

最大エントロピー法による 符号問題の解析

Masahiro Imachi, Y.S.^A and Hiroshi Yoneyama^B

(Nara National College of Technology, JAPAN^A, Saga Univ., JAPAN^B)

2 August '08

Reference: M. Imachi, Y.S. & H. Yoneyama, Prog. Theor. Phys. **115** (2006), 931.

Contents

1. Introduction

1.1 QCD vacuum

1.2 Introduction of θ term

1.3 Lattice field theory (in Euclidean space)

1.4 Phase structure in θ space

1.5 Our study

2. Maximum Entropy Method

2.1 Parameter fitting

2.2 Bayes' theorem

2.3 Maximum entropy method (MEM)

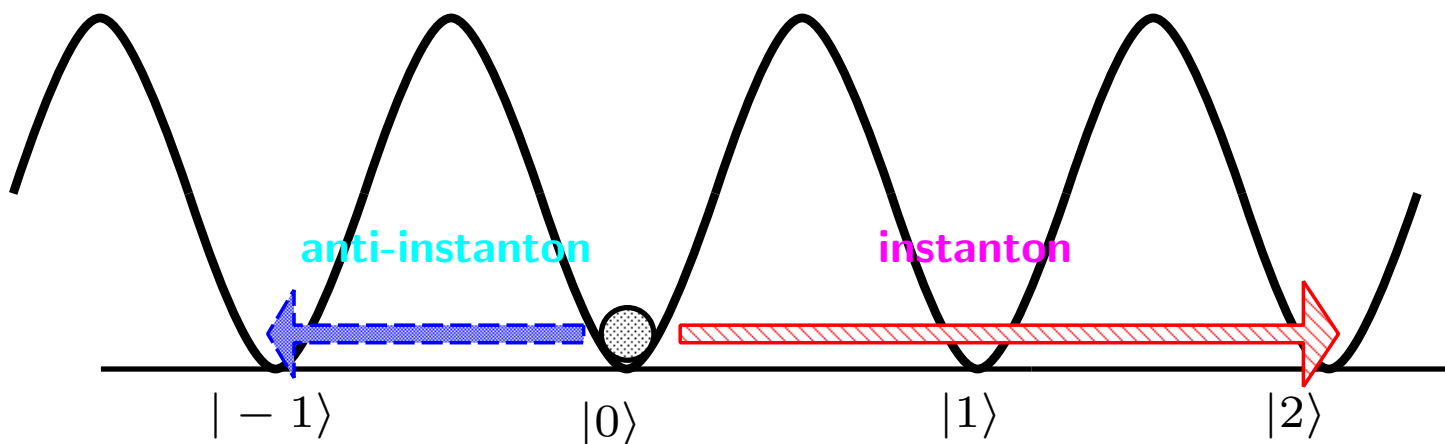
3. Results

4. Summary and Discussion

1. Introduction

- QCD {
- SU(3) gauge symmetry,
 - chiral symmetry and its breaking,
 - asymptotic freedom,
 - confinement,
 - θ -vacuum,
 - ...

1.1 QCD vacuum



- {
- The infinite number of vacua exists with different topology.
 - Instanton connects these vacua.

θ -vacuum : $|\theta\rangle = \sum_{n=-\infty}^{\infty} e^{-in\theta} |n\rangle$, \odot n : winding number

→ action in 4D Euclidean space : $S_\theta = S_{QCD} - i\theta \hat{Q}$.
 θ term

\odot \hat{Q} : Topological charge describes non-trivial topology of QCD.

New parameter “ θ ”

- **chiral anomaly:** (S.J. Adler (1969), J. S. Bell & R. Jackiw (1969))

$$\partial_\mu \langle j_5^\mu(x) \rangle = 2N_f Q, \quad \odot \quad j_5^\mu(x) = \sum_{i=1}^{N_f} \bar{\psi}_i(x) \gamma^\mu \gamma_5 \psi_i(x).$$

- **U(1) problem:** (S. L. Glashow (1967), S. Weinberg (1975), 't Hooft (1976))
mass relation: $m_{\eta'}^2 \gg m_\pi^2$.
decay rate of η meson: $\Gamma(\eta \rightarrow 3\pi) \sim 200 \text{ eV}$.
- **strong CP problem:** (R. J. Crewther et al. (1979), R. D. Peccei & H. R. Quinn (1977), V. Baluni (1979), P. G. Harris et al. (1999).)
 $|\theta| \lesssim \mathcal{O}(10^{-9})$.
- **rich phase structure in θ space:** (G. 't Hooft (1981), J. Cardy & E. Rabinovich (1982), N. Seiberg (1984), G. Shierholtz (1994), J. C. Plefka & S. Samuel (1996), M. Imachi et al. (1996))
SU(N) gauge theory, Z(N) spin model, CP^{N-1} model, \dots

In order to understand non perturbative properties of QCD vacuum,
it is indispensable to analyze QCD action with the θ term.

1.2. Introduction of θ term

ex)

- SU(2) pure gauge theory in Euclidean space

action:
$$S_{\text{Y.M.}} = \frac{1}{4g^2} \int d^4x F_{\mu\nu}^a(x) F_{\mu\nu}^a(x),$$

☺
$$F_{\mu\nu}^a(x) \equiv \partial_\mu A_\nu^a(x) - \partial_\nu A_\mu^a(x) + g\epsilon^{abc} A_\mu^b(x) A_\nu^c(x).$$

→ finite-action configurations: $\lim_{|x| \rightarrow \infty} F_{\mu\nu}(x) = 0.$

☺
$$\begin{cases} \lim_{|x| \rightarrow \infty} A_\mu(x) = 0. \\ \lim_{|x| \rightarrow \infty} A'_\mu(x) = \lim_{|x| \rightarrow \infty} iU(x) \partial_\mu U^\dagger(x). \end{cases}$$

(gauge transformation: $A_\mu(x) \longrightarrow U(x) (A_\mu(x) + i\partial_\mu) U^\dagger(x),$
 ☺ $U(x) = e^{i\theta^a T^a} . ; T^a = \frac{\tau^a}{2} : \text{generator of the group SU(2)})$

↓

$U(x) :$ element of the group SU(2) and a mapping of S^3 into S^3

☺ SU(2) gauge theory in 4D Euclidean space has non trivial topology.

Chern-Simon number:

$$n_{\text{C.S.}}(t) = \frac{1}{8\pi^2} \int d^3x \epsilon^{ijk} \text{tr} [A_i(x) (\partial_j A_k(x) - \frac{2}{3} i A_j(x) A_k(x))] .$$

☺ $n_{\text{C.S.}}(t) = 0, \pm 1, \pm 2, \dots . \delta n_{\text{C.S.}}(t) = 0.$

→ Chern-Simon number is a gauge dependent value.

$\hat{T} |n_{\text{C.S.}}\rangle = |n_{\text{C.S.}} + 1\rangle, \quad \text{☺ } \hat{T}: \text{global gauge transformation}$

topological charge: $Q \equiv \frac{1}{32\pi^2} \int d^4x F_{\mu\nu}^a(x) \tilde{F}_{\mu\nu}^a(x).$
 $\odot \tilde{F}_{\mu\nu}^a(x) \equiv \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} F_{\rho\sigma}^a(x).$

$$Q = \frac{1}{8\pi^2} \int d^4x \partial_\mu \left\{ \epsilon_{\mu\nu\rho\sigma} \text{tr} \left[A_\nu(x) (\partial_\rho A_\sigma(x) - \frac{2}{3} i A_\rho(x) A_\sigma(x)) \right] \right\}$$

$$= \int d^4x \partial_\mu K_\mu(x) = n_{\text{C.S.}}(t = \infty) - n_{\text{C.S.}}(t = -\infty).$$

instanton solution: $\int d^4x \left(F_{\mu\nu}^a(x) \pm \tilde{F}_{\mu\nu}^a(x) \right)^2 \geq 0.$

$$\odot S_{\text{Y.M.}} \geq \frac{8\pi^2}{g^2} |Q|. \Rightarrow S_{\text{Y.M.}}^{(\text{inst})} = \frac{8\pi}{g^2} |Q|.$$

→ Instanton configuration corresponds to tunnelling between different vacuum states, which is a classical solution.

Construction of gauge independent vacuum

$$|\text{vac}\rangle = \sum_{n=-\infty}^{\infty} c_n |n\rangle.$$

$$\hat{T} |\text{vac}\rangle = e^{i\theta} |\text{vac}\rangle.$$

$$l.h.s. = \hat{T} \sum_{n=-\infty}^{\infty} c_n |n\rangle = \sum_{n=-\infty}^{\infty} c_n |n+1\rangle = \sum_{n=-\infty}^{\infty} c_{n-1} |n\rangle = r.h.s.$$

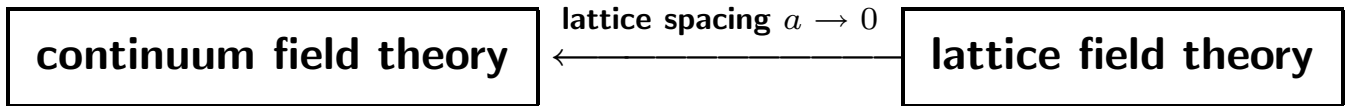
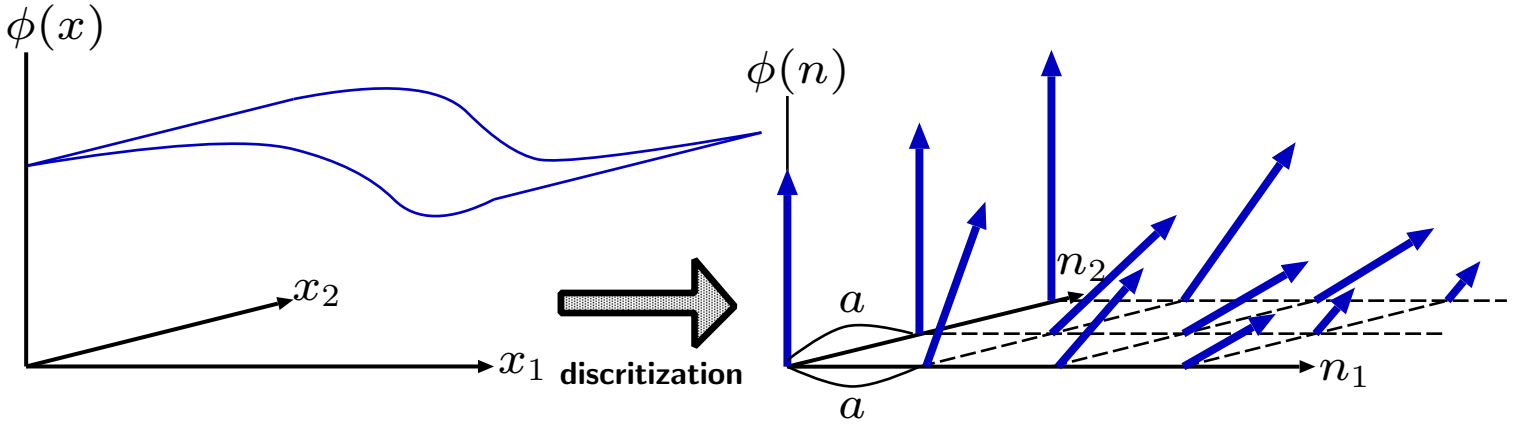
$$\odot c_{n-1} = e^{i\theta} c_n. \Rightarrow c_0 = e^{in\theta} c_n. \text{ Setting } c_0 = 1, \odot c_n = e^{-in\theta}.$$

$$\theta\text{-vacuum: } |\theta\rangle = \sum_{n=-\infty}^{\infty} e^{-in\theta} |n\rangle.$$

\odot Non perturbative effect is important for the vacuum.

1.3 Lattice field theory (in Euclidean space)

analysis of non perturbative effect : $\left\{ \begin{array}{l} \bullet \text{ WKB,} \\ \bullet \text{ } 1/N \text{ expansion,} \\ \bullet \text{ lattice field theory,} \\ \dots \end{array} \right.$



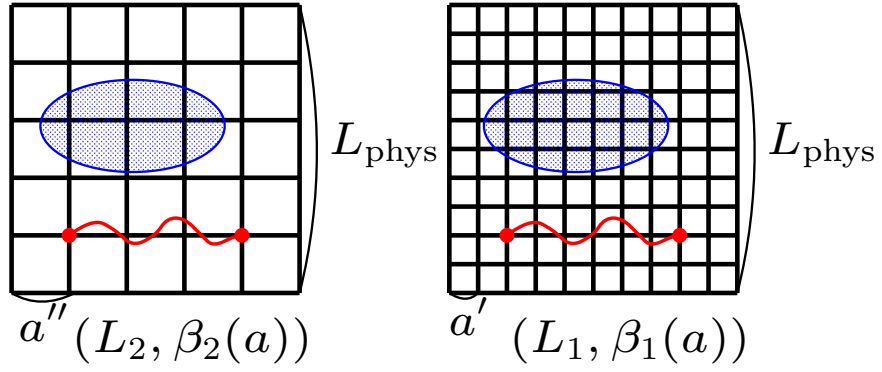
- | | |
|---|---|
| $\left\{ \begin{array}{l} \bullet \text{ field: } \phi(x), \psi(x). \\ \bullet \text{ gauge field: } A_\mu(x). \\ \bullet \text{ action: } S[\phi, \psi, A_\mu]. \end{array} \right.$ | $\left\{ \begin{array}{l} \bullet \text{ lattice field: } \hat{\phi}(n), \hat{\psi}(n). \\ \bullet \text{ link variable: } U_\mu(n). \\ \bullet \text{ lattice action: } S[\hat{\phi}, \hat{\psi}, U_\mu]. \end{array} \right.$ |
|---|---|

→ Which 'lattice world' realizes the continuum field theory?

observable: $\mathcal{O}_{\text{phys}} = a^{-d} \lim_{a \rightarrow 0} \hat{\mathcal{O}}_{\text{lat}}[a] ; [\mathcal{O}_{\text{phys}}] = M^d.$

→ correlation length ξ : $\xi_{\text{phys}} = a \hat{\xi}_{\text{lat}}[a]. \odot \lim_{a \rightarrow 0} \hat{\xi}_{\text{lat}}[a] = \frac{\xi_{\text{phys}}}{a} = \infty.$

\odot The continuum limit will be realized at a 2nd. order phase transition point.



$$\frac{\hat{\mathcal{O}}_{\text{lat}}^{(1)}[a']}{\hat{\mathcal{O}}_{\text{lat}}^{(2)}[a']} = \frac{\hat{\mathcal{O}}_{\text{lat}}^{(1)}[a'']}{\hat{\mathcal{O}}_{\text{lat}}^{(2)}[a'']} \longrightarrow \frac{\hat{\mathcal{O}}_{\text{lat}}^{(1)}[a]}{\hat{\mathcal{O}}_{\text{lat}}^{(2)}[a]} = \frac{a^{-d} \hat{\mathcal{O}}_{\text{lat}}^{(1)}[a]}{a^{-d} \hat{\mathcal{O}}_{\text{lat}}^{(2)}[a]} \simeq \frac{\hat{\mathcal{O}}_{\text{phys}}^{(1)}}{\hat{\mathcal{O}}_{\text{phys}}^{(2)}} \text{ for small } a.$$

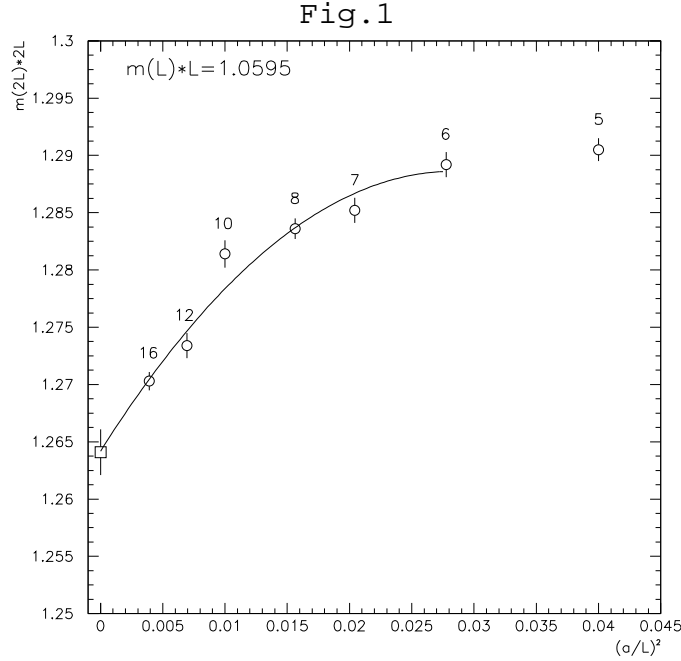


Figure 1: The cut-off dependence of $m(2L)2L$ for given value of $m(L)L = 1.0595$, taken from ref.[1]. The open box shows the extrapolated value.

Monte Carlo simulation

partition function: $Z(\theta) \equiv \langle \theta | e^{-\mathcal{H}t} | \theta \rangle = \sum_{Q=-\infty}^{\infty} \int [DA_\mu]_Q e^{-S+i\theta\hat{Q}}.$

Boltzmann factor: $e^{-S+i\theta\hat{Q}}$: **complex!**



complex action problem/sign problem

c.f.) lattice QCD with finite density, quantum spin model, ...

→ transformation of the partition function

partition function:
$$\mathcal{Z}(\theta) \equiv \frac{\int [DA_\mu] e^{-S[A]+i\theta\hat{Q}[A]}}{\int [DA_\mu] e^{-S[A]}}$$

$$= \frac{\sum_Q \int [DA_\mu] \delta_{Q,\hat{Q}[A]} e^{-S[A]} e^{i\theta Q}}{\int [DA_\mu] e^{-S[A]}}$$

$$= \sum_Q e^{i\theta Q} \frac{\int [DA_\mu] \delta_{Q,\hat{Q}[A]} e^{-S[A]}}{\int [DA_\mu] e^{-S[A]}} \equiv \sum_Q e^{i\theta Q} P(Q).$$

⊙ $P(Q) \equiv \frac{\int [DA_\mu] \delta_{Q,\hat{Q}[A]} e^{-S[A]}}{\int [DA_\mu] e^{-S[A]}}$: topological charge distribution

(

observable: $\langle \mathcal{O} \rangle_\theta = \frac{\int [DA_\mu] \mathcal{O}[A] e^{-S+i\theta\hat{Q}}}{\mathcal{Z}(\theta)} = \frac{\sum_Q P(Q) \langle \mathcal{O} \rangle_Q e^{i\theta Q}}{\mathcal{Z}(\theta)},$

⊙ $\langle \mathcal{O} \rangle \equiv \frac{\int [DA_\mu] \delta_{Q,\hat{Q}[A]} \mathcal{O}[A] e^{-S}}{\int [DA_\mu] \delta_{Q,\hat{Q}[A]} e^{-S}}.$

)

1.4 Phase structures in θ space [2, 3, 4, 5, 6, 7, 8]

- CP^{N-1} model with the θ term

$$\text{CP}^{N-1} \text{ model} \left\{ \begin{array}{l} \bullet \text{ U}(1) \text{ gauge symmetry,} \\ \bullet \text{ asymptotic freedom,} \\ \bullet \text{ confinement,} \\ \bullet \text{ dynamical mass generation,} \\ \bullet \theta\text{-vacuum, } \dots \end{array} \right.$$

$$\text{action: } S = \beta \int d^2x \sum_{\alpha=1}^N |D_\mu z_\alpha(x)|^2; \quad \sum_{\alpha=1}^N |z_\alpha(x)|^2 = 1,$$

$$\odot \quad \beta = \frac{1}{g^2}, \quad D_\mu = \partial_\mu + iA_\mu(x), \quad A_\mu(x) = i \sum_{\alpha=1}^N \bar{z}_\alpha(x) z_\alpha(x).$$

$$\text{topological charge: } \hat{Q} = \frac{i}{2\pi} \int d^2x \epsilon_{\mu\nu} \sum_{\alpha=1}^N D_\mu z_\alpha(x) D_\nu z_\alpha(x).$$

↓ discretization

$$\text{naive lattice action: } S^{(lat)} = \beta \sum_{n,\mu} \left[1 - \sum_{\alpha=1}^N |\bar{z}_\alpha(n) z_\alpha(n + \hat{\mu})|^2 \right].$$

$$\text{geometrical charge: } \hat{Q}^{(lat)} = \frac{1}{2\pi} \sum_{n^*} q(n^*),$$

$$\odot \quad q(n^*) \equiv \frac{1}{2} \sum_{\mu,\nu} \epsilon_{\mu\nu} \{ \theta_{n,\mu} + \theta_{n+\hat{\mu},\nu} - \theta_{n+\hat{\nu},\mu} - \theta_{n,\nu} \} \pmod{2\pi},$$

$$\theta_{n,\mu} \equiv \arg \left\{ \sum_{\alpha=1}^N \bar{z}_\alpha(x) z_\alpha(n + \hat{\mu}) \right\}.$$

partiton function: $\mathcal{Z}(\theta) = \frac{\int [d\bar{z}dz] e^{-S(\bar{z},z) + i\theta \hat{Q}(\bar{z},z)}}{\int [d\bar{z}dz] e^{-S(\bar{z},z)}}$.

↓

complex action problem

→ transformation of the partition function

$$\begin{aligned} \mathcal{Z}(\theta) &= \frac{\sum_Q \int [d\bar{z}dz] \delta_{Q, \hat{Q}(\bar{z},z)} e^{-S(\bar{z},z)} e^{i\theta Q}}{\int [d\bar{z}dz] e^{-S(\bar{z},z)}} \\ &= \sum_Q e^{i\theta Q} \frac{\int [d\bar{z}dz] \delta_{Q, \hat{Q}(\bar{z},z)} e^{-S(\bar{z},z)}}{\int [d\bar{z}dz] e^{-S(\bar{z},z)}} \equiv \sum_Q e^{i\theta Q} P(Q). \end{aligned}$$

⊙ $P(Q) \equiv \frac{\int [d\bar{z}dz] \delta_{Q, \hat{Q}(\bar{z},z)} e^{-S(\bar{z},z)}}{\int [d\bar{z}dz] e^{-S(\bar{z},z)}} : \text{ topological charge distribution}$

$$\left(\odot \int [d\bar{z}dz] \equiv \int d\bar{z}dz \prod_x \delta(|z(x)|^2 - 1) \right)$$

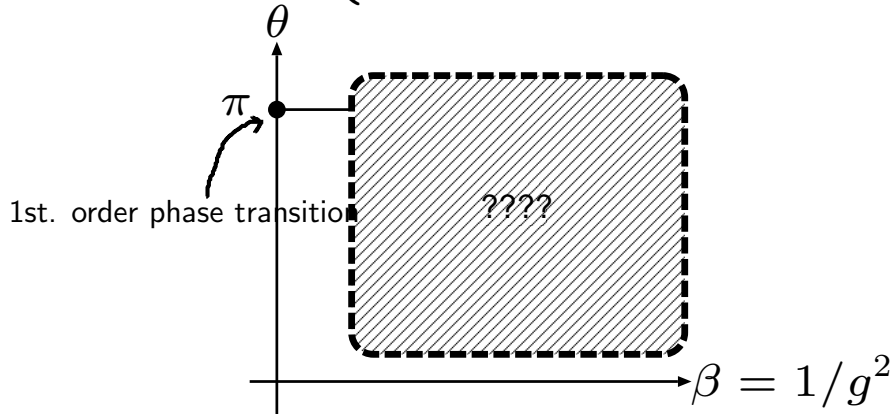
→ **strong coupling regions:** first order phase transition exists at $\theta = \pi$.
(confinement phase → Higgs phase)

weak coupling regions: phase transition point moves toward $\theta = 0$
with $\beta \rightarrow \infty$. **fault!!**

1.5 Our study [8, 9, 10, 11, 12]

model: lattice CP^{N-1} model with fixed point action [13, 14]

⊙ { **strong coupling regions:** first order phase transition exists at $\theta = \pi$.
weak coupling regions: { no phase transition exists for small volumes.
the method breaks down for large volumes.



• Why does the method break down?

⊙ Error of $P(Q)$ disturbs the behavior of the free energy density $f(\theta)$.

⊙ $f(\theta) \equiv -\frac{1}{V} \log \mathcal{Z}(\theta)$, ⊙ V : lattice volume

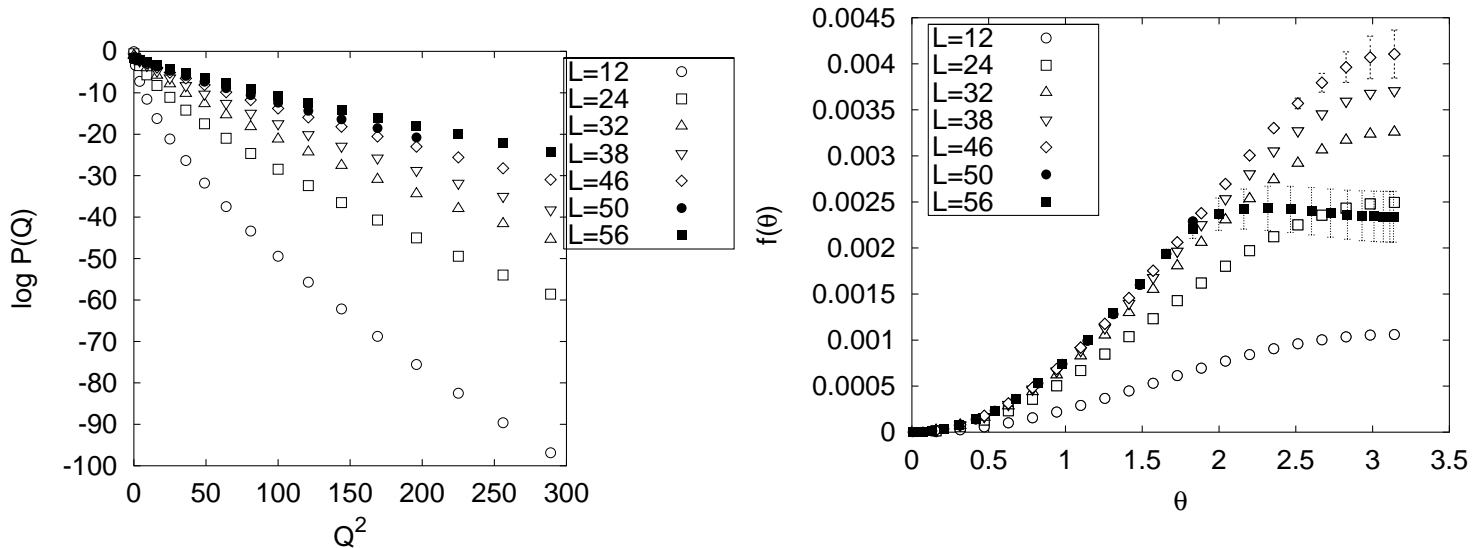


Figure 2: Monte Carlo data of CP^3 model with fixed point action. The coupling constant β is fixed to 3.0 and various lattice sizes L are employed. The number of sweeps is several millions for each case.

Flattening

$$\begin{aligned}
 \underbrace{P_{\text{MC}}(Q)}_{\text{Monte Carlo data}} &= \underbrace{P(Q)}_{\text{true value}} + \underbrace{\Delta P(Q)}_{\text{error}}. \\
 f_{\text{MC}}(\theta) &= -\frac{1}{V} \log \left[\sum_Q P(Q) e^{i\theta Q} + \sum_Q \Delta P(Q) e^{i\theta Q} \right] \\
 &\simeq -\frac{1}{V} \log \left[e^{-Vf(\theta)} + \Delta P(0) \right]. \\
 (\odot \quad |\Delta P(0)| \gg |\Delta P(Q \neq 0)|.) \\
 \odot f_{\text{MC}}(\theta) &\simeq \begin{cases} f(\theta) : & \theta < \theta_f. \\ -\frac{1}{V} \log \Delta P(0) : & \theta \gtrsim \theta_f. \\ & \text{flattening} \end{cases}
 \end{aligned}$$

In the Fourier transform method, the proper behavior of $f(\theta)$ is masked by errors of $P(Q)$ and flattening makes one mislead that a first order phase transition occurs at $\theta = \theta_f$.

- How much statistics are required?

$$\Delta P(0) \lesssim e^{-Vf(\theta)}, \quad \Delta P(0) \propto 1/\sqrt{N}. \quad \Rightarrow \quad \odot \mathcal{N} \gtrsim e^{+V}.$$

Exponentially increasing statistics are needed!

⇓ an alternative way to the Fourier transform

⊙ maximum entropy method (MEM)

Analysis of mock data with MEM

1. Gaussian $P(Q)$ [11]

$$P_G(Q) = Ae^{-\frac{c}{V}Q^2}, \quad \odot A, c, V : \text{ parameters}$$

→ Poisson sum formula

$$\mathcal{Z}_{\text{Poi}}(\theta) = A\sqrt{\frac{\pi V}{c}} \sum_{m=-\infty}^{\infty} e^{-\frac{V}{4c}(\theta-2\pi m)^2}.$$

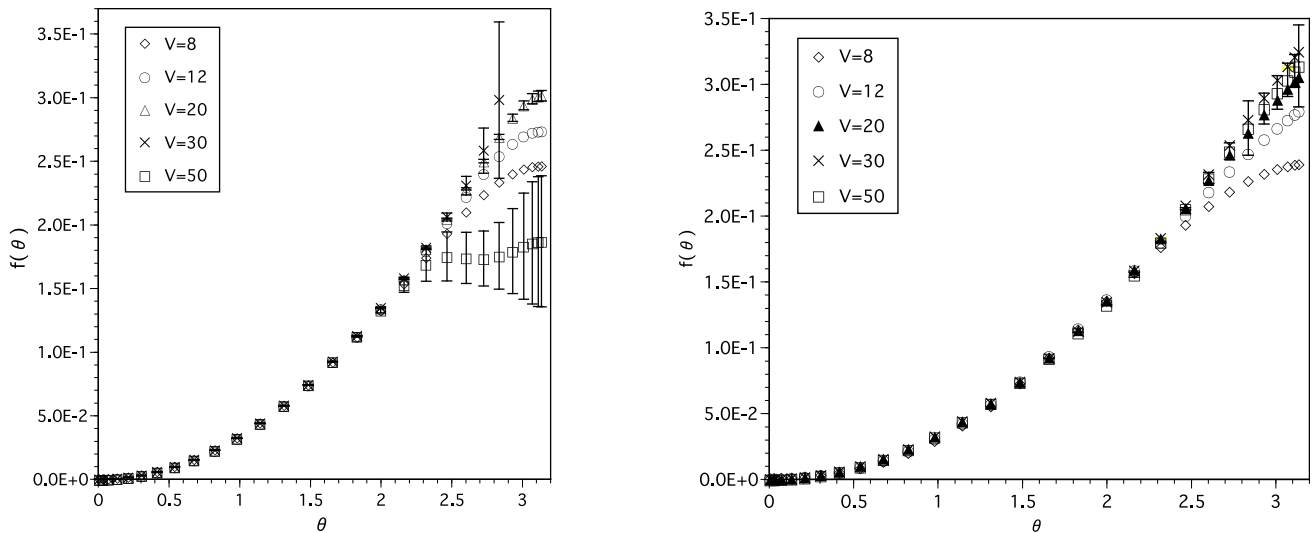


Figure 3: Free energy density $f(\theta)$ for various V . The parameter $c = 7.42$. In left panel, $f(\theta)$ is obtained by numerically Fourier-transforming $P_G(Q)$. In right panel, $f(\theta)$ is obtained by the MEM.

⊙ No flattening behavior is observed in all images obtained by MEM.

→ Does MEM reproduce just a smooth function?

2. $P(Q)$ with singular behavior [12]

$$P_S(Q) = P_G(Q) + p_0 \delta_{Q,0}, \quad \odot p_0 : \text{ parameter}$$

→ Poisson sum formula

$$\mathcal{Z}_S(\theta) = \mathcal{Z}_{\text{Poi}}(\theta) + \tilde{p}_0.$$

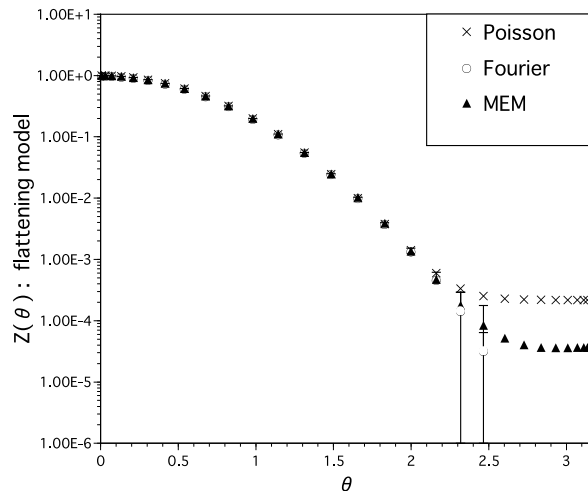


Figure 4: Partition function $\mathcal{Z}(\theta)$ for $V = 50$ and $p_0 = 1.0 \times 10^{-3}$. The Fourier method yields negative values for $\mathcal{Z}(\theta)$ at $\theta \lesssim 2.5$. The true $\mathcal{Z}(\theta)$ (Poisson) is also plotted.

\odot Obtained images reproduce flattening.

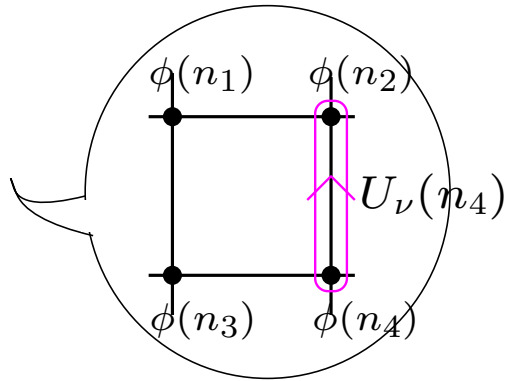
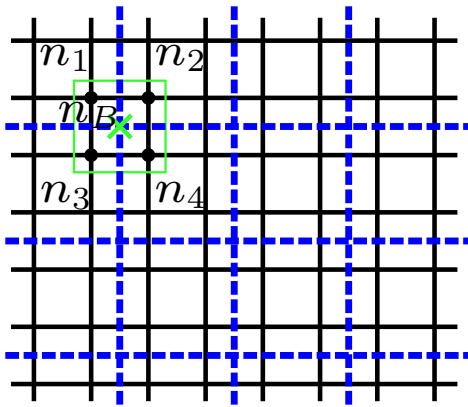
In this study,

- ♠ we apply MEM to Monte Carlo data and calculate the partition function of the CP^{N-1} model with the θ term,
- ♠ we estimate to what extent the most probable image is sensitive to the prior information,
- ♠ we compare the results of the MEM with those of the Fourier transform method, in order to investigate the effectiveness of MEM.

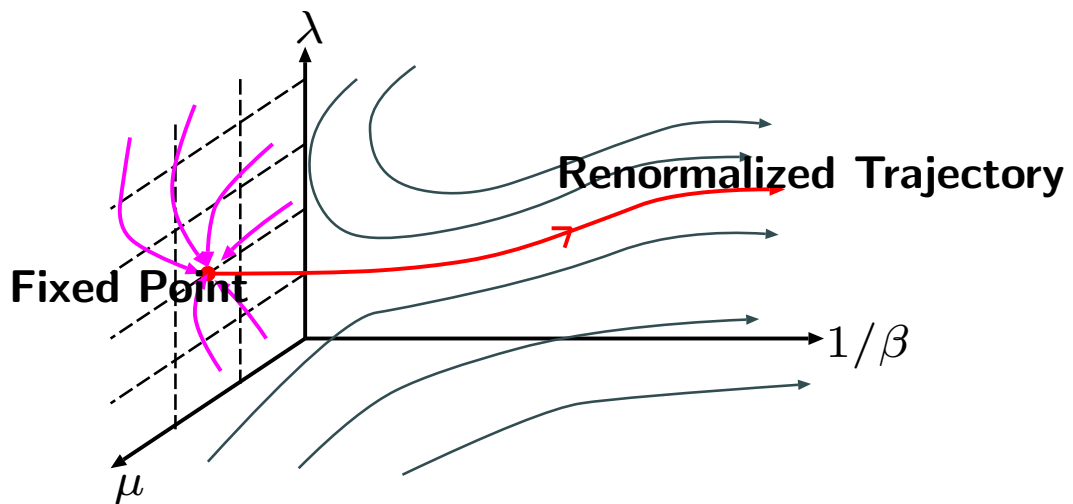
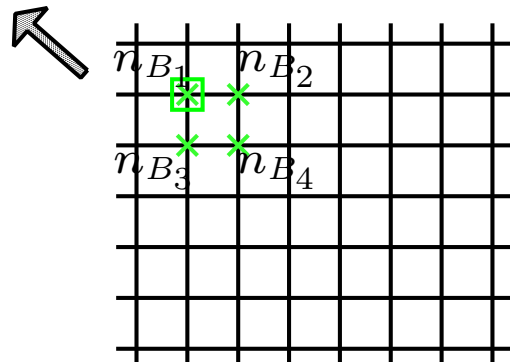
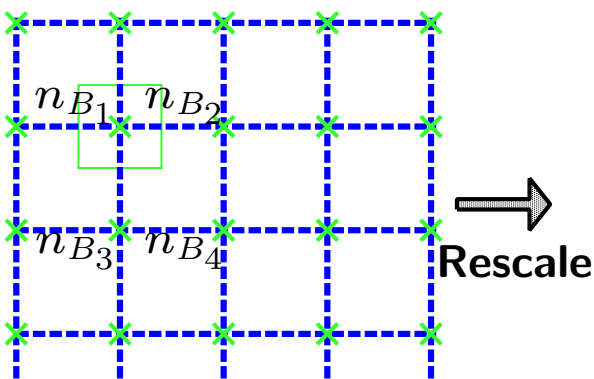
♡ Renormalization group and fixed point action

→ How is the extrapolation to $a \rightarrow 0$ taken?

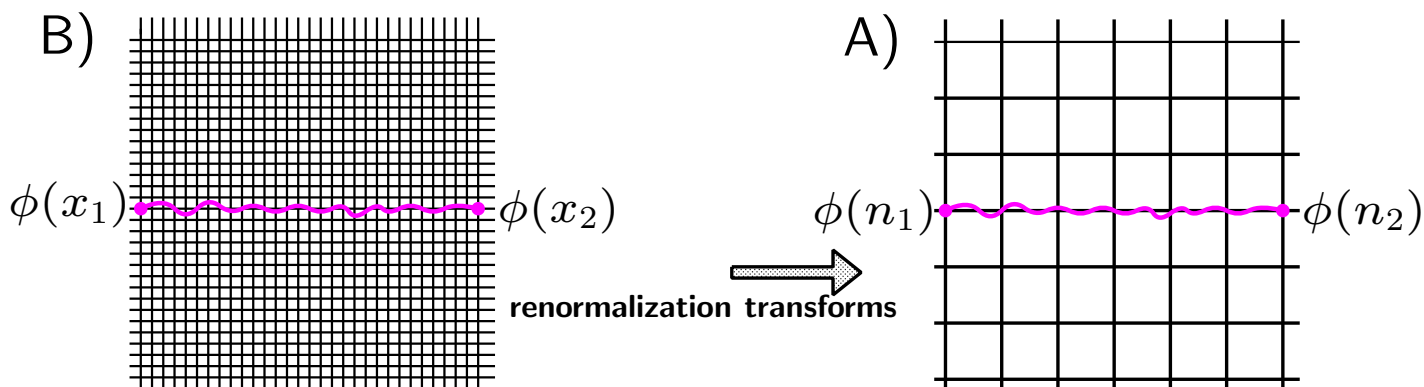
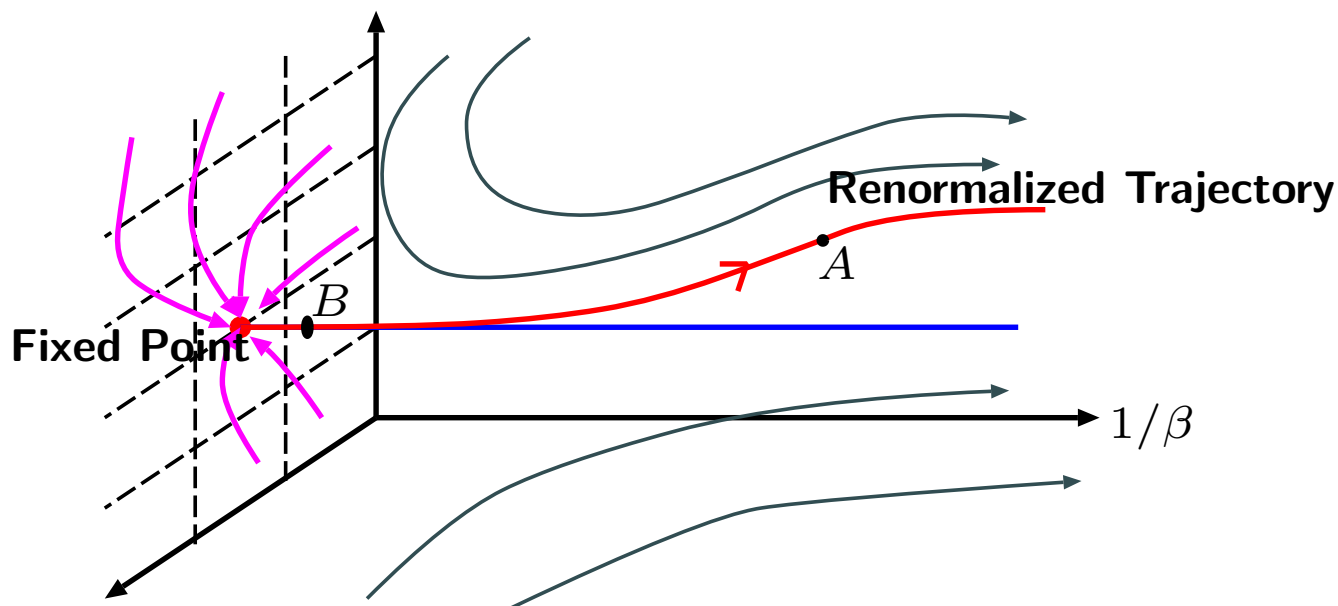
☹ renormalization group



Block transform ↓



♣ fixed point action [13, 14]



renormalization transformation: $e^{\beta' A'[\zeta]} = \int_z e^{-\beta A[z] + \mathcal{T}[z, \zeta]}$.

$\Downarrow \beta \rightarrow \infty$

$A'[\zeta] = \min_z [A[z] + \mathcal{T}[z, \zeta]]$. $\odot \mathcal{T}$: transformation kernel

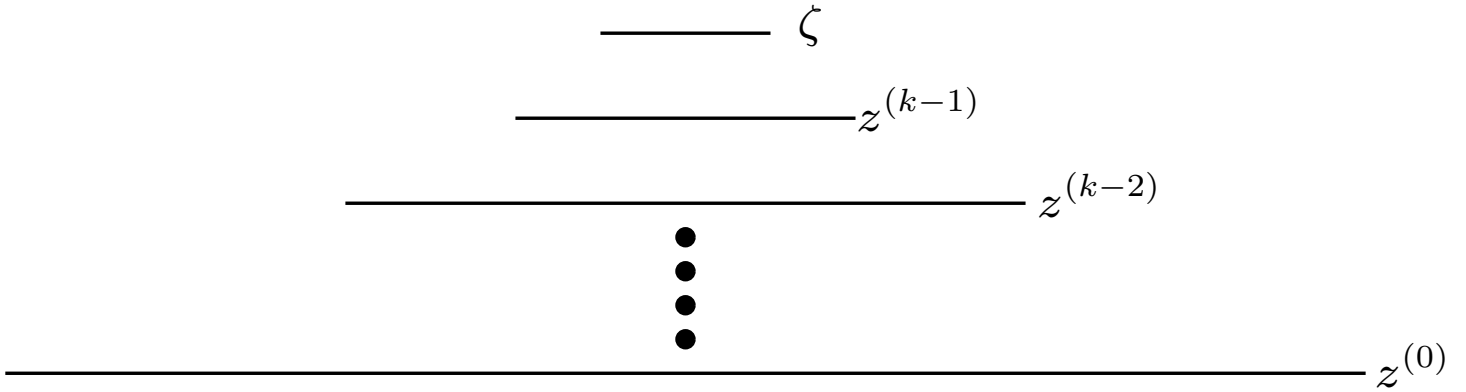
\odot **fixed point equation:** $A_{\text{FP}}[\zeta] = \min_z [A_{\text{FP}}[z] + \mathcal{T}[z, \zeta]]$
on the fixed point.

→ parameterization

$$\begin{aligned}
 A_{\text{FP}}[z] &= \sum_n \sum_r \rho(r) \left(1 - \sum_{\alpha=1}^N N |\bar{z}_\alpha(n) z_\alpha(n+r)|^2 \right) \\
 &+ \sum_{n_1, n_2, n_3, n_4} c(n_1, n_2, n_3, n_4) \left(1 - \sum_{\alpha=1}^N |\bar{z}_\alpha(n_1) z_\alpha(n_2)|^2 \right) \\
 &\quad \times \left(1 - \sum_{\alpha=1}^N |\bar{z}_\alpha(n_3) z_\alpha(n_4)|^2 \right) + \dots
 \end{aligned}$$

→ determination of couplings $\{\rho, c, \dots\}$

$$A^{(k)}[\zeta] = \min_{\{z^{(0)}, z^{(1)}, \dots, z^{(k-1)}\}} \left[A^{(0)}[z^{(0)}] + \mathcal{T}(z^{(0)}, z^{(1)}) + \mathcal{T}(z^{(1)}, z^{(2)}) \right. \\
 \left. + \dots + \mathcal{T}(z^{(k-1)}, \zeta) \right].$$



2. Maximum Entropy Method [15, 16, 17]

2.1 Parameter fitting

inverse Fourier transform: $P_{\text{MC}}(Q) = \int_{-\pi}^{\pi} d\theta \frac{e^{i\theta Q}}{2\pi} \mathcal{Z}(\theta)$.

→ How is $\mathcal{Z}(\theta)$ calculated by use of the inverse Fourier transform?

ex)

• **Least square method:**

$$\chi^2 = \sum_{Q, Q'} \left(P^{(Z)}(Q) - \overline{P_{\text{MC}}(Q)} \right) C_{Q, Q'}^{-1} \left(P^{(Z)}(Q') - \overline{P_{\text{MC}}(Q')} \right).$$

$$\odot \left\{ \begin{array}{l} P^{(Z)}(Q) = \sum_{n=1}^{N_\theta} \frac{e^{i\theta_n Q}}{2\pi} \mathcal{Z}(\theta_n) \equiv \sum_{n=1}^{N_\theta} K(Q, n) \mathcal{Z}_n; \\ \qquad \qquad \qquad N_\theta : \# \text{ of reference points in } \theta \text{ space} \\ \overline{P_{\text{MC}}(Q)} = \frac{1}{N_d} \sum_{l=1}^{N_d} P_{\text{MC}}^{(l)}(Q); \quad N_d : \# \text{ of data set} \\ C_{Q, Q'} : \text{ covariance matrix} \end{array} \right.$$

1. to assume the functional form of $\mathcal{Z}(\theta)$,
2. to calculate parameters characterizing $\mathcal{Z}(\theta)$,
such that χ^2 is minimized.

problems

- (a) necessity for an adequate functional form of $\mathcal{Z}(\theta)$
- (b) **ill-posed problem**: $\# \text{ of } Q \ll \# \text{ of } \theta_n$

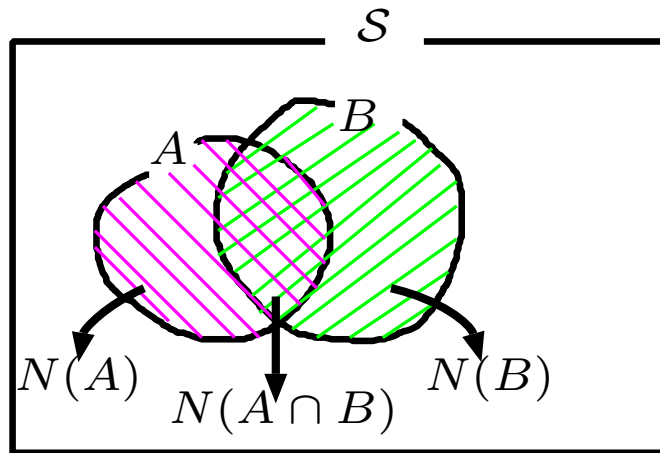
2.2 Bayes' theorem

$p(A)$: probability that an event A occurs

$p(A, B)$: probability that events A and B occur simultaneously

→ **conditional probability:** $p(A|B) \equiv \frac{p(A, B)}{p(B)}$,

probability that an event A occurs under the condition that B occurs



$$p(A) \equiv \frac{N(A)}{N(\mathcal{S})}, \quad p(B) \equiv \frac{N(B)}{N(\mathcal{S})}, \quad p(A, B) = p(B, A).$$

$$p(A|B) = \frac{N(A, B)}{N(B)} = \frac{N(A, B)/N(\mathcal{S})}{N(B)/N(\mathcal{S})} = \frac{p(A, B)}{p(B)}.$$

$$p(A|B) = \frac{p(A, B)}{p(B)} = \frac{p(B, A)}{p(B)} = \frac{p(B|A) \times p(A)}{p(B)}.$$

☺ Bayes' theorem: $p(A|B) = \frac{p(B|A)p(A)}{p(B)}$.

2.3 Maximum Entropy Method (MEM)

MEM

- widely used technique for various field,
- one of the parameter inference based on Bayes' theorem,
- to derive a unique solution by utilizing data and the prior information about the parameters.

ex.)

- lattice QCD (from ref. [17])

$$D(\tau, \mathbf{k}) \underset{\text{correlation function}}{=} \int_0^{+\infty} d\omega K_0(\tau, \omega) \underset{\text{spectral function}}{A(\omega, \mathbf{k})}$$

$$\odot K_0(\tau, \omega) : \text{kernel}, \quad A(\omega, \mathbf{k} = 0) \equiv \omega^2 \rho(\omega).$$

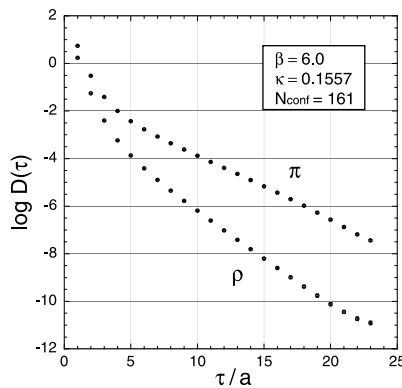


Figure 5: Lattice data of correlation function in the PS and V channels with errors.

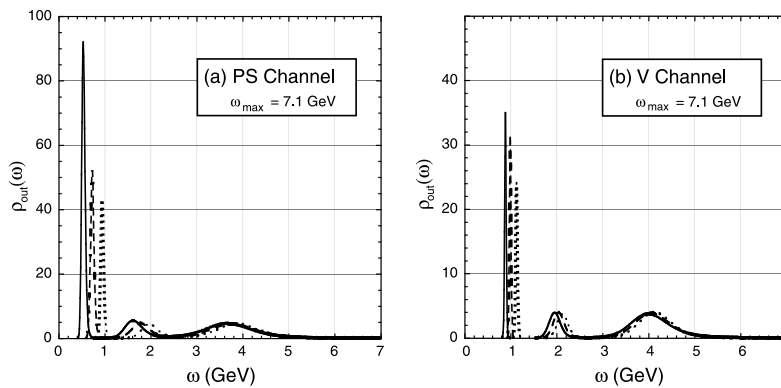


Figure 6: Reconstructed image of spectral functions for the PS (a) and V (b) channels.

posterior probability

posterior probability: probability that $\mathcal{Z}(\theta)$ is realized
when the Monte Carlo data of $\{P_{\text{MC}}(Q)\}$ and
the prior information I are given

$$\begin{aligned} \text{prob}(\mathcal{Z}(\theta)|P_{\text{MC}}(Q), I) &= \frac{\text{prob}(P_{\text{MC}}(Q)|\mathcal{Z}(\theta), I) \times \text{prob}(\mathcal{Z}(\theta)|I)}{\text{prob}(P(Q)|I)} \\ &\propto \underset{\text{likelihood function}}{\text{prob}(P_{\text{MC}}(Q)|\mathcal{Z}(\theta), I)} \times \underset{\text{prior probability}}{\text{prob}(\mathcal{Z}(\theta)|I)}. \end{aligned}$$

☺ prior information $I : \mathcal{Z}(\theta) > 0$.

c.f.) Bayes' theorem: $\text{prob}(A|B, C) = \frac{\text{prob}(B|A, C) \times \text{prob}(A|C)}{\text{prob}(B|C)}$.

$$P_{\text{MC}}(Q) = \int_{-\pi}^{\pi} d\theta \frac{e^{-i\theta Q}}{2\pi} \mathcal{Z}(\theta) \equiv \int_{-\pi}^{\pi} d\theta K(Q, \theta) \mathcal{Z}(\theta).$$

☺ Our task is to calculate the most probable image, $\hat{\mathcal{Z}}(\theta)$,
such that $\text{prob}(\mathcal{Z}(\theta)|P_{\text{MC}}(Q), I)$ is maximized.

$$\text{☺ } \text{prob}(\mathcal{Z}(\theta)|P_{\text{MC}}(Q), m, \alpha, I) \propto \exp \left[-\frac{1}{2}\chi^2 + \alpha S \right].$$

$$\text{☺ } \left\{ \begin{array}{l} S = \sum_n \left[\mathcal{Z}_n - m_n - \mathcal{Z}_n \log \frac{\mathcal{Z}_n}{m_n} \right] : \text{Shannon-Jaynes entropy} \\ m(\theta) : \text{default model that satisfies information } I \\ \alpha : \text{relative weight between } S \text{ and } \underline{\underline{-\frac{1}{2}\chi^2}} \end{array} \right.$$

Procedure for analysis

1. Maximizing $W[\mathcal{Z}]$ to obtain an α -dependent image for a given α .

$$\left. \frac{\delta W[\mathcal{Z}]}{\delta \mathcal{Z}_n} \right|_{\mathcal{Z}=\mathcal{Z}(\alpha)} = \left. \frac{\delta}{\delta \mathcal{Z}_n} \left[-\frac{1}{2} \chi^2 + \alpha S \right] \right|_{\mathcal{Z}=\mathcal{Z}(\alpha)} = 0.$$

☺ $\mathcal{Z}^{(\alpha)}(\theta)$: α -dependent most probable image for a given α

2. Averaging $\mathcal{Z}^{(\alpha)}(\theta)$ over α to obtain the most probable image, $\hat{\mathcal{Z}}(\theta)$.

$$\hat{\mathcal{Z}}_n = \int d\alpha \mathcal{Z}_n^{(\alpha)} \text{prob}(\alpha | P_{\text{MC}}(Q), m, I) \equiv \int d\alpha \mathcal{Z}_n^{(\alpha)} P(\alpha).$$

☺ $\hat{\mathcal{Z}}(\theta)$: **most probable image**

$$\text{☺} \left\{ \begin{array}{l} P(\alpha) \propto g(\alpha) \exp \left[W(\alpha) + \frac{1}{2} \sum_k \log \frac{\alpha}{\alpha + \lambda_k} \right] \\ \quad = g(\alpha) e^{W(\alpha) + \Lambda(\alpha)}. \quad : \text{posterior probability of } \alpha \\ g(\alpha) : \text{prior probability of } \alpha \\ g(\alpha) = \begin{cases} \text{const.} & : \text{Laplace's rule} \\ 1/\alpha & : \text{Jeffrey's rule} \end{cases} \\ \{ \lambda_k \} : \text{eigenvalues of a matrix } \left. \frac{1}{2} \sqrt{\mathcal{Z}_m} \frac{\partial^2 \chi^2}{\partial \mathcal{Z}_m \partial \mathcal{Z}_n} \sqrt{\mathcal{Z}_n} \right|_{\mathcal{Z}=\mathcal{Z}(\alpha)} \end{array} \right.$$

3. Estimating an error as uncertainty of $\hat{\mathcal{Z}}(\theta)$.

$$\langle (\delta \hat{\mathcal{Z}}_n)^2 \rangle = \int d\alpha P(\alpha) \langle (\delta \mathcal{Z}_n^{(\alpha)})^2 \rangle.$$

comment

~the choice of default model $m(\theta) \sim$

The most probable image generally depends on $m(\theta)$.

→ Are all the $m(\theta)$ adequate as long as $m(\theta) > 0$? ⇒ **No !**

(i) to investigate the sensitivity of $\hat{Z}(\theta)$ to the choice on $g(\alpha)$.

posterior probability of α : $P(\alpha) \equiv g(\alpha)e^{W(\alpha)+\Lambda(\alpha)}$.

⇒ We have no information for choosing adequate $g(\alpha)$.

(ii) to calculate the relative error $|\delta \hat{Z}(\theta)| / \hat{Z}(\theta)$,

where $\hat{Z}(\theta)$ with small errors has small uncertainty.

4. Results

Results of Monte Carlo data

model: CP³ model with the fixed point action

analysis: SVD + Newton method with quadruple precision



CP(3)FP ($\beta = 3.0$)					
L	L/ξ	$Q_{\min}-Q_{\max}$	the number of measurements	Fourier	
\vdots	\vdots	\vdots	\vdots	\vdots	\vdots
24	3.4	0-18	10.0M/set		o
32	4.5	0-24	3.0M/set		o
 38	5.3	0-27	5.0M/set		o
46	6.5	0-33	1.0M/set		o
 50	7.0	0-18	30.0M/set		x
56	7.9	0-18	5.0M/set		x

Table 1: Parameter values used in simulations of CP³FP. Here β , ξ and L stand for the inverse coupling constant, correlation length of the time direction and lattice size, respectively. The coupling $\beta = 3.0$. The last column represents whether the Fourier transform works(o) or not(x).

• Non flattening case ($L = 38$)

default model: (i) constant function; $m_c(\theta) = 1.0$.

(ii) Gaussian function; $m_G(\theta) = \exp \left[-\gamma \frac{\log 10}{\pi^2} \theta^2 \right]$.

(\odot) $\gamma = 0.4 \sim 1.2$ increasing by 0.1

ill-posed problem: # of Q : $5 < \#$ of degrees of freedom of $\mathcal{Z}(\theta)$: 28

step1: calculations of most probable image for a given α , $\mathcal{Z}^{(\alpha)}(\theta)$.

\Rightarrow α -dependence of $\mathcal{Z}^{(\alpha)}(\theta)$ is almost invisible for each $m(\theta)$.

step2 & step3: calculations of the most probable image and its error.

\Rightarrow All results of the MEM agree with result of the Fourier transform method within the errors and are independent of $m(\theta)$.

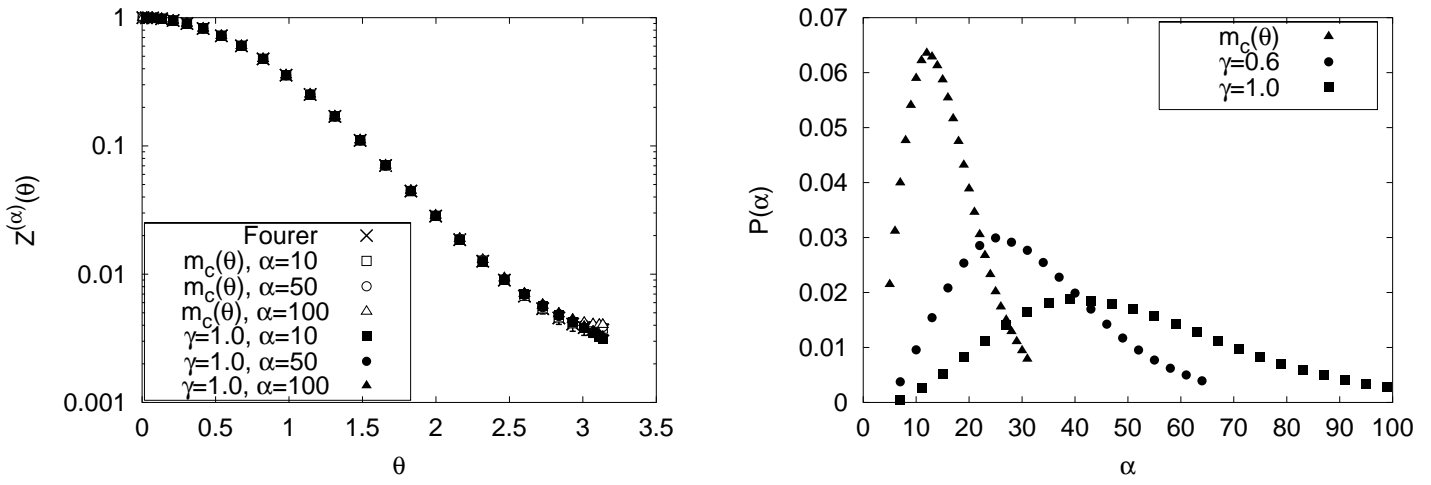


Figure 7: Most probable images for a given α (left panel) and posterior probabilities of α (right panel) in non-flattening case ($L = 38$). As default models, constant, 1.0, and Gaussian function with $\gamma = 0.6$ and 1.0 are used.

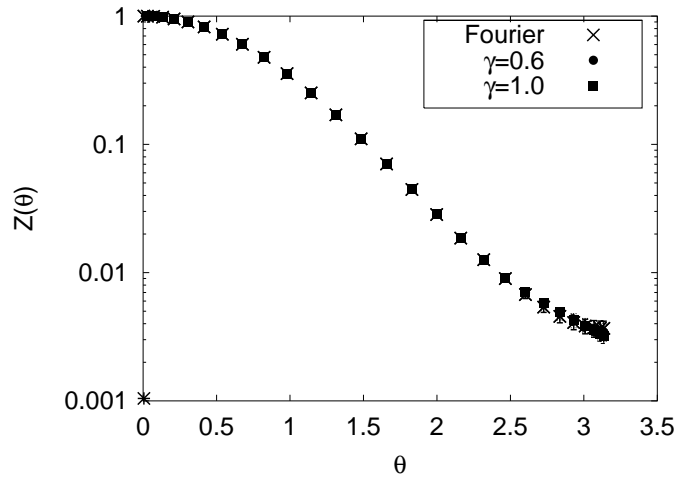


Figure 8: $\hat{Z}(\theta)$ in the non flattening case ($L = 38$). Fourier transform result is also plotted.

$$\odot \hat{Z}_n = \int d\alpha Z_n^{(\alpha)} P(\alpha).$$

\odot All results of the MEM agree with those of the Fourier transform.

- Flattening case ($L = 50$)

default model: (i) Gaussian function; $m_G(\theta) = \exp \left[-\gamma \frac{\log 10}{\pi^2} \theta^2 \right]$.

(☺ $\gamma = 3 \sim 13$ increasing by one)

(ii) $m_{L/50}(\theta) \equiv \hat{\mathcal{Z}}(\theta)$ for smaller volumes

(☺ $L = 24, 32$ and 38)

ill-posed problem: $\#Q : 7 < \#$ degrees of freedom of $\mathcal{Z}(\theta) : 28$

◇ Fourier Transform Method

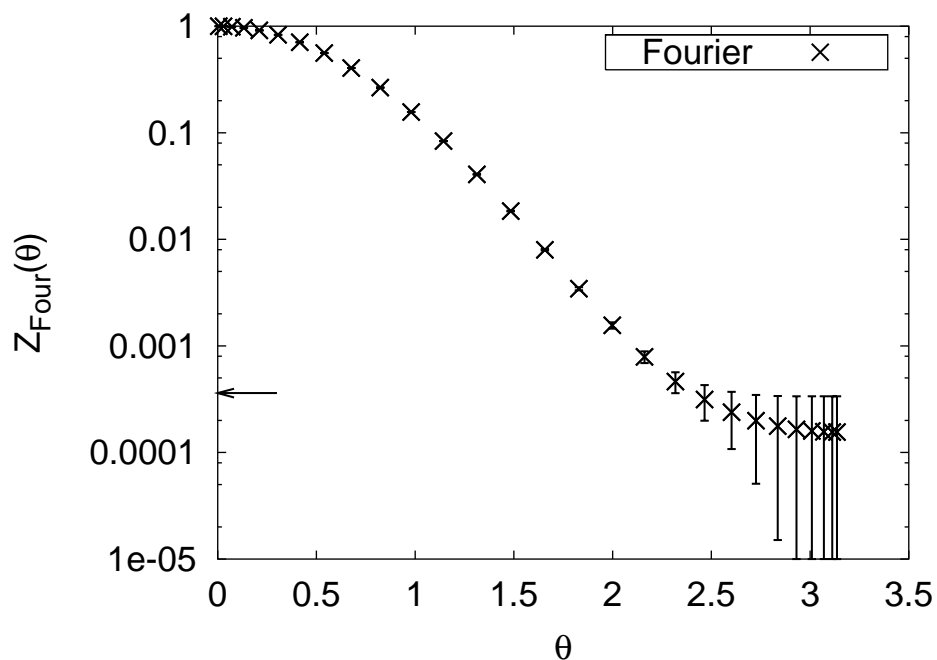


Figure 9: Partition function, $\mathcal{Z}_{\text{Four}}(\theta)$, obtained by the Fourier transform method. The number of measurements is 30.0M/set. Arrow indicates the value of $\epsilon (= 3.610 \times 10^{-4})$.

From the result, we find that $\mathcal{Z}_{\text{Four}}(\theta)$ receives large errors for $\theta > 2.3$
 $(|\delta \mathcal{Z}_{\text{Four}}(\theta)| / \mathcal{Z}_{\text{Four}}(\theta) \gtrsim 0.3)$.

☺ $\epsilon \equiv \sum_Q |\Delta P(Q)|$. : naive error propagation to θ space

◇ MEM

(i) $g(\alpha)$ -dependence of $\hat{\mathcal{Z}}(\theta)$

We systematically investigate to what extent $\hat{\mathcal{Z}}(\theta)$ is sensitive to the choice on $g(\alpha)$, by calculating $\Delta(\theta)$.

$$\Delta(\theta) \equiv \frac{|\hat{\mathcal{Z}}_{\text{Lap}}(\theta) - \hat{\mathcal{Z}}_{\text{Jef}}(\theta)|}{(\hat{\mathcal{Z}}_{\text{Lap}}(\theta) + \hat{\mathcal{Z}}_{\text{Jef}}(\theta))/2}. \quad \left(\begin{array}{l} \hat{\mathcal{Z}}_{\text{Lap}}(\theta) : \hat{\mathcal{Z}}(\theta) \text{ for Laplace's rule} \\ \hat{\mathcal{Z}}_{\text{Jef}}(\theta) : \hat{\mathcal{Z}}(\theta) \text{ for Jeffrey's rule} \end{array} \right)$$

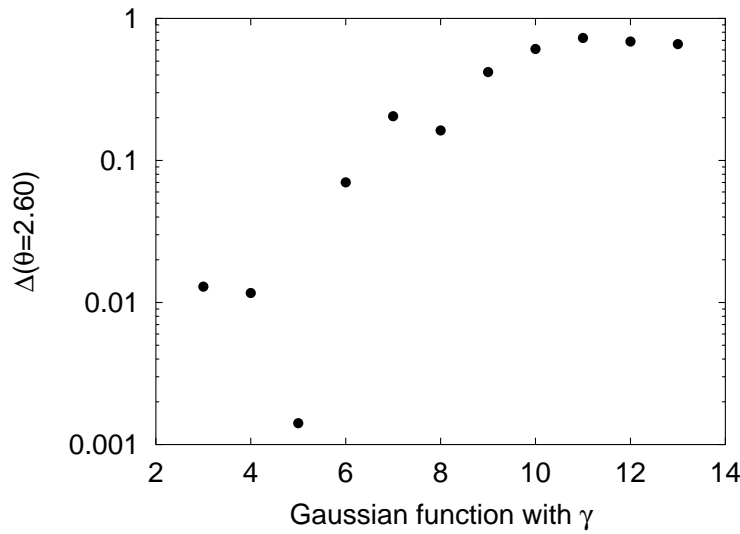


Figure 10: Values of $\Delta(\theta)$ at $\theta = 2.60$. The Gaussian function is used as a default model.

default model	$\Delta(2.31)$	$\Delta(2.60)$	$\Delta(2.83)$	$\Delta(3.14)$
$m_G(\theta)$ with $\gamma = 3.0$	6.28×10^{-3}	1.29×10^{-2}	1.87×10^{-2}	2.22×10^{-2}
$m_G(\theta)$ with $\gamma = 4.0$	6.90×10^{-3}	1.17×10^{-2}	1.50×10^{-2}	1.69×10^{-2}
$m_G(\theta)$ with $\gamma = 5.0$	5.34×10^{-3}	1.41×10^{-3}	1.13×10^{-2}	1.86×10^{-2}
$m_G(\theta)$ with $\gamma = 6.0$	1.29×10^{-2}	7.00×10^{-2}	1.57×10^{-1}	2.18×10^{-1}
$m_G(\theta)$ with $\gamma = 8.0$	1.08×10^{-1}	1.63×10^{-1}	1.85×10^{-1}	1.91×10^{-1}
$m_{24/50}(\theta)$	3.44×10^{-3}	7.64×10^{-3}	1.13×10^{-2}	1.35×10^{-2}
$m_{32/50}(\theta)$	7.27×10^{-3}	1.56×10^{-2}	2.26×10^{-2}	2.68×10^{-2}
$m_{38/50}(\theta)$	1.29×10^{-2}	2.57×10^{-2}	3.59×10^{-2}	4.18×10^{-2}

Table 2: Values of $\Delta(\theta)$ at $\theta = 2.31, 2.60, 2.83$ and 3.14 for various $m(\theta)$.

posterior probability of α : Laplace's rule and Jeffrey's rule are employed.
 $\Rightarrow P(\alpha)$ for Laplace's and Jeffrey's rules behave differently for each case.

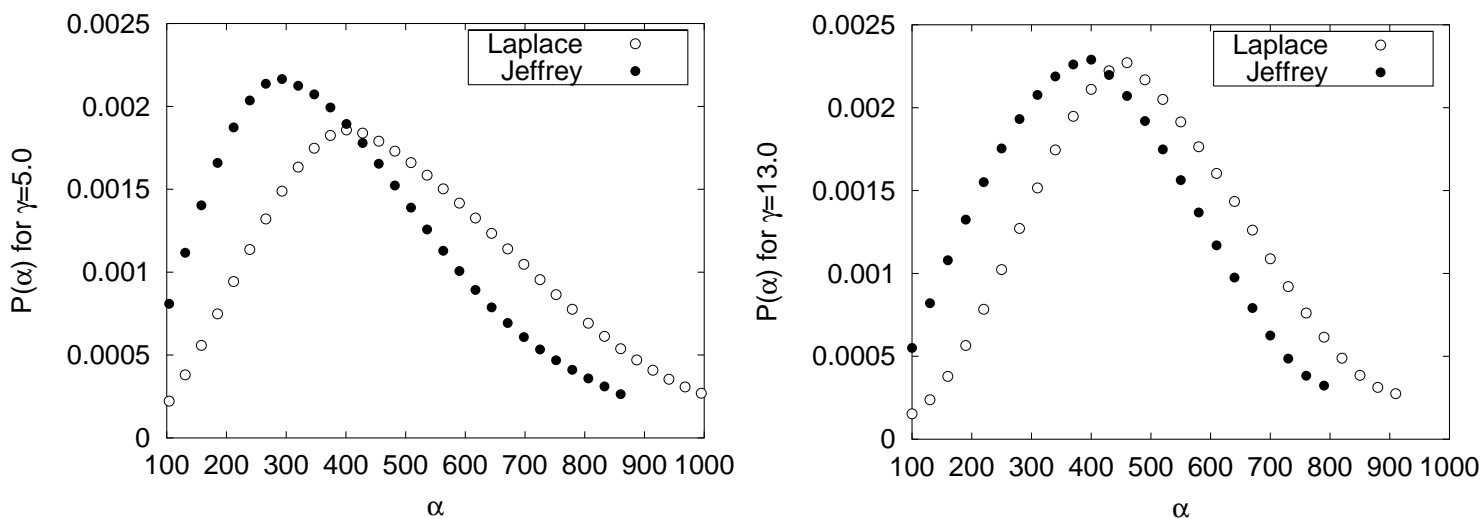


Figure 11: $P(\alpha)$ for Laplace's rule and Jeffrey's rule. The Gaussian functions with $\gamma = 5.0$ (left panel) and $\gamma = 13.0$ (right panel) are used as default models.

most probable image: Laplace's and Jeffrey's rules are employed.
 $\Rightarrow \hat{Z}(\theta)$ with $\gamma = 5.0$ has no $g(\alpha)$ -dependence.

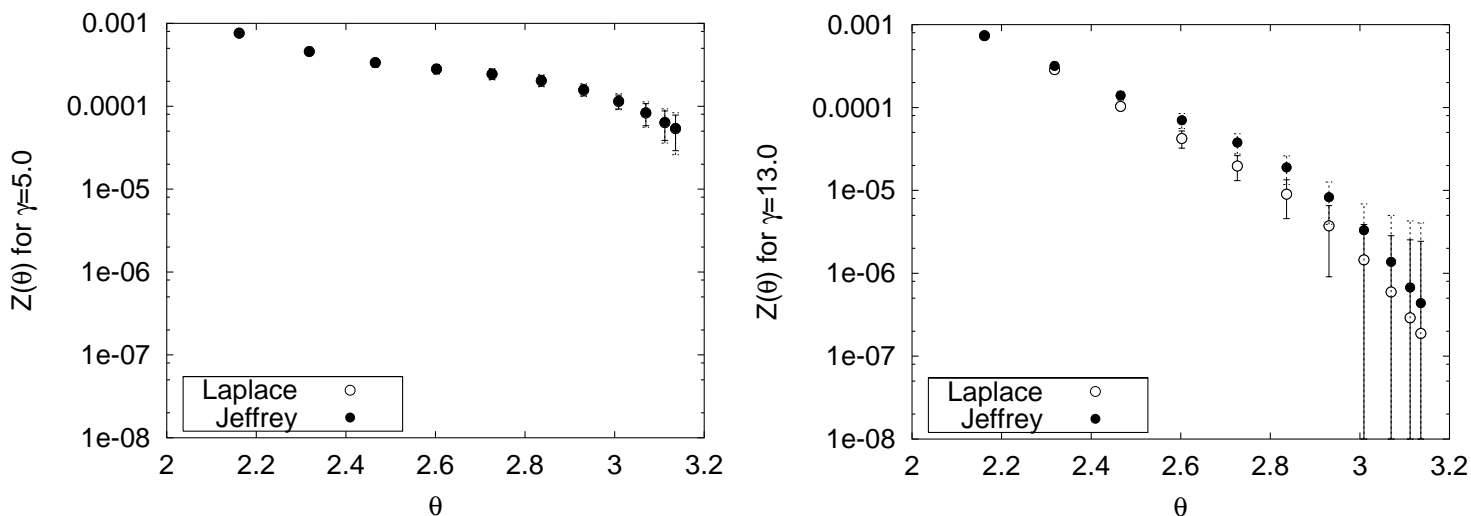


Figure 12: $\hat{Z}(\theta)$ for Laplace's and Jeffrey's rules in $\theta \in [2.0, \pi]$. The same $m_G(\theta)$ are used in the Fig.11. The difference of $\hat{Z}_{\text{Jeff}}(\theta)$ and $\hat{Z}_{\text{Laplace}}(\theta)$ is visible for large θ region in the right panel.

(ii) relative error of $\hat{Z}(\theta)$

We quantitatively investigate the statistical significance of the most probable image, by calculating the relative error, $|\delta \hat{Z}(\theta)| / \hat{Z}(\theta)$.

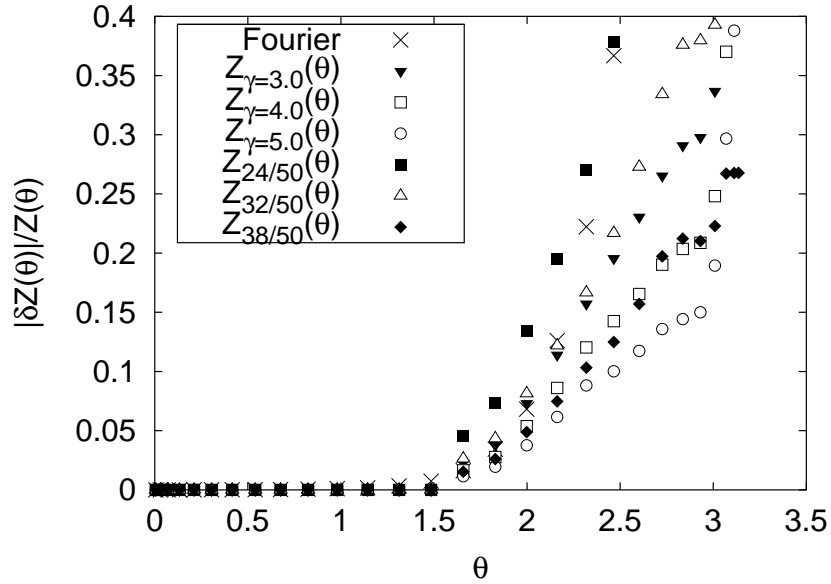


Figure 13: Values of $\frac{|\delta \hat{Z}(\theta)|}{\hat{Z}(\theta)}$ for selected default models. The Fourier result is also plotted (x).

⇒ The relative errors of $\hat{Z}_{\gamma=3.0}(\theta)$, $\hat{Z}_{\gamma=4.0}(\theta)$, $\hat{Z}_{\gamma=5.0}(\theta)$ and $\hat{Z}_{38/50}(\theta)$ are smaller than that of Fourier transform method, the error of MEM means uncertainty of image, though.

◇ Final images (in Laplace's rule)

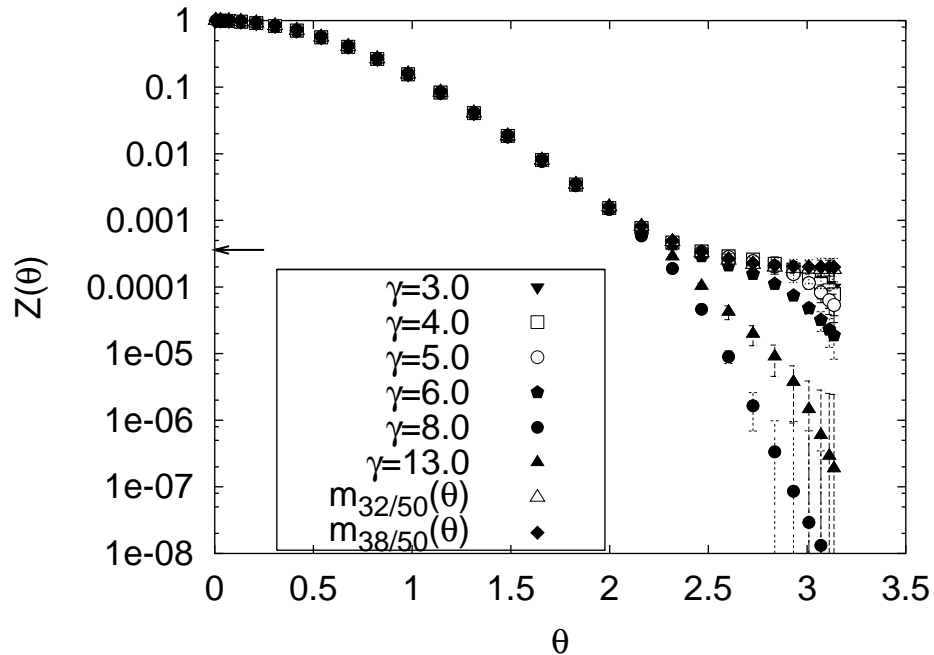


Figure 14: Final images for some $m(\theta)$. Arrow indicates the value of ϵ .

☺ All results of MEM

have small errors compared to
the result of the Fourier transform
(see Fig.9),

&

have little default model dependence
for $\theta \lesssim 3.0$.

Statistical fluctuation of the final image

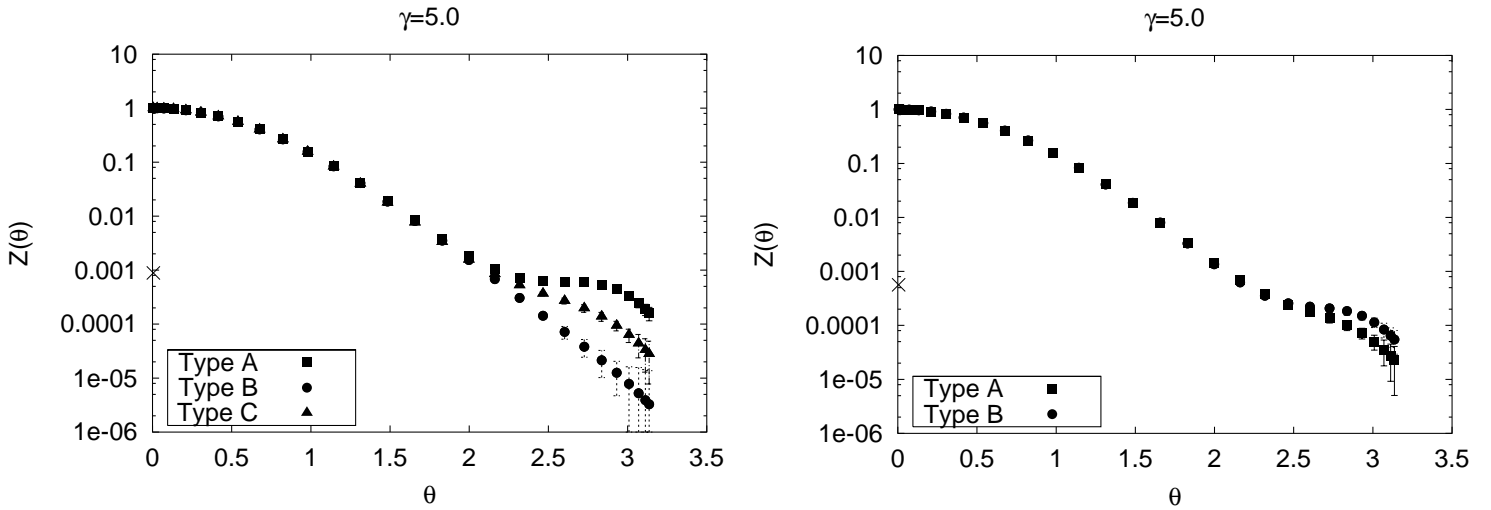


Figure 15: Final images $\hat{\mathcal{Z}}(\theta)$ for $L = 50$ with 5.0M/set(left panel) and 15.0M/set(right panel). Gaussian default model with $\gamma = 5.0$ is used. While $\hat{\mathcal{Z}}(\theta)$ for a large θ region in left panel are not stable against statistical fluctuation, $\hat{\mathcal{Z}}(\theta)$ in right panel are stable.

Final images selected by constraint

constraint: $\frac{|\delta \hat{\mathcal{Z}}(\theta)|}{\hat{\mathcal{Z}}(\theta)} \lesssim 0.2$, for example.

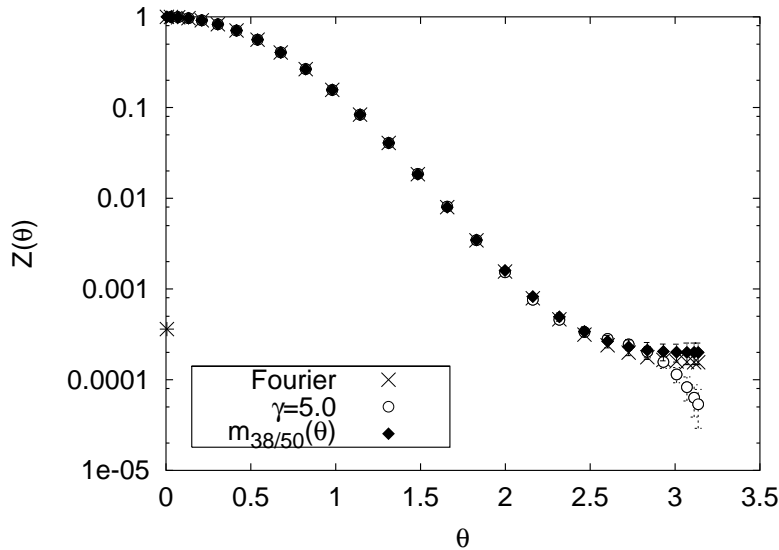


Figure 16: Values of $\hat{\mathcal{Z}}(\theta)$. The Gaussian function with $\gamma = 5.0$ and $\hat{\mathcal{Z}}(\theta)$ for $L = 38$ are used as default models. No default model dependence is seen for $\theta < 3.0$. $\mathcal{Z}(\theta)$ obtained by the Fourier transform method is also plotted.

4. Summary and Discussion

Summary

- phase structure in θ space
→ **model:** CP^3 model with the θ term



sign problem

$$P(Q) \xrightarrow{\text{Fourier transform}} \mathcal{Z}(\theta)$$

- ☹️ { **strong coupling regions:** first order phase transition exists at $\theta = \pi$.
weak coupling regions: { no phase transition exists for small volumes.
the method breaks down for large volumes.



flattening

$$P(Q) \xleftarrow{\text{inverse Fourier transform}} \mathcal{Z}(\theta)$$

MEM

- widely used technique for various fields,
- parameter inference based on Bayes' theorem,
- to derive a unique solution by utilizing data and the prior information.

♣ Non flattening cases ($L = 24, 32, 38$)

☺ All the results of the MEM are almost in agreement with those of the Fourier transform and are independent of $m(\theta)$.

♣ Flattening case ($L = 50$)

- ☺ The MEM gives the most probable image with reasonably small errors for a wide range of θ .
- ☺ $\hat{Z}(\theta)$ depends strongly on $m(\theta)$ for the region where the value of $\hat{Z}(\theta)$ is smaller than ϵ .

Discussions

- Which image $\hat{Z}(\theta)$ is the most probable one?
⇒ **estimator:** $\Delta(\theta)$? $|\delta \hat{Z}(\theta)| / \hat{Z}(\theta)$?
- How is $m(\theta)$ chosen ? ⇒ ref.[18]
- It could be worthwhile to study LFT with the finite density in terms of the MEM.

References

- [1] M. Lüscher, P. Weisz & U. Wolff, Nucl. Phys. **B359** (1991), 221.
- [2] G. 't Hooft, Nucl. Phys. **B190 [FS3]** (1981), 455.
- [3] J. L. Cardy & E. Rabinovici, Nucl. Phys. **B205 [FS5]** (1982), 1.
J. L. Cardy, Nucl. Phys. **B205 [FS5]** (1982), 17.
- [4] U. -J. Wiese, Nucl. Phys. **B318** (1989), 153.
W. Bietenholz, A. Pochinsky & U. -J. Wiese, Phys. Rev. Lett **75** (1995), 4524.
- [5] N. Seiberg, Phys. Rev. Lett. **53** (1984), 637.
- [6] S. Olejnik & G. Schierholz, Nucl. Phys. **B** (Proc. Suupl) **34** (1994), 709.
G. Schierholz, Nucl. Phys. **B** (Proc. Suppl) **37A** (1994), 203.
- [7] J. C. Plefka & S. Samuel, Phys. Rev. **D56** (1997), 44.
- [8] A. S. Hassan, M. Imachi, N. Tsuzuki & H. Yoneyama, Prog. Theor. Phys. **95** (1995), 175.
- [9] M. Imachi, S. Kanou & H. Yoneyama, Prog. Theor. Phys. **102** (1999), 653.
- [10] R. Burkhalter, M. Imachi, Y. S. & H. Yoneyama, Prog. Theor. Phys. **106** (2001), 613.
- [11] M. Imachi, Y. S. & H. Yoneyama, Prog. Theor. Phys. **111** (2004), 387.
- [12] M. Imachi, Y. S. & H. Yoneyama, hep-lat/0506032.
- [13] P. Hasenfratz & F. Niedermayer, Nucl. Phys. **B414** (1994), 785.
- [14] R. Burkhalter, Phys. Rev. **D54** (1996), 4121.
- [15] R. K. Bryan, Eur. Biophys. J. **18** (1990), 165.
- [16] M. Jarrell & J. E. Gubernatis, Phys. Rep. **269** (1996), 133.
- [17] M. Asakawa, T. Hatsuda & Y. Nakahara, Prog. Part. Nucl. Phys. **46** (2001), 459.
- [18] A. Caticha & R. Preuss, MaxEnt'03 (2003), arXiv:physics/0312131.