolved?

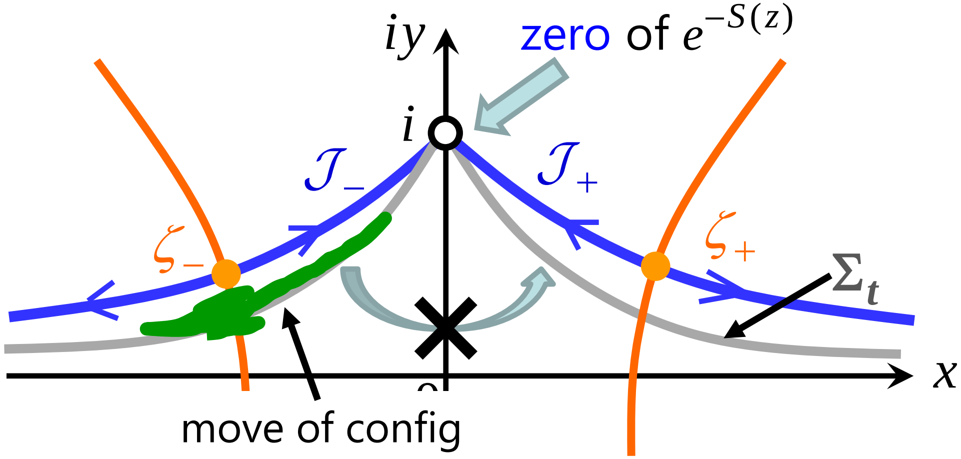
NO!

Ergodicity problem in thimble methods

large flow time t ⇒ relaxation of oscillatory integral ⇒ **Sign problem** resolved? **NO!**

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

E.g. $e^{-S(x)} = e^{-\beta x^2/2}(x-i)^\gamma \ \left(\beta \gg 1,\ \gamma \in \mathbb{Z}_{>0}\right)$ $\left(\begin{array}{l}\text{finite-density QCD :}\\ e^{-S[A]} = e^{-S_{\rm YM}[A]}\mathrm{Det}\,K[A]\end{array}\right)$



- 2 crit pts : $\zeta_\pm$
- 2 thimbles : $\mathcal{J}_\pm$

$\oplus$

- 1 zero of $e^{-S(z)}$: $z_* = i$

zero at $z_* = i$ ⇔ $\mathrm{Re}\,S(z) = +\infty$ at $z_* = i$

⇔ [∞ potential barrier on Σ_T] ⇔ configs cannot move

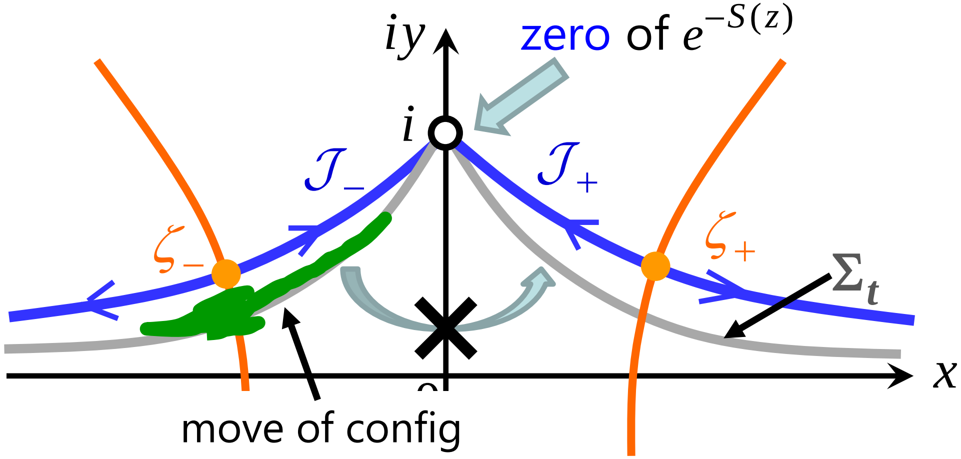
Ergodicity problem in thimble methods

large flow time t ⇨ relaxation of oscillatory integral ⇨ **Sign problem** resolved?

NO!

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

E.g. $e^{-S(x)} = e^{-\beta x^2/2}(x-i)^\gamma \;\; (\beta \gg 1,\; \gamma \in \mathbb{Z}_{>0})$ (finite-density QCD : $e^{-S[A]} = e^{-S_{\mathrm{YM}}[A]}\mathrm{Det}\, K[A]$)



- 2 crit pts : $\zeta_\pm$
- 2 thimbles : $\mathcal{J}_\pm$

$\oplus$

- 1 zero of $e^{-S(z)}$: $z_* = i$

zero at $z_* = i$ ⇔ $\mathrm{Re}\, S(z) = +\infty$ at $z_* = i$

⇔ [∞ potential barrier on Σ_T] ⇔ configs cannot move

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

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

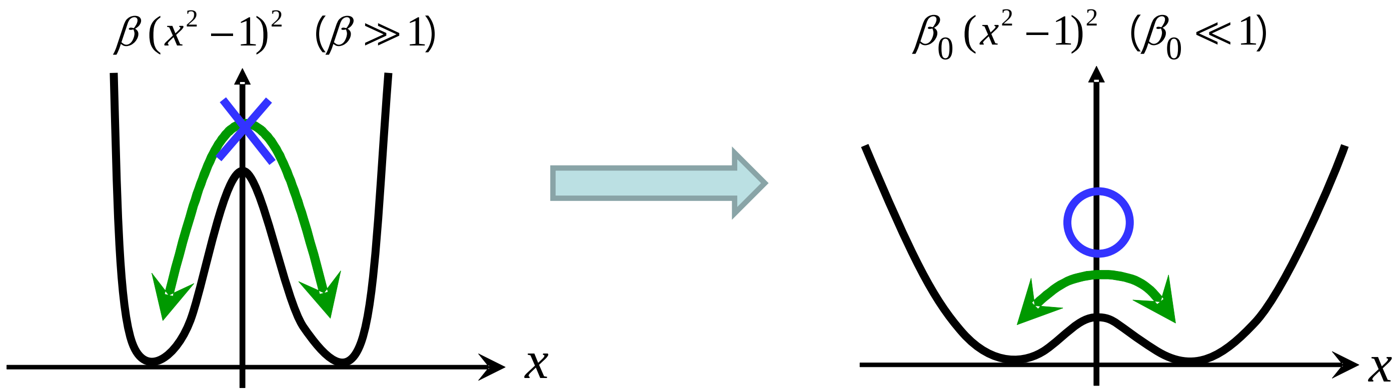
[MF-Umeda 2017] **using the flow time as a temp param**

Idea of tempering

[Marinari-Parisi 1992]

Suppose that the action $S(x;\beta)$ gives an ergodicity problem which disappears at a different value of β (say β_0)

Example： $S(x;\beta)=\beta\,(x^2-1)^2$ with $\beta \gg 1$



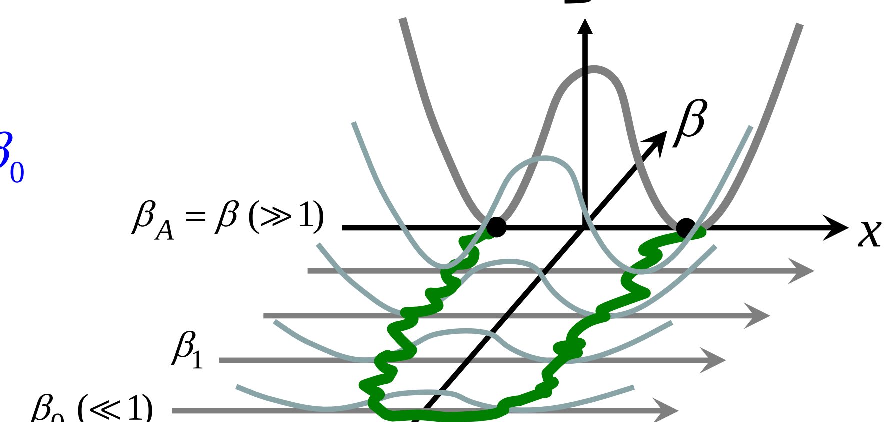
transition is difficult
due to the high potl barrier

transition is easy

Extend the config space by inserting multiple replicas between $\beta\,(\equiv \beta_A)$ and β_0

Easy transitions through detours

(Extended config space has AdS geometry **[MF-Matsumoto-Umeda 2018]**)



【parallelization】 "parallel tempering" or "replica exchange Monte Carlo"

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

Tempered Lefschetz thimble method

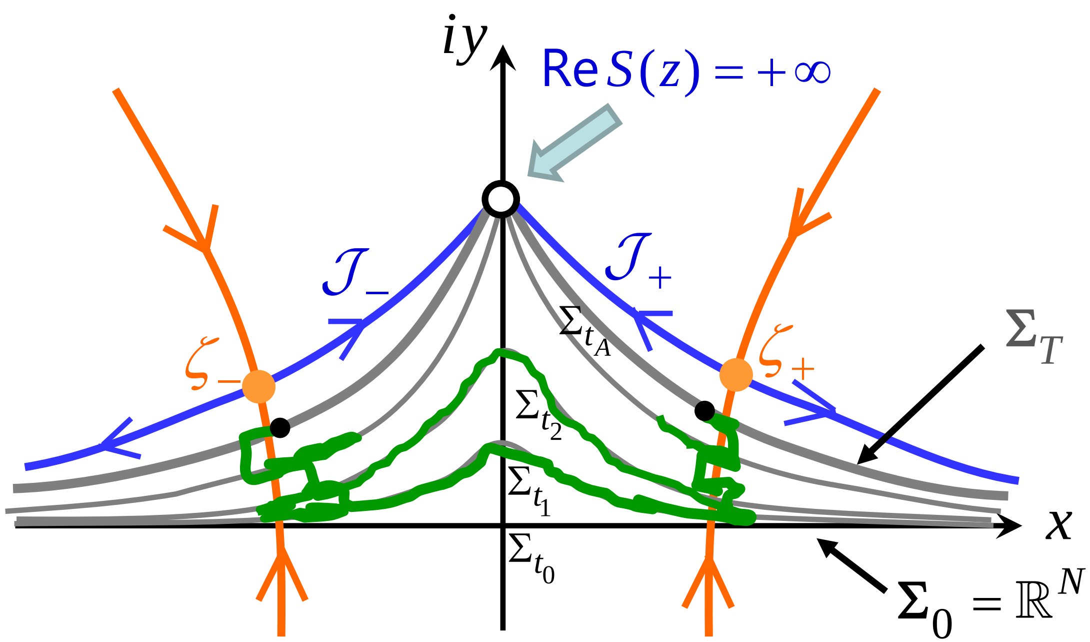
[Fukuma-Umeda 1703.00861]

■ TLT method

(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 $\{(t_a, x)\}$

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



Sign and ergodicity problems are solved simultaneously !

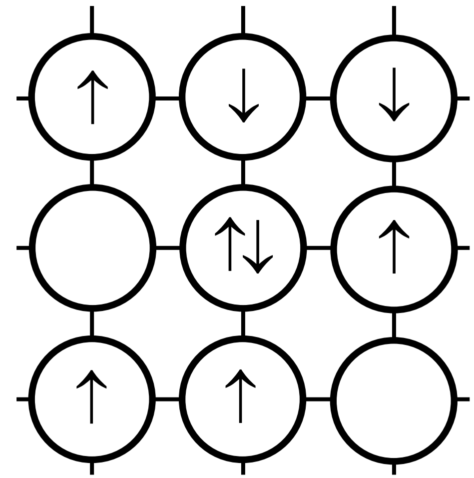
Hubbard model (1/4)

■ Hubbard model   toy model for electrons in a solid **[Hubbard 1963]**

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

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

$$\begin{cases} n_{\mathbf{x},\sigma} \equiv c^\dagger_{\mathbf{x},\sigma} c_{\mathbf{x},\sigma} \\ \kappa\,(>0) : \text{hopping parameter} \\ \mu : \text{chemical potential} \\ U\,(>0) : \text{on-site repulsive potential} \end{cases}$$



$(N_S$: # of sites$)$

- Quantum Monte Carlo (discretized imaginary time : $\beta = N_\tau \epsilon$)

  Trotter decomposition + bosonization (HS transformation)

$$\Rightarrow \quad Z_{\beta,\mu} \equiv \mathrm{tr}\, e^{-\beta H}$$

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

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

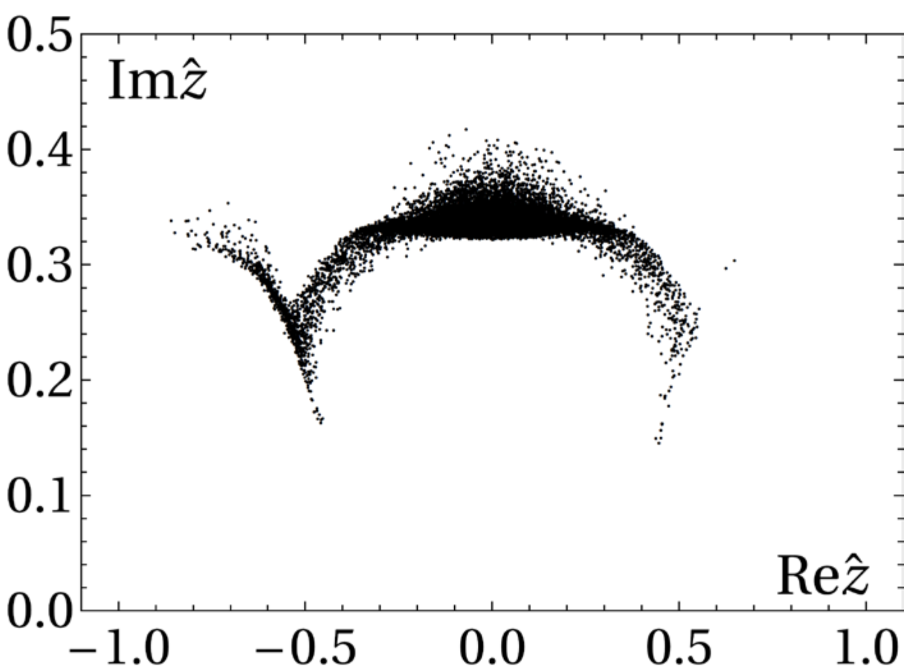
Hubbard model (2/4)

[MF-Matsumoto-Umeda 1906.04243]

scattered plot of flowed configs at $T = 0.5$ $(\beta\mu = 5)$

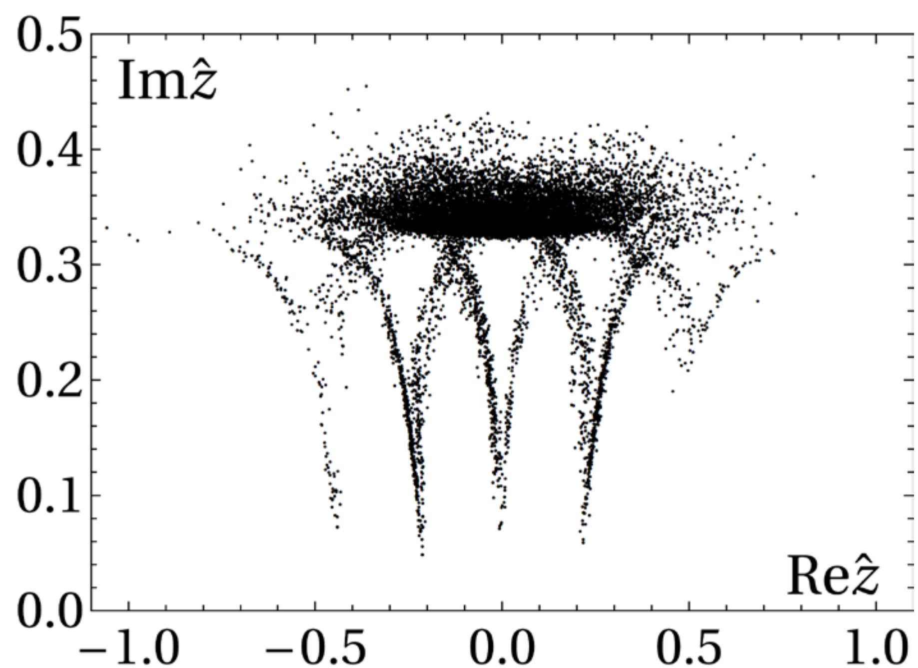
(projected on a plane $\hat{z} = (1/N)\sum_i z^i$)

w/o temp



stuck to a small # of thimbles

w/ temp

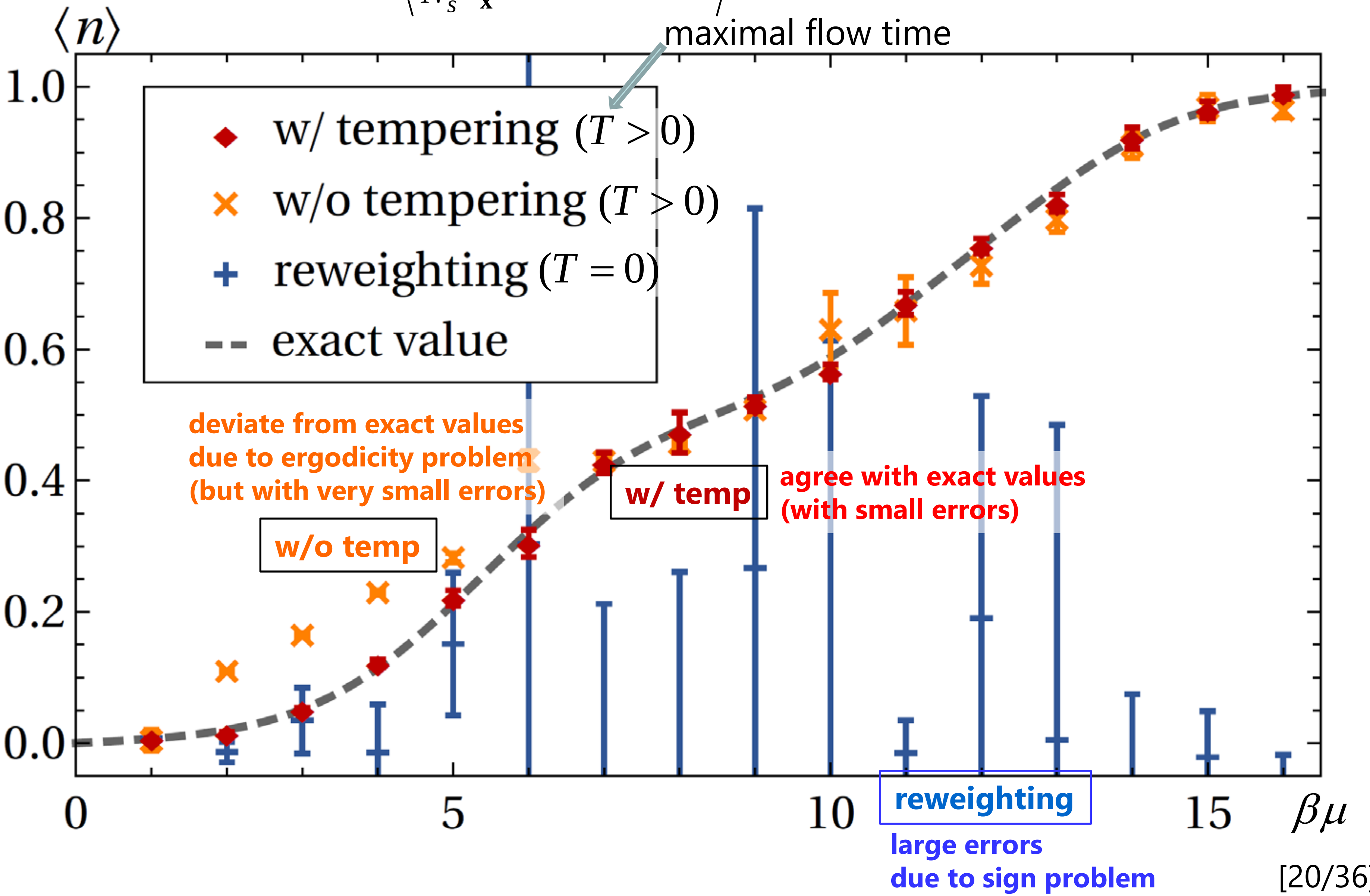


distributed widely
over many thimbles

Hubbard model (3/4)

[MF-Matsumoto-Umeda 1906.04243]

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



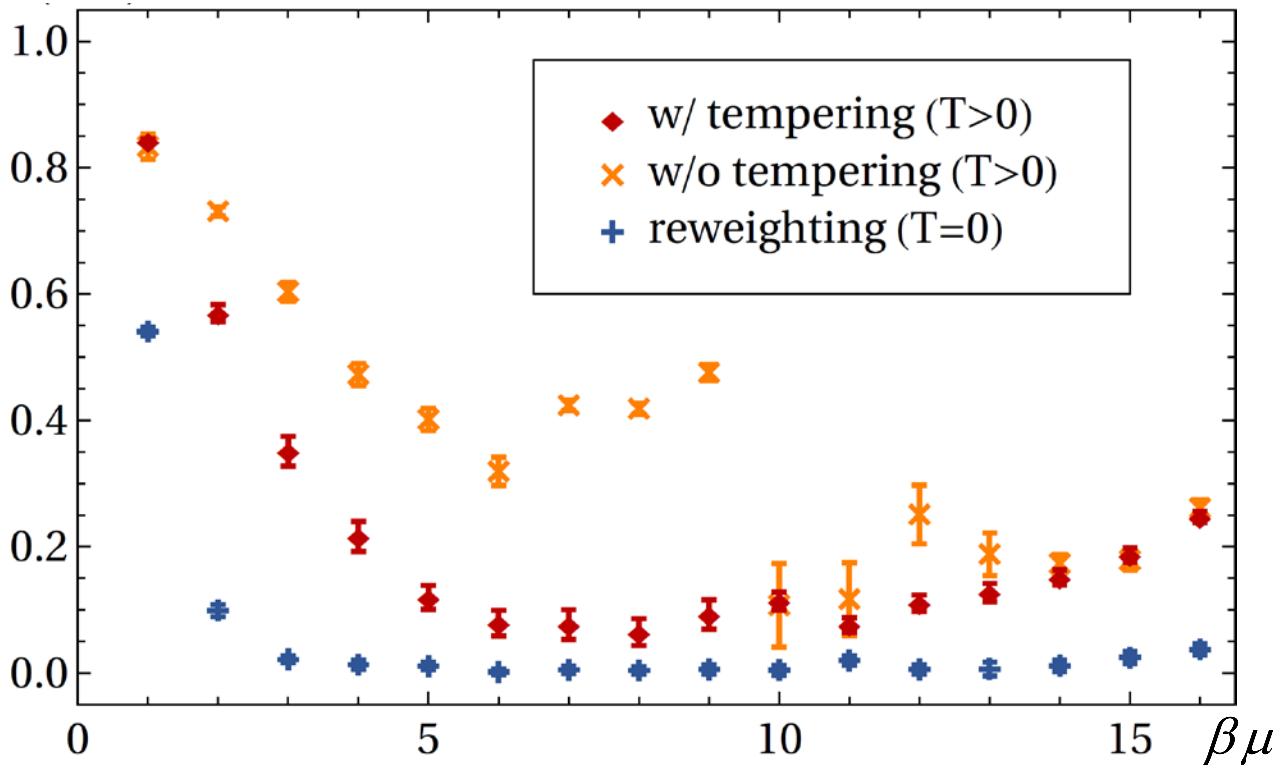
Hubbard model (4/4)

[MF-Matsumoto-Umeda 1906.04243]

average phase factor

$$\left(\langle \mathcal{O}(x) \rangle = \frac{\langle e^{i\theta(z)}\, \mathcal{O}(z) \rangle_{\Sigma_T}}{\langle e^{i\theta(z)} \rangle_{\Sigma_T}} \right)$$

$$\left| \langle e^{i\theta(z)} \rangle_{\Sigma_T} \right|$$



When only a single (or very few) thimble is sampled by mistake, the average phase factor can take a larger value (due to the lack of cancellations among different thimbles)

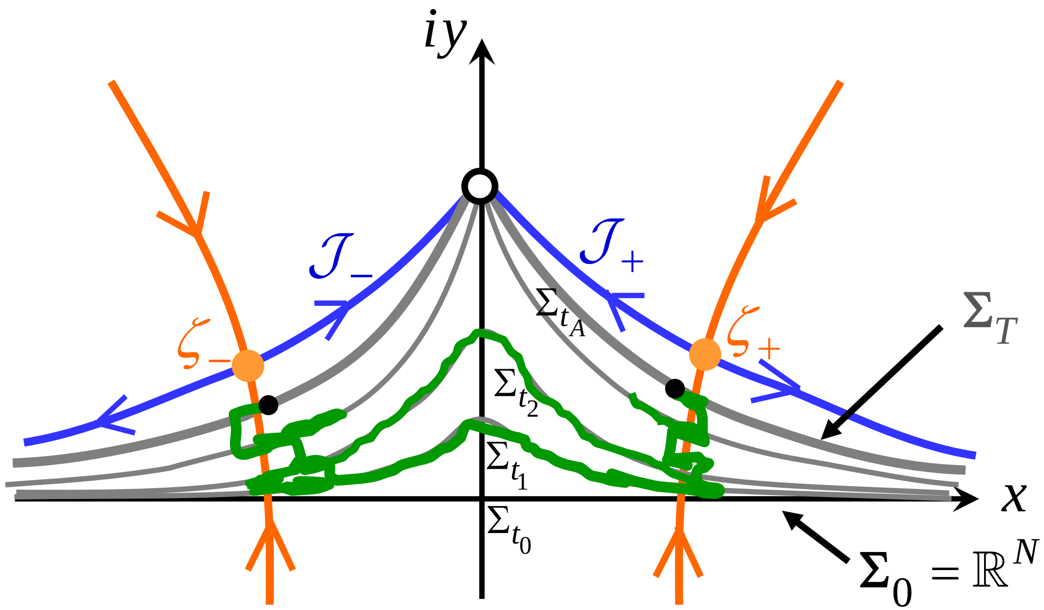
Plan

1. Introduction
2. Sign problem and various approaches
3. Complex Langevin method
4. (Generalized) Lefschetz thimble method
5. Tempered Lefschetz thimble (TLT) method
6. Worldvolume Hybrid Monte Carlo (WV-HMC) method
7. Application to various models
8. Summary and outlook

Pros and cons of the original TLT method

■ <u>TLT method</u> **[MF-Umeda 2017]**

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



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

<u>Cons</u> : 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

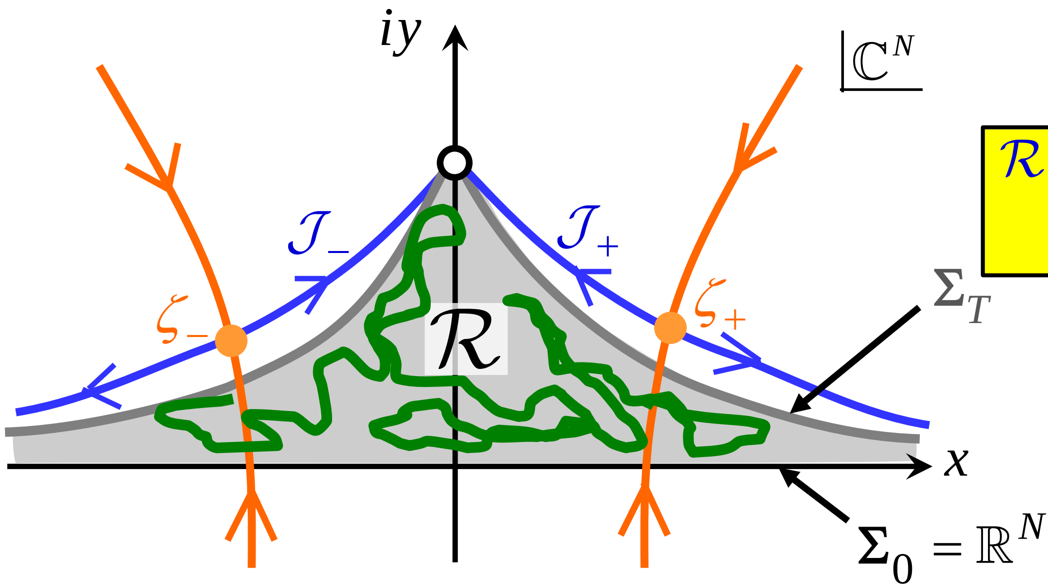
Worldvolume HMC (1/2)

[MF-Matsumoto 2012.08468]

■ <u>Worldvolume Hybrid Monte Carlo</u> (WV-HMC)

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" $\mathbb{C}^N = \mathbb{R}^{2N}$

$$\begin{pmatrix} \text{orbit of particle} & \to & \text{worldline} \\ \text{orbit of string} & \to & \text{worldsurface} \\ \text{orbit of surface (membrane)} & \to & \text{worldvolume} \end{pmatrix}$$

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

$\oplus$ 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 HMC (2/2)

[MF-Matsumoto 2012.08468]

■ mechanism

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

⇦ t-independent

⇦ t-independent

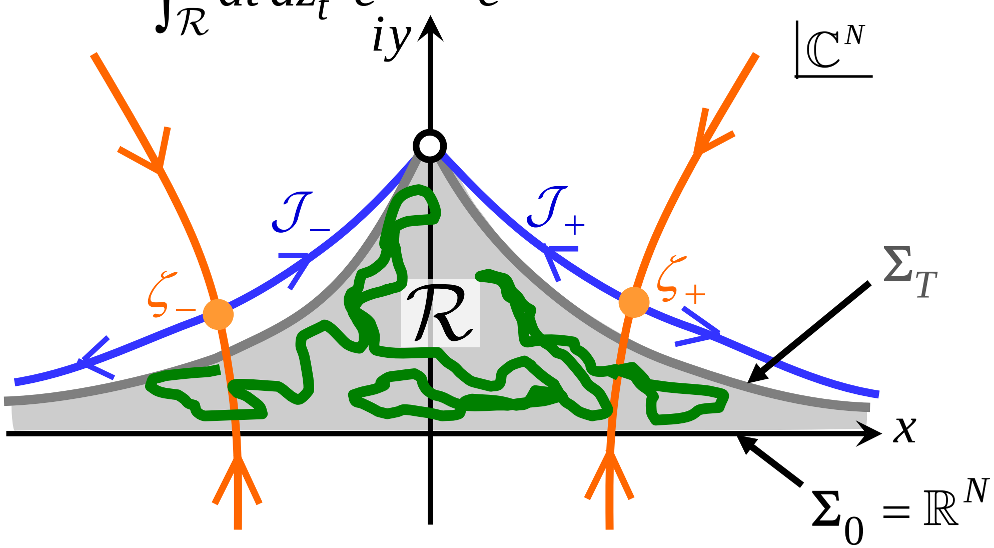
$$= \frac{\int_0^T dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\; e^{-S(z_t)} \mathcal{O}(z_t)}{\int_0^T dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\; e^{-S(z_t)}}$$

$(W(t)$: arbitrary fcn$)$

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

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

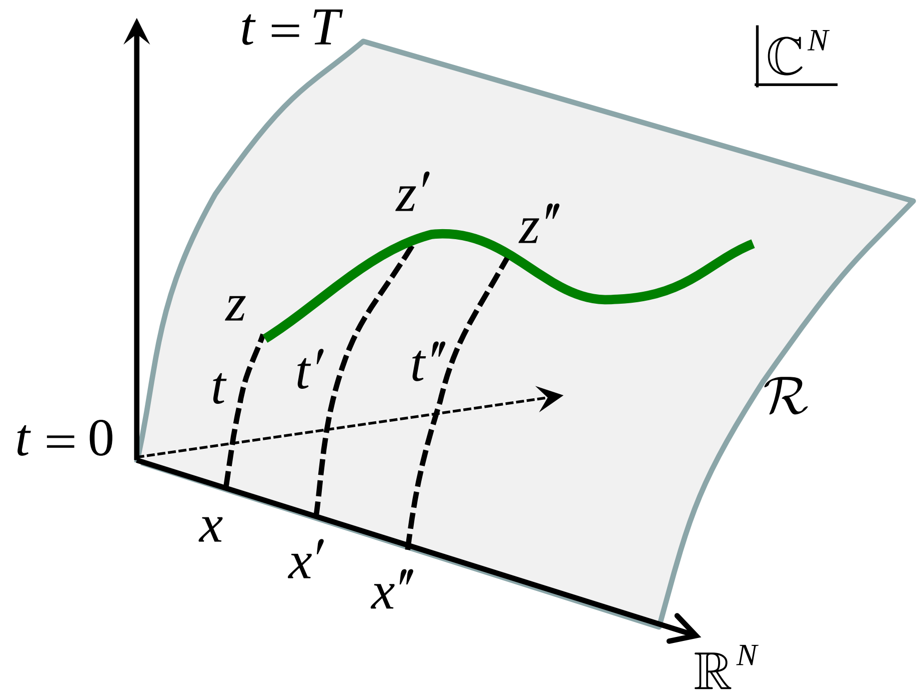
$\Leftarrow$ path integral over the worldvolume $\mathcal{R}$



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

Two pictures in WV-HMC

[MF-Matsumoto 2012.08468]



(1) **<u>Target-space picture</u>**

[MF-Matsumoto 2012.08468]

sample: $\{z, z', z'', \ldots\}$

(2) **<u>Parameter-space picture</u>**

[MF-Matsumoto 2012.08468]

[Fujisawa et al. 2112.10519]

sample: $\{(t, x), (t', x'), (t'', x''), \ldots\}$

At first sight, (2) may seem simpler,
but actually (1) is faster and more solid as an algorithm

We employ (1) <u>target-space picture</u>

Computational cost of WV-HMC

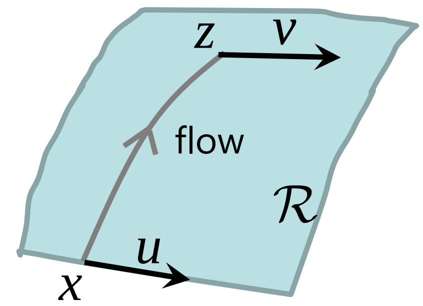
[MF-Matsumoto 2012.08468]
[MF-Matsumoto-Namekawa, Lattice2022]

$z=(z^i)\in\mathbb{C}^N \quad (N\propto V$: DOF)

<u>Configuration flow</u> $\quad \dot{z}_i=\overline{\partial_i S(z)} \quad \Rightarrow O(N)$

<u>Vector flow</u> $\quad \dot{v}_i=\overline{\partial_i\partial_j S(z)\,v_j} \quad \Rightarrow O(N^2)$ [when $\partial_i\partial_j S(z)$ is dense]

$\Rightarrow O(N)$ [when $\partial_i\partial_j S(z)$ is sparse (local field case)]



<u>RATTLE</u>
$$\begin{cases} z'=z+\Delta s\,\pi-\Delta s^2\,\overline{\partial V(z)}-\lambda \\ \pi_{1/2}=(z'-z)/\Delta s \\ \tilde{\pi}'=\pi_{1/2}-\Delta s\,\overline{\partial V(z')} \\ \pi'=\Pi'_{\mathcal{R}}\tilde{\pi}' \end{cases}$$

$\big(V(z)=\mathrm{Re}S(z)+W(t(z))\big)$

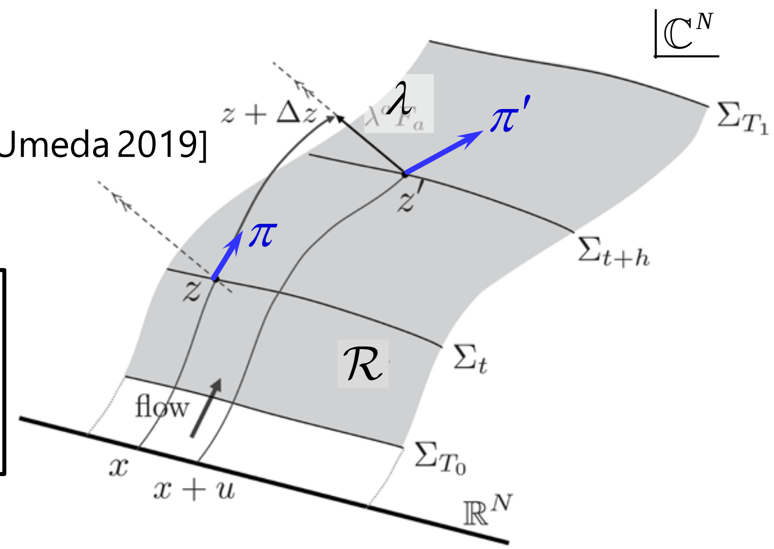
cf) RATTLE on a single thimble $\mathcal{J}=\Sigma_\infty$ [Fujii et al. 2013]
RATTLE on Σ_t [Alexandru@Lattice2019, MF-Matsumoto-Umeda 2019]

$\lambda\in N_z\mathcal{R}$ is determined s.t. $z'\in\mathcal{R}$

⇔

For given $z=z_t(x)$ and π,
find $h\in\mathbb{R}$, $u\in\mathbb{R}^N$, $\lambda\in N_z\mathcal{R}$
s.t. $z_t(x)+\Delta s\,\pi-\Delta s^2\,\overline{\partial V(z)}-\lambda=z_{t+h}(x+u)$

This can be solved by Newton's method
with BiCGStab for linear inversion
(which requires only config/vector flows) $\Rightarrow O(N)$



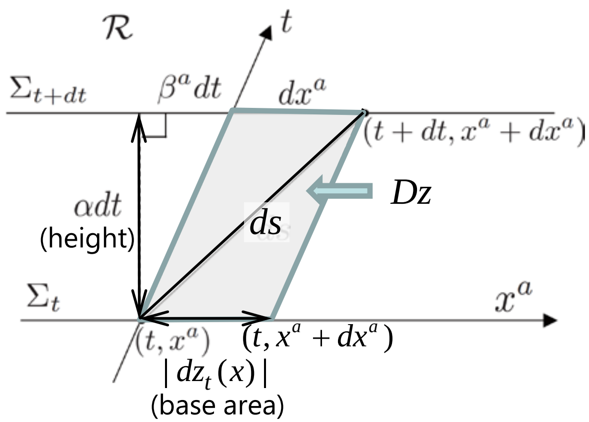
Comput cost at each MD step is expected to be $O(N)$ for local field theory (with the absence of fermion determinant)

Appendix: Details of WV-HMC (1/2)

[MF-Matsumoto 2012.08468]

■ Preparation

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



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

= vol element Dz of the induced metric

ADM decomposition

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

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

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

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

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

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

Appendix: Details of WV-HMC (2/2)

[MF-Matsumoto 2012.08468]

■ Algorithm

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

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

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

$$\left(e^{i\varphi(z)} \equiv \frac{\det J}{|\det J|} \right)$$

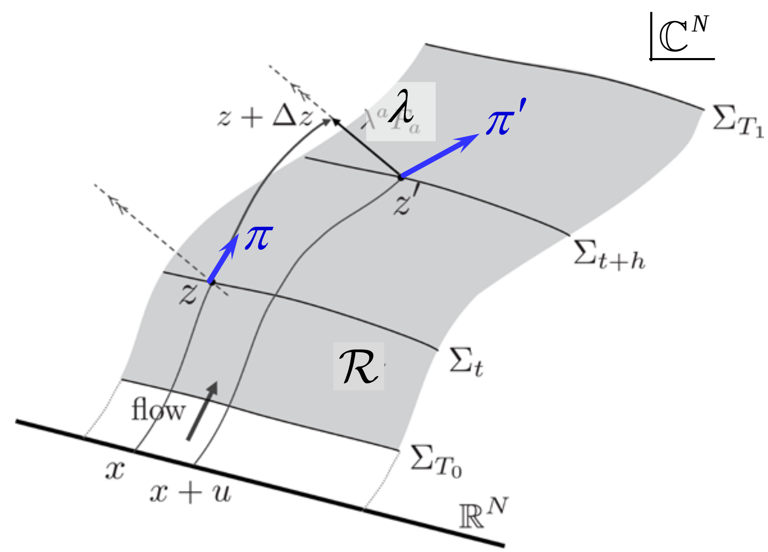
HMC on a constrained space **[Andersen 1983, Leimkuhler-Skeel 1994]**

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

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

$\lambda \in N_z\mathcal{R}$ and $\lambda' \in N_{z'}\mathcal{R}$ are determined s.t.

$$\begin{cases} z' \in \mathcal{R} \\ \pi' \in T_{z'}\mathcal{R} \end{cases}$$



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

Plan

1. Introduction
2. Sign problem and various approaches
3. Complex Langevin method
4. (Generalized) Lefschetz thimble method
5. Tempered Lefschetz thimble (TLT) method
6. Worldvolume Hybrid Monte Carlo (WV-HMC) method
7. Application to various models
8. Summary and outlook

Successfully applied to ...

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

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

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

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

— complex scalar field at finite density **[MF-Namekawa 2022, in preparation]** **(WV-HMC)**

So far always successful for any models when applied,
though the system sizes are not yet very large (DOF $N \lesssim 10^4$)

Successfully applied to ...

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

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

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

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

— complex scalar field at finite density **[MF-Namekawa 2022, in preparation]** **(WV-HMC)**

So far always successful for any models when applied,
though the system sizes are not yet very large (DOF $N \lesssim 10^4$)

Chiral random matrix model (1/2)

[MF-Matsumoto 2012.08468]

■ finite density QCD

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

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

⇩ toy model

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

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

(quantum field replaced by a matrix incl spacetime DOF)

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

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

$\left(\text{DOF} : N = 2n^2 \ \Leftrightarrow\ 4L^4 (N_c^2 - 1)\right)$

■ role as an important benchmark model

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

Chiral random matrix model (2/2)

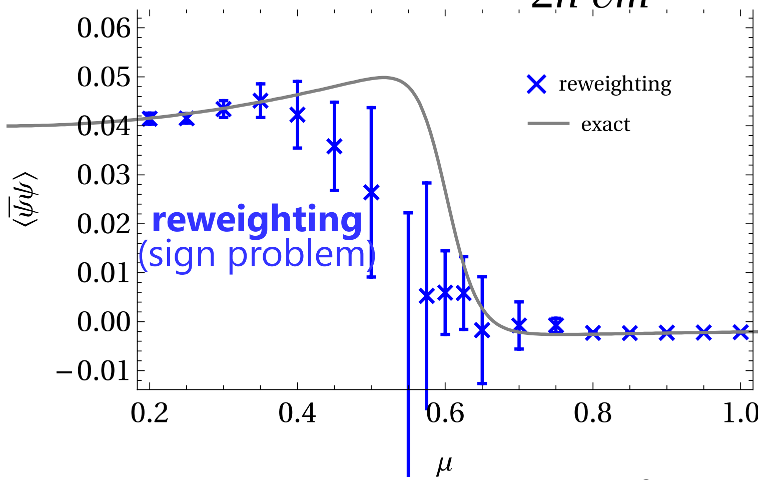
[MF-Matsumoto 2012.08468]

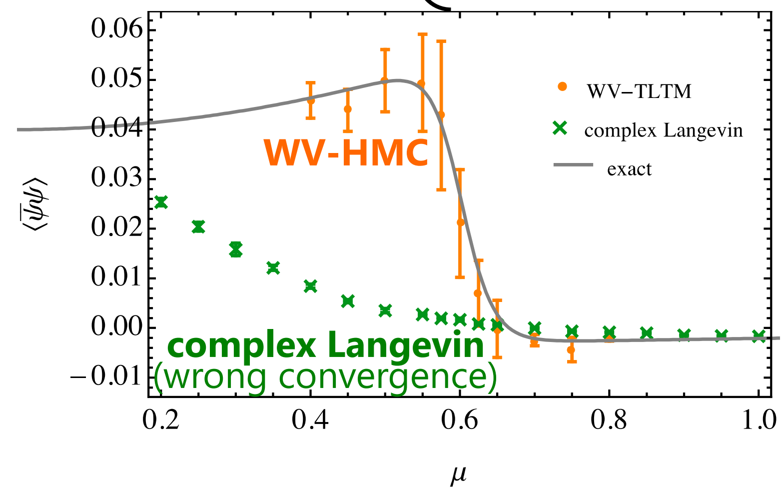
matrix size : $n=10$ (DOF : $N=200$)

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

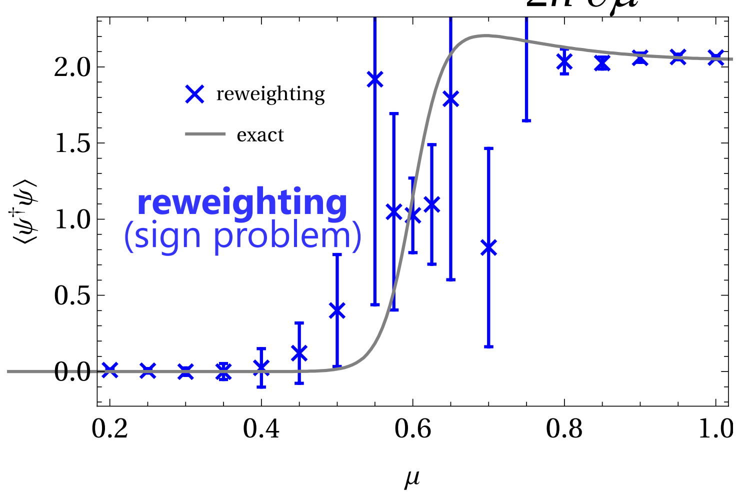
sample size
reweighting : 10k
complex Langevin :10k
WV-TLTM : 4k-17k

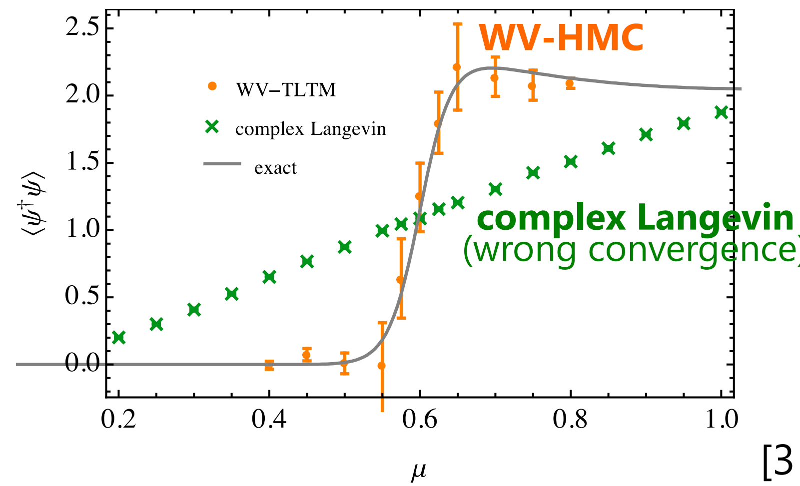
chiral condensate $\langle\bar{\psi}\psi\rangle \equiv \dfrac{1}{2n}\dfrac{\partial}{\partial m}\ln Z_{\text{Steph}}$ $[m=0.004,\ T=0]$





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





Finite-density complex scalar field (1/3)

$\varphi(x) = \frac{1}{\sqrt{2}}[\varphi_1(x) + i\varphi_2(x)]$: complex scalar field

Continuum action

(x_0 : Euclidean time)

$$S(\varphi) = \int d^d x \left[\partial_\nu \varphi^* \partial_\nu \varphi + m^2 \varphi^* \varphi + \lambda (\varphi^* \varphi)^2 + \mu (\varphi^* \partial_0 \varphi - \partial_0 \varphi^* \varphi) \right]$$
$$\simeq \int d^d x \left[(\partial_\nu \varphi^* + \mu \delta_{\nu,0} \varphi^*)(\partial_\nu \varphi - \mu \delta_{\nu,0} \varphi) + m^2 |\varphi|^2 + \lambda |\varphi|^4 \right]$$

Lattice action [Aarts 0810.2089]

$$S(\varphi) = \sum_n \left[(2d + m^2) |\varphi_n|^2 + \lambda |\varphi_n|^4 - \sum_{\nu=0}^{d-1} (e^{\mu \delta_{\nu,0}} \varphi_n^* \varphi_{n+\nu} + e^{-\mu \delta_{\nu,0}} \varphi_n \varphi_{n+\nu}^*) \right]$$

Introducing (ξ_n, η_n) with $\varphi_n = \frac{1}{\sqrt{2}}(\xi_n + i\eta_n)$, we have

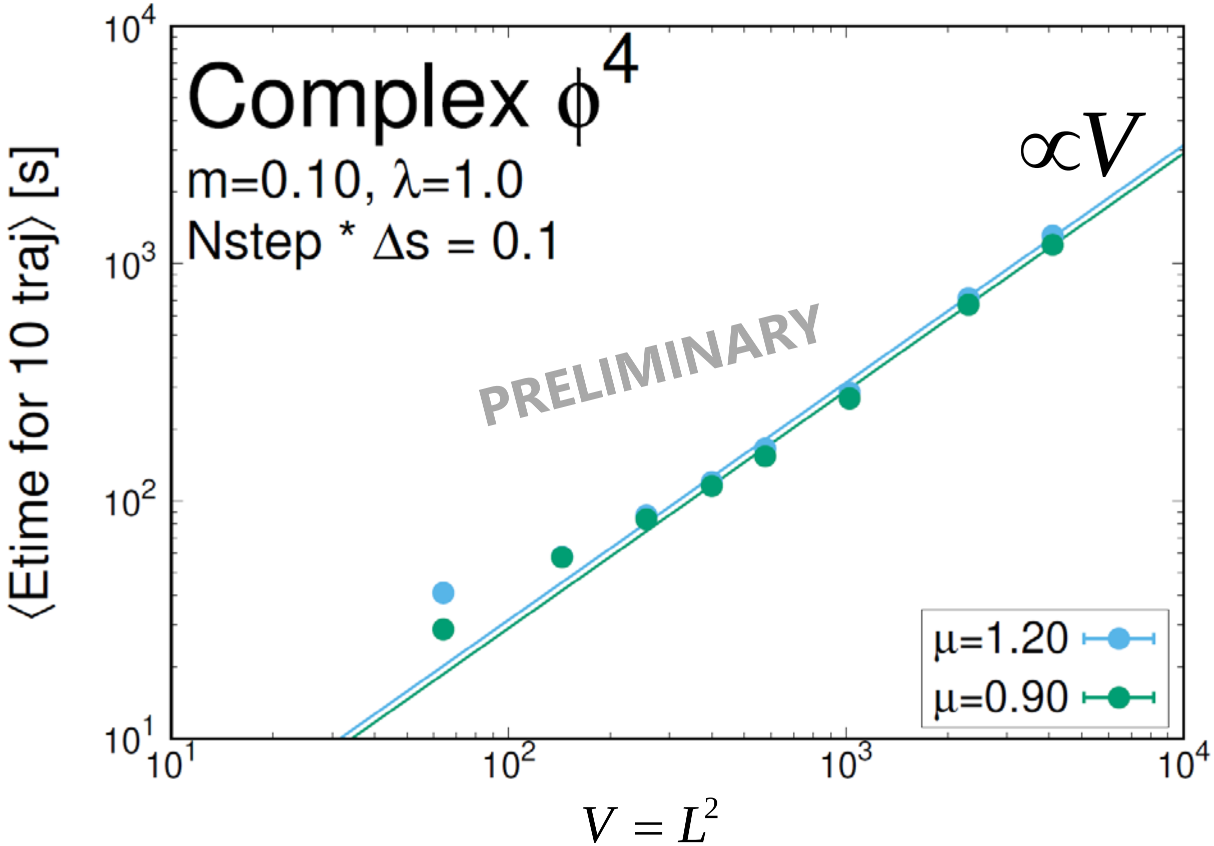
$$S(\xi,\eta) = \sum_n \left[\begin{array}{l} \frac{2d + m^2}{2}(\xi_n^2 + \eta_n^2) + \frac{\lambda}{4}(\xi_n^2 + \eta_n^2)^2 - \sum_{i=1}^{d-1} (\xi_{n+i}\xi_n + \eta_{n+i}\eta_n) \\ - \cosh\mu (\xi_{n+0}\xi_n + \eta_{n+0}\eta_n) - i \sinh\mu (\xi_{n+0}\eta_n - \eta_{n+0}\xi_n) \end{array} \right]$$

We complexify $(\xi,\eta) \in \mathbb{R}^{2V}$ to $(z,w) \in \mathbb{C}^{2V}$ with the flow equation

$$\dot{z}_n = [\partial S(z,w)/\partial z_n]^*, \quad \dot{w}_n = [\partial S(z,w)/\partial w_n]^* \quad \left(\begin{array}{l} V: \text{ lattice volume} \\ \quad \Rightarrow N = 2V \end{array} \right)$$

■ Computational scaling in 2D

[MF-Matsumoto-Namekawa, Lattice2022]



computed on Yukawa21
(@YITP, Kyoto Univ)

⇨ to be run on Fugaku

The figure clearly shows that the comput cost scales as $O(V) = O(N)$
$(N = 2V)$

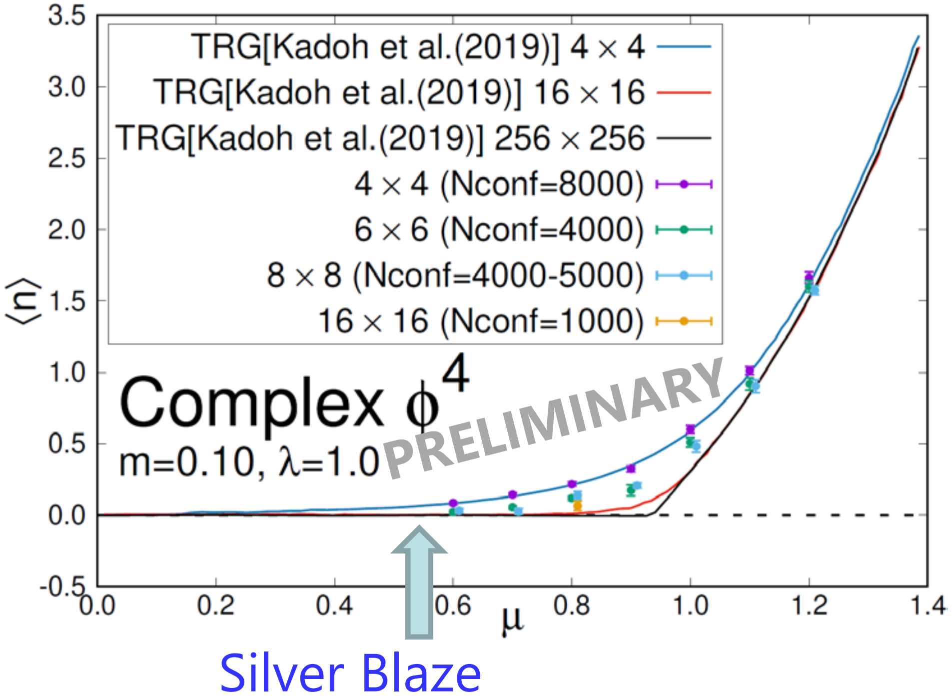
(NB: The scaling will become $O(V^{1.25})$
if we reduce the MD stepsize as $\Delta s \propto V^{-1/4}$
to keep the same amount of acceptance for increasing volume)

Finite-density complex scalar field (3/3)

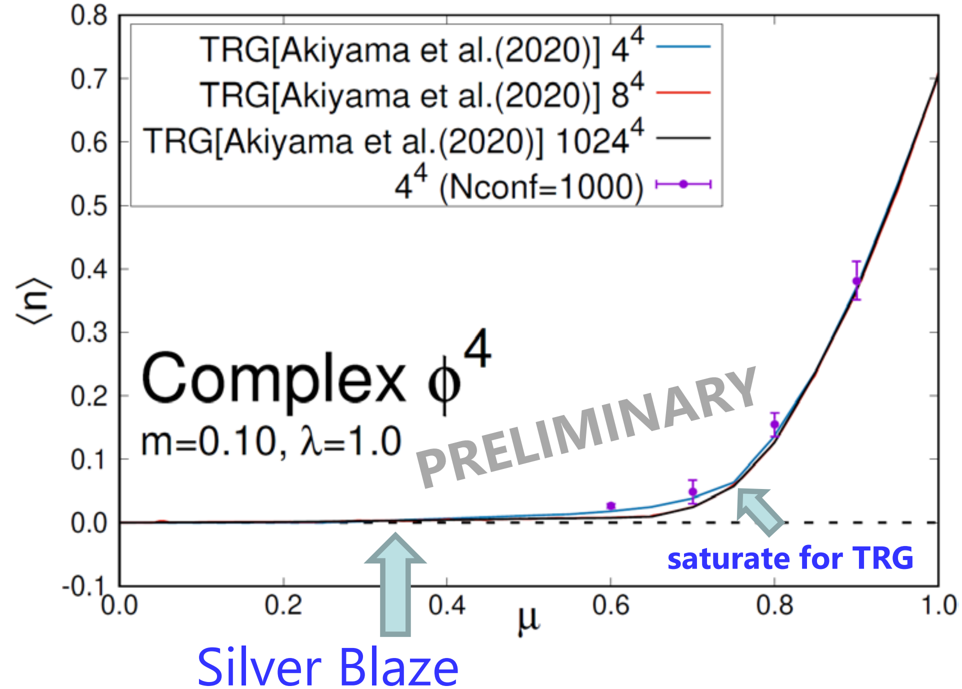
[MF-Namekawa, in preparation]

■Comparison with TRG method **[TRG (2D): Kadoh et al. 1912.13092]**
[TRG (4D): Akiyama et al. 2005.04645]

2D



4D



- WV-HMC gives results consistent with TRG at small volumes
- It should be interesting to investigte behaviors at large 4D volumes where TRG starts suffering from the systematic errors due to the introduction of D_{cut}

Yang-Mills theory has become a target

WV-HMC also works for group manifolds

[MF, in preparation]

[MF, in preparation]

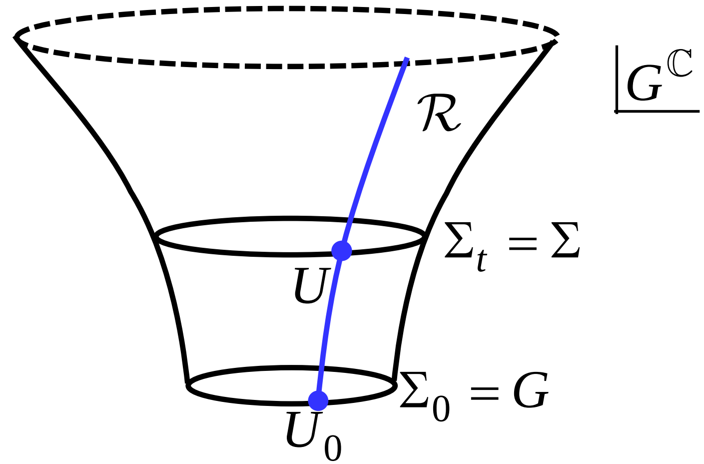
[MF-Kanamori-Namekawa, ongoing]

[Example]

1-site model with a pure imaginary coupling:

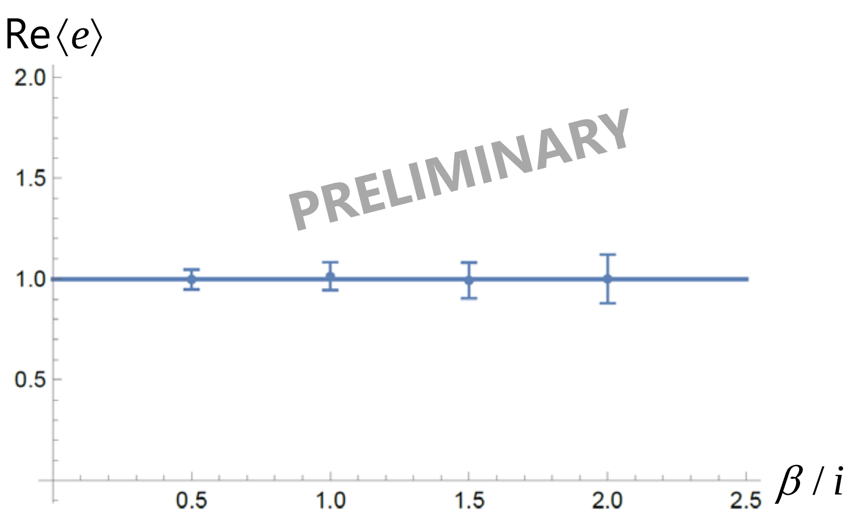
$$S(U) \equiv \beta\, e(U) \equiv \frac{\beta}{2N}\,\mathrm{tr}\left(2 - U - U^{-1}\right)$$

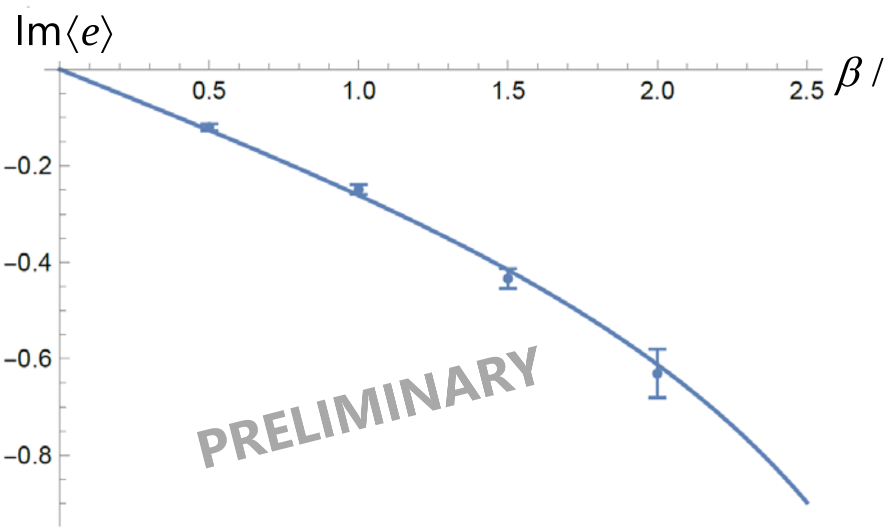
$$\left(U \in G = SU(2)\,;\;\; \beta \in i\,\mathbb{R}\right)$$



analytic result: $\langle e \rangle = 1 - I_2(\beta) / I_1(\beta)$

numerical result (WV-HMC):





This algorithm can be applied to Yang-Mills theory straightforwardly

[MF-Kanamori-Namekawa, ongoing]

Plan

1. Introduction
2. Sign problem and various approaches
3. Complex Langevin method
4. (Generalized) Lefschetz thimble method
5. Tempered Lefschetz thimble (TLT) method
6. Worldvolume Hybrid Monte Carlo (WV-HMC) method
7. Application to various models
8. Summary and outlook

Summary and outlook

■ Summary

▼ WV-HMC seems to be a promising method esp. for local field theories

- Allows a coding such as to enjoy the locality
- Does not have a wrong convergence problem (as in the complex Langevin)
- Can increase the precision to arbitrarily high order by increasing the sample size

  (No systematic errors such as those caused by introducing D_{cut} in TRG)

"Power of Monte Carlo"

■ Outlook

▼ Application to QCD

- WV-HMC for a path integration on a group manifold **[MF, in preparation]**
- WV-HMC for pure Yang-Mills with finite θ **[MF-Kanamori-Namekawa, ongoing]**
- WV-HMC for finite-density QCD **[MF-Kanamori-Namekawa-..., ongoing]**

▼ Further improvement of algorithm **[MF, MF-Matsumoto-Namekawa-..., ongoing]**

▼ Combining various algorithms

(e.g.) path optimization method and/or TRG (non-MC) **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 those in the early universe, heavy ion collision experiments, ...

Thank you.