The basics and applications of the tempered Lefschetz thimble method for the numerical sign problem

Masafumi Fukuma (Dept Phys, Kyoto Univ)

@Lattice 2021 (online)

Based on work with

Nobuyuki Matsumoto, Yusuke Namekawa and Naoya Umeda
(RIKEN/BNL) (YITP, 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" [arXiv:2107.06858]

1. Introduction

Numerical sign problem

Numerical sign problem:

has prevented the first-principles analysis of physically important systems

Examples

(1) QCD at finite density

(2) Solid state systems (using QMC)

- strongly correlated electron systems
- frustrated classical/quantum spin systems

(3) Real-time dynamics of quantum fields

(4) QCD with finite θ

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, 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]**

(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)
4. Worldvolume-TLTM (WV-TLTM)
5. Summary and outlook

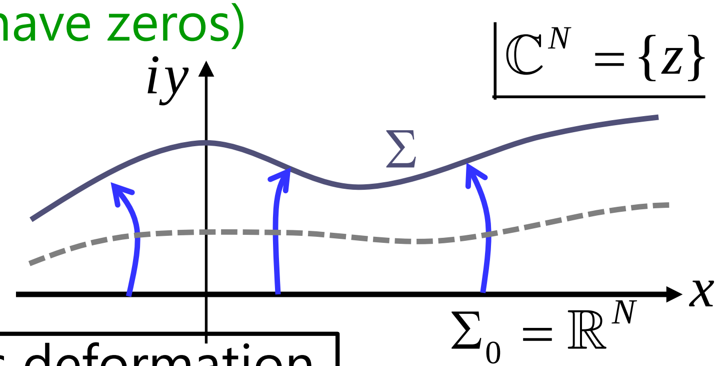
2. Lefschetz thimble method

Basic idea of the thimble method (1/2)

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

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

Integral does 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)

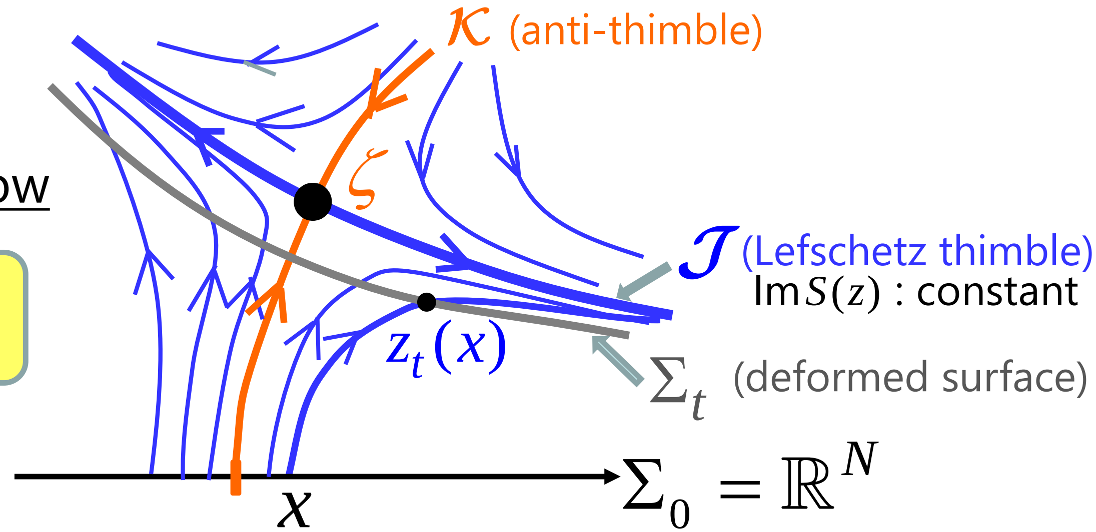
$$\langle \mathcal{O}(x) \rangle \equiv \frac{\int_{\Sigma_0} dx\, e^{-S(x)}\, \mathcal{O}(x)}{\int_{\Sigma_0} dx\, e^{-S(x)}} = \frac{\int_{\Sigma} dz\, e^{-S(z)}\, \mathcal{O}(z)}{\int_{\Sigma} dz\, e^{-S(z)}}$$

severe sign problem

sign problem will be significantly reduced if 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

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

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

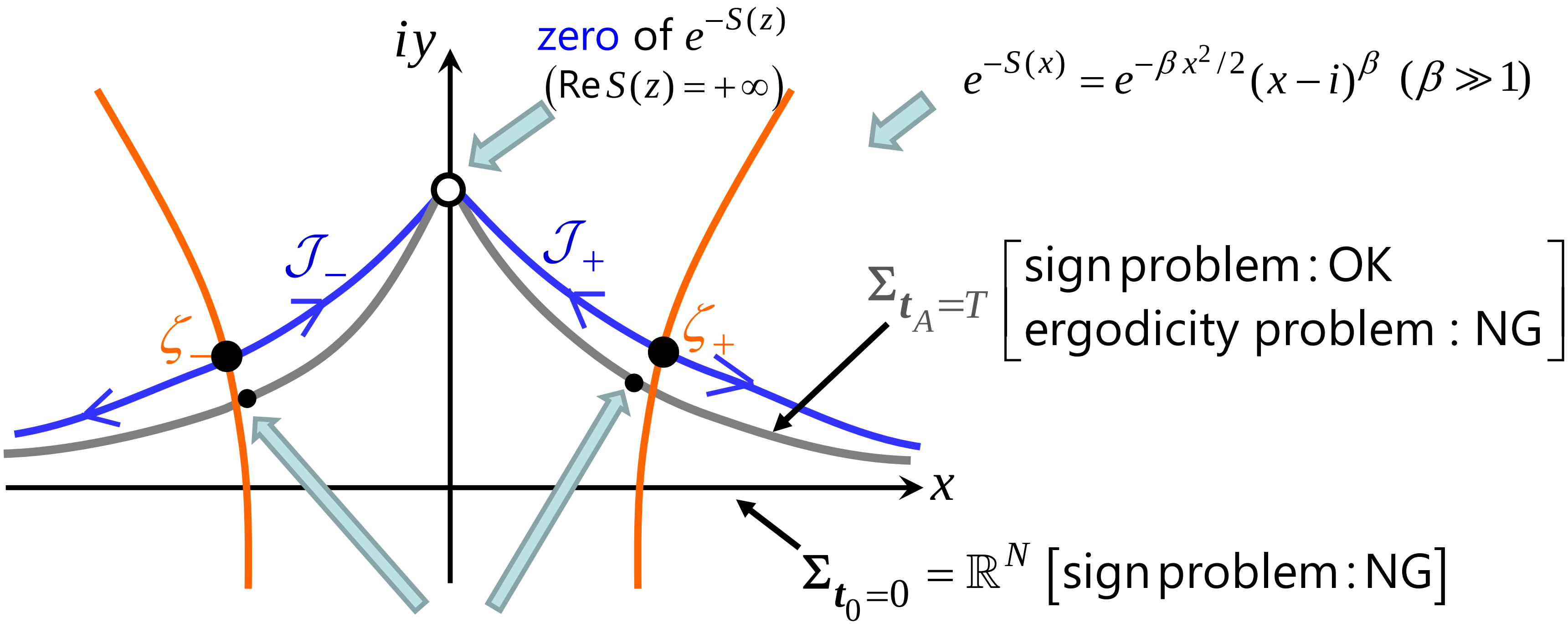
3. Tempered Lefschetz thimble method (TLTM)

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]

temper the system with the flow time

Tempered Lefschetz thimble method (TLTM)

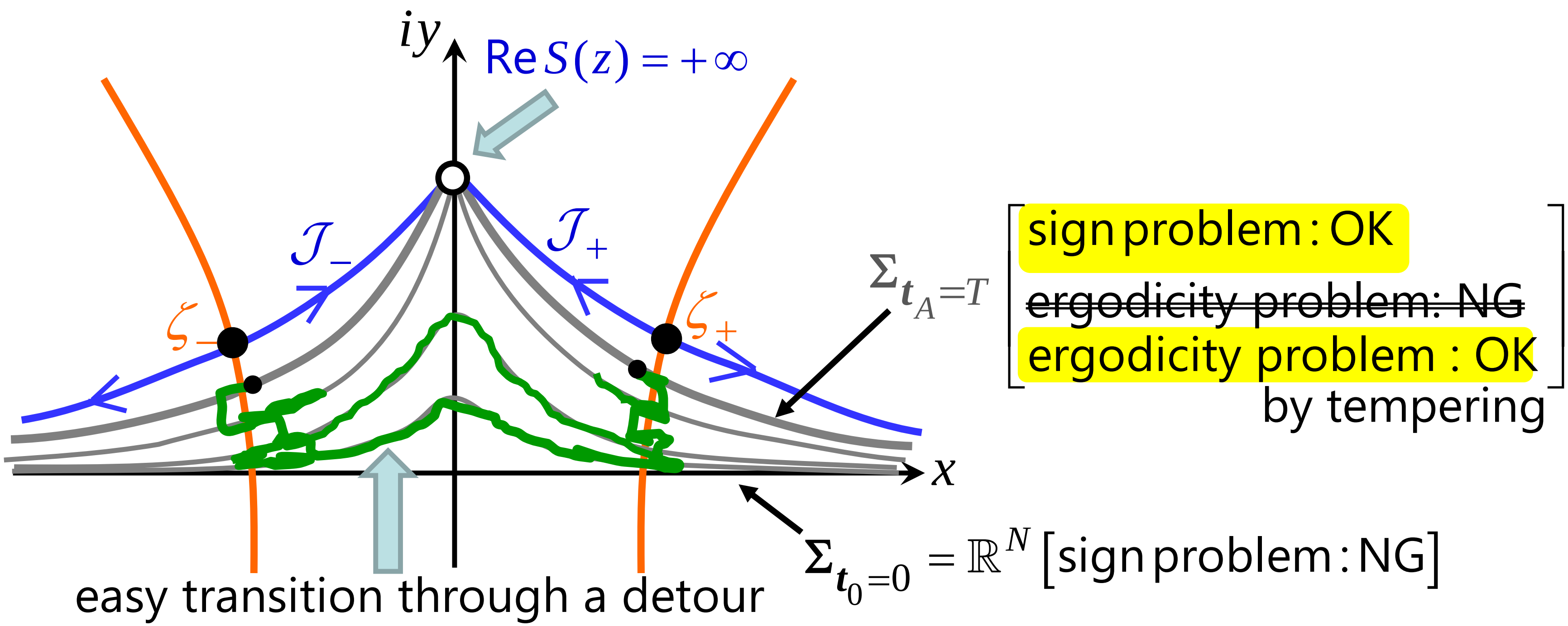
[Fukuma-Umeda 1703.00861]

■ 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) Estimate observables with a sample on Σ_T



Sign and ergodicity problems solved simultaneously !

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 ≤ 200)

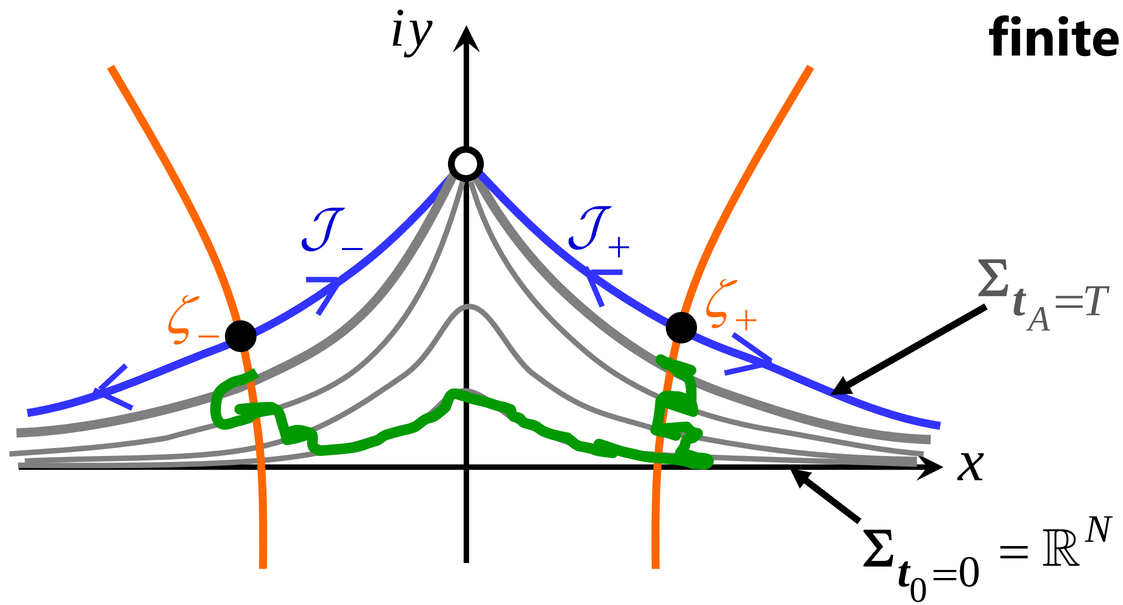
4. Worldvolume TLTM (WV-TLTM)

Pros and cons of 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



Pros : can be applied to any systems 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

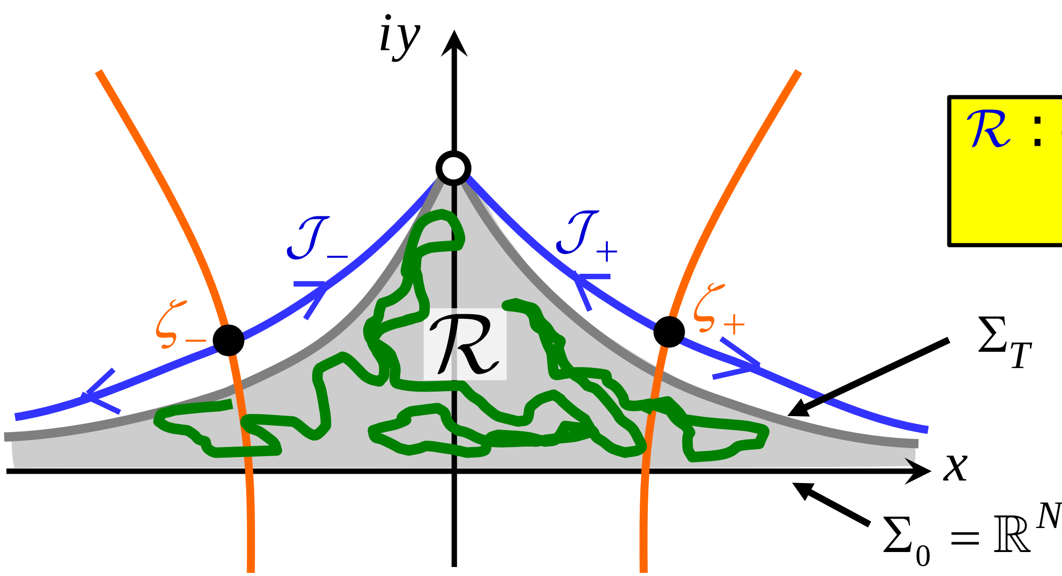
Idea of WV-TLTM (1/2)

[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 → worldline
orbit of string → worldsurface
orbit of surface (membrane) → 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
- 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

Idea of 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

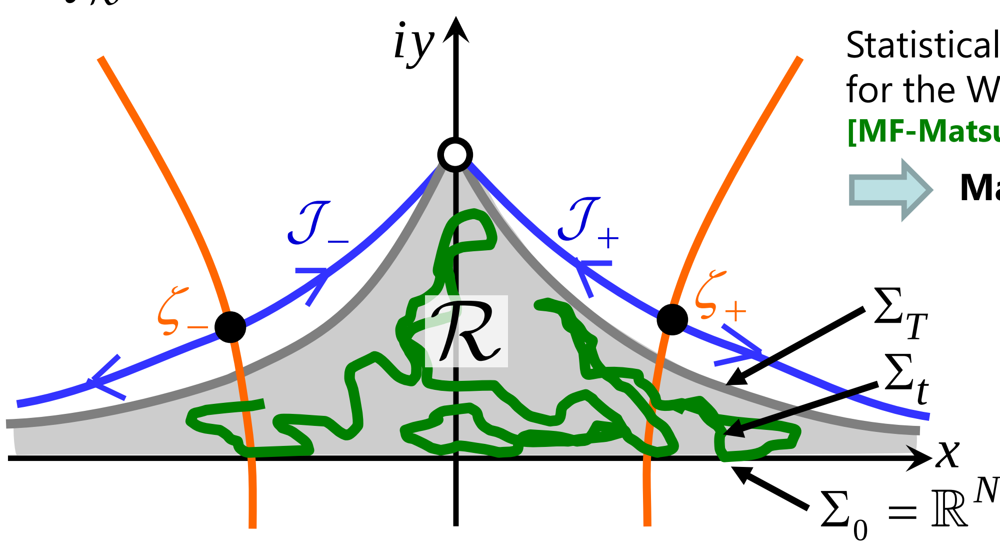
$$= \frac{\int_0^T dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\, e^{-S(z)}\, \mathcal{O}(z)}{\int_0^T dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\, e^{-S(z)}}$$

($W(t)$: arbitrary function)

(chosen s.t. the distribution in t direction is almost uniform)

$$= \frac{\int_{\mathcal{R}} dt\, dz_t\; g(z)\, \mathcal{O}(z)}{\int_{\mathcal{R}} dt\, dz_t\; g(z)}$$

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



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

⇒ **Matsumoto's talk (Wednesday)**

HMC on the worldvolume

■ Algorithm

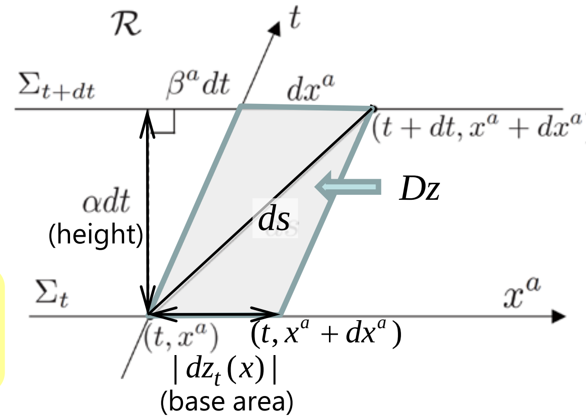
[MF-Matsumoto 2012.08468]

ADM decomposition of the induced metric :

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

⇨ vol element of $\mathcal{R}$: $Dz = \alpha\, dt\, |dz_t(x)| = \alpha\, |\det J|\, dt dx$

⇨ rewt factor : $A(z) \equiv \dfrac{dt\, dz_t}{Dz} e^{-i \mathrm{Im} S(z)} = \alpha^{-1}(z) \dfrac{\det J}{|\det J|} e^{-i \mathrm{Im} S(z)}$



potential $V(z) \equiv \mathrm{Re} S(z) + W(t(z))$

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

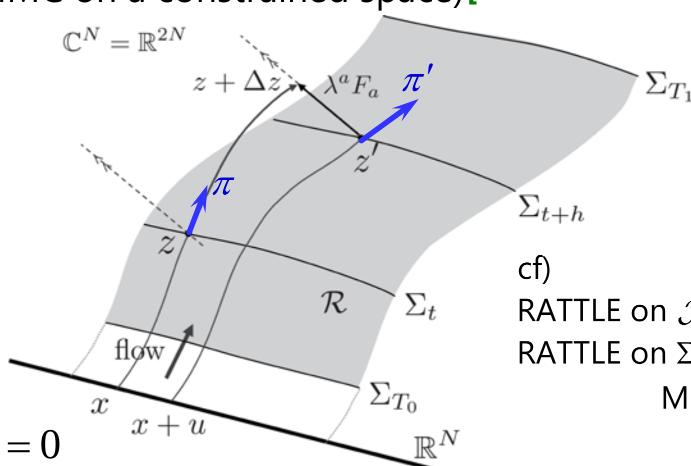
$$\left(\langle f(z) \rangle_{\mathcal{R}} \equiv \frac{\int_{\mathcal{R}} Dz\, e^{-V(z)} f(z)}{\int_{\mathcal{R}} Dz\, e^{-V(z)}} \right)$$

$\langle f(z) \rangle_{\mathcal{R}}$: estimated with RATTLE (HMC on a constrained space) **[Andersen 1983, Leimkuhler-Skeel 1994]**

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

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]

Application: chiral random matrix model (1/2)

[MF-Matsumoto 2012.08468]

■ finite density QCD

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

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

⇩ toy model

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

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

(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]**

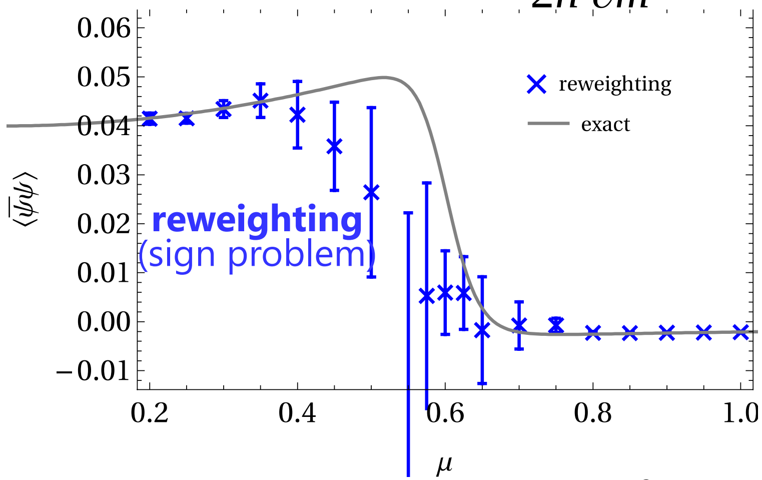
Application: chiral random matrix model (2/2)

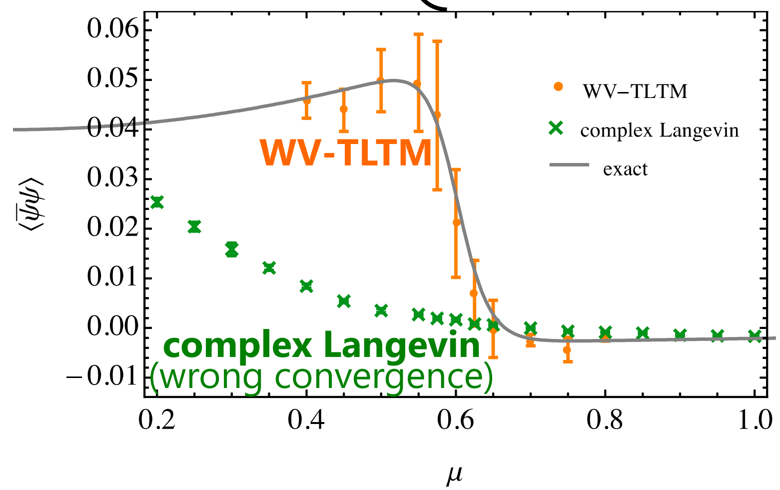
matrix size : $n = 10$ (DOF : $N = 200$) **[MF-Matsumoto 2012.08468]**

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

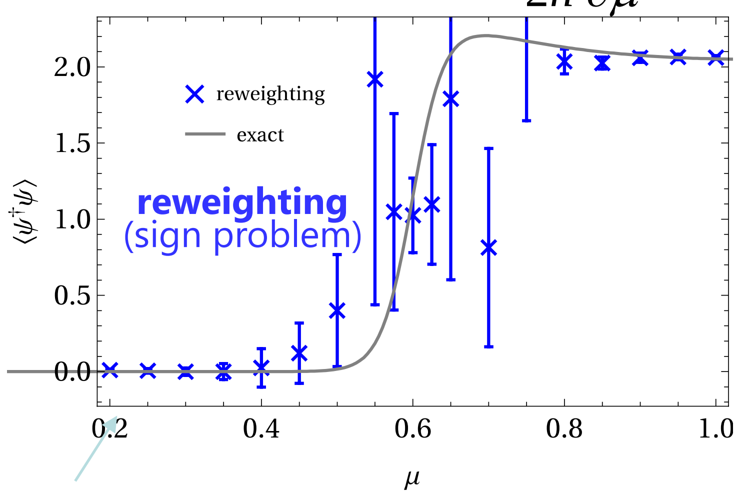
sample size
reweighting : 10k
complex Langevin :10k
WV-TLTM : 4k-17k

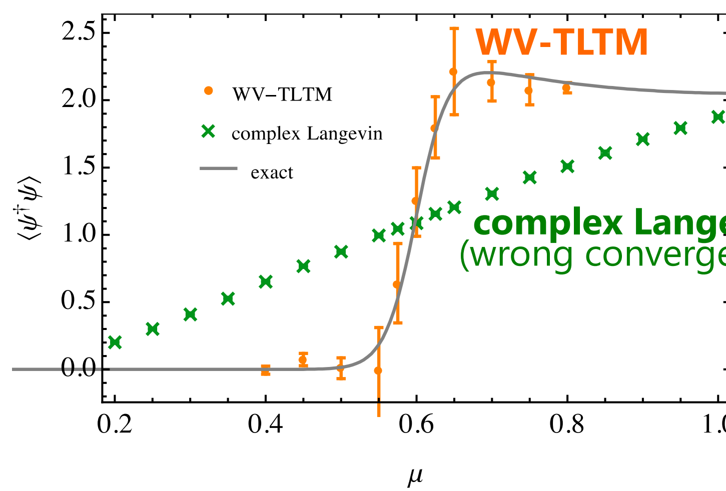
chiral condensate $\langle \bar{\psi}\psi \rangle \equiv \frac{1}{2n}\frac{\partial}{\partial m}\ln Z_{\mathrm{Steph}}$ $[m = 0.004,\ T = 0]$





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





5. Summary and outlook

Summary and outlook

■ Summary

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

- the sign and ergodicity problems are solved simultaneously

▼ TLTM has been successfully applied to various models

(yet only to toy models 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, 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.