Parallel tempering algorithm for integration over Lefschetz thimbles

Naoya Umeda (Kyoto Univ.)

w/ Masafumi Fukuma (Kyoto Univ.)

@APCTP, 22 Sep 2017

"Discrete approaches to the Dynamics of Fields and Space-Time"

based on : [PTEP 2017 (2017) no.7, 073B01 (arXiv:1703.00861)]

and work in progress

Motivation: sign problem

Monte Carlo calculation (real action)

$$\langle \mathcal{O}(x)\rangle \equiv \frac{1}{Z}\int dx\, \mathcal{O}(x) e^{-s(x)} \quad \left(Z \equiv \int dx e^{-S(x)}\right)$$

$$= \int dx\, p(x)\mathcal{O}(x) \quad \left(p(x) \equiv \frac{1}{Z} e^{-S(x)}\right)$$

$$\simeq \frac{1}{N}\sum_{i=1}^{N} \mathcal{O}(x_i)$$

But, if $S(x)$ is complex,
$p(x)$ can no longer be regarded as probability distribution.

Reweighting

$$\langle \mathcal{O}(x)\rangle = \frac{\langle e^{-i\mathrm{Im}S(x)}\mathcal{O}(x)\rangle_{\mathrm{Re}S(x)}}{\langle e^{-i\mathrm{Im}S(x)}\rangle_{\mathrm{Re}S(x)}}$$

$$\langle \mathcal{O}(x)\rangle_{S_R} \equiv \frac{1}{Z_R}\int dx \mathcal{O}(x) e^{-\mathrm{Re}S(x)}$$

(local) sign problem:
$e^{-i\mathrm{Im}S(x)}$ oscillates if $|\mathrm{Im}S(x)|$ become large.

$\Rightarrow$ complex Langevin method, Lefschetz thimble method, …

Lefschetz thimble method (1/4)

Idea: take better integration contours in complex plane

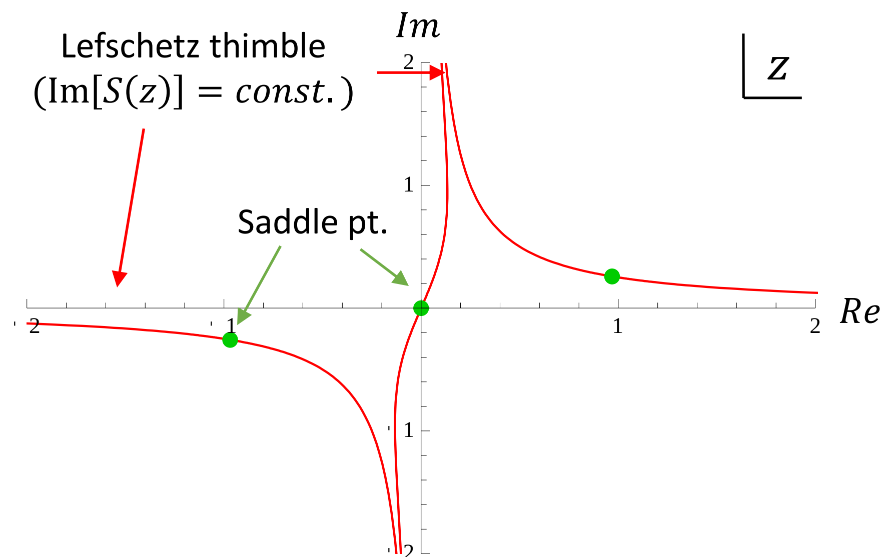
$$x = (x_i) \in \mathbb{R}^N \to z = (z_i) \in \mathbb{C}^N$$

<u>Lefschetz thimble</u> [Cristoforetti *et. al.* 2012]

$$\left\{ z(t) \middle| \frac{\partial z(t)}{\partial t} = \overline{\frac{\partial S(z(t))}{\partial z}}, z(t=-\infty) = z^{(0)} \right\}$$

Saddle pt.

- N –dimensional submanifold in $\mathbb{C}^N$
- Defined by anti-holomorphic gradient flow
- $\mathrm{Im}S(z) = \mathrm{const.}$ on each thimble



Ex.) $S(z) = -\frac{c}{2}z^2 + \frac{1}{4}z^4 \quad (c = \frac{\sqrt{3}}{2} + \frac{1}{2}i)$

Transform the integration contour:

$$Z = \int_{\mathbb{R}^N} dx\, e^{-S(x)}$$

$$= \sum_\sigma n_\sigma \int_{J_\sigma} dz\, e^{-S(z)}$$

(J_σ: Lefschetz thimble, n_σ: intersection number)

$$= \sum_\sigma n_\sigma e^{-i\mathrm{Im}S(z_\sigma)} \int_{J_\sigma} dz\, e^{-\mathrm{Re}S(z)}$$

$\Rightarrow$ the sign problem can be avoided.

Lefschetz thimble method (2/4)

[Alexandru *et. al.* (2016)]

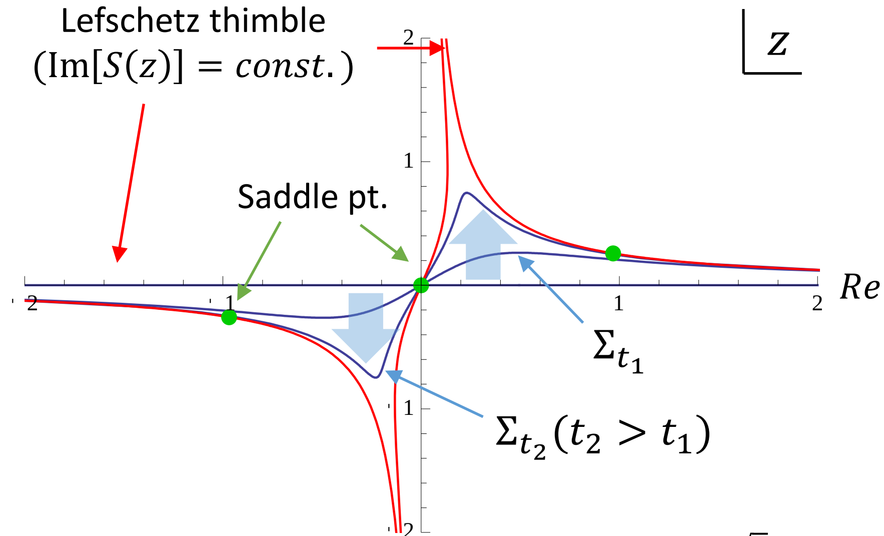
We can use the gradient flow to change the variables:

$$\frac{\partial z_i(x;t)}{\partial t} = \overline{\frac{\partial S(z(x;t))}{\partial z_i}}$$

$$\frac{\partial J_{ij}(x;t)}{\partial t} = \overline{\frac{\partial^2 S\big(z(x;t)\big)}{\partial z_i \partial z_k} J_{kj}(x;t)}$$

$$z_i(x;t=0) = x_i, \qquad J_{ij}(x;t=0) = \delta_{ij}$$

$$\left(J_{ij} = \frac{\partial z_i}{\partial x_j} \text{: Jacobian} \right)$$



Ex.) $S(z) = -\frac{c}{2}z^2 + \frac{1}{4}z^4 \quad \left(c = \frac{\sqrt{3}}{2} + \frac{1}{2}i\right)$

$$Z = \int_{\mathbb{R}^N} dx\, e^{-S(x)}$$

$$= \int_{\Sigma_t} dz\, e^{-S(z)}$$

$$= \int_{\mathbb{R}^N} dx\, \det J(x;t) e^{-S(z(x;t))}$$

$$\left(S_{\mathrm{eff}}(x;t) \equiv \mathrm{Re} S\big(z(x;t)\big) - \log|\det J| \right)$$

$$= \int_{\mathbb{R}^N} dx\, e^{-S_{\mathrm{eff}}(x;t)} e^{i[\arg\det J - \mathrm{Im} S(z(x;t))]}$$: effective action + residual phase

No need to compute intersection numbers

Lefschetz thimble method (3/4)

$$Z = \int_{\mathbb{R}^N} dx\, e^{-S_{\mathrm{eff}}(x;t)} e^{i[\arg\det J - \mathrm{Im}S(z(x;t))]}$$

Effective action

Incorporate into operators (reweighting)

Correlation functions can be estimated by using S_{eff}:

$$\langle \mathcal{O}(x) \rangle = \frac{1}{Z}\int dx\, e^{-S(x)} = \frac{\langle e^{i[\arg\det J - \mathrm{Im}S(z(x;t))]} \mathcal{O}(z(x;t)) \rangle_{\mathrm{eff}}}{\langle e^{i[\arg\det J - \mathrm{Im}S(z(x;t))]} \rangle_{\mathrm{eff}}}$$

This is t-independent.

⇒ The sign problem can be avoided for sufficiently large t.

((empirically) the residual phase do not cause additional problems.)

Lefschetz thimble method (3/4)

<u>However, there are some problems.</u>

- For large t, configurations will be trapped on one thimble.

➡ We can avoid this by introducing tempering algorithm.

[Fukuma-NU (2017)], [Alexandru *et. al.* (2017)]

- If $\frac{\mathrm{d}^2 S}{dx^2} = 0$ (Hesse matrix have 0 eigenvalues) on saddle points,
  Lefschetz thimble is not well-defined.
  (i.e. the thimbles can not be regarded as an N-dimensional manifold.)

  However, the antiholomorphic gradient flow is well-defined even in this case.

  ➡ We can use LTM without any change.

[Fukuma-NU (work in progress)]

____ : topics of my talk

Plan of talk

- Introduction
- Parallel tempering algorithm for LTM
  - Simulated tempering
  - Parallel tempering
  - Parallel tempering for LTM
- Examples
  - (0+1)D massive Thirring model
  - One matrix model with quartic potential
- Summary

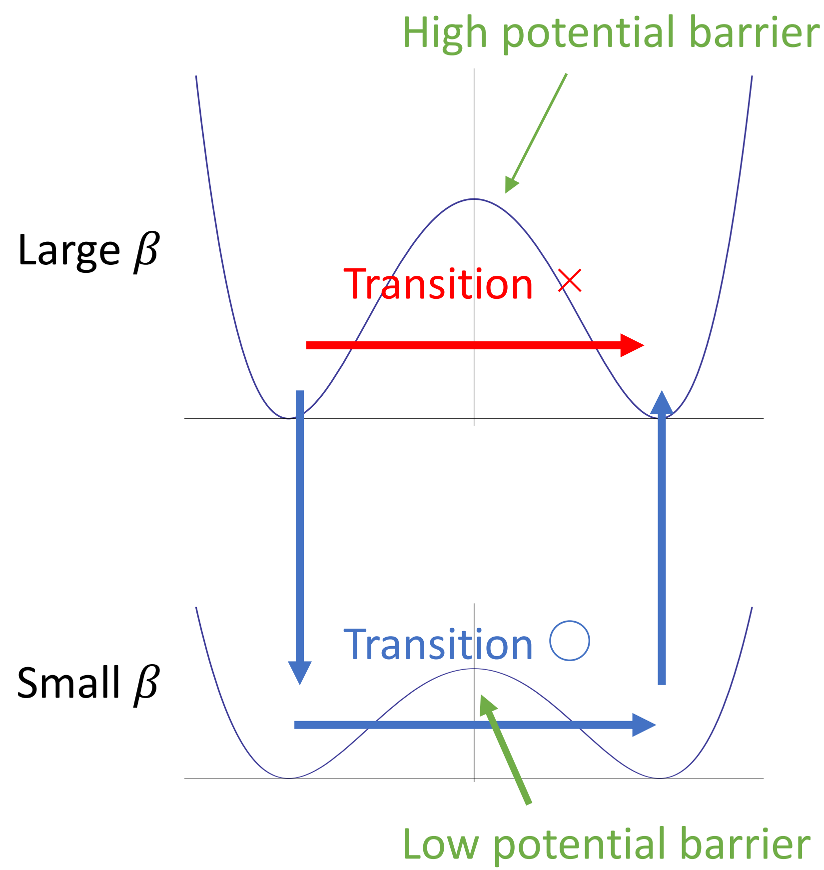
Multimodal problem

(Before reviewing the parallel tempering, I review the simulated tempering method.)

$\left(S(x) = \beta V(x)\right)$

Multimodal problem

If the action have potential barriers, the configuration is trapped at a local minimum.



The idea of the tempering algorithm

Expand configuration space:

$$x \to X \equiv (x, \beta)$$

→ we can evade the potential barrier passing through the configs with smaller β.

Simulated tempering (2/2)

Simple approach: simulated tempering

Expand configuration space: $x \to \mathrm{X} \equiv (x, \beta)$

(i.e. regard β as an additional dynamical variable.)

where, β can take discrete variables: $\beta \in \{\beta_0, \beta_1, \ldots, \beta_A\}$ $(\beta_0 > \beta_1 > \cdots > \beta_A)$

($\Rightarrow$ total equilibrium distribution: $\mathcal{P}_{eq}(x, a) = \frac{1}{A+1} e^{-S(x;\beta_a)}/Z(\beta_a)$)

Algorithm

Define two transition matrix:

$$\begin{cases} P_1: X = (x, \beta_a) \to X' = (x', \beta_a) & \text{(ordinary transition matrix for fixed } \beta = \beta_a) \\ P_2: X = (x, \beta_a) \to X' = (x, \beta_{a'}) & \\ \quad \text{Accept/reject with probability } \min\{1, \mathcal{P}_{eq}(x, a')/\mathcal{P}_{eq}(x, a)\} & \end{cases}$$

and apply P_1 and P_2 alternately.

(Finally we calculate correlation functions by using a subset of configurations $(x, \beta = \beta_0)$.)

Problem

We must (at least roughly) estimate $Z(\beta_a)$ a priori. ⇒ *Parallel tempering*

Parallel tempering

Parallel tempering

Consider replicas of configuration spaces: $X = \{x_a\}_{a=0,1,\cdots,A}$
The configuration of replica a explores with β_a
and be exchanged among replicas at fixed intervals.

Algorithm

Define two transition matrix:

$$\begin{cases} P_1: X = (x, \beta_a) \to X' = (x', \beta_a) \quad \text{(ordinary transition matrix for fixed } \beta = \beta_a\text{)} \\ P_2: \text{exchange configurations between replica } a_1 \text{ and } a_2 \\ \qquad \text{with probability } \min\{1, e^{-\Delta S}\} \end{cases}$$

$$\left(\Delta S \equiv S(x_{a_1}, \beta_{a_1}) + S(x_{a_2}, \beta_{a_2}) - S(x_{a_1}, \beta_{a_2}) - S(x_{a_2}, \beta_{a_1})\right)$$

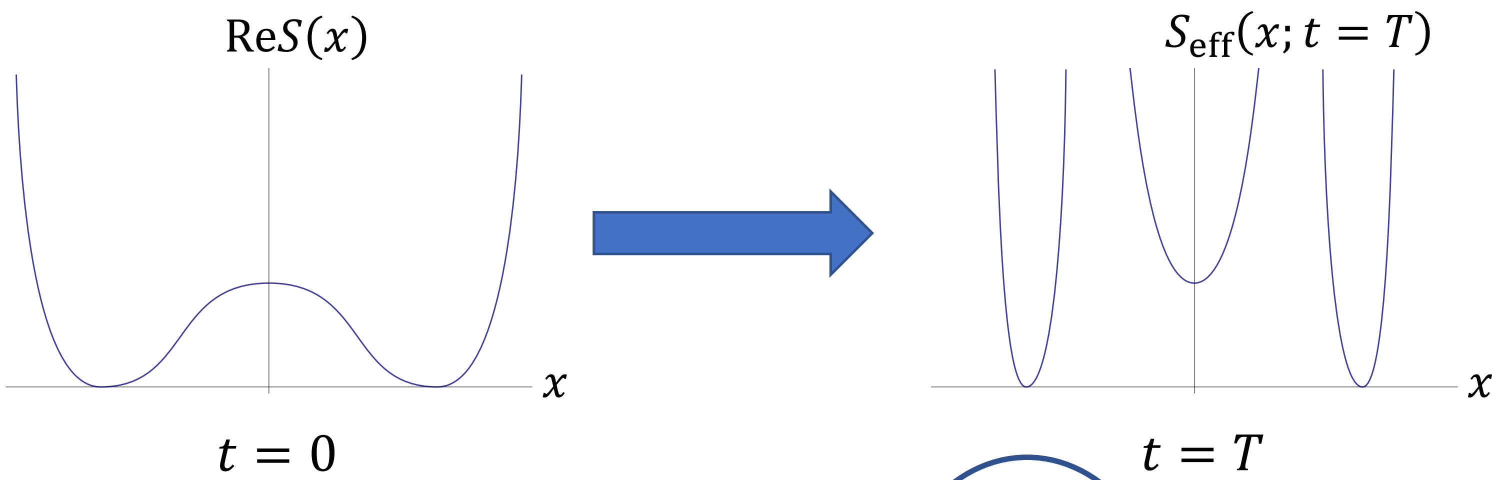
and apply P_1 and P_2 alternately.

We need not to estimate $Z(\beta_a)$.
($\Rightarrow$ All we need is to choose the set of $\{\beta_a\}$.)

Parallel tempering for LTM (1/2)

Multimodal problem in Lefschetz Thimble Method

If we take the flow time ($= T$) large,
there are infinitely (or exponentially) high potential barrier in $S_{\mathrm{eff}}(x; t = T)$:



Flow time ($= T$)	large	small	middle
Sign problem	○	×	△
Multimodal problem	×	○	△

← Previous approach [Alexandru *et. al.* (2016)]

Need fine-tuning

Parallel tempering for LTM (2/2)

Our approach:

We introduce parallel tempering algorithm
by regarding the flow time t as a tempering parameter.

(Consider several number of flow time $t_a (t_0 = T > t_1 > \cdots > t_A = 0)$
$\Rightarrow$ parallel tempering)

➡ No need to fine-tune. (All we need is to take sufficiently large T (and A).)

Flow time $(= T)$	large	small	middle	Parallel tempering
Sign problem	○	×	△	○
Multimodal problem	×	○	△	○

Plan of talk

- Introduction
- Parallel tempering algorithm for LTM
  - Simulated tempering
  - Parallel tempering
  - Parallel tempering for LTM
- Examples
  - <u>(0+1)D massive Thirring model</u>
  - One matrix model with quartic potential
- Summary

(0+1)D massive Thirring model (1/3)

Example 1 : (0+1)-dimensional massive Thirring model at finite density

This model is defined from (1+1)-dimensional massive Thirring model by dimensional reduction:

$$Z = \int_{\mathrm{PBC}} [d\phi(\tau)] \int_{\mathrm{ABC}} [d\bar{\psi}(\tau) d\psi(\tau)]\, e^{-S[\phi,\bar{\psi},\psi]}$$

$$S[\phi,\bar{\psi},\psi] = \int_0^\beta d\tau \left[-\bar{\psi}(\gamma^0(\partial_0 + i\phi + \mu) + m)\psi + \frac{1}{2g^2}\phi^2\right] \qquad \left(\gamma^0 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\right)$$

Discretize: $\phi(\tau) \to \phi_n \ (n = 1, \cdots, N)$

$$Z = \int d\phi\, e^{-S(\phi)}$$

$$S(\phi) = \frac{1}{2g^2} \sum_n \left[1 - \frac{1}{2}\left(e^{i\phi_n} + e^{-i\phi_n}\right)\right] - \log\det D$$

$$D_{nl}(\phi) = \frac{1}{2}(e^{i\phi_n+\mu}\delta_{n+1,l} - e^{-i\phi_{n-1}-\mu}\delta_{n-1,l} - e^{i\phi_N+\mu}\delta_{n,N}\delta_{l,1} + e^{-i\phi_N-\mu}\delta_{n,1}\delta_{l,N}) + m\delta_{n,l}$$

- Sign problem occurs for $\mu \neq 0$
- Analytically solvable even for finite N
- More than two thimbles can have non-negligible contributions to Z

(0+1)D massive Thirring model (2/4)

This model can be solved analytically even for finite N:

$$Z = \int d\phi \, e^{-S(\phi)}$$

$$S(\phi) = \frac{1}{2g^2}\sum_n \left[1 - \frac{1}{2}\left(e^{i\phi_n} + e^{-i\phi_n}\right)\right] - \log\det D$$

$$D_{nl}(\phi) = \frac{1}{2}(e^{i\phi_n+\mu}\delta_{n+1,l} - e^{-i\phi_{n-1}-\mu}\delta_{n-1,l} - e^{i\phi_N+\mu}\delta_{n,N}\delta_{l,1} + e^{-i\phi_N-\mu}\delta_{n,1}\delta_{l,N}) + m\delta_{n,l}$$

$$Z = \frac{e^{-N\alpha}}{2^{N-1}}\left[\cosh(N\mu)\, I_1^N(\alpha) + \rho_+ I_0^N(\alpha)\right]$$

$$\left(\alpha \equiv \frac{1}{2g^2}, \quad I_n\text{: modified Bessel function of 1st kind} \quad 2\rho_\pm \equiv \left(\sqrt{m^2+1}+m\right)^N \pm \left(\sqrt{m^2+1}-m\right)^N\right)$$

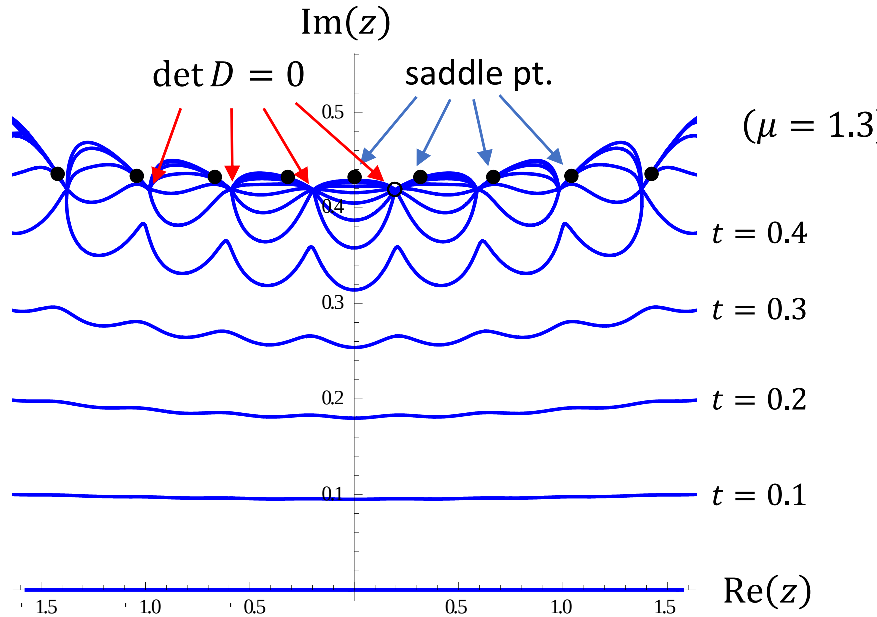
Chiral condensation

$$\langle \bar{\chi}\chi \rangle = \langle \frac{1}{N}\mathrm{tr} D^{-1}(U) \rangle = \frac{\rho_- I_0^N(\alpha)}{\sqrt{m^2+1}\left[\cosh(N\mu) I_1^N(\alpha) + \rho_+ I_0^N(\alpha)\right]}$$

→ We calculate the chiral condensation numerically by using LTM.

(0+1)D massive Thirring model (3/4)

We calculate the region $0 < \mu < 2$ with $N = 8, g^2 = \frac{1}{2}$, and $m = 1$.



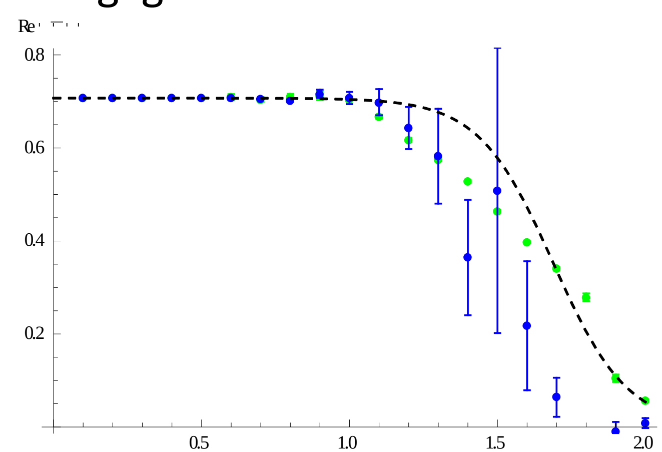
The integration contours stick to thimbles when $T \gtrsim 1$.

➡ We set $T = 2$.

The main contribution is that from the saddle point $\mathrm{Re}(z) \equiv Re\left(\frac{1}{N}\sum_n z_n\right) = 0$.
However, the neighboring saddle points have non-negligible contributions.

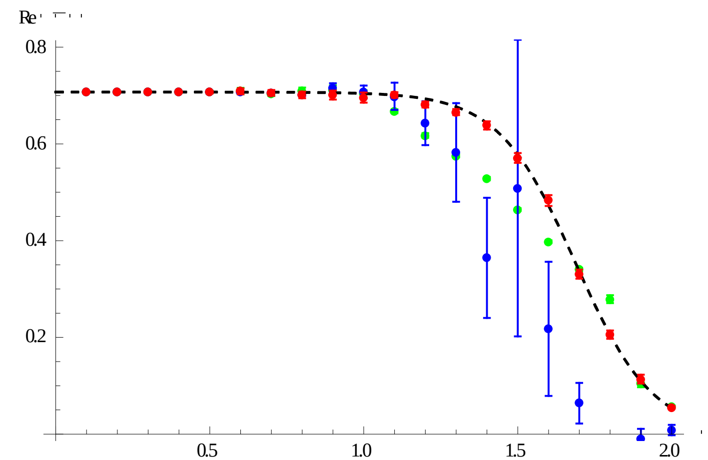
⇒ ordinary algorithm gives wrong results.

blue: T=0 (reweighting)
green: T=2 (without PT)
dotted: analytic solution



(0+1)D massive Thirring model (4/4)

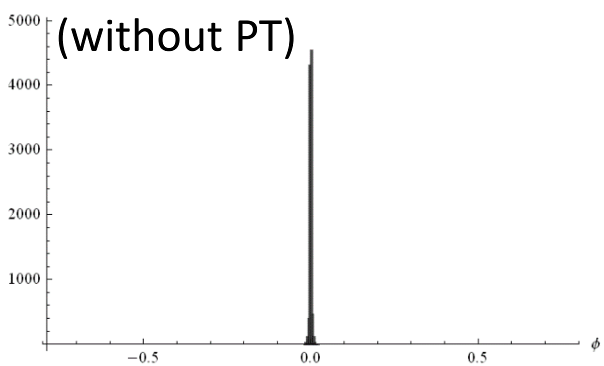
We set $\{t_a\} = \{t_0 = 0, t_1 = 0.1, \cdots, t_{19} = 1.9, t_{20} = 2.0\}$
and introduce parallel tempering algorithm.

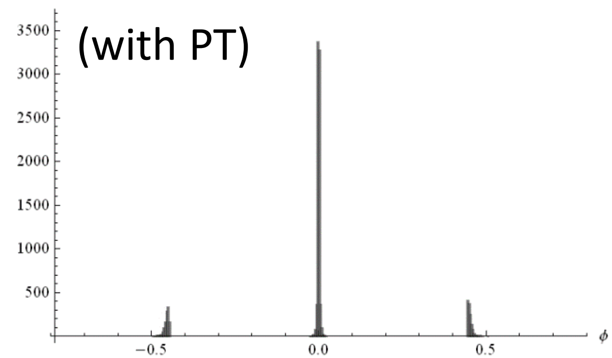


blue: T=0 (reweighting)
green: T=2 (without PT)
red: T=2 (with PT)
dotted: analytic solution

⇒ The results agree with the analytic solution.

The histogram of $\phi = \frac{1}{N}\sum_n \phi_n$





$(\mu = 1.3)$

Several thimbles contribute correctly.

Plan of talk

- Introduction
- Parallel tempering algorithm for LTM
  - Simulated tempering
  - Parallel tempering
  - Parallel tempering for LTM
- Examples
  - (0+1)D massive Thirring model
  - <u>One matrix model with quartic potential</u>
- Summary

One matrix model

Example 2 (one-matrix model with quartic potential)

$$S(M) = N\beta \, \mathrm{trV(M)}$$

$$V(x) = -\frac{c}{2}x^2 + \frac{1}{4}x^4$$

(M: $N \times N$ Hermitian matrix, $c = e^{i\theta}$)

- sign problem occurs if $c \in \mathbb{C}$.
- Analytically solvable in $N \to \infty$ limit.
- Continuous symmetry ($SU(N)$) ➡ Lefschetz thimble is not well-defined.

However, antiholomorphic gradient flow is well-defined even without Lefschetz thimble.

$\Rightarrow$ LTM can work in this case.

Furthermore, there can be a 3-cut solution for $c \in \mathbb{C}$.

↑

The topic of the rest of my talk

Matrix model (real, 1/2)

First I review the real case of this model.

$$S(M) = N\beta \,\mathrm{tr}V(M)$$
$$V(x) = -\frac{c}{2}x^2 + \frac{1}{4}x^4 \quad (M: N \times N \text{ Hermitian matrix, } c = \pm 1)$$

<u>$N \to \infty$ limit</u>

- Eigenvalues distribute continuously $\Rightarrow \rho(x)$
- We can solve this model by using the resolvent $\omega(z) \equiv \left\langle \frac{1}{N} \mathrm{tr} \left(\frac{1}{z-M} \right) \right\rangle = \int dx \frac{\rho(x)}{z-x}$

EOM give quadratic equation of $\omega(z)$:

$$\omega(z)^2 - \beta V'(z)\omega(z) - \beta Q(z) = 0$$

$$\left(Q(z) \equiv \left\langle \frac{1}{N} \mathrm{tr} \left(\frac{V'(z) - V'(M)}{z - M} \right) \right\rangle \text{ : polynomial of } z \right)$$

$$\Rightarrow \omega(z) = \frac{\beta}{2} V'(z) - \sqrt{\text{polynomial}}$$

$V(z)$: 4th order $\Rightarrow$ (polynomial): 6th order

Matrix model (real, 2/2)

Properties of $\omega(z)$

$$\omega(z) \to \frac{1}{z} \quad (|z| \to \infty)$$
$$\omega(x+i\epsilon) - \omega(x-i\epsilon) = 2\pi i\rho(x)$$

Eigenvalues distribute in the potential $V(x) = -\frac{c}{2}x^2 + \frac{1}{4}x^4$

Two types of solution

1-cut solution: $\omega(z) = \frac{\beta}{2}\left(V'(z) - \sqrt{\left(z^2 + \frac{a^2}{2} - c\right)^2 (z^2 - a^2)}\right) \quad \left(a^2 \equiv \frac{2}{3}c\left(1 - \sqrt{1 + \frac{\beta}{12c^2}}\right)\right)$

2-cut solution: $\omega(z) = \frac{\beta}{2}\left(V'(z) - \sqrt{z^2\left(z^2 - c + \frac{2}{\sqrt{\beta}}\right)\left(z^2 - c - \frac{2}{\sqrt{\beta}}\right)}\right)$

- $c = -1$: only 1-cut solution is realized.
- $c = +1$: 1-cut solution is realized for $\beta < 4$,
  2-cut solution is realized for $\beta > 4$.

Especially, 3rd-order phase transition occurs at $\beta = \beta_c = 4$.

Matrix model (complex, 1/4)

Complex case

$$S(M) = N\beta\, \mathrm{trV(M)}$$
$$V(x) = -\frac{c}{2}x^2 + \frac{1}{4}x^4 \quad (M: N\times N \text{ Hermitian matrix, } c = e^{i\theta})$$

- The cut of $\omega(z)$ is no longer on the real axis.
- However, we can find the form of resolvent $\omega(z) = \frac{\beta}{2}(V'(z) - \sqrt{\mathrm{polynomial}})$.

6th order polynomial: in general there should be the 3-cut solution.

We can again solve the model by two steps:

1. Assume the number of cuts
2. Determine the form of $\omega(z)$ by $\omega(z) \to \frac{1}{z}$ $(|\mathrm{z}| \to \infty)$.

Matrix model (complex, 2/4)

Complex case

$$S(M) = N\beta\, \mathrm{trV(M)}$$
$$V(x) = -\frac{c}{2}x^2 + \frac{1}{4}x^4 \quad (M: N\times N \text{ Hermitian matrix, } c = e^{i\theta})$$

result

1-cut solution: $\omega(z) = \frac{\beta}{2}\left(V'(z) - \sqrt{\left(z^2 + \frac{a^2}{2} - c\right)^2 (z^2 - a^2)}\right) \quad \left(a^2 \equiv \frac{2}{3}c\left(1 - \sqrt{1 + \frac{\beta}{12c^2}}\right)\right)$

2-cut solution: $\omega(z) = \frac{\beta}{2}\left(V'(z) - \sqrt{z^2\left(z^2 - c + \frac{2}{\sqrt{\beta}}\right)\left(z^2 - c - \frac{2}{\sqrt{\beta}}\right)}\right)$

3-cut solution: $\omega(z) = \frac{\beta}{2}\left(V'(z) - \sqrt{z^6 - 2cz^4 + \left(c^2 - \frac{\beta}{4}\right)z^2 - u(\beta, c)}\right)$

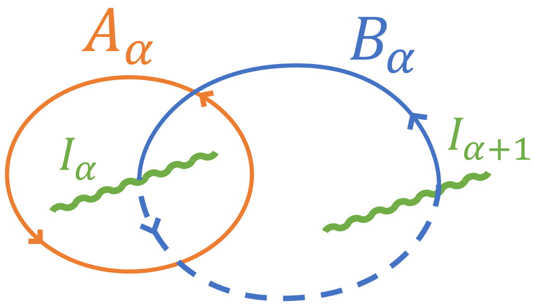
There remains 1 (complex) parameter.

Matrix model (complex, 3/4)

The parameter $u(\beta, c)$ can be determined as following:

David's discussion [David 1992]

The real part of free energy should be minimized
with constraint that
filling fraction $n_\alpha \in [0,1]$ (especially $n_\alpha \in \mathbb{R}, \Sigma_\alpha n_\alpha = 1$)



Filling fraction (for cut I_α) : $n_\alpha = \int_{A_\alpha} \frac{dz}{2\pi i} \omega(z)$

Minimize Re(free energy) ⇔ Re(chemical potential μ_α) = μ (α-indep.)

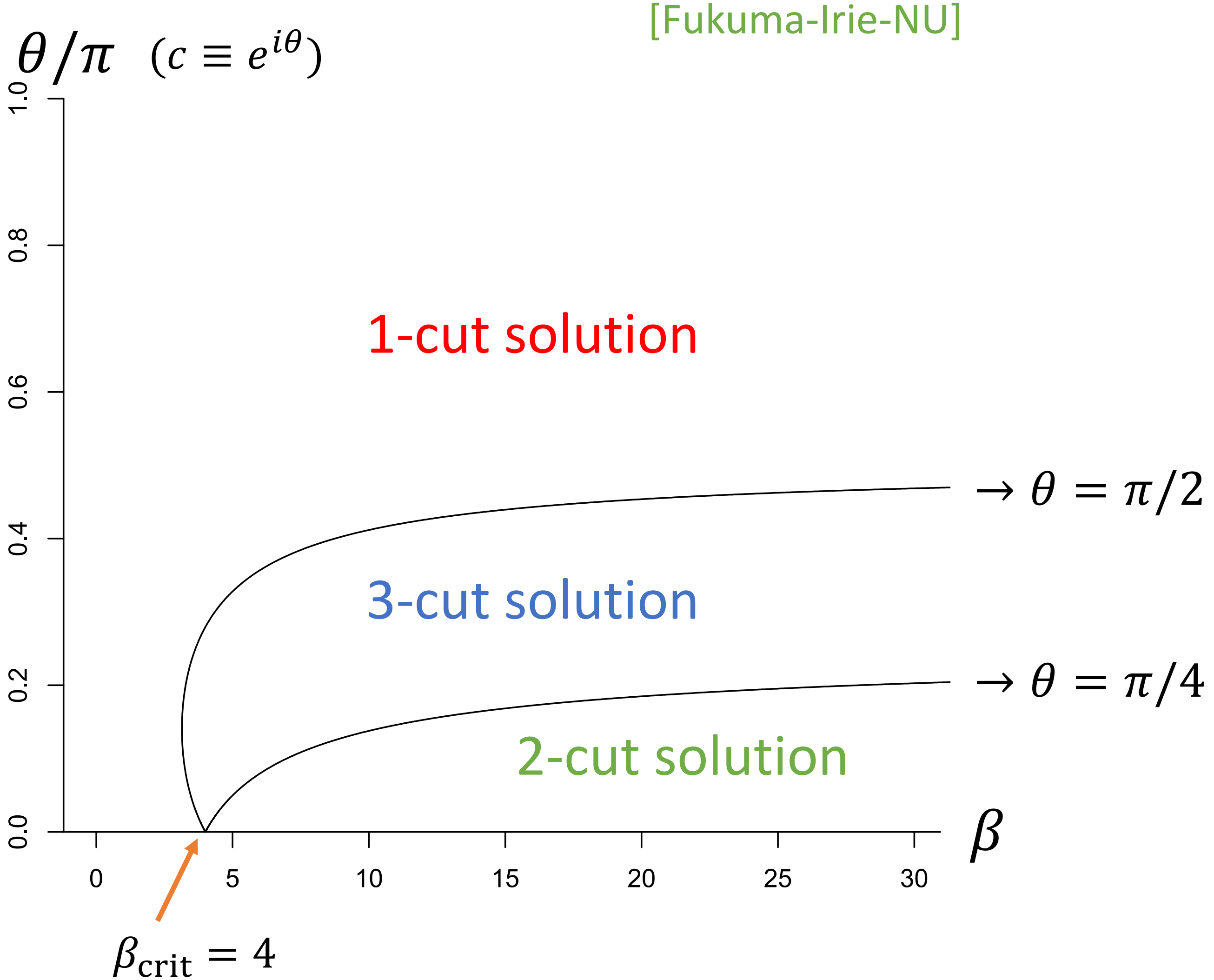
Difference between chemical potentials: $\mu_{\alpha+1} - \mu_\alpha = \int_{B_\alpha} \frac{dz}{2\pi i} \omega(z)$

$\text{Im}\left[\int_{A_\alpha} \frac{dz}{2\pi i} \omega(z)\right] = 0$, $\text{Re}\left[\int_{B_\alpha} \frac{dz}{2\pi i} \omega(z)\right] = 0$ ⇒ $u(\beta, c)$ is determined.

Phase diagram (prospection)

The phase diagram is prospected as follows:

[Fukuma-Irie-NU]



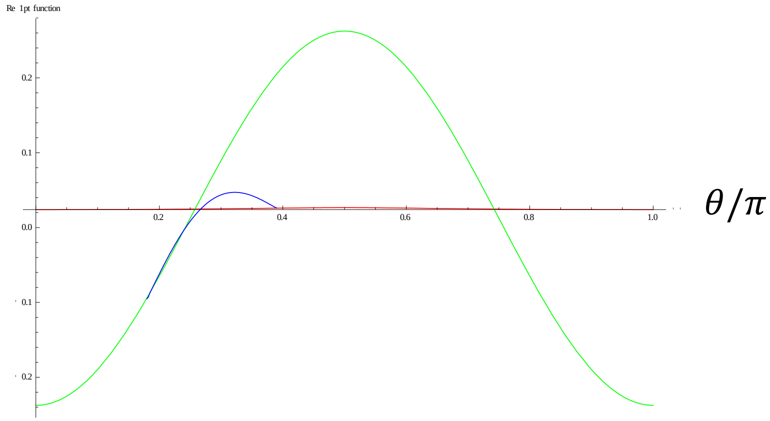
Matrix model (complex, 4/4)

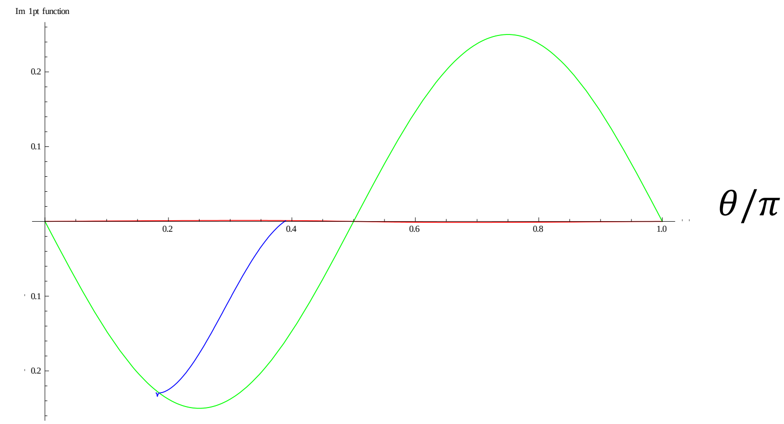
Red: 1-cut soln.
Green: 2-cut soln.
Blue: 3-cut soln.
$(c \equiv e^{i\theta})$

On $\beta = 20$ line, three types of solution should be appear.

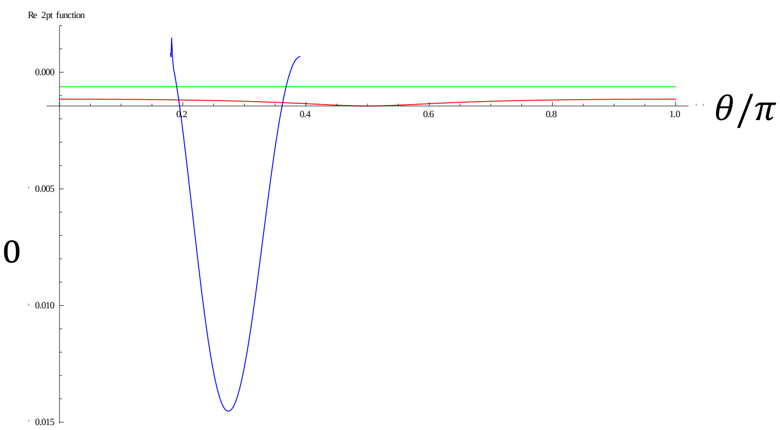
Re | Im

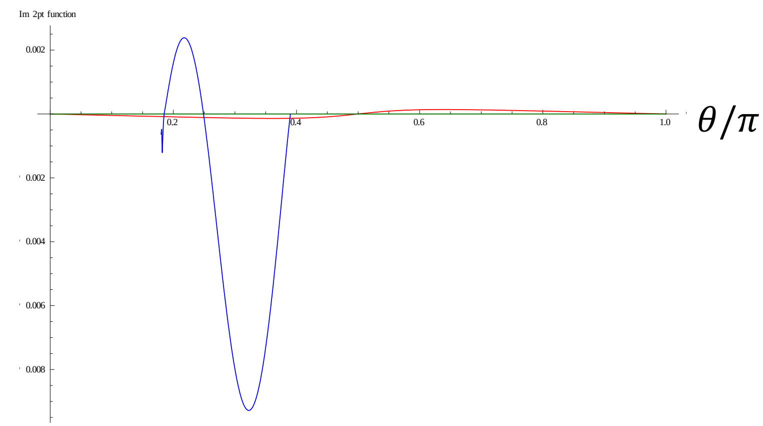
$\left.\frac{dF}{d\beta}\right|_{\beta=20}$





$\left.\frac{d^2F}{d\beta^2}\right|_{\beta=20}$





In particular, (probably) 3rd-order phase transitions occur at $\theta \sim 0.2\pi, 0.4\pi$.

Numerical results (1/2)

We calculate $\frac{dF}{d\beta} = \langle \frac{1}{N} \mathrm{tr} V(M) \rangle$ of this model by using LTM.

Parameters

- $N = 6$
- $\beta = 20$, $\theta \in [0, \pi]$ $(c \equiv e^{i\theta})$

  ➡ Sign problem

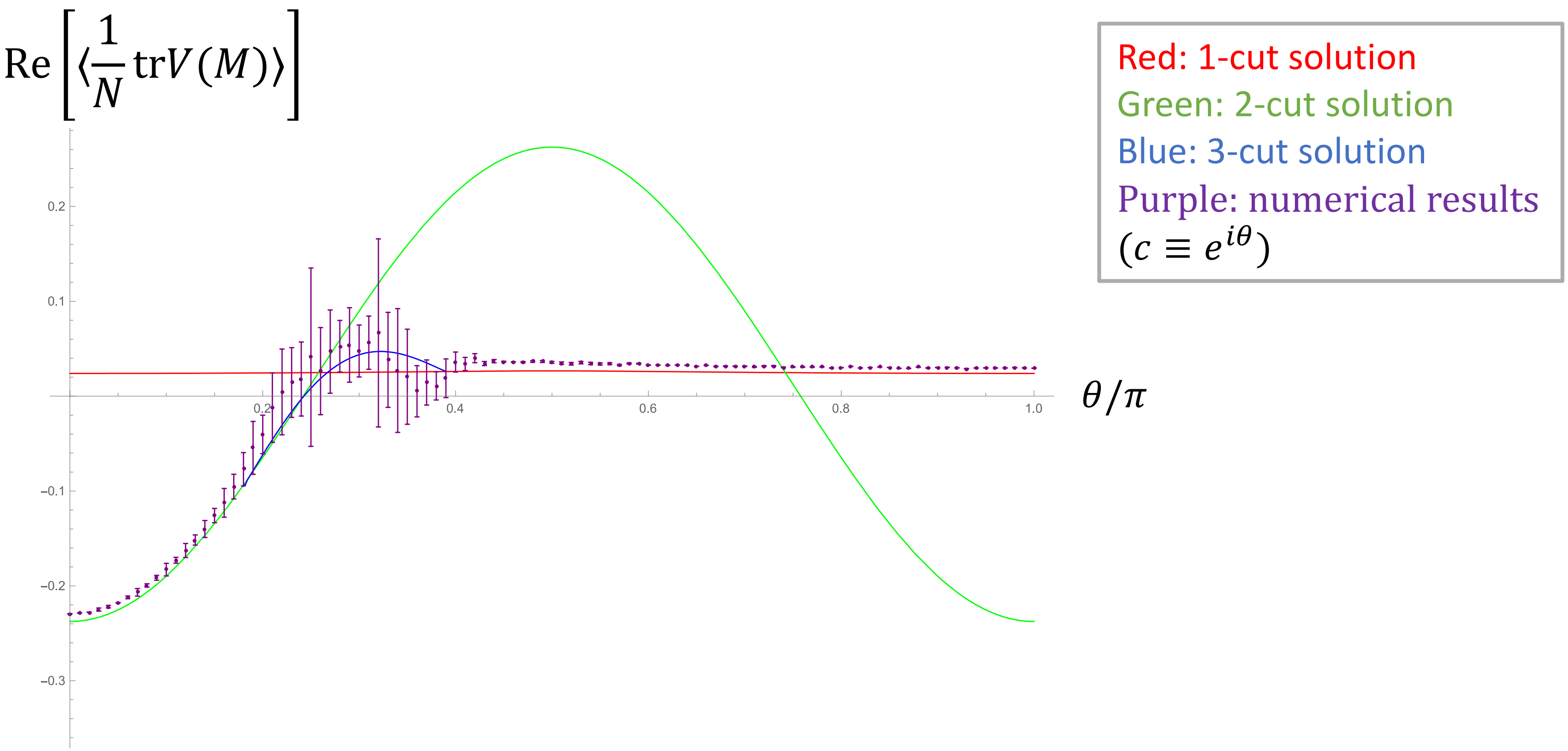
- We introduce two types of tempering parameter: (β, t)

  set of β: $\{\beta\} = \{5, 10, 15, 20\}$, set of t: $\{t\} = \left\{0, \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1\right\}$
- Dynamical variables: $N \times N$ Hermitian matrix (not diagonalized)

  ➡ The models have $SU(N)$ symmetry.

Numerical results (2/2)

Results



Numerical results correctly agree with analytical solutions.

Summary

We introduce the parallel tempering algorithm to LTM by regarding the flow time t as a tempering parameter.

⇒ This method is quite versatile:
We can calculate *any* models (in principle) without the fine-tuning of t.

(i.e. we can calculate the models
- which have multi-thimble contributions.
- which have continuous symmetry.)

Future problems

- LTM is Costly
  Computational cost $\propto$ $(\text{degrees of freedom})^3$
- Residual sign problem
- Global sign problem