Worldvolume tempered Lefschetz thimble method and its error estimation

Nobuyuki Matsumoto
(Kyoto Univ→ RIKEN BNL)

in collaboration with Masafumi Fukuma (Kyoto Univ), Yusuke Namekawa(YITP)

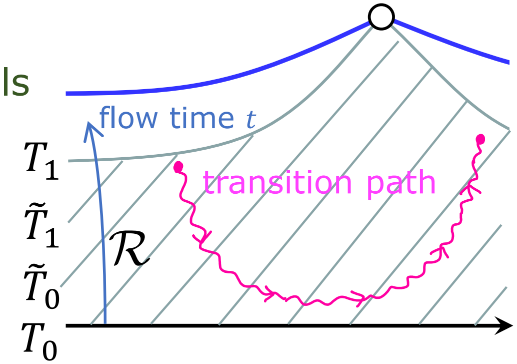
LATTICE2021 7/27/2021

Based on the papers:

- Worldvolume tempered Lefschetz thimble method:
  Fukuma-NM PTEP2021 [2012.08468]
- Statistical analysis method for the worldvolume hybrid Monte Carlo algorithm
  Fukuma-NM-Namekawa 2107.06858

Worldvolume tempered Lefschetz thimble method (WV-TLTM) Fukuma-NM 2012.08468
= an algorithm toward solving the sign problem

→ Fukuma-san's talk on 7/29 for details



Features:

- Solves the sign problem without additional multimodality.
  (Continuous change of the flow time in HMC updates.)

- Computational cost: (WV-TLTM) ≪ (conventional TLTM).

applied to various systems (all successful)
0+1 Thirring model [Fukuma-Umeda 1703.00861],
2D Hubbard model [Fukuma-NM-Umeda 1906.04243, 1912.13303]
Stephanov model, frustrated spin system, …

We confirm its effectiveness in the application to the Stephanov model.
toy model of finite density QCD

We also discuss the statistical analysis in WV-TLTM: Fukuma-NM-Namekawa 2107.06858

- Give justification for using the subset of configurations in $[\tilde{T}_0, \tilde{T}_1] \subset [T_0, T_1]$.

- Effective sample size is kept constant under this restriction:

$$N_{\rm conf}^{\rm eff} = N_{\rm conf}/\tau_{\rm int} \ : \ T_0,\, T_1\text{-independent}$$

This statistical analysis method becomes useful
for future large-scale computations with the WV-TLTM.

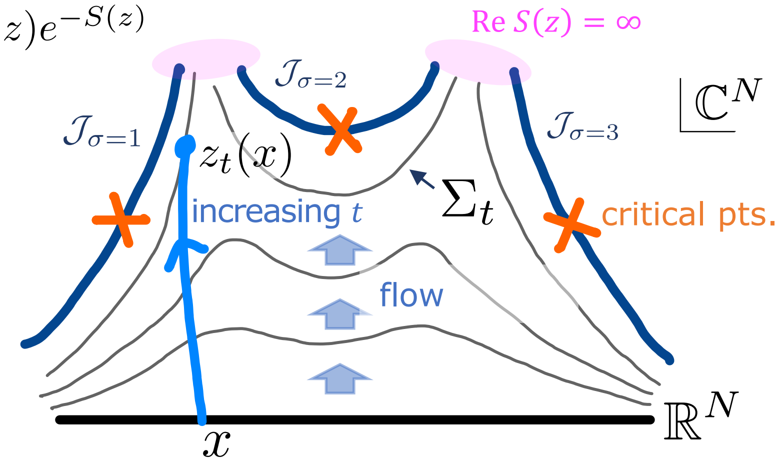
We deform the integration surface into $\mathbb{C}^N$ to reduce the oscillation:

Witten 2010, Cristoforetti et al. 2012, Alexandru et al. 2015

- Complexify variables: $\underset{\mathbb{R}^N}{\overset{\cap}{x}} \to \underset{\mathbb{C}^N}{\overset{\cap}{z}}$ $\begin{cases} \mathcal{O}(x) \to \mathcal{O}(z) \\ \mathcal{O}(x)e^{-S(x)} \to \mathcal{O}(z)e^{-S(z)} \end{cases}$

- Integrate the gradient flow:

$$\begin{cases} \frac{d}{dt} z_t = [\partial S(z_t)]^* \\ z_{t=0} = x \end{cases} \Rightarrow \text{define a surface } \Sigma_t \equiv z_t(\mathbb{R}^N)$$



Cauchy's theorem

$$\int_{\mathbb{R}^N} dx\, e^{-S(x)}\mathcal{O}(x) \overset{\downarrow}{=} \int_{\Sigma_t} dz_t\, e^{-S(z_t)}\mathcal{O}(z_t)$$ (for entire functions $e^{-S(z)}, e^{-S(z)}\mathcal{O}(z)$)

<u>Crucial property</u>

$\Sigma_t \xrightarrow{t\to\infty} \cup_\sigma \mathcal{J}_\sigma$ $\quad \mathcal{J}_\sigma$: Lefschetz thimble, on which $\mathrm{Im}\, S(z) = \mathrm{Im}\, S(z_\sigma)$ $(z \in \mathcal{J}_\sigma)$

∴ sign problem is reduced at large t

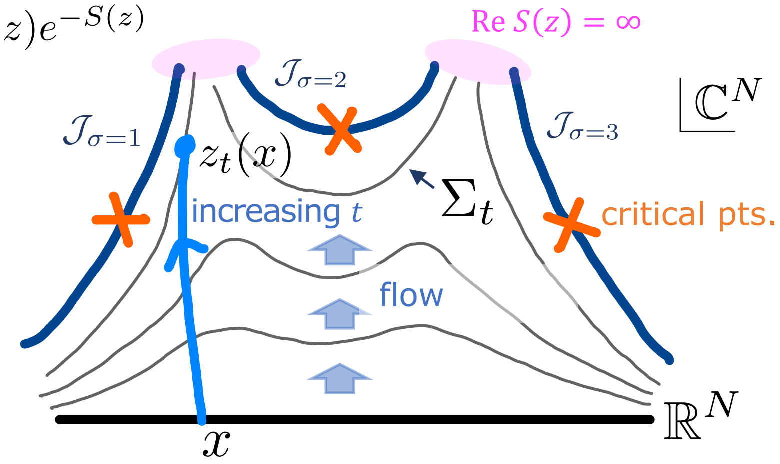
However, $\mathcal{J}_\sigma$'s are separated by regions where $\mathrm{Re}\, S(z) = \infty$.

∴ need to take care of the ergodicity at large t

We deform the integration surface into $\mathbb{C}^N$ to reduce the oscillation:

Witten 2010, Cristoforetti et al. 2012, Alexandru et al. 2015

- Complexify variables: $\underset{\mathbb{R}^N}{\overset{\cap}{x}} \to \underset{\mathbb{C}^N}{\overset{\cap}{z}}$ $\begin{cases} \mathcal{O}(x) \to \mathcal{O}(z) \\ \mathcal{O}(x)e^{-S(x)} \to \mathcal{O}(z)e^{-S(z)} \end{cases}$



- Integrate the gradient flow:

$$\begin{cases} \frac{d}{dt} z_t = [\partial S(z_t)]^* \\ z_{t=0} = x \end{cases}$$ define a surface $\Sigma_t \equiv z_t(\mathbb{R}^N)$

Cauchy's theorem

$$\int_{\mathbb{R}^N} dx\, e^{-S(x)} \mathcal{O}(x) \stackrel{\downarrow}{=} \int_{\Sigma_t} dz_t\, e^{-S(z_t)} \mathcal{O}(z_t)$$ [for entire functions $e^{-S(z)}, e^{-S(z)}\mathcal{O}(z)$]

<u>Crucial property</u>

$\Sigma_t \xrightarrow{t \to \infty} \cup_\sigma \mathcal{J}_\sigma$ $\mathcal{J}_\sigma$: Lefschetz thimble, on which $\mathrm{Im}\, S(z) = \mathrm{Im}\, S(z_\sigma)$ $(z \in \mathcal{J}_\sigma)$

dilemma

$\therefore$ sign problem is reduced at large t

However, $\mathcal{J}_\sigma$'s are separated by regions where $\mathrm{Re}\, S(z) = \infty$.

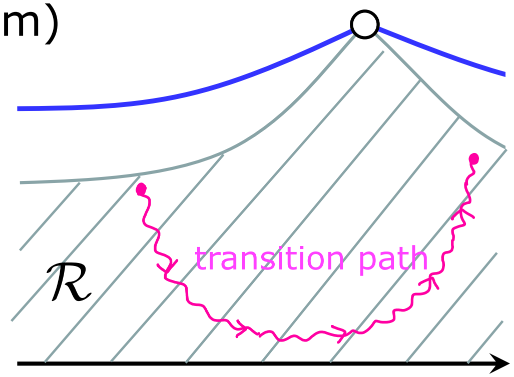
$\therefore$ need to take care of the ergodicity at large t

WV-TLTM (2/2) Fukuma-NM 2012.08468

Proposal: Further integrate over t (thanks to Cauchy's theorem)

$$\langle \mathcal{O}(x) \rangle = \frac{\int_{T_0}^{T_1} dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\, e^{-S(z_t)} \mathcal{O}(z_t)}{\int_{T_0}^{T_1} dt\, e^{-W(t)} \int_{\Sigma_t} dz_t\, e^{-S(z_t)}}$$

integration over the worldvolume $\mathcal{R} \equiv \cup_{t=T_0}^{T_1} \Sigma_t$



$W(t)$: tuned to make t distributed uniformly

$$= \frac{\int_{\mathcal{R}} Dz\, e^{-[W(t(z)) + \mathrm{Re}\, S(z)]} \left(\frac{dt\, dz_t}{Dz}\right) e^{-i\, \mathrm{Im}\, S(z)} \mathcal{O}(z)}{\int_{\mathcal{R}} Dz\, e^{-[\underbrace{W(t(z)) + \mathrm{Re}\, S(z)}_{\equiv V(z)}]} \underbrace{\left(\frac{dt\, dz_t}{Dz}\right) e^{-i\, \mathrm{Im}\, S(z)}}_{\equiv A(z)}}$$

$V(z)$: potential

$A(z)$: reweighting factor

Dz: invariant volume element

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

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

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

<u>HMC on $\mathcal{R}$</u> RATTLE: Andersen 1983, Leimkuhler-Skeel 1994

RATTLE in Lefschetz thimble methods:
cf. Fujii-Honda-Kato-Kikukawa-Komatsu 2013,
Alexandru (LATTICE 2019), Fukuma-NM-Umeda 2019

- Sign problem tamed without additional multimodality.
- No need to care acceptance in the t-direction.
- No need to calculate $J_t \equiv \partial z_t / \partial x$ in generating configurations (using iterative methods).

$e^{i\varphi} = \det J_t / |\det J_t|$ is required only in the operator evaluation

Stephanov model: a toy model of finite density QCD Stephanov 1996, Halasz et al. 1998

Finite density QCD (for simplicity, 1 flavor and zero temperature)

$$Z_{\rm QCD} = \int [dA]\, e^{\int \frac{1}{2g^2} {\rm tr} F_{\mu\nu}^2} \det \begin{bmatrix} m & \sigma_\mu(\partial_\mu + A_\mu) + \mu \\ \sigma_\mu^\dagger(\partial_\mu + A_\mu) + \mu & m \end{bmatrix}$$

$\downarrow$ replace the gauge field with a random matrix

Stephanov model (1 flavor and zero temperature)

$$Z_{\rm Steph} \equiv \int (d^2 W)\, e^{-n\, {\rm tr} W^\dagger W} \det \begin{bmatrix} m & iW + \mu \\ iW^\dagger + \mu & m \end{bmatrix}$$

(W : complex $n \times n$ matrix, $n \propto {\rm vol}(\mathbb{R}^4)$)

- Reproduces qualitative features of finite density QCD Shuryak-Verbaarschot 1992, Berbenni-Bitsch et al. 1998

Chiral condensate

$$\langle \bar\psi \psi \rangle \equiv \frac{1}{2n} \frac{\partial}{\partial m} \log Z_{\rm Steph}$$

- $\langle \bar\psi\psi \rangle \neq 0$ at $n \to \infty$, even when $m \to 0$
- We can discuss chiral phase transition

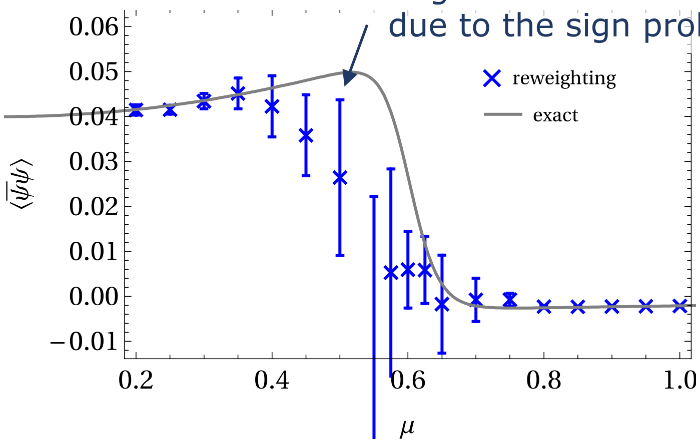
Good benchmark for testing the effectiveness of WV-TLTM!

Result for $n = 10\ (N = 200)$ $\left[m = 0.004 \right]$

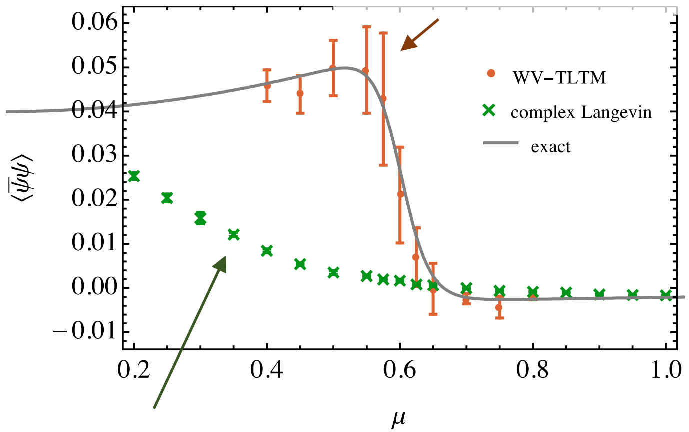
GPU is used for linear operations
~70 replicas appear to be required in the conventional TLTM with parallel tempering
Swendsen-Wang 1986, Geyer 1991

<u>Chiral condensate</u> $\langle \bar{\psi}\psi \rangle = \frac{1}{2n}\frac{\partial}{\partial m}\log Z_{\rm Steph}$

large errors
due to the sign problem

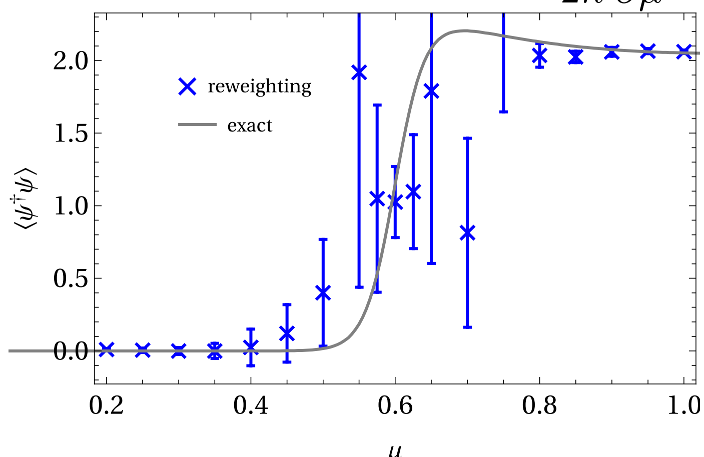


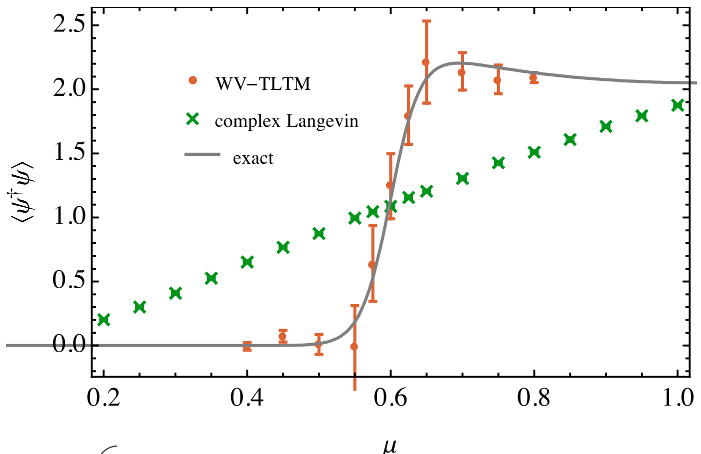
WV-TLTM reproduces
exact values with controlled errors



"wrong convergence problem" [Bloch et al. 2018]

<u>Baryon number density</u> $\langle \psi^\dagger\psi \rangle \equiv \frac{1}{2n}\frac{\partial}{\partial \mu}\log Z_{\rm Steph}$





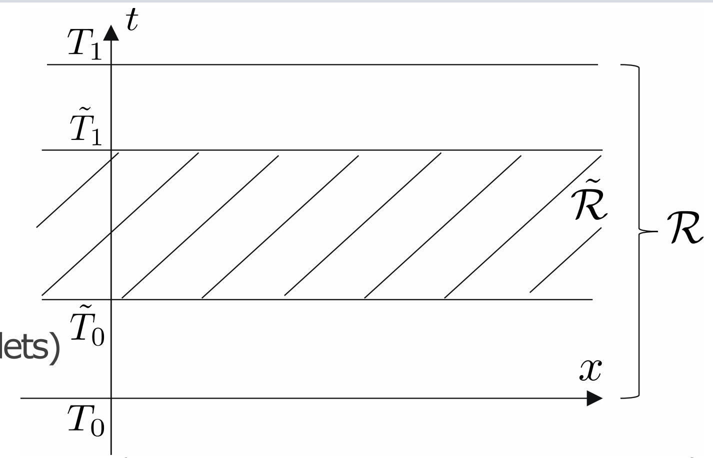
Successful evaluation at the DOF
hard to be reached with the conventional TLTM

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

Statistical analysis for the WV-TLTM (1/3)

Fukuma-NM 2012.08468,
Fukuma-NM-Namekawa 2107.06858

Choice of T_0, T_1: $\left(\mathcal{R} = \cup_{t=T_0}^{T_1} \Sigma_t \right)$

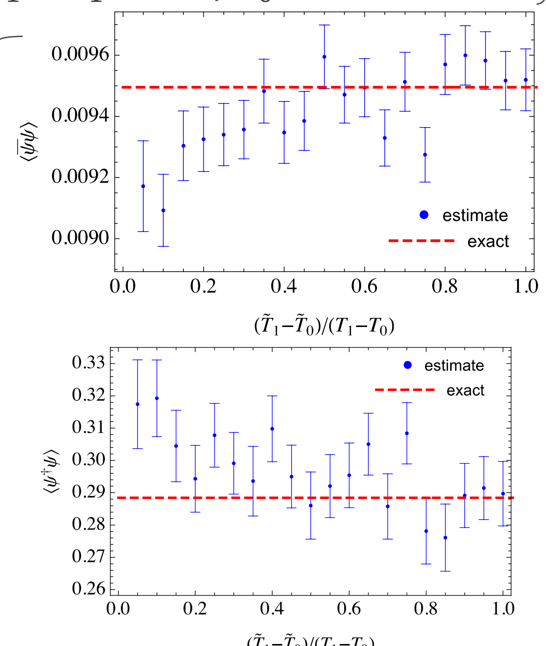


- Generating configurations:
  - T_0 : sufficiently small to recover ergodicity
    (Generically $T_0 = 0$ suffices.)
  - T_1 : sufficiently large to reduce oscillations
    (Behavior drastically improves after reaching zeros of fermion dets)

- Estimation:
  take subsample in a subregion $[\tilde{T}_0, \tilde{T}_1] \subset [T_0, T_1]$.
  for various $\tilde{T}_0$, $\tilde{T}_1$.

  We then can
  - check if N_{conf} is large enough
    (by confirming $\tilde{T}_0$, $\tilde{T}_1$-independence of $\langle \mathcal{O}(x) \rangle$
    thanks to Cauchy's theorem.)
  - exclude the two extreme regions if necessary:
    - small t: large sign fluctuations
    - large t: large autocorrelations and
      possible systematic errors
      due to complicated geometries

(application to the Stephanov model
$N_f = 1, n = 2, \mu = 0.6, \tau = 0,$
$m = 0.004, N_{\mathrm{conf}} = 500\mathrm{k}$
$\tilde{T}_1 = T_1$: fixed, $\tilde{T}_0$:varied)

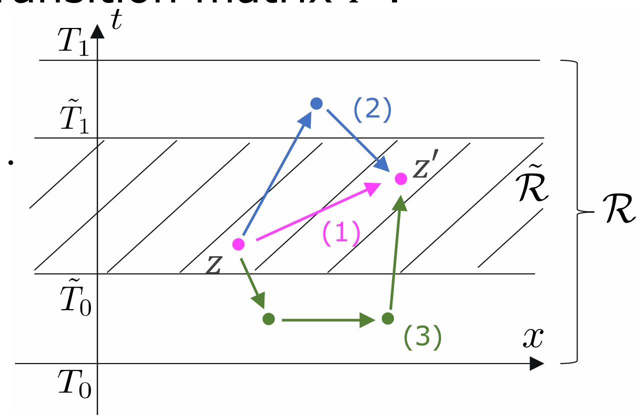


<u>Justification</u> (for applying ordinary statistical methods to the subset of configurations)

- The retrieved configs are actually generated by the transition matrix $\tilde{P}$:

$$\tilde{P}(z'|z) = \underbrace{P(z'|z)}_{(1)} + \underbrace{\int_{\tilde{\mathcal{R}}^c} dw\, P(z'|w)P(w|z)}_{(2)} + \underbrace{\int_{\tilde{\mathcal{R}}^c} dw_2\, dw_1\, P(z'|w_2)P(w_2|w_1)P(w_1|z)}_{(3)} + \cdots$$

(P: transition matrix of the original process)



- $\tilde{P}$ satisfies the detailed balance:

$$\tilde{P}(\tilde{z}'|\tilde{z})e^{-S(\tilde{z})} = \tilde{P}(\tilde{z}|\tilde{z}')e^{-S(\tilde{z}')}$$

- ergodicity of $P \Longrightarrow$ ergodicity of $\tilde{P}$

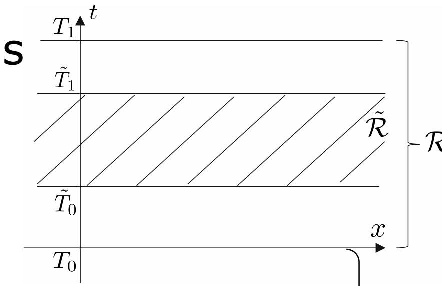
➡ We can apply ordinary statistical methods also to the retrieved configs.

Statistical analysis for the WV-TLTM (3/3) Fukuma-NM-Namekawa 2107.06858

Possible concern about retrieving: decrease of the number of configs

⇒ compensated by decrease of autocorrelation.

"Effective sample size" $N_{\rm conf}^{\rm eff} \equiv N_{\rm conf}/\tau_{\rm int}$ is indep of $\tilde{T}_0$, $\tilde{T}_1$.



This can be proved rigorously when $\underbrace{\tau_{\rm int}(t) \simeq 1}_{\text{purely random}}$.

- Probability for a configuration to appear in $[\tilde{T}_0, \tilde{T}_1]$:

$$p = \frac{\tilde{T}_1 - \tilde{T}_0}{T_1 - T_0}$$

$W(t)$ is tuned so that the distribution of t is uniform

∴ Integrated autocorrelation time for the subsample:

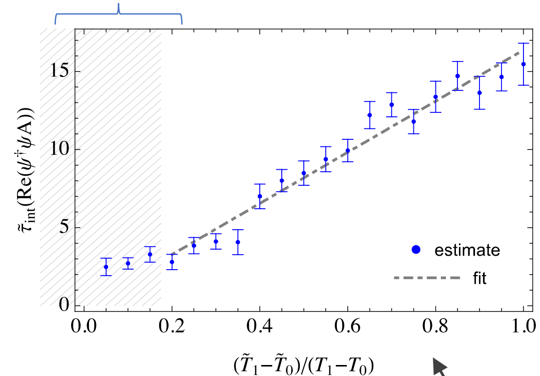
$$\tilde{\tau}_{\rm int} = p\,\tau_{\rm int}$$

- The sample size is reduced with the same ratio.

⇒ the above statement holds

$\tilde{\tau}_{\rm int}$ saturated at $\simeq 1$

checked



$\tilde{T}_1 = T_1$: fixed
$\tilde{T}_0$:varied

Consequence:

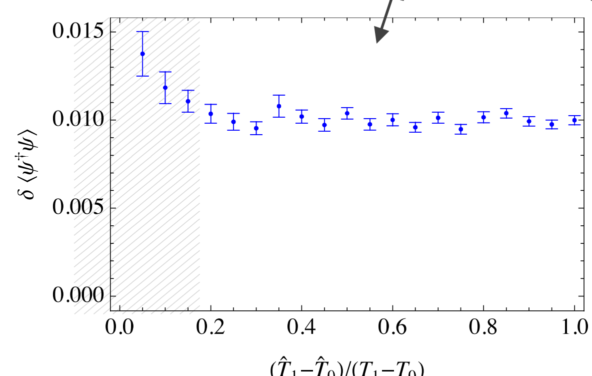
The statistical error is indep of $\tilde{T}_0$, $\tilde{T}_1$:

$$\delta\langle\mathcal{O}\rangle = \frac{\sigma_{\mathcal{O}}}{\sqrt{N_{\rm conf}^{\rm eff}}}$$

: $\tilde{T}_0$, $\tilde{T}_1$-independent

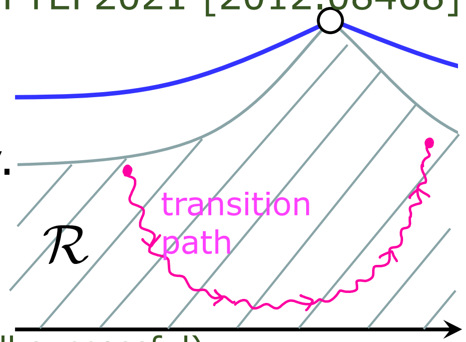
$\sigma_{\mathcal{O}}$: constant

checked



We can now safely apply the WV-TLTM to large-scale computations with this statistical analysis method!

- Worldvolume tempered Lefschetz thimble method (WV-TLTM)
  = an algorithm toward solving the sign problem   Fukuma-NM PTEP2021 [2012.08468]



Features:

- Solves the sign problem without additional multimodality.
  (Continuous change of the flow time in HMC updates.)
- Computational cost: (WV-TLTM) $\ll$ (conventional TLTM).

applied to various systems (all successful)
0+1 Thirring model [Fukuma-Umeda 1703.00861],
2D Hubbard model [Fukuma-NM-Umeda 1906.04243, 1912.13303]
Stephanov model, frustrated spin system, …

- Applied to the Stephanov model and confirmed the effectiveness.

- Justified using the subset of configurations in estimations.

Fukuma-NM-Namekawa 2107.06858

Future outlook

- Apply our algorithm to large scale computations of physical systems.

  finite density QCD, strongly-correlated electron systems, frustrated spin systems, and real-time QM,…

- Develop the algorithm further, e.g.,
  - method to compute the phase of the Jacobian, $e^{i\varphi}$, with low cost
  - efficient way to determine the weight function $W(t)$ [present: multicanonical method]

Thank you.