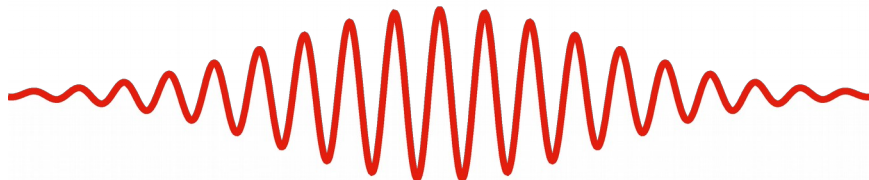


Path Optimization Using Neural Network in Field Theories

Akira Ohnishi¹, Yuto Mori², Kouji Kashiwa³

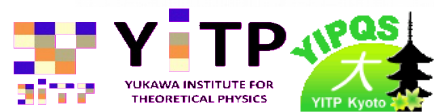
*1. Yukawa Inst. for Theoretical Physics, Kyoto U.,
2. Dept. Phys., Kyoto U., 3. Fukuoka Inst. Tech.*

*International Workshop on the
The Sign Problem in QCD and Beyond,
September 10-14, Bielefeld, Germany*



SIGN '18

International Workshop on the Sign Problem in QCD and Beyond



Collaborators

Akira Ohnishi¹, Yuto Mori², Kouji Kashiwa³

1. Yukawa Inst. for Theoretical Physics, Kyoto U.,

2. Dept. Phys., Kyoto U., 3. Fukuoka Inst. Tech.



Y. Mori
(grad. stu.)



K. Kashiwa



AO (10 yrs ago)

1D integral: Y. Mori, K. Kashiwa, AO, PRD 96 ('17), 111501(R) [arXiv:1705.05605]

ϕ^4 w/ NN: Y. Mori, K. Kashiwa, AO, PTEP 2018 ('18), 023B04 [arXiv:1709.03208]

Lat 2017: AO, Y. Mori, K. Kashiwa, EPJ Web Conf. 175 ('18), 07043 [arXiv:1712.01088]

NJL thimble: Y. Mori, K. Kashiwa, AO, PLB 781('18),698 [arXiv:1705.03646]

PNJL w/ NN: K. Kashiwa, Y. Mori, AO, arXiv:1805.08940.

0+1D QCD: Y. Mori, K. Kashiwa, AO, in prep.

PNJL with vector int. using NN: K. Kashiwa, Y. Mori, AO, in prep.

Introduction

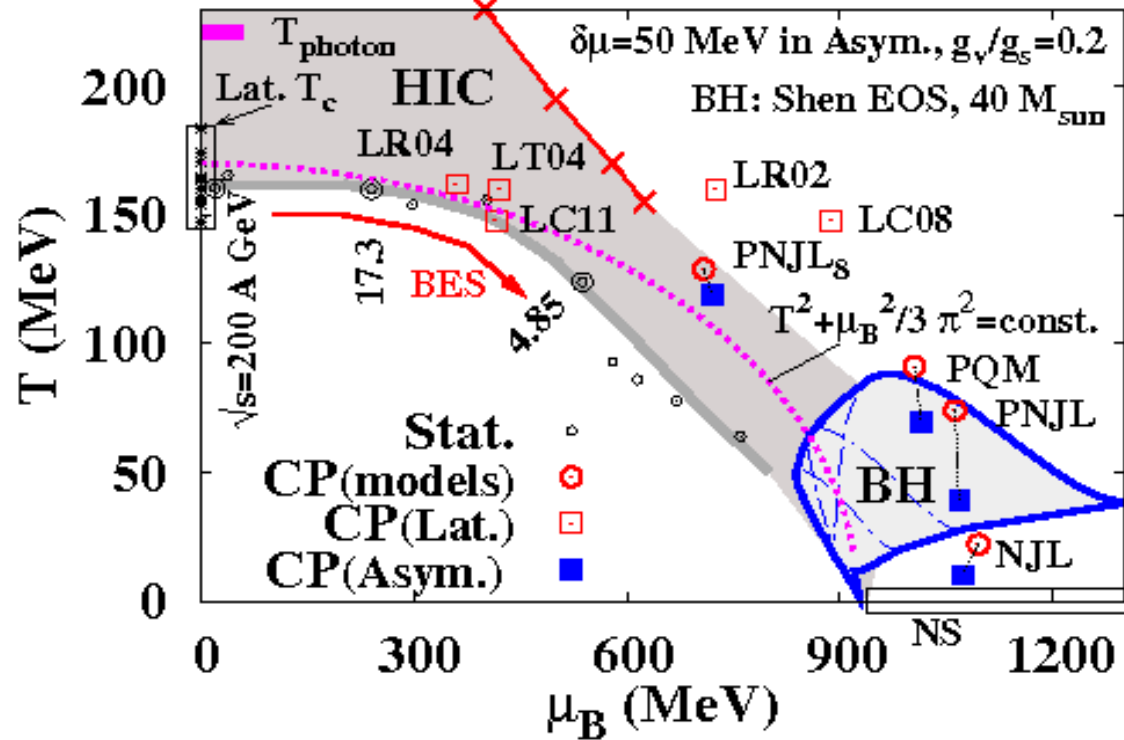
■ Sign problem for complex actions

- Grand challenge in theor. phys.
- Largest obstacle to explore QCD phase diagram

■ Approaches

- Taylor expansion, Analytic cont., Canonical, **Strong coupling LQCD, ...**

AO, PTPS193('12)1.

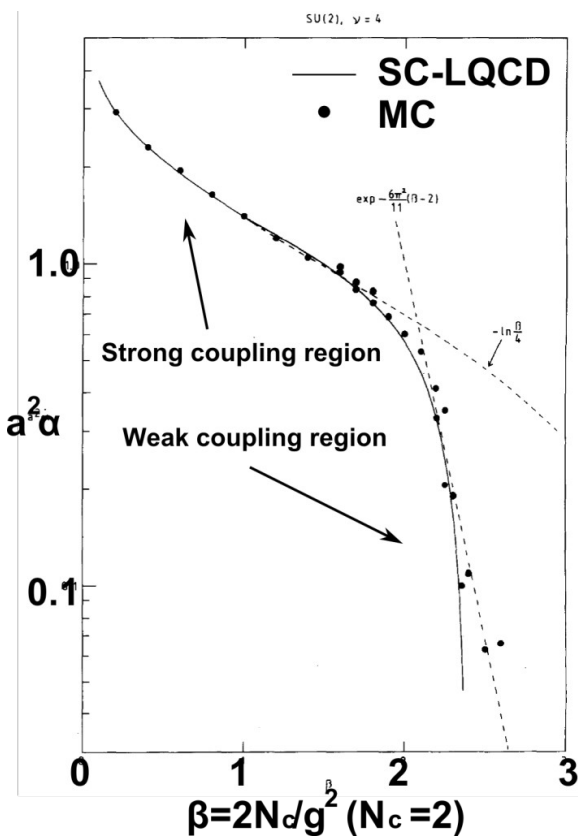


- Complex Langevin method (CLM)
Parisi ('83), Klauder ('83), Aarts et al. ('11), Nagata et al. ('16)
- Lefschetz thimble method (LTM)
E. Witten ('10), Cristoforetti et al. ('12), Fujii et al. ('13)
- Generalized LTM (GLTM)
A. Alexandru, et al., ('16)

Complexified variable methods

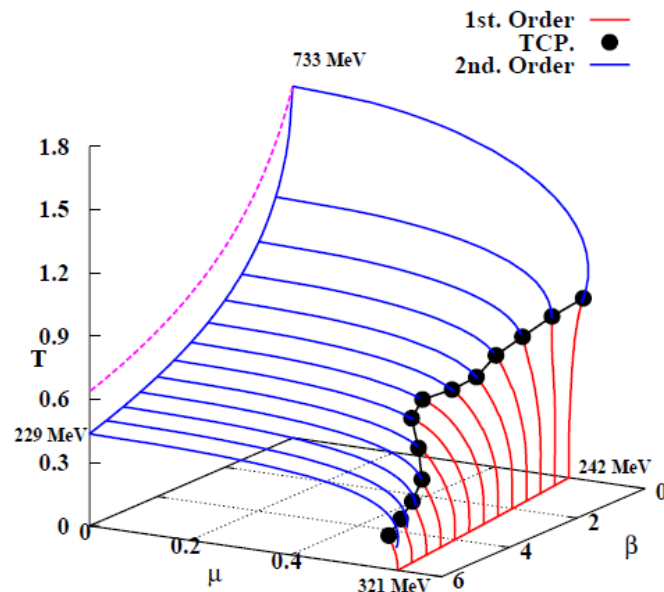
Strong Coupling Lattice QCD

Pure YM \rightarrow Area Law



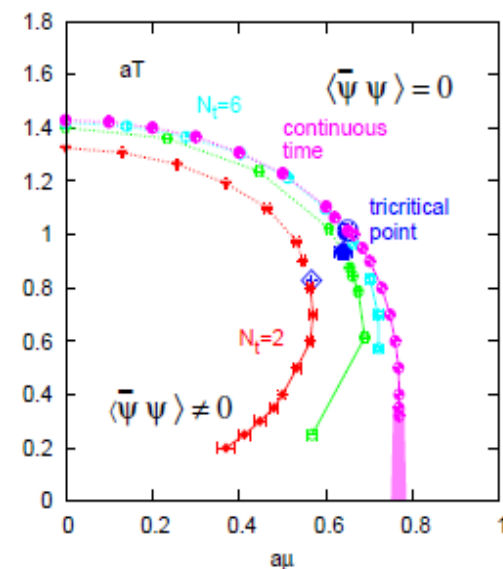
Wilson ('74), Creutz ('80), Munster ('80, '81), Lottini, Philipsen, Langelage's ('11)

Phase diagram (mean field)



Kawamoto ('80), Kawamoto, Smit ('81), Damagaard, Hochberg, Kawamoto ('85), Ilgenfritz, Kripfganz ('85), Bilic, Karsch, Redlich ('92), Fukushima ('03); Nishida ('03), Kawamoto, Miura, AO, Ohnuma ('07). Miura, Nakano, AO, Kawamoto ('09), Nakano, Miura, AO ('10), Miura, Kawamoto, Nakano, AO ('17)

Lattice MC simulations (Fluctuations)

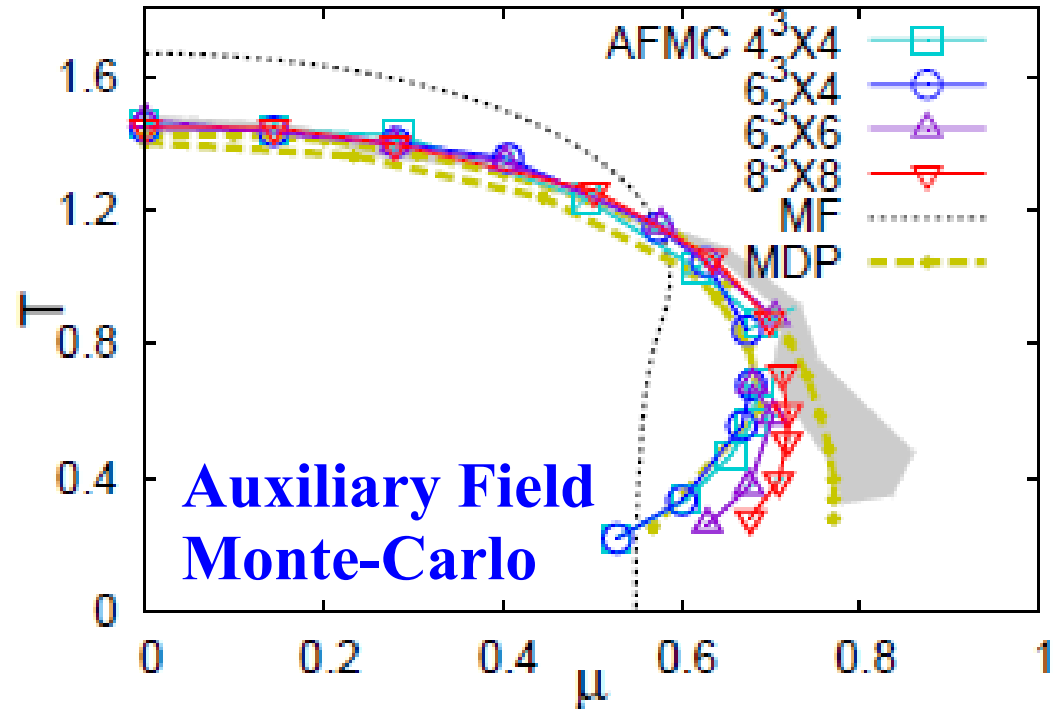
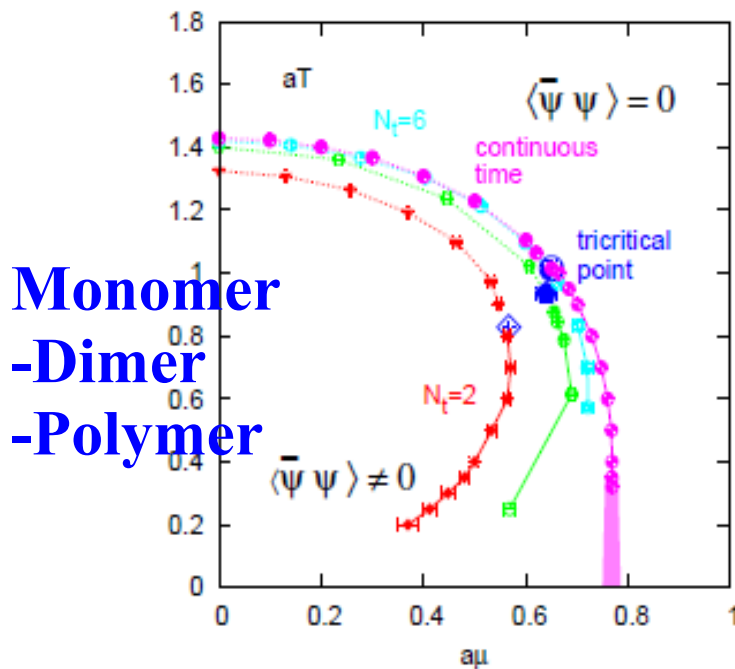


Rossi, Wolff ('84), Mutter, Karsch ('89), de Forcrand, Fromm ('10), de Forcrand, Unger ('11), AO, Ichihara, Nakano ('12), Ichihara, Nakano, AO ('14), de Forcrand, Langelage, Philipsen, Unger ('14), Ichihara, Morita, AO ('15)

Phase Diagram in Strong Coupling Lattice QCD

Aoki ('84), Rossi, Wolff ('84), Damagaard, Hochberg, Kawamoto ('85), Mutter, Karsch ('89), Bilic, Karsch, Redlich ('92), Fukushima ('03), Nishida ('03), Kawamoto, Miura, AO, Ohnuma ('07), Miura, Nakano, AO, Kawamoto ('09), Nakano, Miura, AO ('10), de Forcrand, Fromm ('10), de Forcrand, Unger ('11), Misumi, Nakano, Kimura, AO('12), Misumi, Kimura, AO('12), AO, Ichihara, Nakano ('12), Tomboulis ('13), Ichihara, Nakano, AO ('14), Ichihara, Morita, AO ('15)

- Integrate links first, and fermions later → Milder sign prob.
- Two indep. methods give consistent phase diagram in the *Strong Coupling Limit*.



de Forcrand, Fromm ('10), de Forcrand, Langelage, Philipsen, Unger ('14)

T.Ichihara, AO, T.Z.Nakano, PTEP2014,123D02.



Phase Diagram Away from the Strong Coupling Limit

■ Sampling at finite coupling ($g < \infty$ or $2N_c/g^2 > 0$) is difficult !

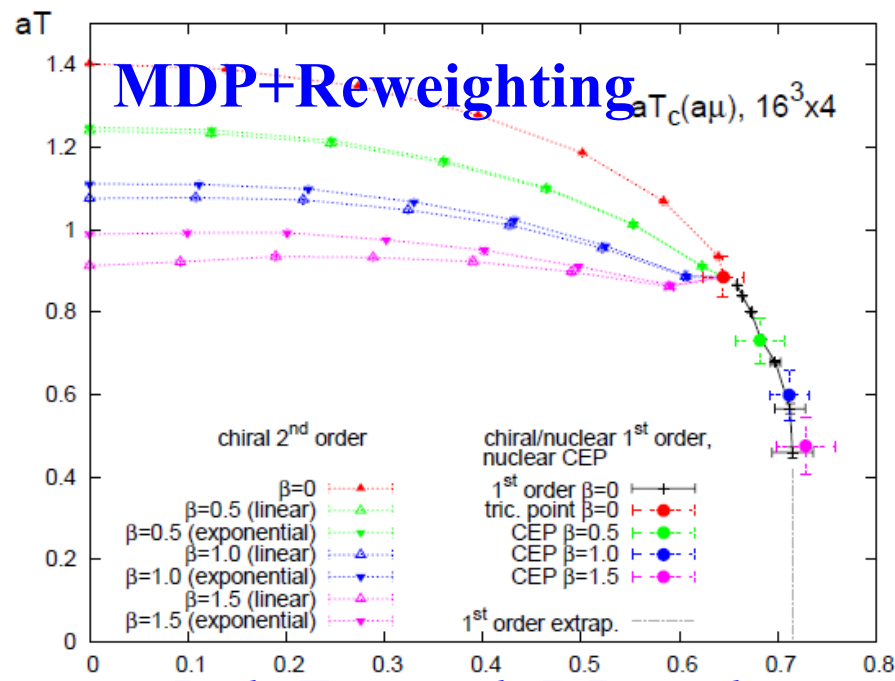
● Monomer-Dimer-Polymer sim. → Reweighting from SCL

P. de Forcrand, J. Langelage, O. Philipsen, W. Unger ('14)

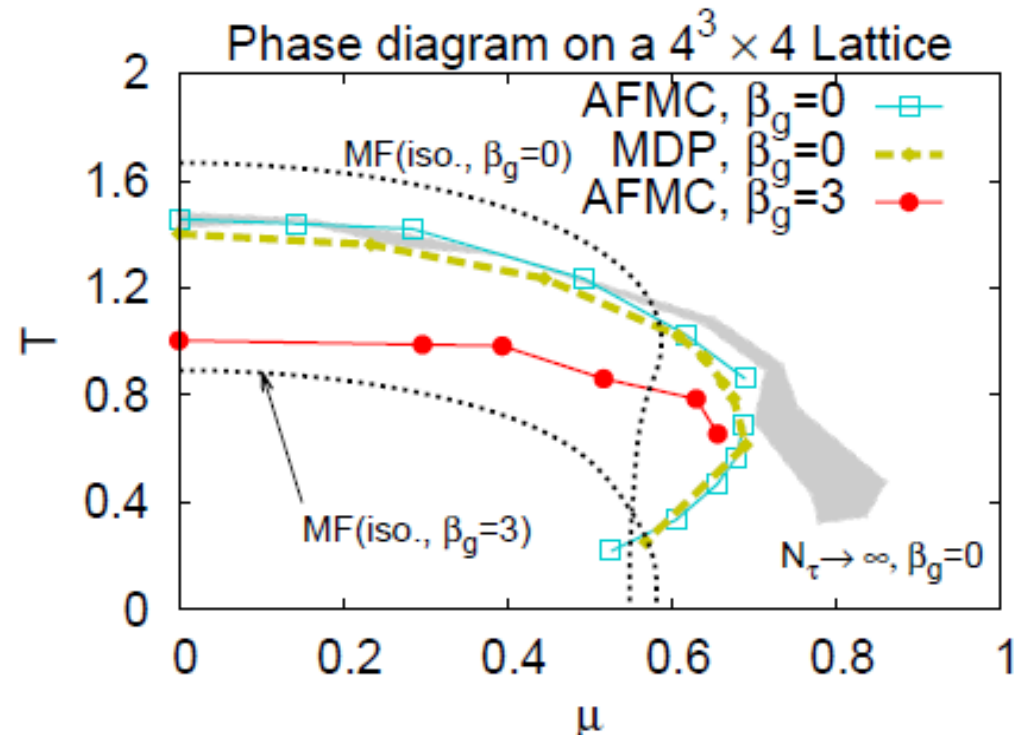
● Auxiliary Field Monte-Carlo

→ Direct sampling is possible at small μ/T , but fails at large μ

AO, T. Ichihara (Lattice 2015)



P. de Forcrand, J. Langelage, O. Philipsen, W. Unger ('14)



AO, Ichihara (Lattice2015)

Why is direct sampling difficult at finite $1/g^2$?
→ Sign problem from Repulsive Interaction

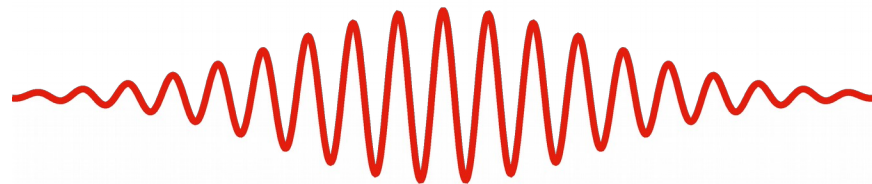
and also from the Polyakov loop ?
(U. Wenger on Mon.)

One of the origins of the sign problem = Repulsion

- Repulsive 4-Fermi interaction \rightarrow Sign prob. via Hubbard transf.

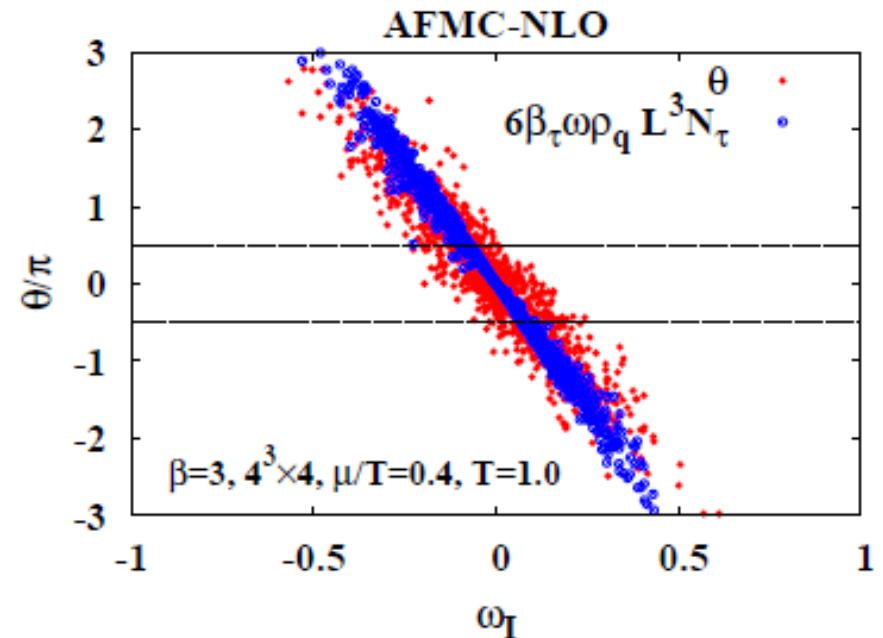
$$\exp \left[-G^2 (q^\dagger q)^2 \right] = \int d\omega e^{-\omega^2/2 - iG(q^\dagger q)\omega}$$

- With small fluctuation in density $q^\dagger q \sim \rho$
 - Phase $\sim G\omega\rho \times L^3 \rightarrow$ Serious sign prb.



SIGN '18

International Workshop on the Sign Problem in QCD and Beyond



One of the origins of the sign problem = Repulsion

- Repulsive 4-Fermi interaction \rightarrow Sign prob. via Hubbard transf.

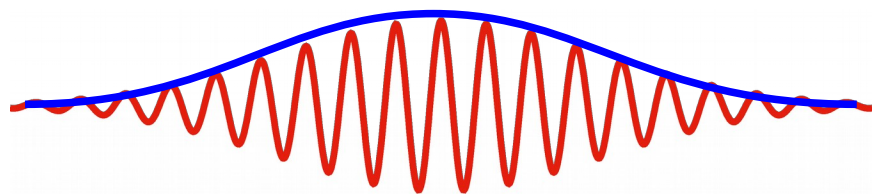
$$\exp \left[-G^2 (q^\dagger q)^2 \right] = \int d\omega e^{-\omega^2/2 - iG(q^\dagger q)\omega}$$

- With small fluctuation in density $q^\dagger q \sim \rho$

- Phase $\sim G\omega\rho \times L^3 \rightarrow$ Serious sign prb.

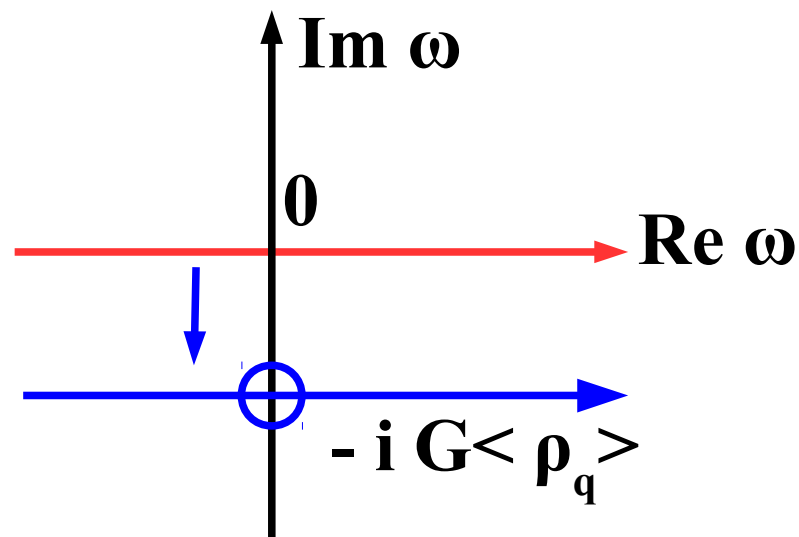
- Shifted ω should kill sign prb.

$$-\frac{\omega^2}{2} - iG\omega\rho_q = -\frac{1}{2}(\omega + iG\rho_q)^2 - G^2\rho_q^2$$



SIGN '18

International Workshop on the Sign Problem in QCD and Beyond



Strong Coupling Lattice QCD action (1)

■ Lattice QCD action (unrooted staggered fermion)

$$S = \frac{1}{2} \sum_{x,\mu} [\eta_\mu(x) \bar{\chi}_x U_\mu(x) \chi_{x+\hat{\mu}} - \eta_\mu^{-1}(x) \bar{\chi}_{x+\hat{\mu}} U_\mu^\dagger(x) \chi_x] + m_0 \sum_x \bar{\chi}_x \chi_x - \frac{2N_c}{g^2} \sum_P \text{ReTr} U_P \quad (\eta_\nu(x) = (-1)^{x_0+\dots+x_{\nu-1}} e^{\mu\delta_{\nu 0}})$$

■ Eff. Action in the Strong Coupling Limit

Damgaard, Kawamoto, Shigemoto ('84),

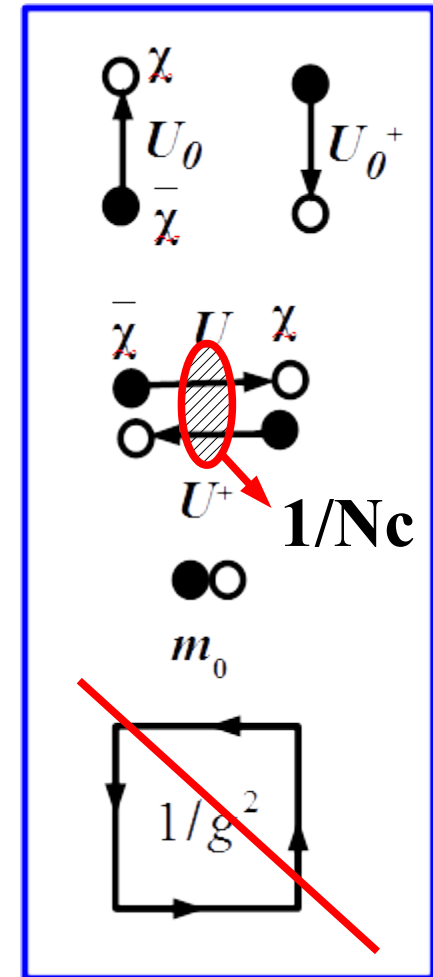
Jolicoeur, Kluberg-Stern, Lev, Morel, Petersson ('84).

- Spatial link integral → Four-Fermi int. (LO in 1/d expansion)

$$S_{\text{eff}}^{(\text{SCL})} = \frac{1}{2} \sum_x [V_x^+ - V_x^-] + m_0 \sum_x M_x - \frac{1}{4N_c} \sum_{x,j} M_x M_{x+\hat{j}}$$

$$V_x^+ = \bar{\chi}_x e^\mu U_0(x) \chi_{x+\hat{0}}, \quad V_x^- = \bar{\chi}_{x+\hat{0}} e^{-\mu} U_0^\dagger(x) \chi_x, \quad M_x = \bar{\chi}_x \chi_x$$

SCL → (Scalar-Pseudoscalar) 4-Fermi int.



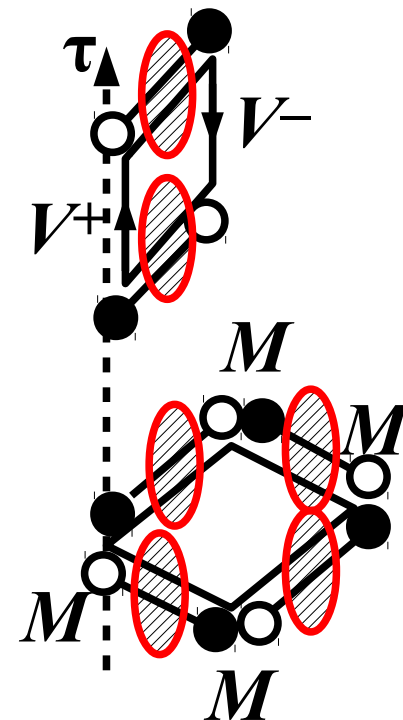
Strong Coupling Lattice QCD action (2)

■ $1/g^2$ correction

Faldt, Petersson ('86), Miura, Nakano, AO, Kawamoto ('09)

$$S_{\text{eff}}^{(\text{NLO})} = S_{\text{eff}}^{(\text{SCL})} + \frac{\beta_\tau}{2} \sum_{x,j} \left[V_x^+ V_{x+\hat{j}}^- + V_x^- V_{x+\hat{j}}^+ \right] - \beta_s \sum_{x,k,j,k \neq j} M_x M_{x+\hat{j}} M_{x+\hat{j}+\hat{k}} M_{x+\hat{k}}$$

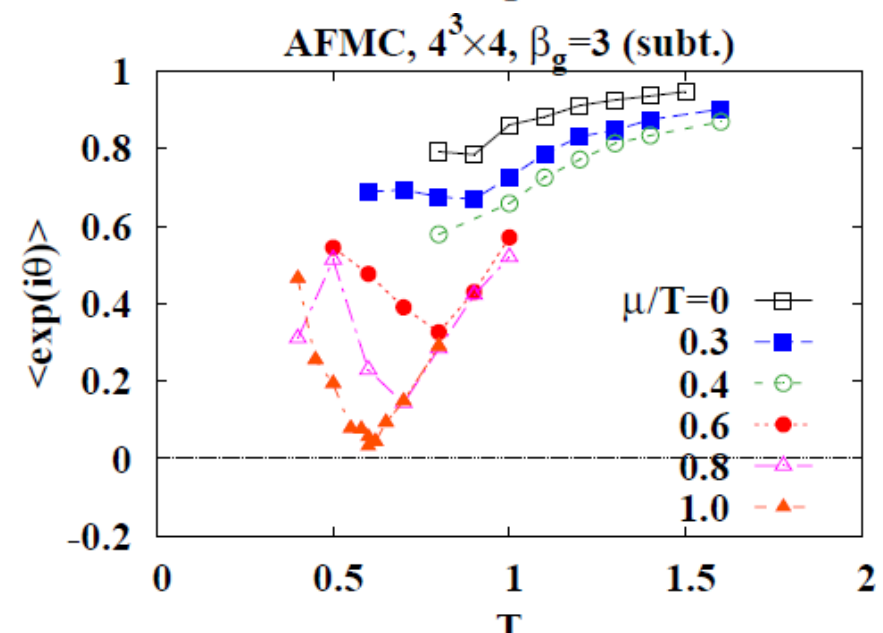
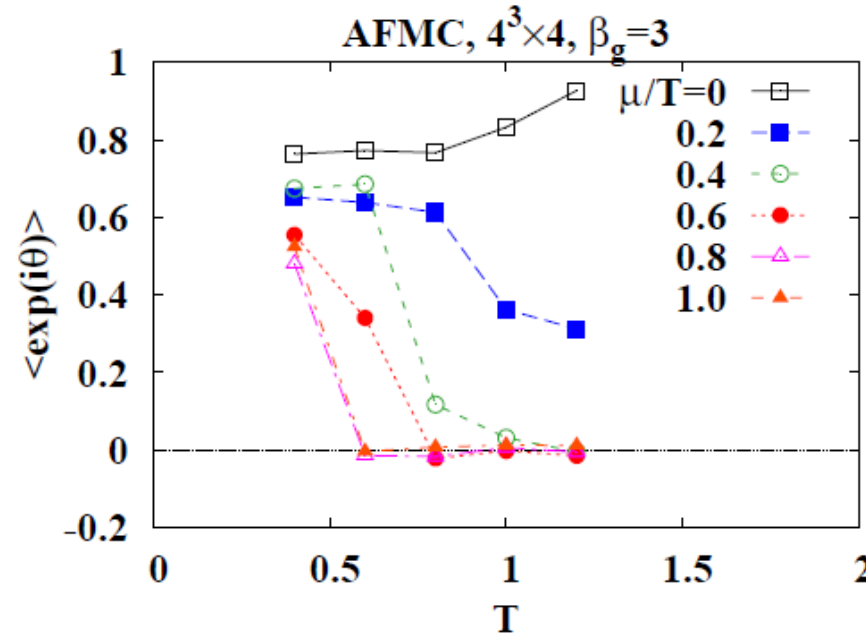
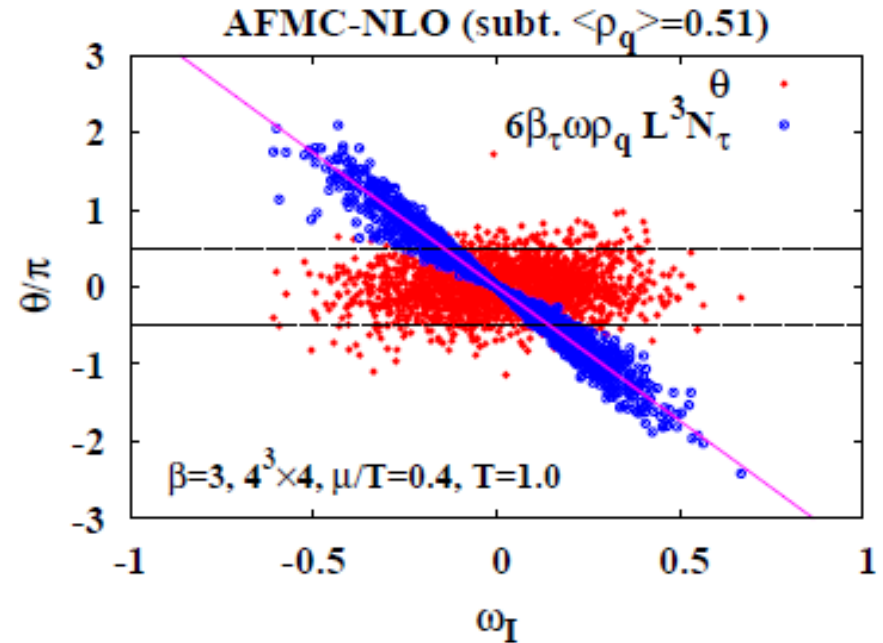
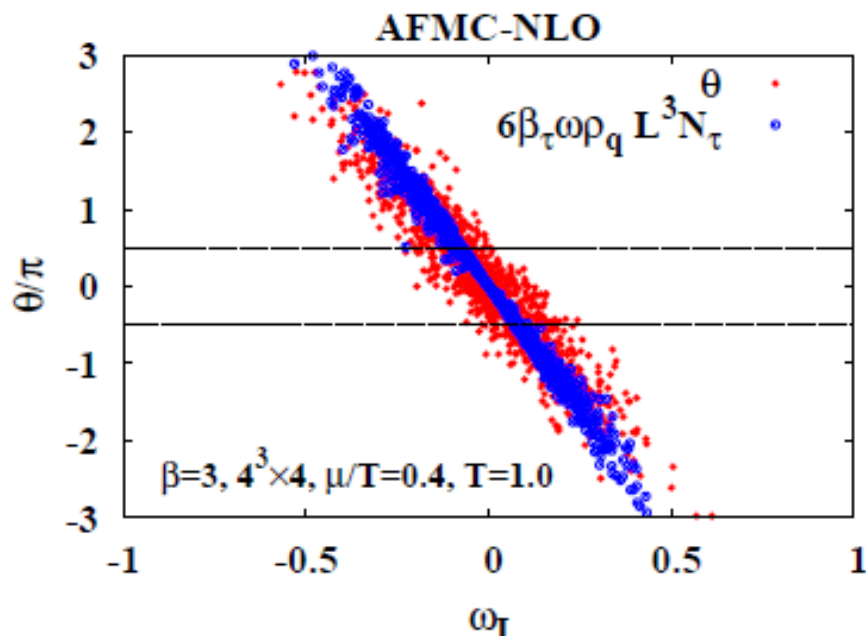
$$\beta_\tau = 1/2 N_c^2 g^2, \quad \beta_s = 1/16 N_c^4 g^2$$



Plaquet ($1/g^2$) corr. generates

- *Density-Density (VV) 4-Fermi Repulsion*
- *8 Fermi attraction*

Imaginary Part Shifted AFMC works !



AO, Ichihara (Lattice2015)

*When density fluctuation is small,
auxiliary field shift of the imaginary part by constant
works well.*

*However, when density fluctuation is large
(e.g. around the 1st order phase transition),
simple shift does not work.*

*Configuration dep. shift of the imaginary part
→ Complexified variable methods
(LTM, CLM, GTM, POM, ...)*

- Introduction

- Path Optimization Method

Y. Mori, K. Kashiwa, AO, PRD 96 ('17), 111501(R) [arXiv:1705.05605]

*AO, Y. Mori, K. Kashiwa, EPJ Web Conf. 175 ('18), 07043 [arXiv:1712.01088]
(Lattice 2017 proceedings)*

- Application to field theory using neural network

Y. Mori, K. Kashiwa, AO, PTEP 2018 ('18), 023B04 [arXiv:1709.03208]

- Application to gauge theory: 1-dimensional QCD

→ **Talk by Yuto Mori (next talk)**

- Application to effective models with phase transition

K. Kashiwa, Y. Mori, AO, arXiv:1805.08940 and in prep.

- Discussions

- Summary

Path Optimization Method

*Y. Mori, K. Kashiwa, AO, PRD 96 ('17), 111501(R) [arXiv:1705.05605]
AO, Y. Mori, K. Kashiwa, EPJ Web Conf. 175 ('18), 07043 [arXiv:1712.01088]
(Lattice 2017 proceedings)*

Lefschetz thimble method

*E. Witten ('10), Cristoforetti et al. (Aurora) ('12),
Fujii et al. ('13), Alexandru et al. ('16).
[Alexandru (Tue)]*

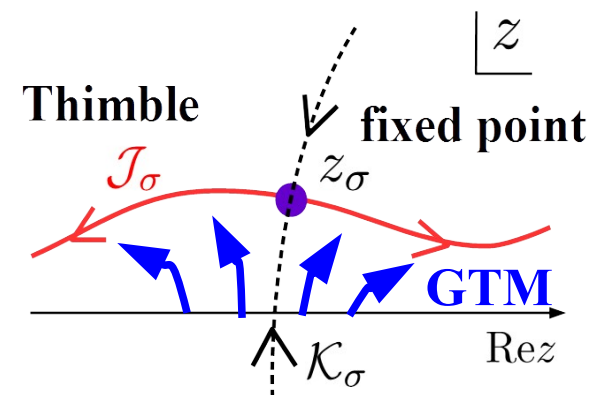
- Solving the flow eq. from a fixed point σ
→ Integration path (thimble)

Note: $\text{Im}(S)$ is constant on one thimble

$$\mathcal{J}_\sigma : \frac{dz_i(t)}{dt} = \overline{\left(\frac{\partial S}{\partial z_i} \right)} \rightarrow \frac{dS}{dt} = \sum_i \left| \frac{\partial S}{\partial z_i} \right|^2 \in \mathbb{R}, \quad \mathcal{C} = \sum_\sigma n_\sigma \mathcal{J}_\sigma$$

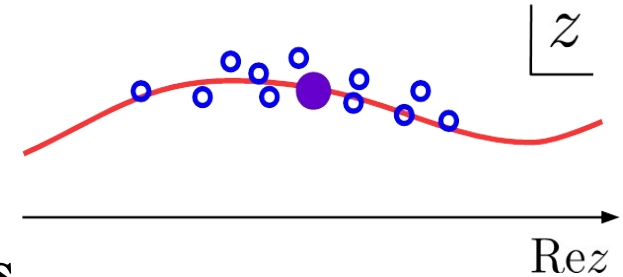
- Problem:

- Phase from the Jacobian (residual. sign pr.),
- Different Phases of Multi-thimbles (global sign pr.),
- Numerical Cost,
- Stokes phenomena, ...



Complex Langevin method

*Parisi ('83), Klauder ('83), Aarts et al. ('11),
Nagata et al. ('16); Seiler et al. ('13), Ito et al. ('16).
[Sexty (Mon), Stamatescu (Tue)]*



- Solving the complex Langevin eq. → Configs.

$$\frac{dz_i}{dt} = -\frac{\partial S}{\partial z_i} + \eta_i(t) (\eta_i : \text{White noise}), \quad \langle \mathcal{O}(x) \rangle = \langle \mathcal{O}(z) \rangle$$

- No sign problem.

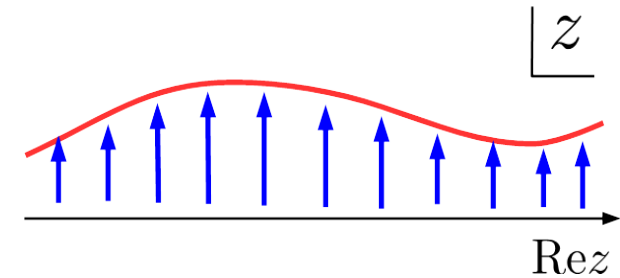
- Problem:

- CLM can give converged but wrong results,
and we cannot know if it works or not in advance.

Path optimization method

*Mori et al. ('17), AO, Mori, Kashiwa (Lattice 2017),
Mori et al. ('18), Kashiwa et al. ('18);
Alexandru et al. ('17 (Learnifold), '18 (SOMMe), '18),
Bursa, Kroyter ('18).*

[Alexandru(Tue), Mori, Lawrence, Kroyter (Wed)]



- Integration **path** is **optimized** to evade the sign problem, i.e. to enhance the average phase factor.

$$\text{APF} = \langle e^{i\theta} \rangle_{\text{pq}} = \int dx e^{-S} / \int dx |e^{-S}| = \mathcal{Z} / \mathcal{Z}_{\text{pq}}$$

Sign Problem \rightarrow Optimization Problem

- Cauchy(-Poincare) theorem: the partition fn. is invariant if
 - the Boltzmann weight $W=\exp(-S)$ is holomorphic (analytic),
 - and the path does not go across **the poles and cuts of W** .
- At Fermion det.=0, S is singular but W is not singular
 - Problem: quarter/square root of Fermion det.

Cost Function and Optimization

- **Cost function:** a measure of the seriousness of the sign problem.

$$\begin{aligned}\mathcal{F}[z(x)] &= \frac{1}{2} \int dx \left| e^{i\theta(x)} - e^{i\theta_0} \right|^2 \left| J(x) e^{-S(z(x))} \right| \\ &= |\mathcal{Z}| \left(\left| \langle e^{i\theta} \rangle_{\text{pq}} \right|^{-1} - 1 \right) = \mathcal{Z}_{\text{pq}} - |\mathcal{Z}| \\ &\quad [\theta = \arg(Je^{-S}), \theta_0 = \arg(\mathcal{Z})]\end{aligned}$$

- **Optimization:** the integration path is optimized to minimize the Cost Function.
(via Gradient Descent or Machine Learning)

- **Example: One-dim. integral \rightarrow Complete set**

$$z(x) = x + iy(x), \quad y(x) = \sum_n c_n H_n(x)$$

$$\mathcal{Z} = \int dx J(x) e^{-S(z(x))}, \quad J(x) = \frac{dz(t)}{dx}$$

Benchmark test: 1 dim. integral

■ A toy model with a serious sign problem

J. Nishimura, S. Shimasaki ('15)

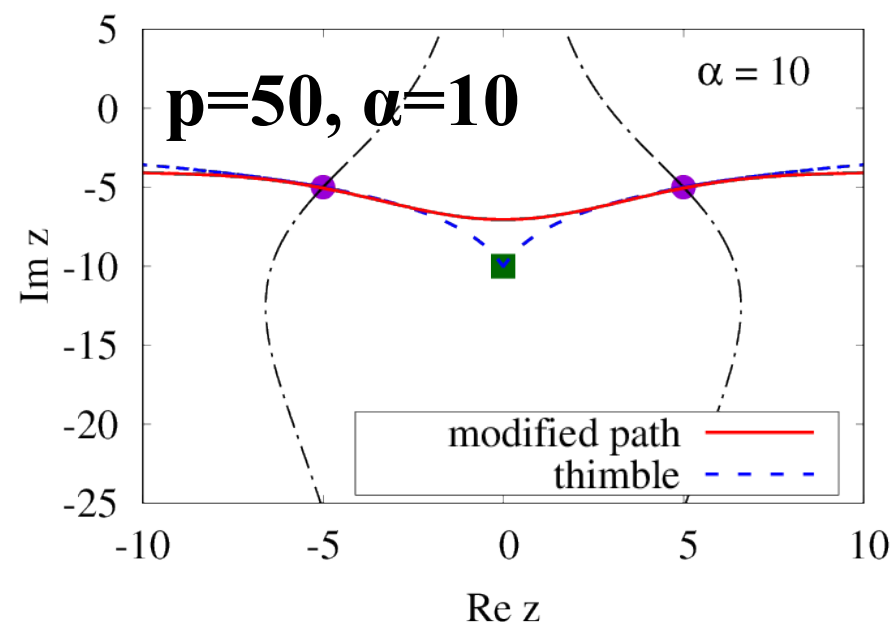
$$\mathcal{Z} = \int dx (x + i\alpha)^p \exp(-x^2/2)$$

- Sign prob. is serious with large p and small $\alpha \rightarrow$ CLM fails

■ Path optimization

$$y(x) = c_1 \exp(-c_2^2 x^2/2) + c_3, \quad J = 1 + i dy/dx$$

- Gradient Descent optimization
- Optimized path ~ Thimble around Fixed Points

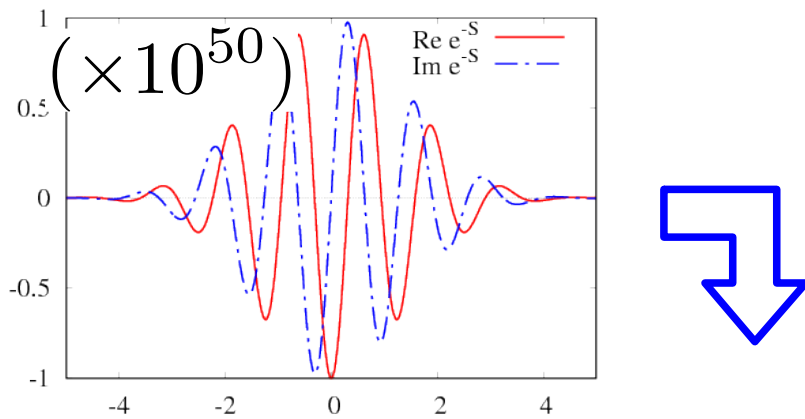


Mori, Kashiwa, AO ('17); AO, Mori, Kashiwa (Lat 2017)

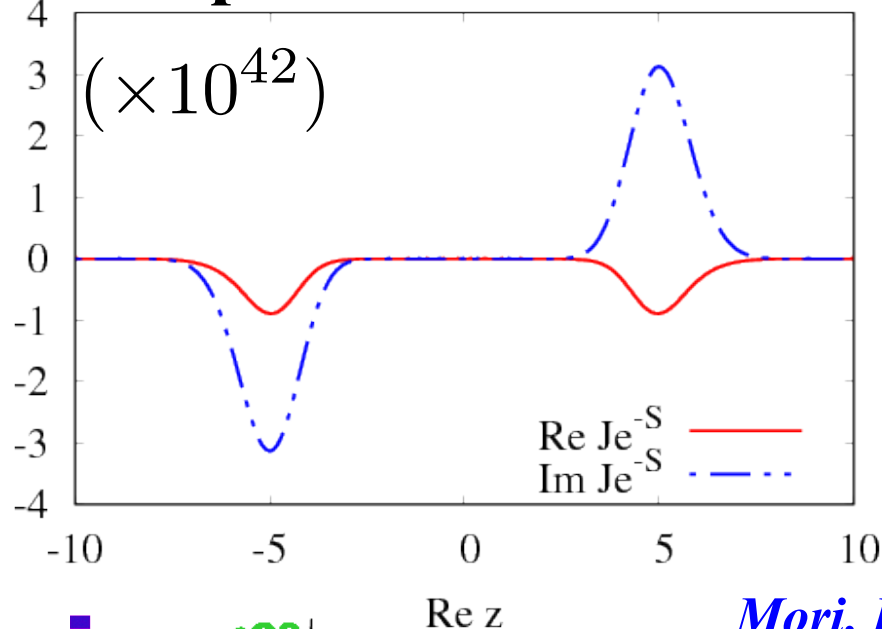
Benchmark test: 1 dim. integral

Stat. Weight $J e^{-S}$

On Real Axis

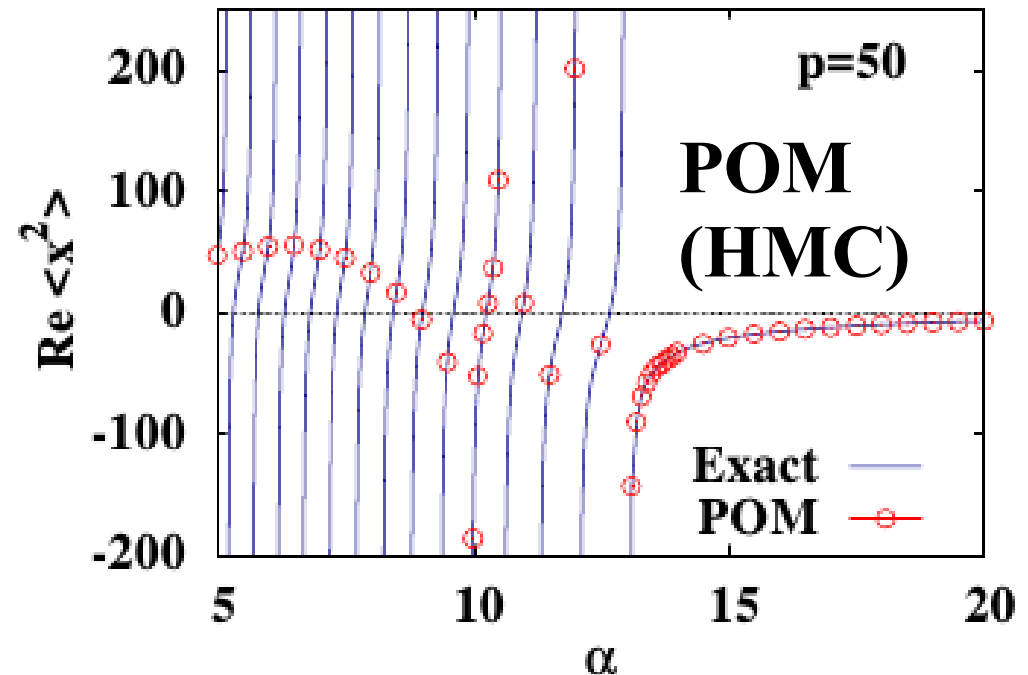
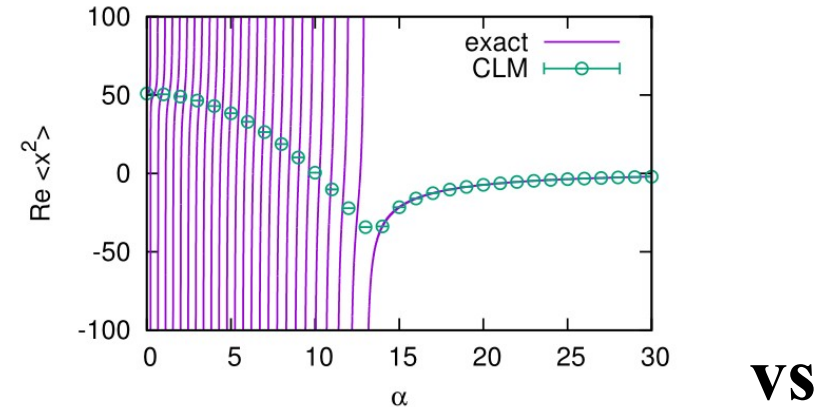


On Optimized Path



Observable

CLM *Nishimura, Shimasaki ('15)*



Mori, Kashiwa, AO ('17); AO, Mori, Kashiwa (Lat 2017)

*Now it's the time to apply POM
to field theories !*

Application of POM to Field Theory

- Preparation & variation of trial fn. is tedious in multi-D systems

$$z_i(x) = x_i + i \sum_{n_1, n_2, \dots} c_i(n_1 n_2 \dots) H_{n_1}(x_1) H_{n_2}(x_2) H_{n_3}(x_3) \dots$$

- Neural network

- Combination of linear and non-linear transformation.

$$a_i = g(\underline{W}_{ij}^{(1)} x_j + \underline{b}_i^{(1)}) \quad \text{parameters}$$

$$f_i = g(\underline{W}_{ij}^{(2)} a_j + \underline{b}_i^{(2)})$$

$$z_i = x_i + i(\underline{\alpha}_i f_i(x) + \underline{\beta}_i)$$

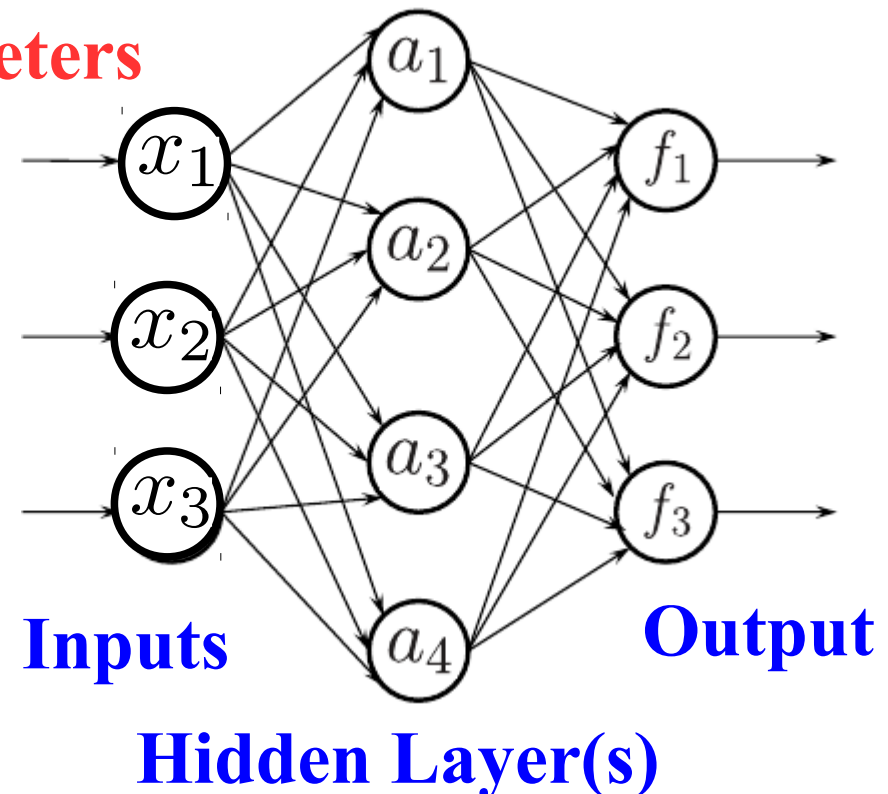
$$g(x) = \tanh x \quad (\text{activation fn.})$$

- Universal approximation theorem

Any fn. can be reproduced
at (hidden layer unit #) $\rightarrow \infty$

G. Cybenko, MCSS 2 ('89) 303

K. Hornik, Neural Networks 4('91) 251



Optimization of many parameters

- Stochastic Gradient Descent method, E.g. ADADELTA algorithm
M. D. Zeiler, arXiv:1212.5701

Grad. Desc. :
 $dc_i/dt = -\partial\mathcal{F}/\partial c_i$

par. in (j+1)th step

$$c_i^{(j+1)} = c_i^{(j)} - \eta v_i^{(j+1)}$$

Learning rate

mean sq. ave. of v

$$v_i^{(j+1)} = \frac{\sqrt{s_i^{(j)} + \epsilon}}{\sqrt{r_i^{(j+1)} + \epsilon}} F_i^{(j)}$$

decay rate

mean sq. ave. of F

$$r_i^{(j+1)} = \gamma r_i^{(j)} + (1 - \gamma)(F_i^{(j)})^2$$

$$s_i^{(j+1)} = \gamma s_i^{(j)} + (1 - \gamma)(v_i^{(j+1)})^2$$

gradient evaluated in MC (batch training)

$$F_i = \partial\mathcal{F}/\partial c_i$$

Cost fn.

Machine learning
 ~ Educated algorithm
 to generic problems

Hybrid Monte-Carlo with Neural Network

Initial Config. on Real Axis

$$\text{HMC } H(x, p) = \frac{p^2}{2} + \text{Re}S(z(x))$$

Jacobian

→ via Metropolis judge

Do k = 1, N_epoch

Do j = 1, N_conf/N_batch

Mini-batch training of Neural Network

$$c_i^{(j+1)} = c_i^{(j)} - \eta v_i^{(j+1)}$$

**Grad. wrt parameters
(N_batch configs.)**

$$F_i = \frac{1}{N_{\text{batch}}} \sum_n \partial \mathcal{F}(n) / \partial c_i$$

Enddo

New N_conf configs. by HMC

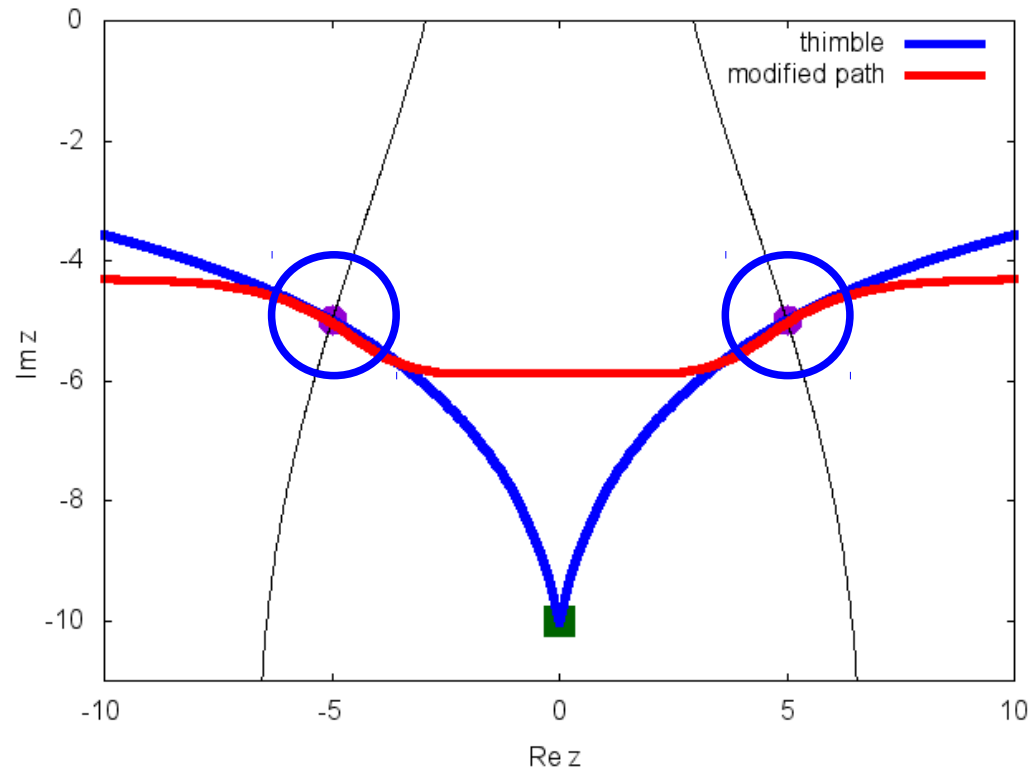
$$H(x, p) = \frac{p^2}{2} + \text{Re}S(z(x))$$

Enddo

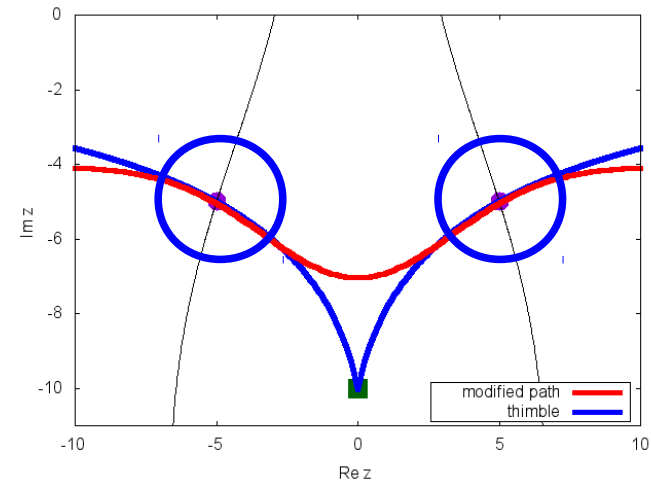
N_batch ~ 10, N_conf ~ 10,000, N_epoch ~ (10-20)

Optimized Path by Neural Network

Neural Network



Gaussian +Gradient Descent



*Optimized paths are different,
but both reproduce thimbles around the fixed points !*

AO, Mori, Kashiwa (Lat 2017)

Comments in order

- **Neural Network in POM → Unsupervised Learning**
 - We do not prepare “Teacher data” with problems and answers, but we prepare “Training data”.
 - When APF increases, we can judge that “NN learned”.
c.f. Learnifold: Machine learning of GTM manifold
(Alexandru (Tue))
- **Sign problem ~ (almost) “NP hard” problem**
 - Once the optimized path (answer) is given, we can examine the result (average phase factor) in MC simulations.

*Application to complex ϕ^4 theory
using neural network*

Y. Mori, K. Kashiwa, AO, PTEP 2018 ('18), 023B04 [arXiv:1709.03208]

Complex ϕ^4 theory at finite μ

Complex ϕ^4 theory

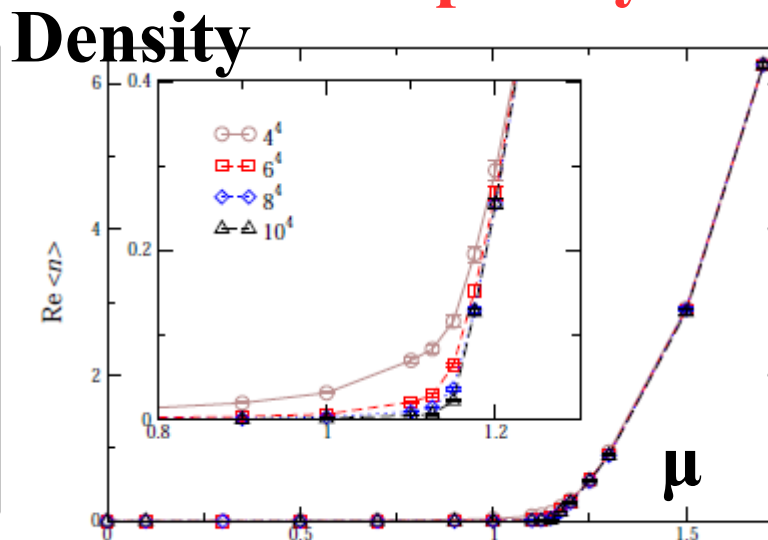
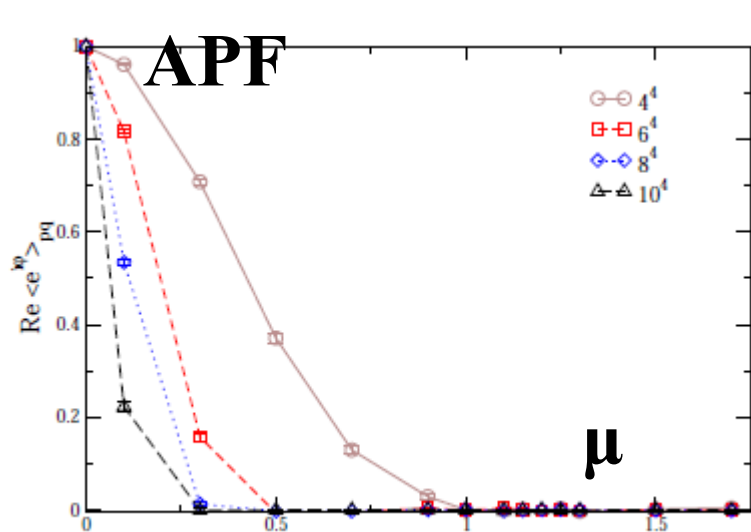
$$\mathcal{L} = \partial_\mu \phi^* \partial^\mu \phi - m^2 \phi^* \phi - \lambda (\phi^* \phi)^2$$

Action on Euclidean lattice at finite μ .

$$S = \sum_x \left[\frac{(4 + m^2)}{2} \phi_{a,x} \phi_{a,x} + \frac{\lambda}{4} (\phi_{a,x} \phi_{a,x})^2 - \phi_{a,x} \phi_{a,x+\hat{1}} - \cosh \mu \phi_{a,x} \phi_{a,x+\hat{0}} + i \epsilon_{ab} \sinh \mu \phi_{a,x} \phi_{b,x+\hat{0}} \right] \left(\phi = \frac{1}{\sqrt{2}} (\phi_1 + i \phi_2) \right)$$

complex

Complexify



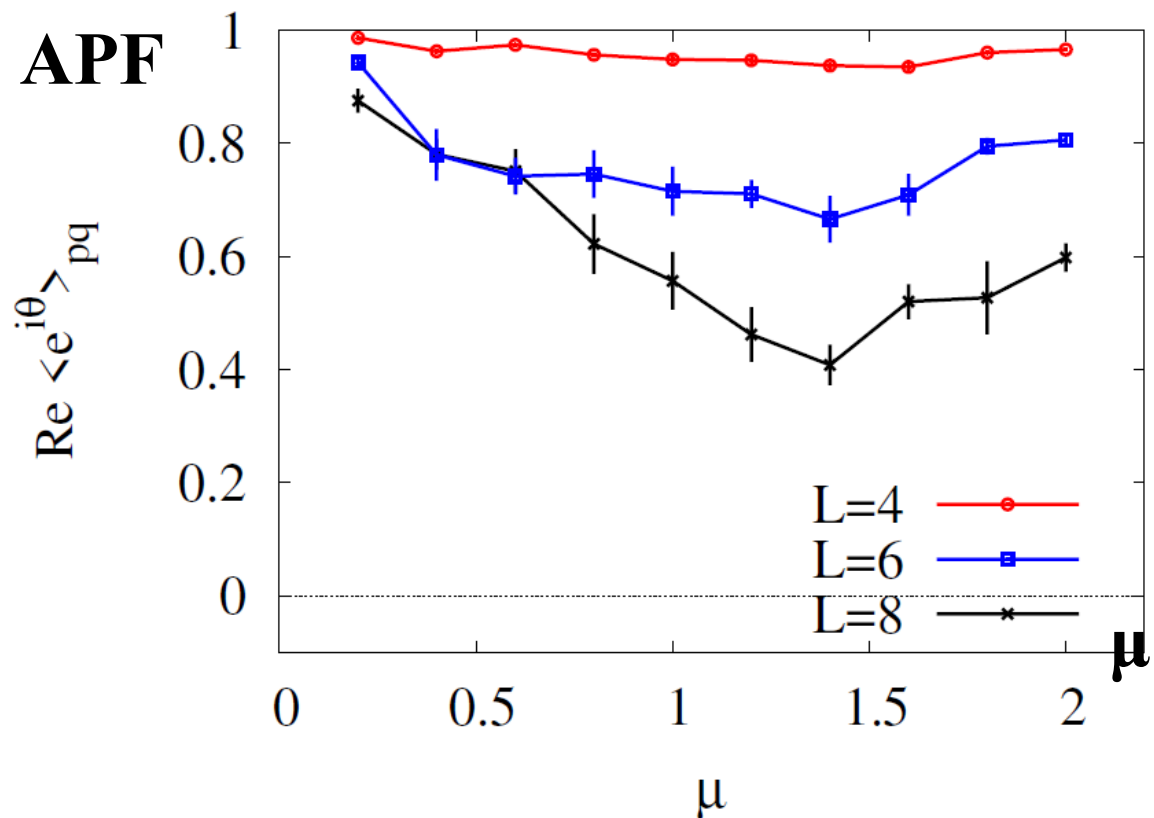
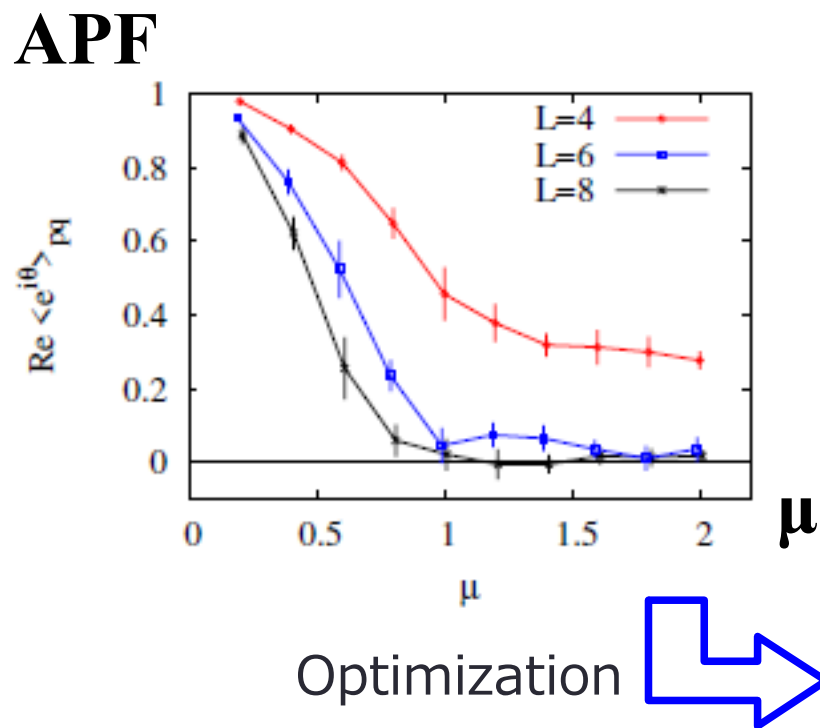
Complex
Langevin
& Lefschetz
thimble
work.

μ G. Aarts, PRL102('09)131601; H. Fujii, et al., JHEP 1310 (2013) 147

POM result (1): Average phase factor

■ POM for 1+1D ϕ^4 theory

- $4^2, 6^2, 8^2$ lattices, $\lambda=m=1$
- $\mu_c \sim 0.96$ in the mean field approximation
- Enhancement of the average phase factor after optimization.



Y. Mori, K. Kashiwa, AO, PTEP 2018 ('18), 023B04 [arXiv:1709.03208]

POM result (2): Density

- Results on the real axis
Small average phase factor, Large errors of density
- On the optimized path
Finite average phase factor, Small errors

Mean Field App.

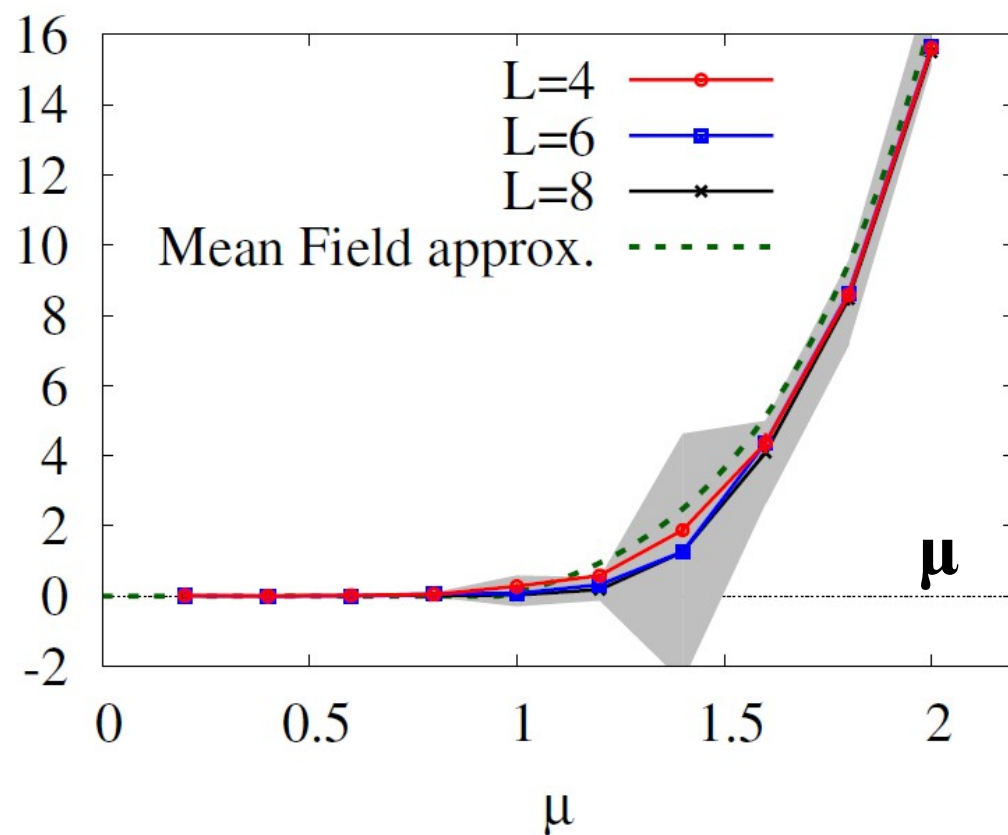
$$\frac{S}{V} = \left(1 + \frac{m^2}{2} - \cosh \mu \right) \phi^2 + \frac{\lambda}{4} \phi^4 ,$$

$$n = \phi^2 \sinh \mu ,$$

$$\phi_{\text{stat.}}^2 = \begin{cases} 0 & (|\mu| < \mu_c) , \\ \frac{2}{\lambda} (\cosh \mu - 1 - \frac{m^2}{2}) & (|\mu| \geq \mu_c) , \end{cases}$$

Density

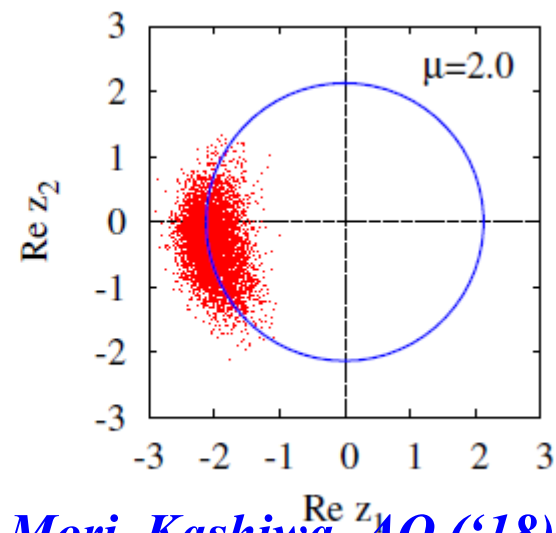
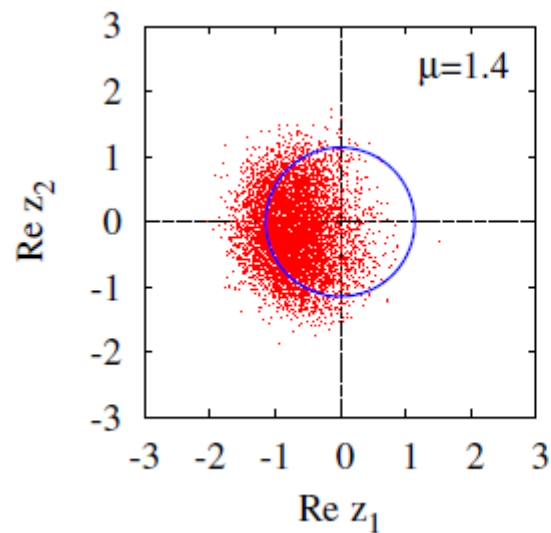
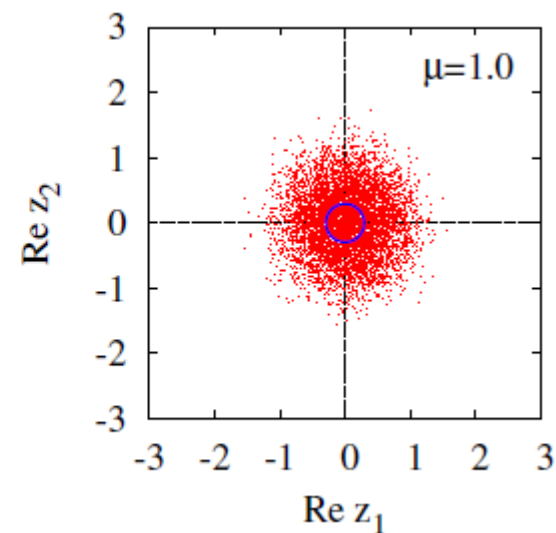
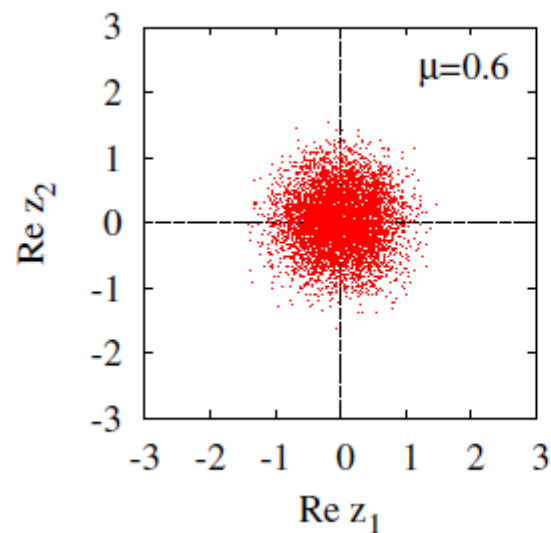
Re <n>



Mori, Kashiwa, AO ('18)

POM result (3): Configurations

- Updated configurations after optimization
→ sampled around the mean field results
- Global U(1) symmetry in (φ_1, φ_2) is broken(*)
by the optimization
or by the sampling.



* This does not contradict the Elitzur's theorem.

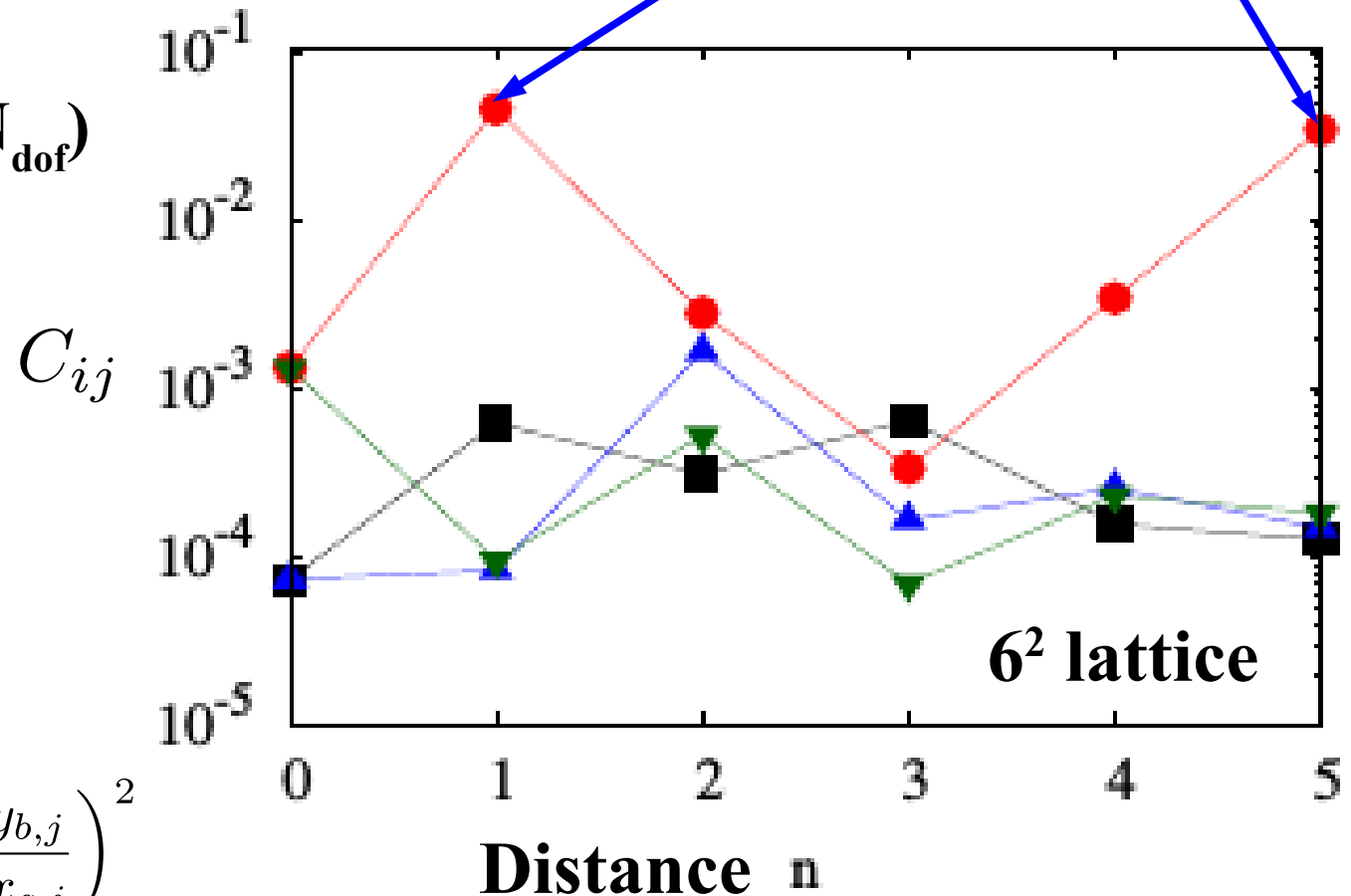
Mori, Kashiwa, AO ('18)

Which y 's should be optimized ?

- Correlation btw (z_1, z_2) of temporal nearest neighbor sites are strong. Other correlations $\sim 10^{-2}$ times smaller

$$\text{Im}(S) = \sum_x \epsilon_{ab} \sinh \mu \phi_{a,x} \phi_{b,x+\hat{0}}$$

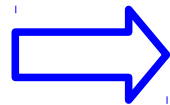
- Hope to reduce the cost to be $O(N_{\text{dof}})$



$$C_{ij} \equiv \left(\frac{\partial y_{a,i}}{\partial x_{b,j}} \right)^2 + \left(\frac{\partial y_{b,j}}{\partial x_{a,i}} \right)^2$$

Y. Mori, Master thesis

*Application to Gauge Theory:
1 dimensional QCD*



Next Talk (by Yuto Mori)

Y. Mori, K Kashiwa, AO, in prep.

Application to Effective Models with Phase Transition

*Y. Mori, K. Kashiwa, AO, PLB 781('18),698 [arXiv:1705.03646]
K. Kashiwa, Y. Mori, AO, arXiv:1805.08940 and in prep.*

Complexified Variable Methods with Phase Transition

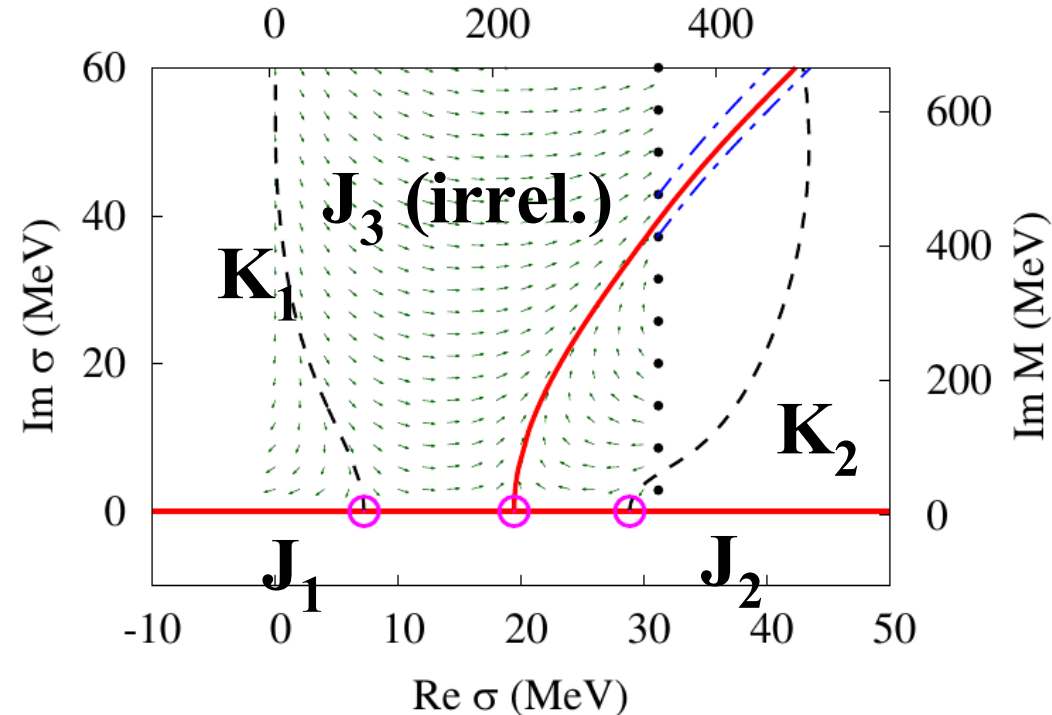
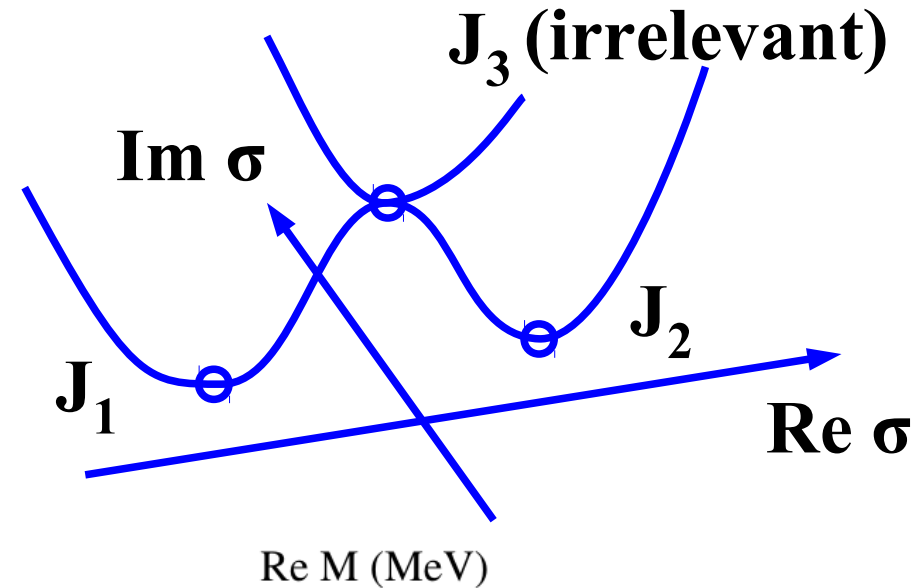
Phase transition

- Transition of the dominant weight from a thimble to another thimble.
- Thimble structure would be complicated.

Thimbles in Nambu-Jona-Lasinio model *Mori et al. ('18b)*

- Complicated structure with many cuts from log.

Does POM work in Eff. models with P.T. ?



*Y. Mori, K. Kashiwa, AO,
PLB 781('18),698 [arXiv:1705.03646]*

Application to PNJL

- Polyakov-loop implemented NJL model with homogeneous auxiliary field ansatz of $(\sigma, \pi, \Phi, \bar{\Phi})$.

$$\Gamma_{\text{PNJL}} = \int d^4x \left[\bar{q}(-i \not{D} + m_0 - \mu\gamma_0)q - G \left\{ (\bar{q}q)^2 + (\bar{q}i\vec{\tau}q)^2 \right\} \right]$$

$$+ \frac{V}{T} \mathcal{V}_{\Phi}(\Phi, \bar{\Phi}) \simeq \frac{V}{T} [\mathcal{V}_{\text{NJL}} + \mathcal{V}_{\Phi}] \quad \text{Pol. loop pot.}$$

$$\mathcal{V}_{\text{NJL}} = -2N_f \int \frac{d^3p}{(2\pi)^3} \left[N_c \underline{E_p} - N_c \sqrt{\mathbf{p}^2 + m_0^2} + T \log(\underline{f^- f^+}) \right]$$

$$+ G(\sigma^2 + \vec{\pi}^2) \quad \sigma, \pi \quad \sigma, \pi, \Phi, \bar{\Phi}$$

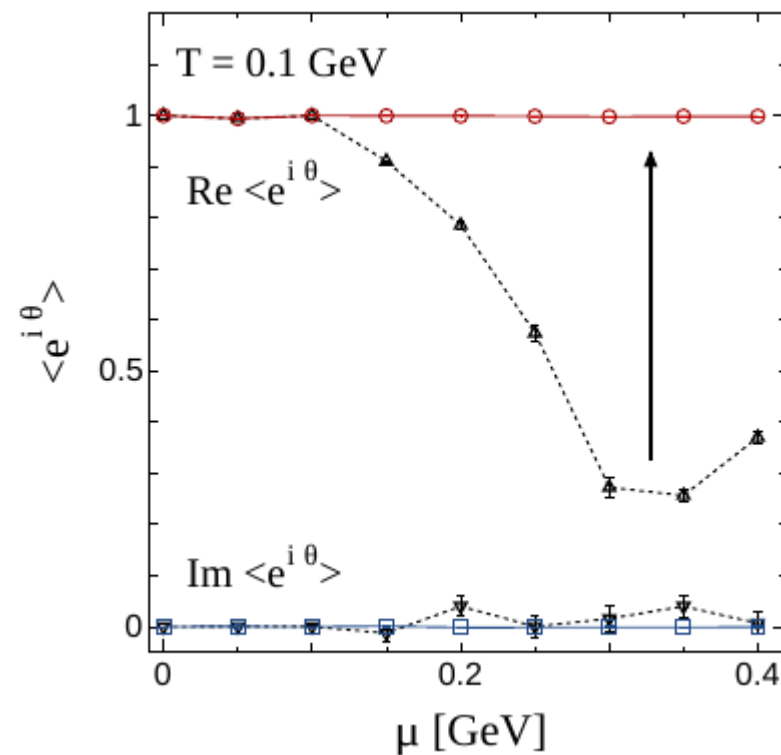
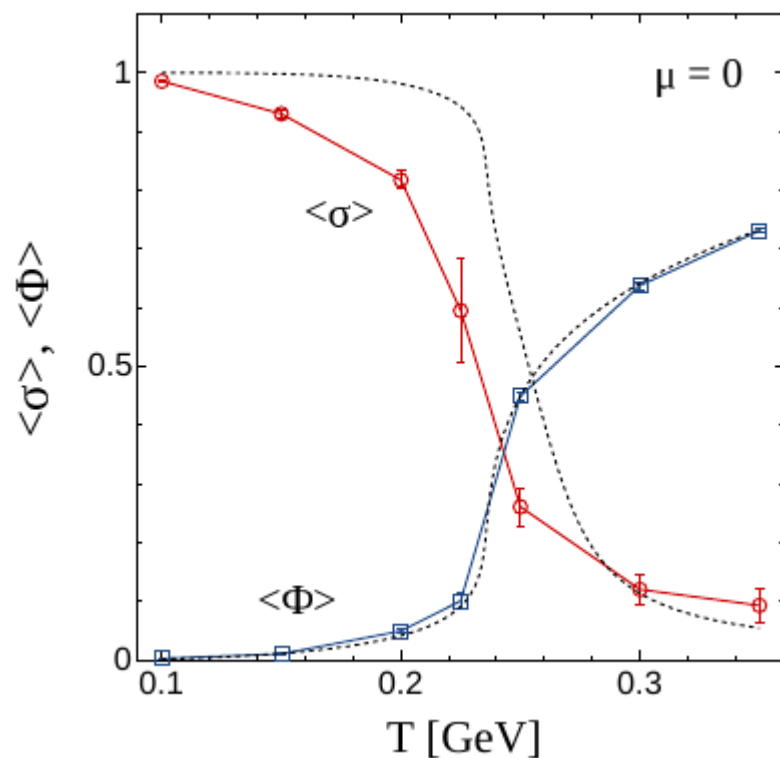
- Action (Γ) has the imaginary part.
- The sign problem exists at finite volume.
(Mean field becomes exact at infinite volume.)

K. Kashiwa, Y. Mori, AO, arXiv:1805.08940.

Application to PNJL

■ Path optimization for (A_3, A_8) with $V=k/T^3$, $k=64$.

- Transition from Confined NG phase to Deconfined Wigner phase
- Ave. Phase Factor ~ 0.25 around P.T. $\rightarrow \sim 1$ after optimization



K. Kashiwa, Y. Mori, AO, arXiv:1805.08940.

PNJL with Vector Coupling

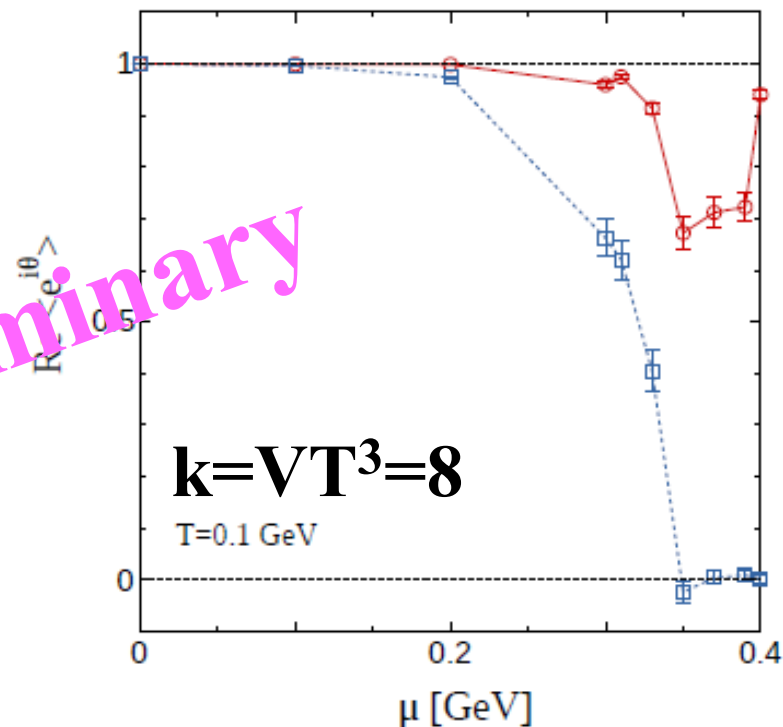
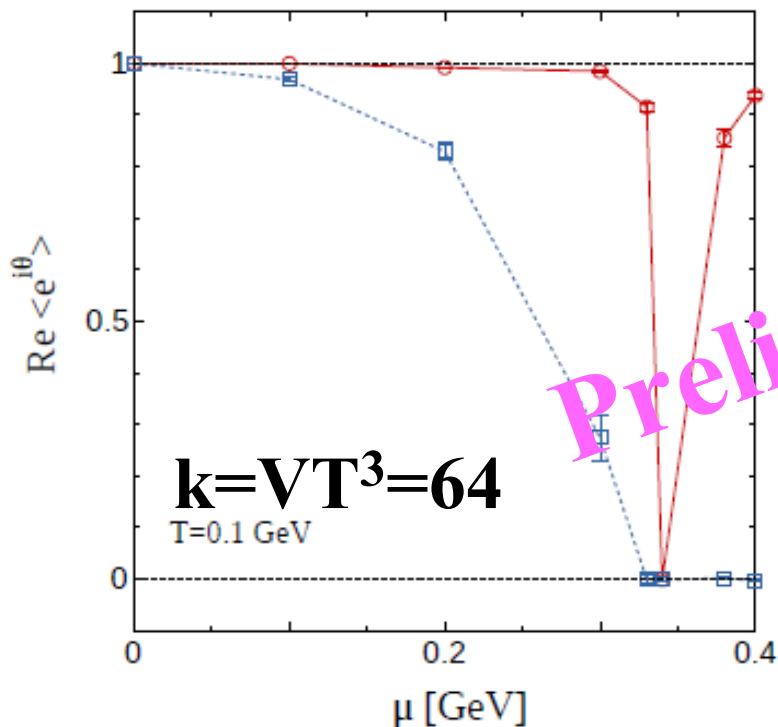
■ Action of PNJL model with vector coupling

$$\Gamma_{\text{PNJL}_V} = \Gamma_{\text{PNJL}} + G_v \int d^4x [(\bar{q}\gamma_0 q)^2 - (\bar{q}\gamma_i q)^2]$$

$$\simeq -\log \det D + \int d^4x (G(\sigma^2 + \vec{\pi}^2) + G_v \omega_\mu^2) + \frac{V}{T} \mathcal{V}_\Phi(\Phi, \bar{\Phi})$$

$$D = -i \not{D} + m_0 + 2G\sigma - \gamma_0(\mu - 2iG_v \omega_4) - 2iG\gamma_5 \vec{\pi} \cdot \vec{\tau} + 2G_v \gamma_i \omega_i$$

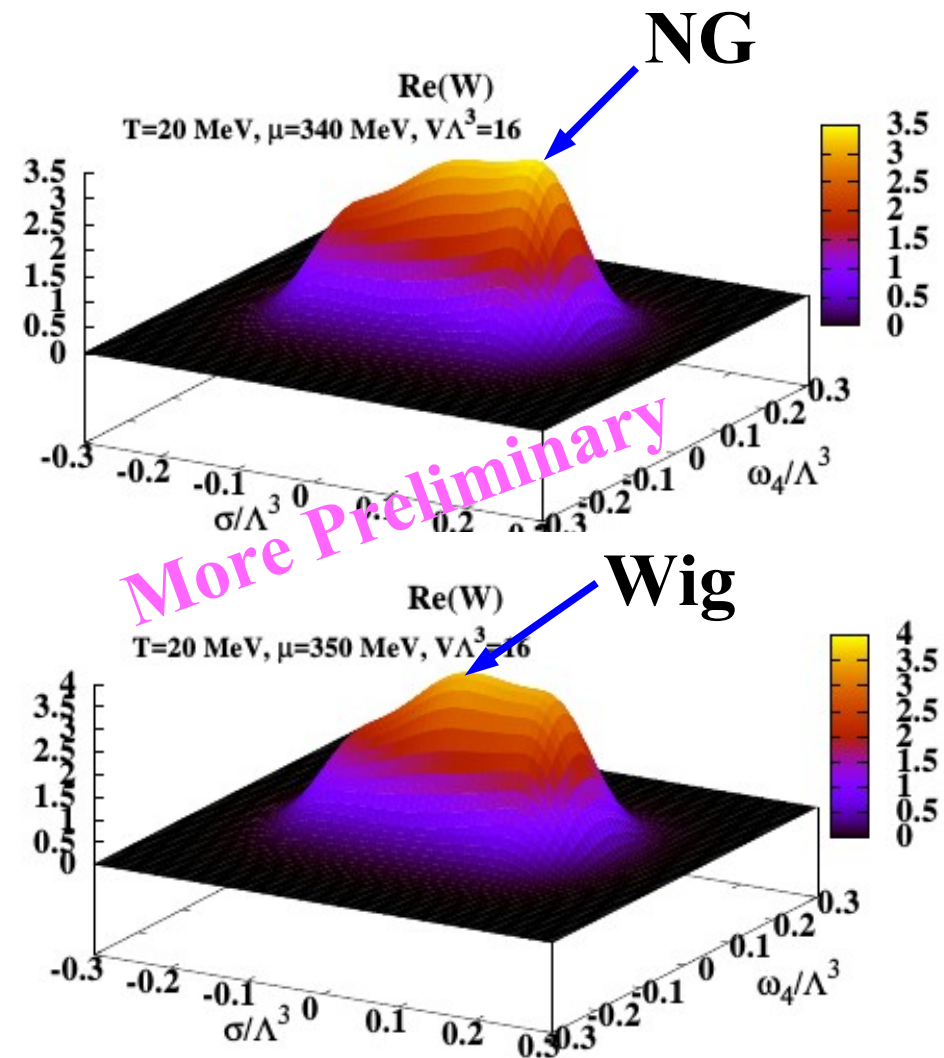
Wick rotated vector field \rightarrow Complex Action



K. Kashiwa, Y. Mori, AO, in prep.

Do we describe two thimbles ?

- It seems yes.
- Statistical weight in NJL ν on (σ, ω_4) plain after optimization on the 2D mesh.
→ Three thimbles
- Transition from the Nambu-Goldstone phase to the Wigner phase occurs at $\mu=(340-350)$ MeV
- In the $V \rightarrow \infty$ limit, this corresponds to the phase transition. (We may need exchange MC or different tempering.)



Discussions

Frequently Asked Questions

- How many parameters do you have ?

→ Many ;) For generic trial function ($V = \#$ of variables)

$$y_i = y_i(x_1, x_2, \dots, x_V)$$

$$N_{\text{par}} = (N_{\text{layer}} + 1) \times V \times (N_{\text{unit}} + 1) + 2V$$

- How about the numerical cost ?

→ A lot ;) Derivative of J with respect to parameters cost most.

$$\frac{\partial J}{\partial c_i} = J \frac{\partial J_{jk}^{-1}}{\partial z_l} \frac{\partial z_l}{\partial c_i} \rightarrow \mathcal{O}(V^3)$$

- It is still polynomial.

Does the sign problem becomes “P” problem?

→ No. The average phase factor is still $\exp(-\# V)$.

If extrapolation is possible from finite V , we have a hope.

- How can we reduce the cost ? → Next page

How can we reduce the numerical cost ?

- Restrict the function form of $y(x)$.

- Imaginary part is a function of its real part.

E.g. Alexandru, Bedaque, Lamm, Lawrence, PRD97('18)094510

Thirring model, 1+1D QED

$$y_i = f(x_i), f(x) = \lambda_0 + \lambda_1 \cos x$$

- Nearest neighbor site

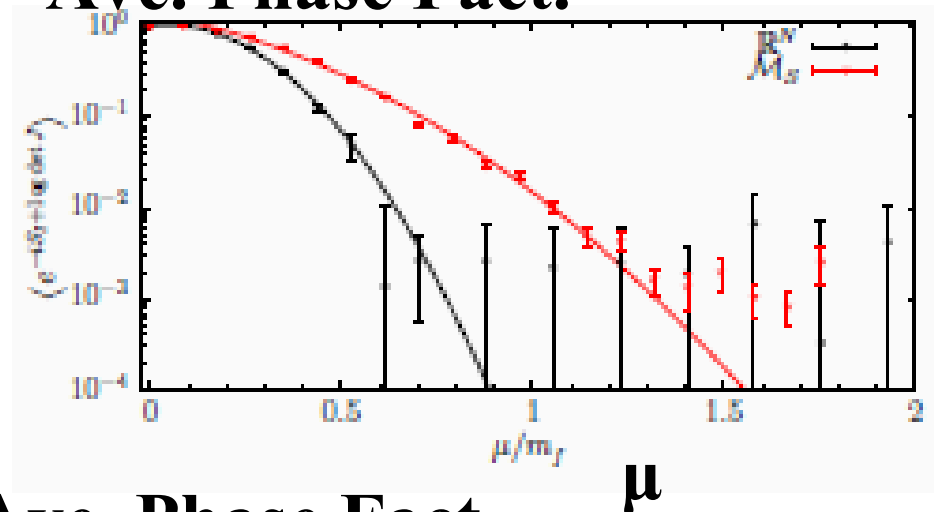
F. Bursa, M. Kroyter, arXiv:1805.04941

0+1 D ϕ^4 theory

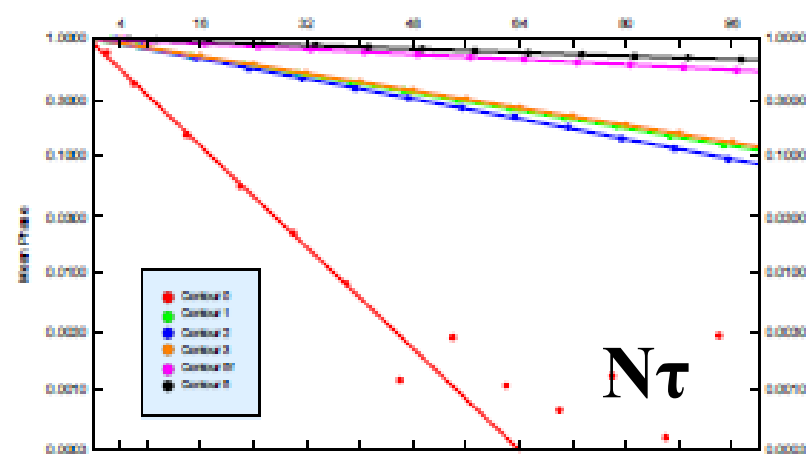
Translational inv. + U(1) sym.

$$y_{a,i} = \frac{\varepsilon_{ab} x_{a,i+1}}{1 + x_{1,i}^2 + x_{2,i}^2}$$

Ave. Phase Fact.

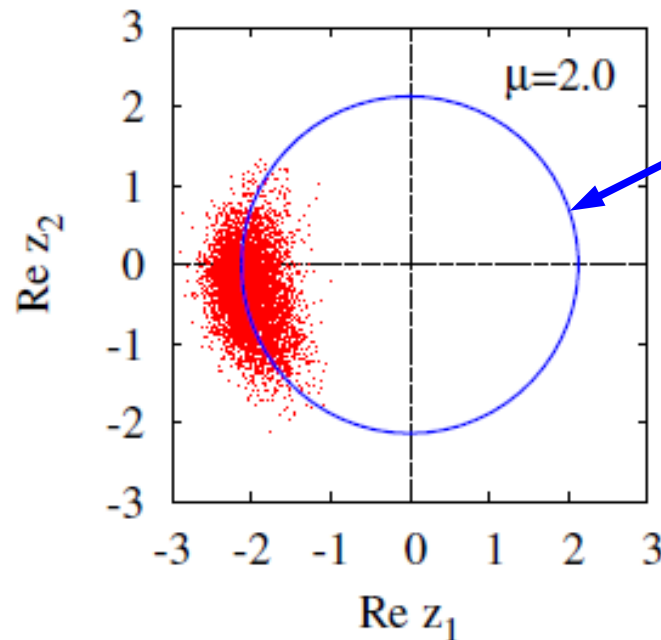


Ave. Phase Fact.



Frequently Asked Questions (cont.)

- What happens when we have 10^{10} fixed points ?
 - In that case we should give up. (My answer @ Lattice 2017)
 - If those fixed points are connected by the symmetry, we may be able to perform path optimization.



Mean field results
= Degenerate fixed points
(All have the same θ .)

If they have different complex phases, the global sign problem emerges and the partition function would be almost zero.

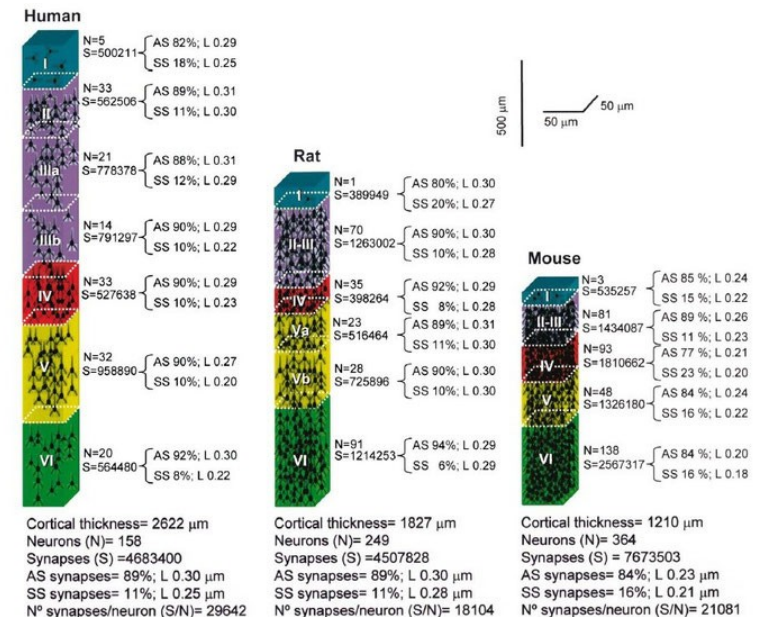
E.g. H. Fujii, S. Kamata, Y. Kikukawa, arXiv:1710.08524

Summary

- The sign problem is a grand challenge in theoretical physics, and appears in many fields of physics. Complexified variable methods (CLM, LTM, GTM, POM) would be promising to evade the sign problem.
- **Path optimization** with the use of the **neural network** is demonstrated to work in **field theories** having many variables.
 - 1+1D ϕ^4 theory at finite μ (neural network)
 - 0+1D QCD w/ fermions (2D mesh, neural network)
 - 3+1D homogeneous PNJL (neural network)
- Neural network (single hidden layer) is the simplest device of machine learning, and it helps us to **generate and optimize generic multi-variable functions**, $y_i = y_i(\{x\})$.
- **Gauge fields (link variables or Polyakov loop)** and **Repulsive interaction** makes the sign problem more severe.

Prospect

- Path optimization in 3+1 D field theories would require reduction of numerical cost.
 Imaginary part
 = f (real parts of same point and nearest neighbor points)
 may be a good guess.
- Deep learning (# of hidden layers > 3) may be helpful to explore complex paths, which human beings (~ 7 layers) cannot imagine, while “Understanding” the results of machine learning need to be done by human beings (at present).

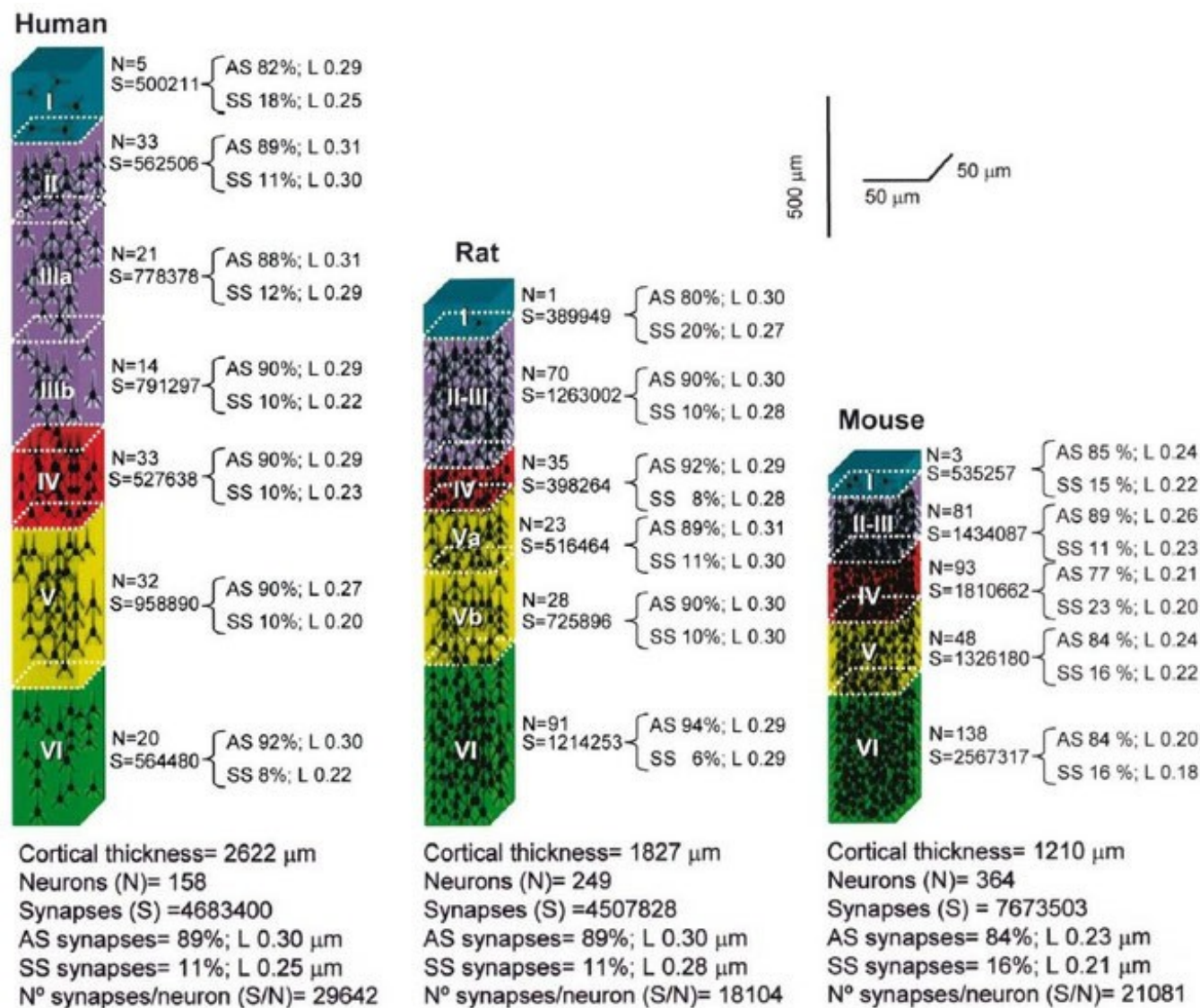


Defelipe 2011a (Review). The evolution of the brain, the human nature of cortical circuits, and intellectual creativity. Front Neuroanat 5, 29.

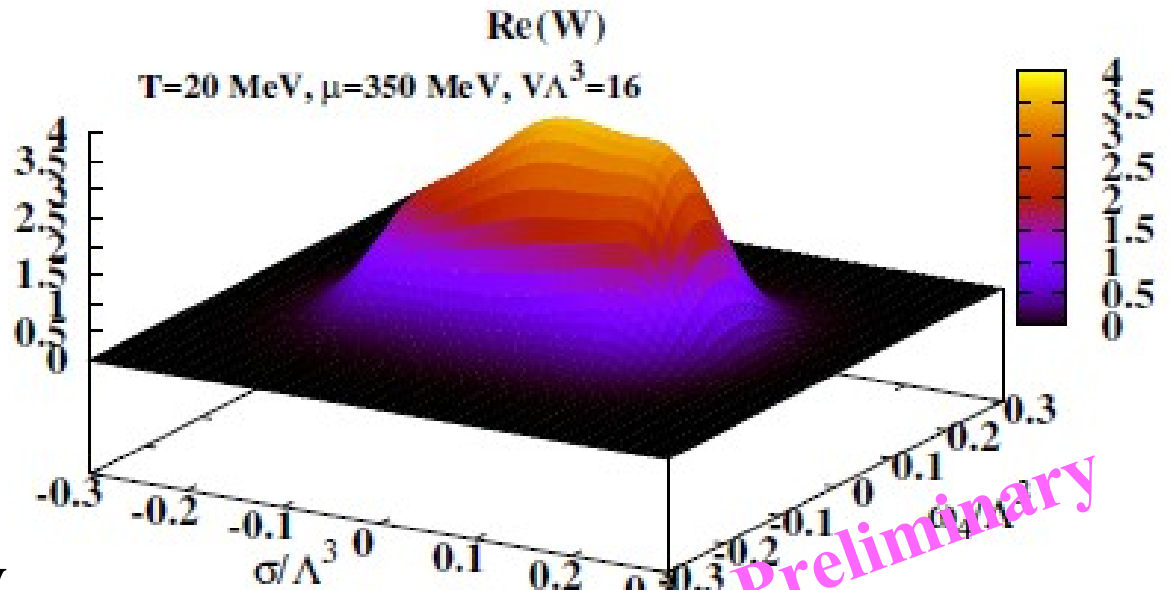
Thank you for your attention !

Layers in Human Brain

- Defelipe 2011a (Review). The evolution of the brain, the human nature of cortical circuits, and intellectual creativity. *Front Neuroanat* 5, 29.

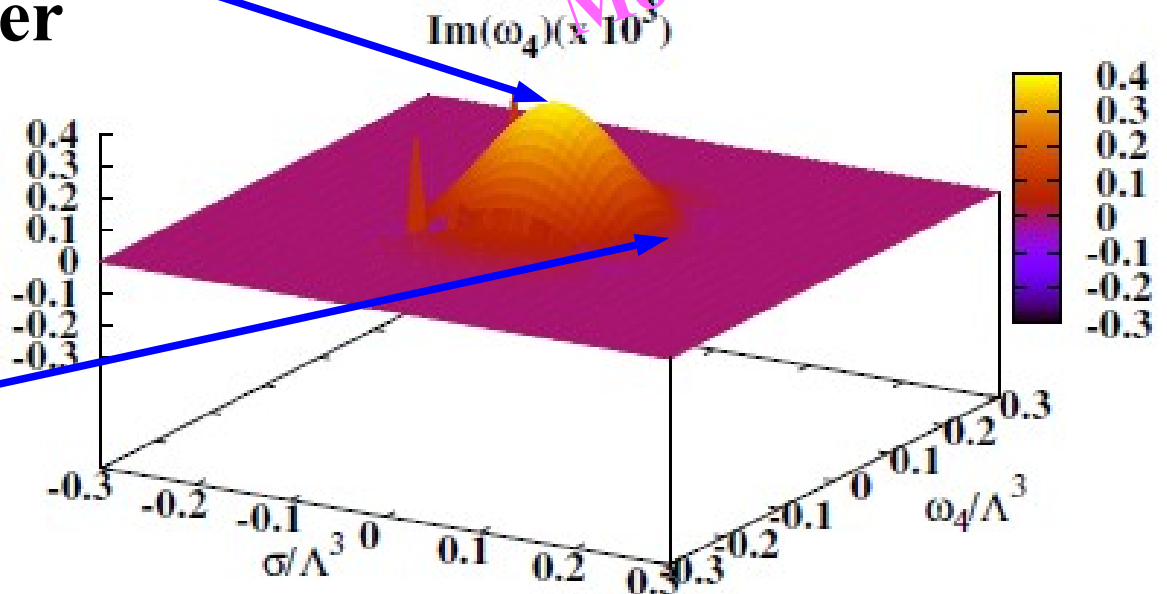


Config. dep. vector field shift ?



High density
Quark Matter

Low density
Nuclear Matter



More Preliminary

*Application to Gauge Theory:
1 dimensional QCD*

Y. Mori, K Kashiwa, AO, in prep.

0+1 dimensional QCD

0+1 dimensional QCD (1 dim. QCD)

with one species of staggered fermion on a 1xN lattice

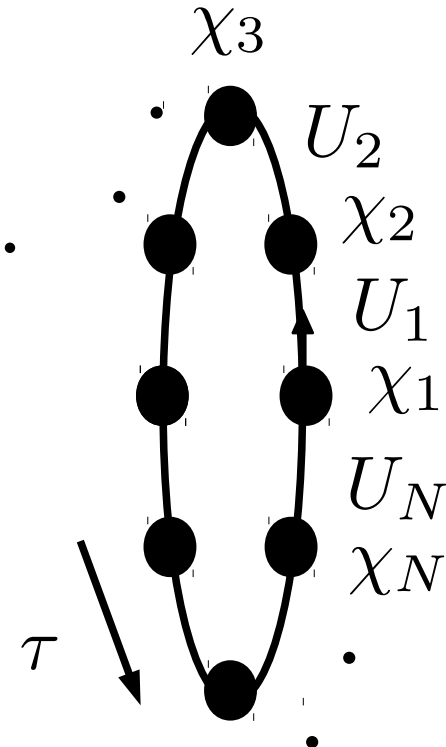
$$S = \frac{1}{2} \sum_{\tau} (\bar{\chi}_{\tau} e^{\mu} U_{\tau} \chi_{\tau+\hat{0}} - \bar{\chi}_{\tau+\hat{0}} e^{-\mu} U_{\tau}^{-1} \chi_{\tau}) + m \sum_{\tau} \bar{\chi}_{\tau} \chi_{\tau} = \frac{1}{2} \bar{\chi} D \chi$$

$$\mathcal{Z} = \int \mathcal{D}U \det D[U] = \int dU \det \left[X_N + (-1)^N e^{\mu/T} U + e^{-\mu/T} U^{-1} \right]$$

$$X_N = 2 \cosh(E/T), \quad E = \operatorname{arcsinh} m, \quad U = U_1 U_2 \cdots U_N, \quad T = 1/N$$

*Bilic('88), Ravagli('07), Aarts('10, CLM), Bloch('13, subset),
Schmidt('16, LTM), Di Renzo('17, LTM)*

- A toy model, but the actual source of QCD sign prob.
- Studied well in the context of strong coupling LQCD
*E.g. Miura, Nakano, AO, Kawamoto('09, '09, '17),
de Forcrand, Langelage, Philipsen, Unger ('14)*



1 dim. QCD in diagonal gauge

■ Diagonal gauge

$$U = (e^{iz_1}, e^{iz_2}, e^{iz_3}) \quad (z_1 + z_2 + z_3 = 0)$$

$$\mathcal{Z} = \int dU e^{-S(U)} = \int_{\mathbb{R}} dx_1 dx_2 H(x) e^{-S(x)} = \int_{\mathbb{C}} dz_1 dz_2 H(z) e^{-S(z)}$$

$$= \int dx_1 dx_2 \det \left(\frac{\partial z_a}{\partial x_b} \right) \left[\frac{8}{3\pi^2} \prod_{a < b} \sin^2 \left(\frac{z_a - z_b}{2} \right) \right] \left[\prod_a (X_N + 2 \cos(z_a - i\mu)) \right]$$

Jacobian

Haar measure

exp(-S)

■ Path optimization (t: fictitious time)

→ $y(x_1, x_2)$ itself is the parameter on the (x_1, x_2) mesh point

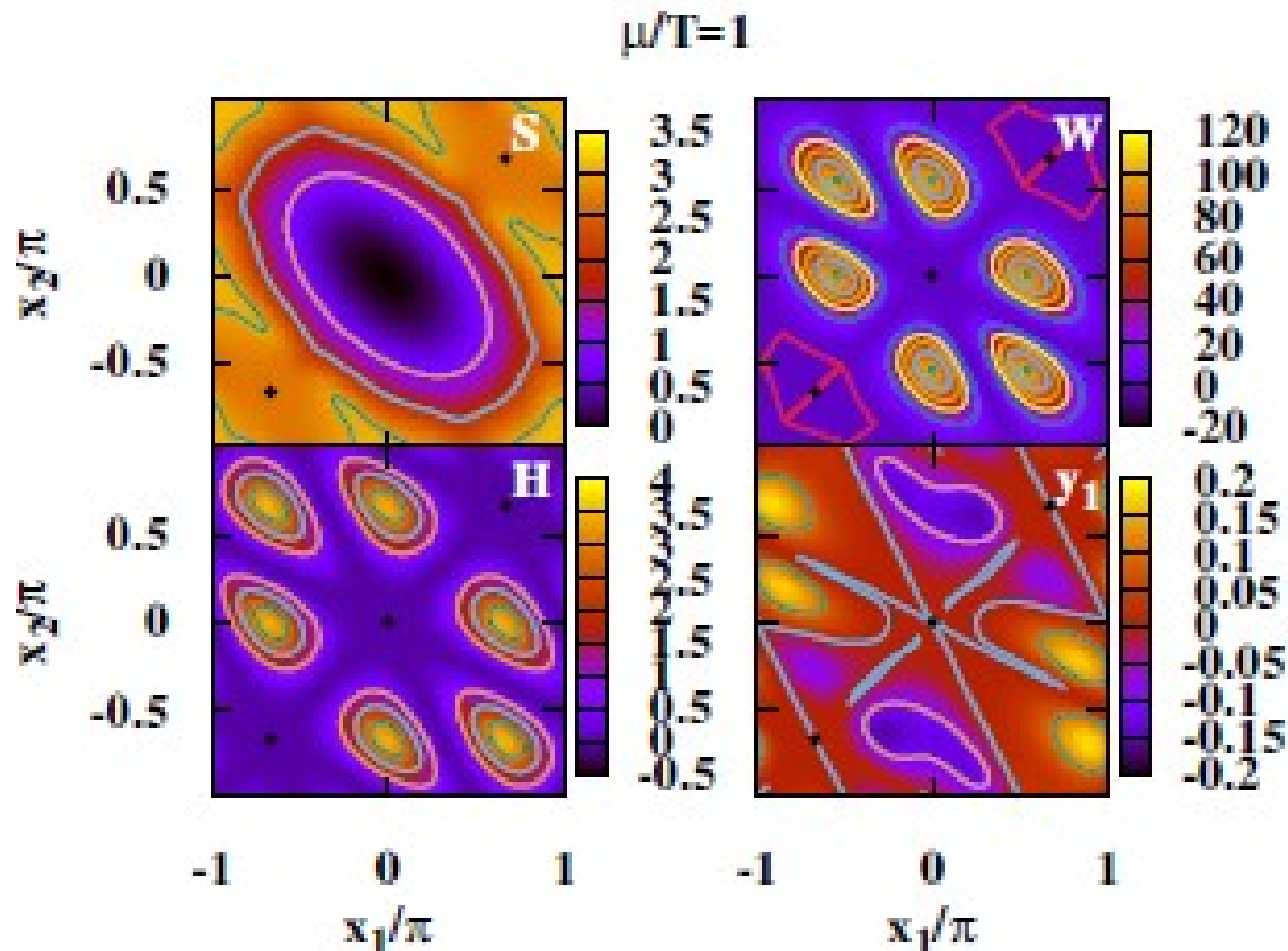
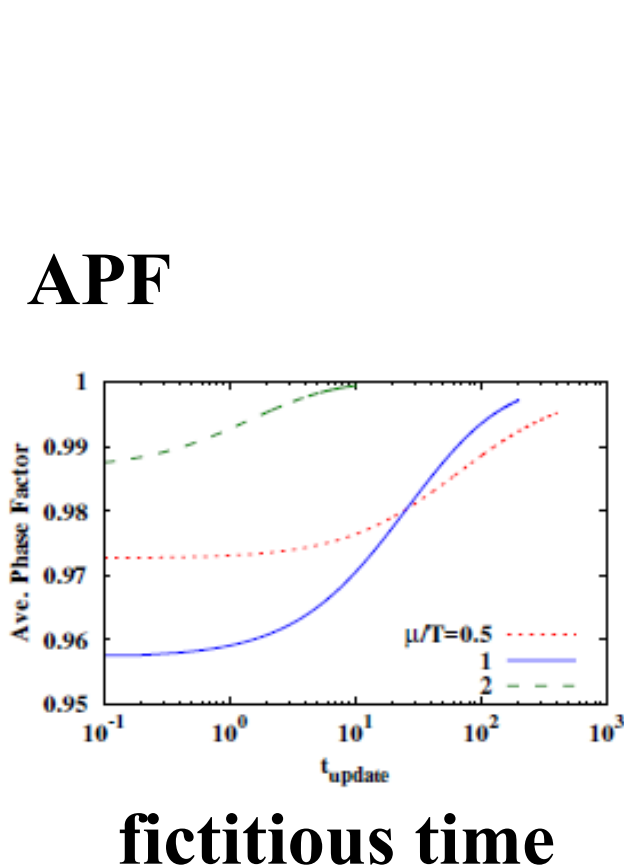
$$z_i = x_i + iy_1, \quad y_i = y_i(x_1, x_2)$$

$$\frac{dy_i}{dt} = -\frac{\partial \mathcal{Z}_{pq}}{\partial y_i}, \quad \mathcal{Z}_{pq} = \int dx_1 dx_2 |JH e^{-S}|$$

Path Opt. of 1 dim. QCD in diagonal temporal gauge

■ Path optimization

- Average phase factor > 0.99 \rightarrow Easily achieved
- $\exp(-S)$ and Haar Measure \rightarrow “six pads” *Schmidt+('16, LTM)*



Mori, Kashiwa, AO, in prep.

1 dim. QCD with Hybrid MC

■ Concern...

- Six pads are separated by the Haar measure barrier.

$$\text{Symmetry : } S(-z) = (S(z^*))^*, z_i \leftrightarrow z_j (i, j = 1, 2, 3)$$

Do we need exchange MC or different tempering ?

E.g. Fukuma, Matsumoto, Umeda ('17)

■ Hybrid Monte-Carlo in 1 dim. QCD

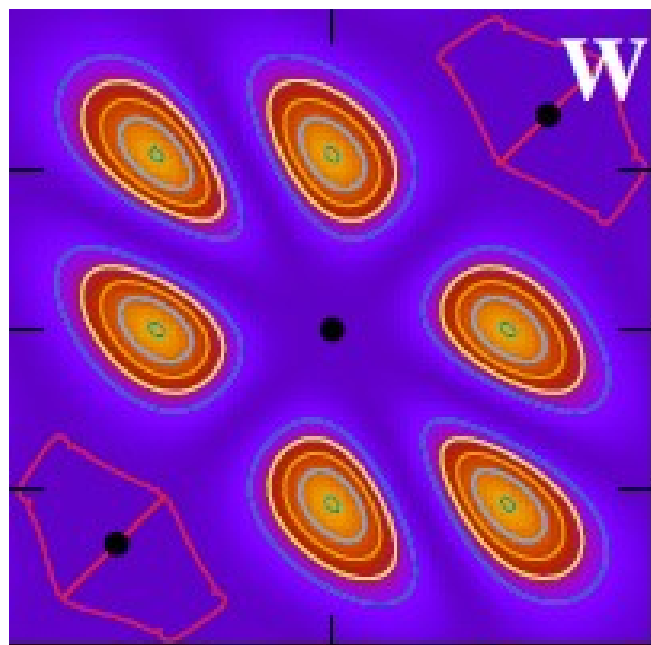
$$U \rightarrow \mathcal{U}(U) = U \prod_{a=1}^{N_c^2-1} e^{-y_i \lambda_i / 2}, \quad H = \frac{P^2}{2} + \text{Re}(S(\mathcal{U}(U)))$$

SL(3)

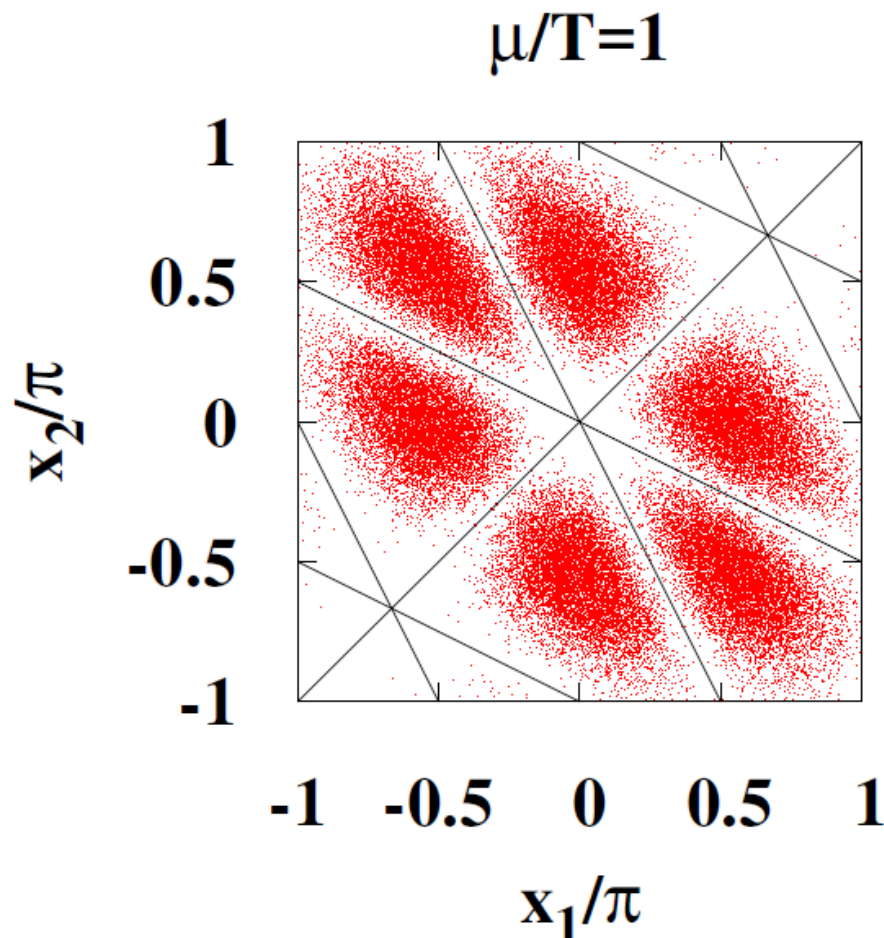
- 8 variables → path optimization using Neural Network

1 dim. QCD with Hybrid MC

- HMC + diagonalization of the link
→ All six pads are visited, and no Ex. MC needed.



Mesh point + Grad. Desc.



HMC + NN

Mori, Kashiwa, AO, in prep.