New criterion for correctness in the complex Langevin method - an application to finite density QCD

Keitaro Nagata(KEK)
Hideo Matsufuru (KEK), Jun Nishimura (KEK, Sokendai),
Shinji Shimasaki (Keio Univ.)

Reference)
Matsufuru, KN, Nishimura, Shimasaki, work in progress
KN, Nishimura, Shimasaki, 1606.07627

1. Introduction

Recent progress of complex Langevin method(CLM) enables us to study systems of complex action, e.g QCD at finite density

[Aarts, Seiler, Stamatescu (‘10), Seiler, Sexty, Stamatescu (’13), Sexty (’13), Fodor, Katz, Sexty, Torok (’15), ...]

Crucial question:

How do we distinguish whether results in CLM are correct or wrong ?

Revisiting the argument for justification of CLM leads us to

a new criterion of correctness = probability of drift terms

[KN, Nishimura, Shimasaki, 1606.07627, Shimasaki talk in Lat'16]

New criterion - probability of drift terms

✓**legitimate**

- obtained from theoretical argument

✓**clear**

- signal is qualitative

✓**cheap**

- no additional calculation required

Outline of this talk

- **Framework: condition for probability of drift**

- **Demonstration of effectiveness of the new criterion in a solvable system**
  - **2d SU(2) Yang-Mills**
  - [see 1606.07627 , Shimasaki talk in Lat'16 for applications to one-variable models and RMT]

- **Applications of the criterion to finite density QCD**
  - **we study QCD at finite density using CLM and determine the reliable range of chemical potential using the new criterion.**

Complex Langevin method

- Langevin equation with finite stepsize

$$z^{(\eta)}(t+\epsilon) = z^{(\eta)}(t) + \epsilon v(z) + \sqrt{\epsilon}\underline{\eta(t)}$$

Gaussian white noise

drift term

- Expectation value of an observable O

$$\Phi(t) = \left\langle \mathcal{O}\left(x^{(\eta)}(t) + iy^{(\eta)}(t)\right)\right\rangle_{\eta} = \int dx\, dy\, \mathcal{O}(x+iy)\, P(x,y;t)$$

$$P(x,y;t) = \left\langle \prod_k \delta\left(x_k - x_k^{(\eta)}(t)\right)\delta\left(y_k - y_k^{(\eta)}(t)\right)\right\rangle_{\eta} .$$

ε-evolution with finite stepsize

- ε-evolution (for holomorhpic observable)

$$\Phi(t+\epsilon) = \int dx\, dy\, \mathcal{O}(x+iy)\, P(x,y;t+\epsilon)$$

: … : = ordering operator (put derivative right)

$$= \sum_{n=0}^{\infty} \frac{1}{n!}\epsilon^n \int dx\, dy\, \{: \tilde{L}^n : \mathcal{O}(x+iy)\}\, P(x,y;t)$$

$$\tilde{L} = (v(z) + \partial_z)\partial_z$$

- *The probability distribution should fall-off faster than any power law at large drift*
- If the integrals converge for any n, we can take ε->0 limit (sufficiently smaller than the convergence radius)

$$\Phi(t+\epsilon) = \int dxdy\{(1+\epsilon\tilde{L})\mathcal{O}(x+iy)\}P(x,y;t) + O(\epsilon^2)$$

KN, Nishimura, Shimasaki, 1606.07627

finite time(τ)-evolution

- Repeating ε-time evolution nu-times

$$\Phi(t+\nu\epsilon) = \int dxdy\{(1+\epsilon\tilde{L})^{\nu}\mathcal{O}(x+iy)\}P(x,y;t) + O(\nu\epsilon^2)$$

- ε->0 limit $\tau = \nu\epsilon$ fixed

$$\Phi(t+\tau) = \sum_{n=0}^{\infty} \frac{1}{n!}\tau^n \int dx\,dy\left\{\tilde{L}^n\,\mathcal{O}(z)\right\}P(x,y;t)$$

- the probability should fall off faster than any power of drift at large v

- the relation to the original path-integral is shown in terms of induction [see sec 2.3 in 1606.07627]

KN, Nishimura, Shimasaki, 1606.07627

2. Applications to Gauge theories ~ framework

CLM for gauge theory with S ∈ C

- Complex Langevin method
  - complexify link variables $U_{n,\mu} \in \mathrm{SU(N)} \to \mathcal{U}_{n,\mu} \in \mathrm{SL}(N,\mathbb{C})$
  - action and observables $S(U) \to S(\mathcal{U})$
  - Langevin equation for link variables

$$\mathcal{U}_{n,\mu}(t+\epsilon) = e^{iX}\mathcal{U}_{n,\mu}(t),$$

t : Langevin time

$$X = \sum_a \lambda_a [(\mathcal{D}_{an\mu}S)\epsilon + \sqrt{2\epsilon}\eta_{an\mu}]$$

generators: e.g. Gell-Mann matrices for SU(3)

drift terms

Gaussian white noise(we consider only real noise)

Gauge cooling to stabilize CL simulations

[Seiler, Sexty, Stamatescu ('12)]

- complexified gauge transformation after each Langevin step

$$\mathcal{U}_{n,\mu} \to g_n \mathcal{U}_{n,\mu} g_{n+\hat{\mu}}^{-1} \qquad g_n \in \mathrm{SL}(N, \mathbb{C})$$

- chose g_n so that it suppresses a norm
  - unitarity norm (we concentrate on u. norm in this work)

$$\mathcal{N}_{\mathrm{u}} = \frac{1}{N_V} \sum_{n,\mu} \mathrm{tr}[\mathcal{U}_{n,\mu}^{\dagger} \mathcal{U}_{n,\mu} + (\mathcal{U}_{n,\mu}^{\dagger})^{-1} \mathcal{U}_{n,\mu}^{-1} - 2]$$

H.c. is taken after complexification

  - Justification of the complex Langevin method with the gauge cooling procedure [KN, Nishimura, Shimasaki, 1508.02377]

The probability of the drift terms

- We introduce the probability of drift terms p(u)

$$p(u;t) = \int \mathcal{D}\mathcal{U} \sum_{n,\mu} \delta\left(u - u_{n,\mu}(\mathcal{U})\right) P(\mathcal{U};t)$$

where

$$u_{n,\mu} = \sqrt{\frac{1}{N_c^2 - 1} \sum_a v_{an\mu} v_{an\mu}^{\dagger}}$$

$$v_{an\mu} = \mathcal{D}_{an\mu} S, \left(a = 1, 2, \cdots N_c^2 - 1\right)$$

Just drift terms
No additional cost !

If the probability p(u;t) falls off exponentially or faster, then the equality holds between the Langevin time average of an observable and its physical expectation value.

3. Application to 2d SU(2) YM theory

2d SU(2) Yang-Mills theory

- Framework

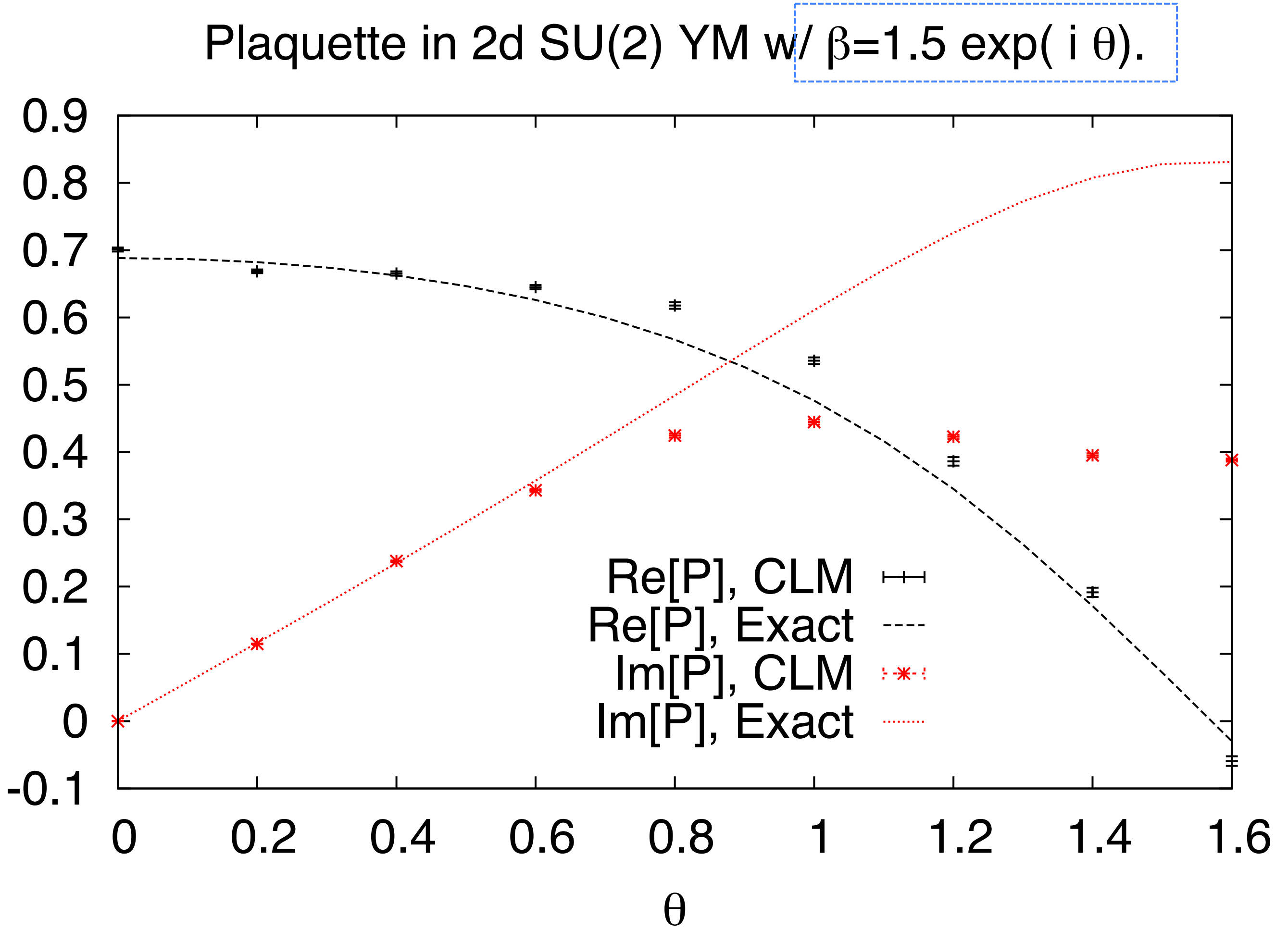
$$Z = \int \mathcal{D}U e^{-S}$$

$$S = -\frac{\beta}{2N} \sum_x \mathrm{tr}[U_{01}(x) + U_{01}^{-1}(x)]$$

  - exactly solvable using character expansion
  - sign problem for complex β
  - CLM works for small Im[β], while it fails for large Im[β] [Makino et al.('15)]

Plaquette

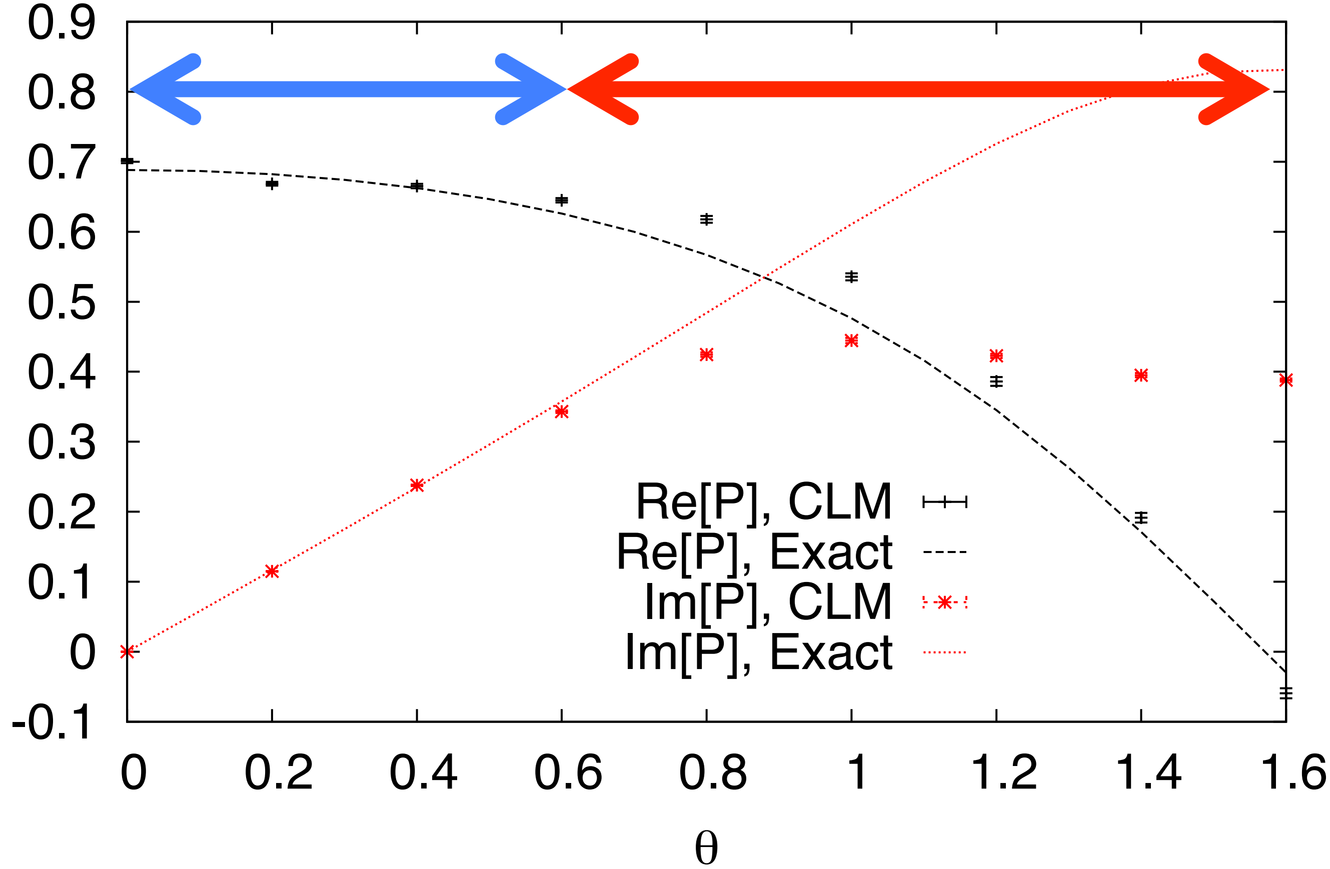
Setup
- lattice size 4^4
- $\varepsilon = 10^{-5}$, t=100
- one g.c. with unitarity norm



Plaquette

Plaquette in 2d SU(2) YM w/ β=1.5 exp(i θ).

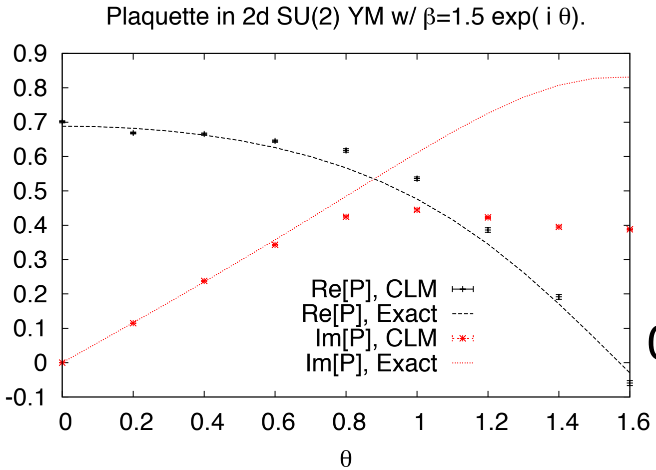
Setup
- lattice size 4^4
- $\varepsilon = 10^{-5}$, t=100
- one g.c. with unitarity norm

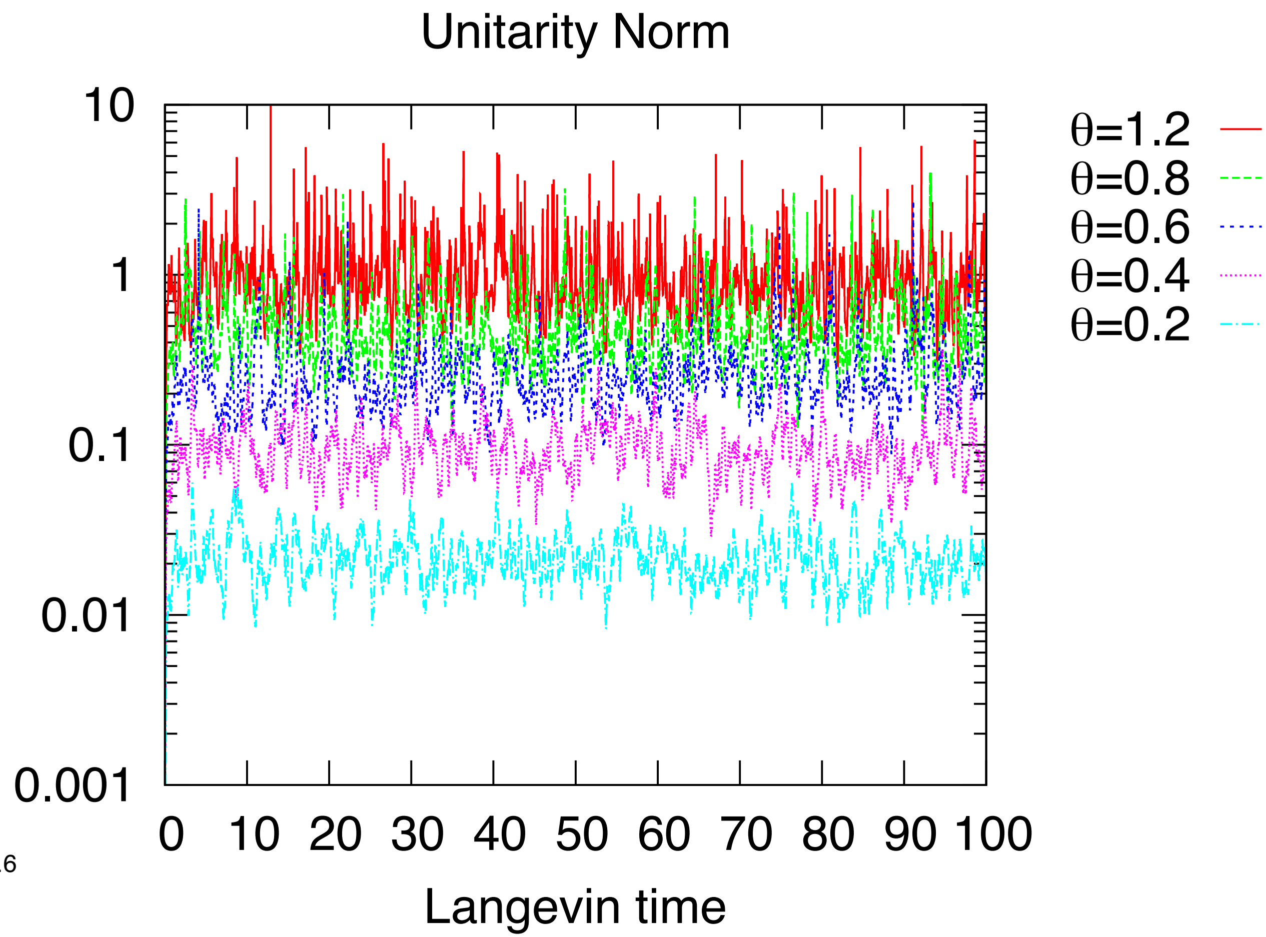


- $\theta \lesssim 0.6$: agreement with exact result
- $\theta \gtrsim 0.6$: deviation from exact result

[Makino et al ('15)]

Unitarity norm

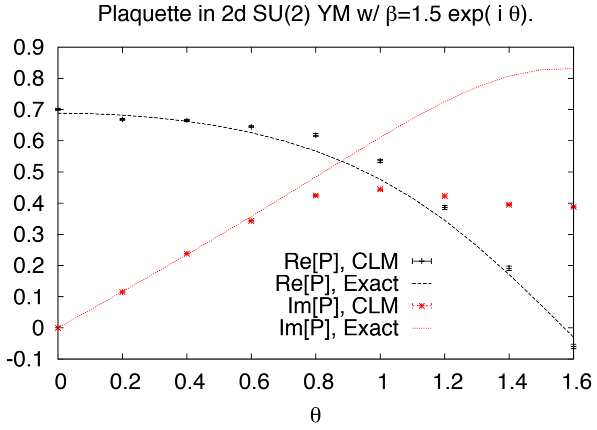


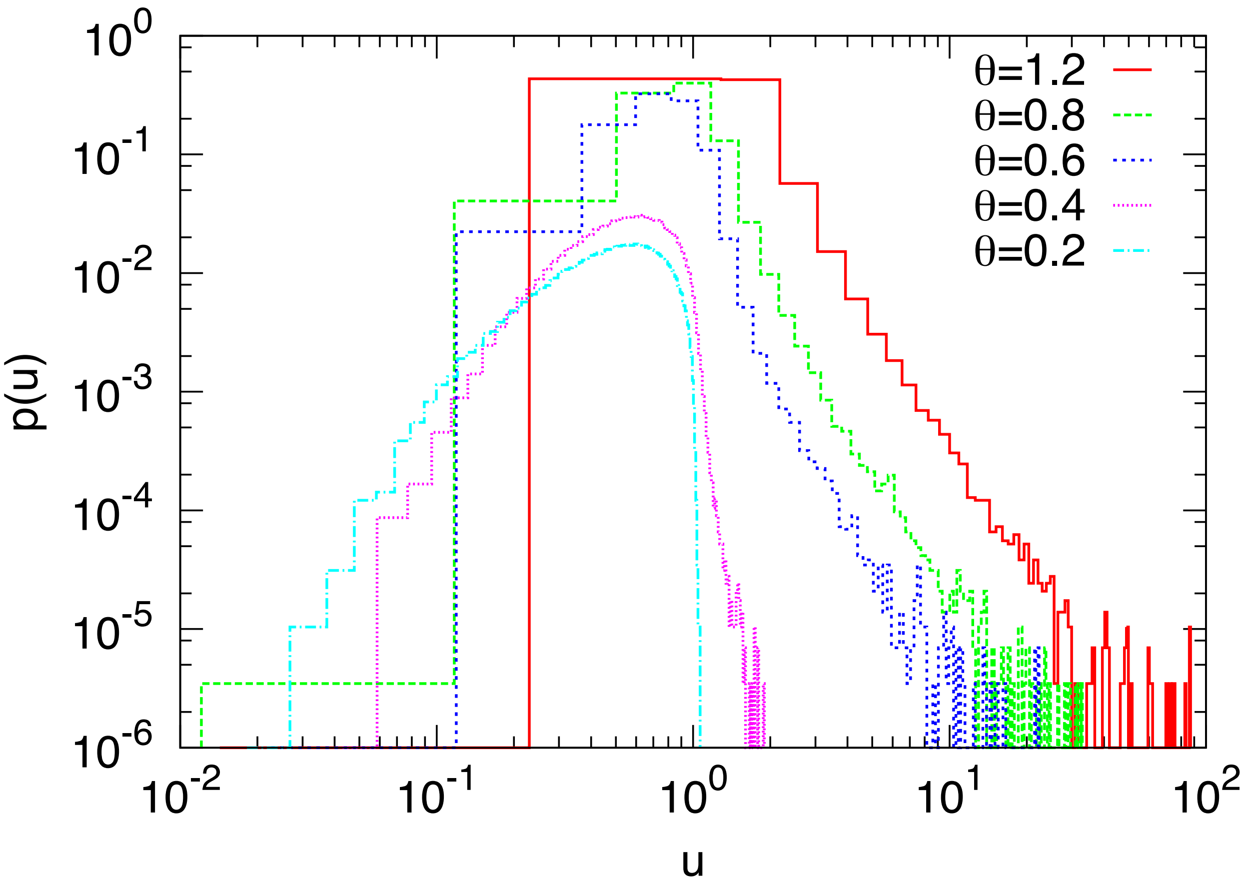


- CLM can fail even if the unitarity norm is under control

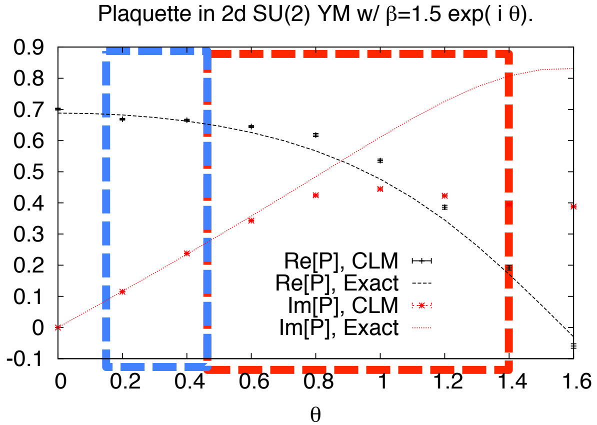
[Makino et al ('15)]

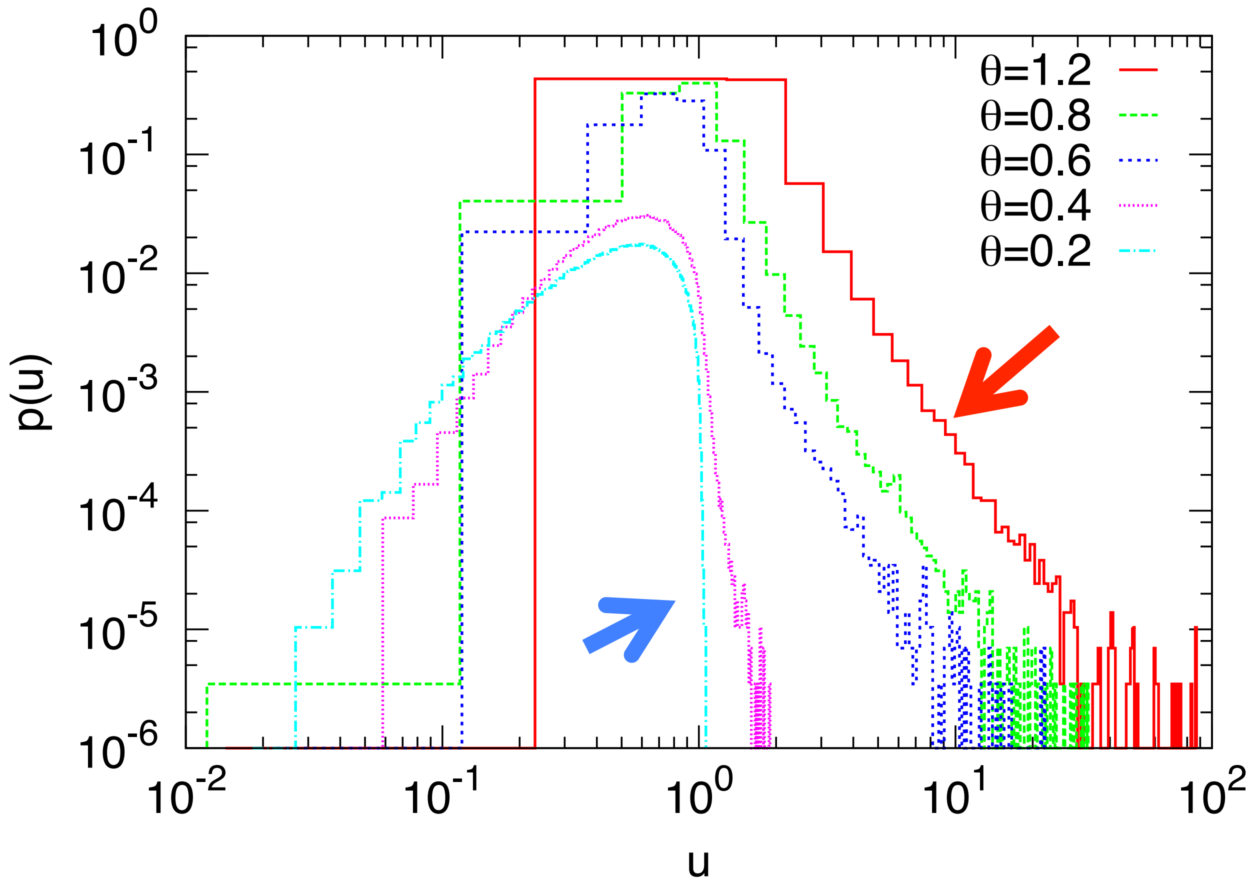
Probability of drift terms





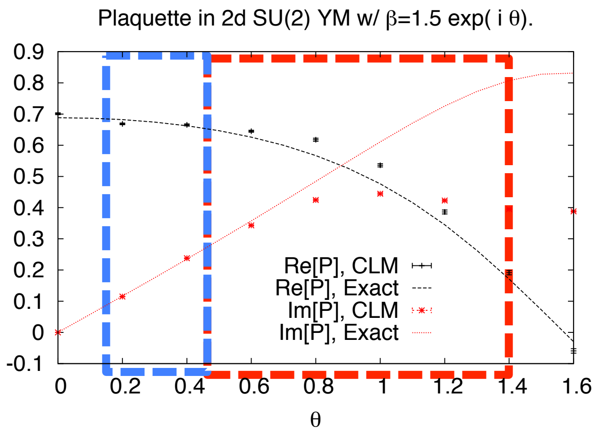
Probability of drift terms

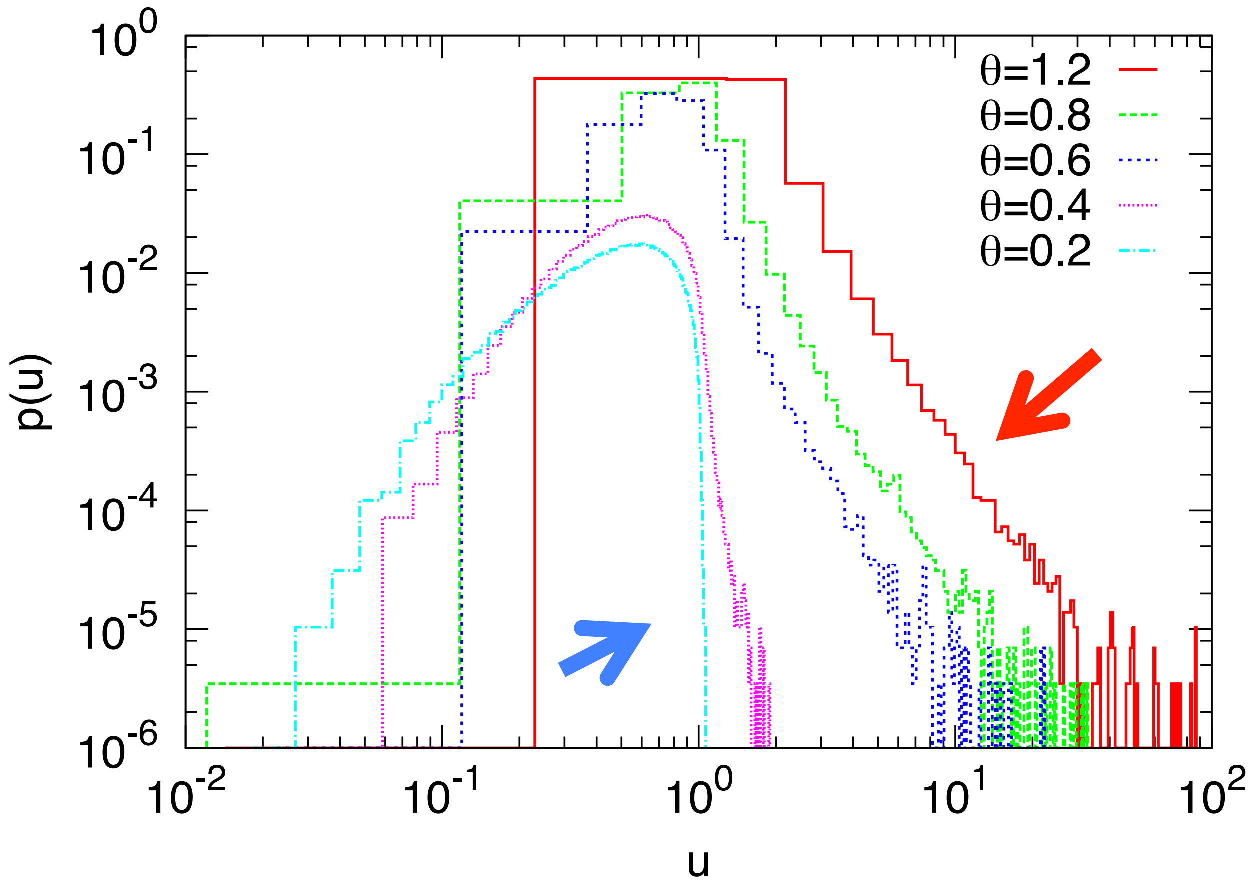




- θ = 0.2, 0.4 : CLM successful : p(u) falls off exponentially or faster
- θ ≳ 0.6 : CLM fails : power law

Probability of drift terms





Correctness of the results can be distinguished by the probability of drift terms.

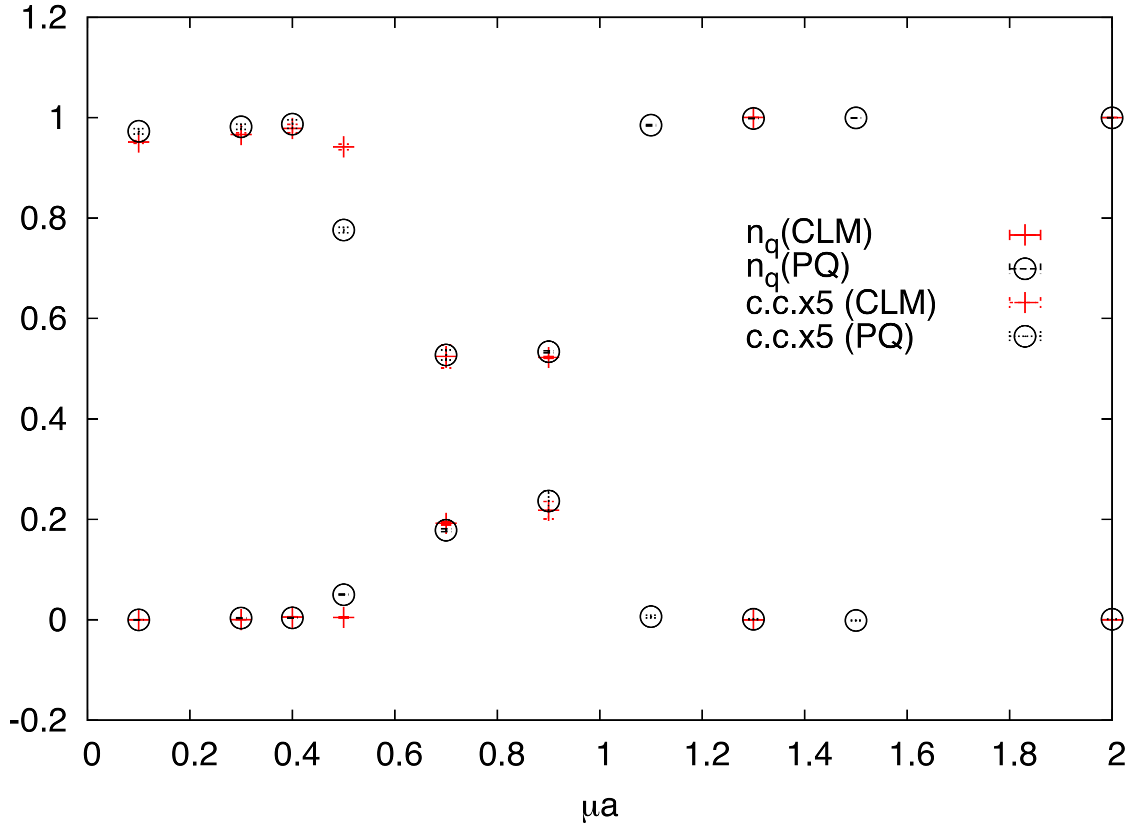
4. Applications to QCD at finite density

Setup

- We consider finite density QCD at low T with light quarks
  - lattice size: 4^3x8
  - Nf=4 staggered fermion with m a=0.05 (we keep mass small to see singular drift problem)
- Langevin setup
  - Langevin time: t = 10~20 with fixed $\varepsilon = 10^{-4}$
  - gauge cooling: 10~20 times
  - we use bilinear noise method with Kogut-Sinclair type improvement [Sinclair's talk Lat'15]

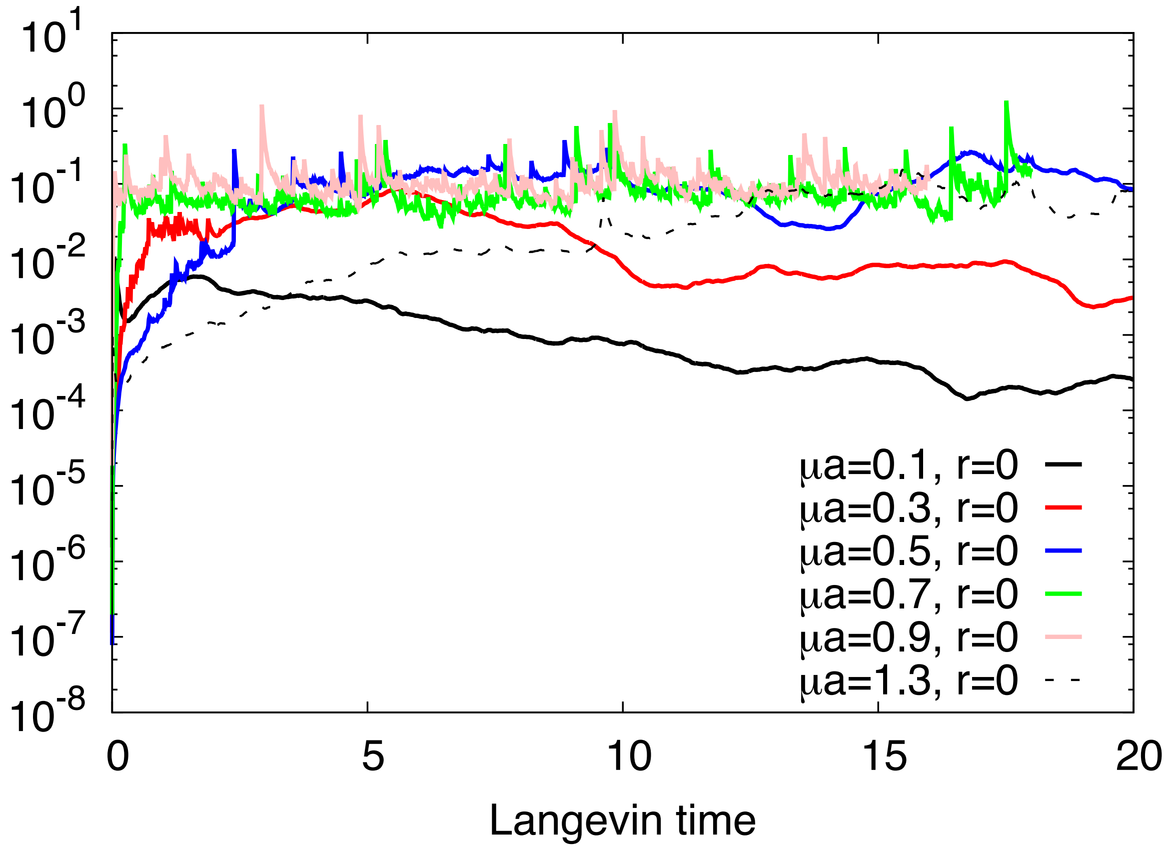
- Results are preliminary: t=20 may not be sufficient to confirm the tail part of the probability of the drift terms.

Quark number & chiral condensate (CLM vs PQ)



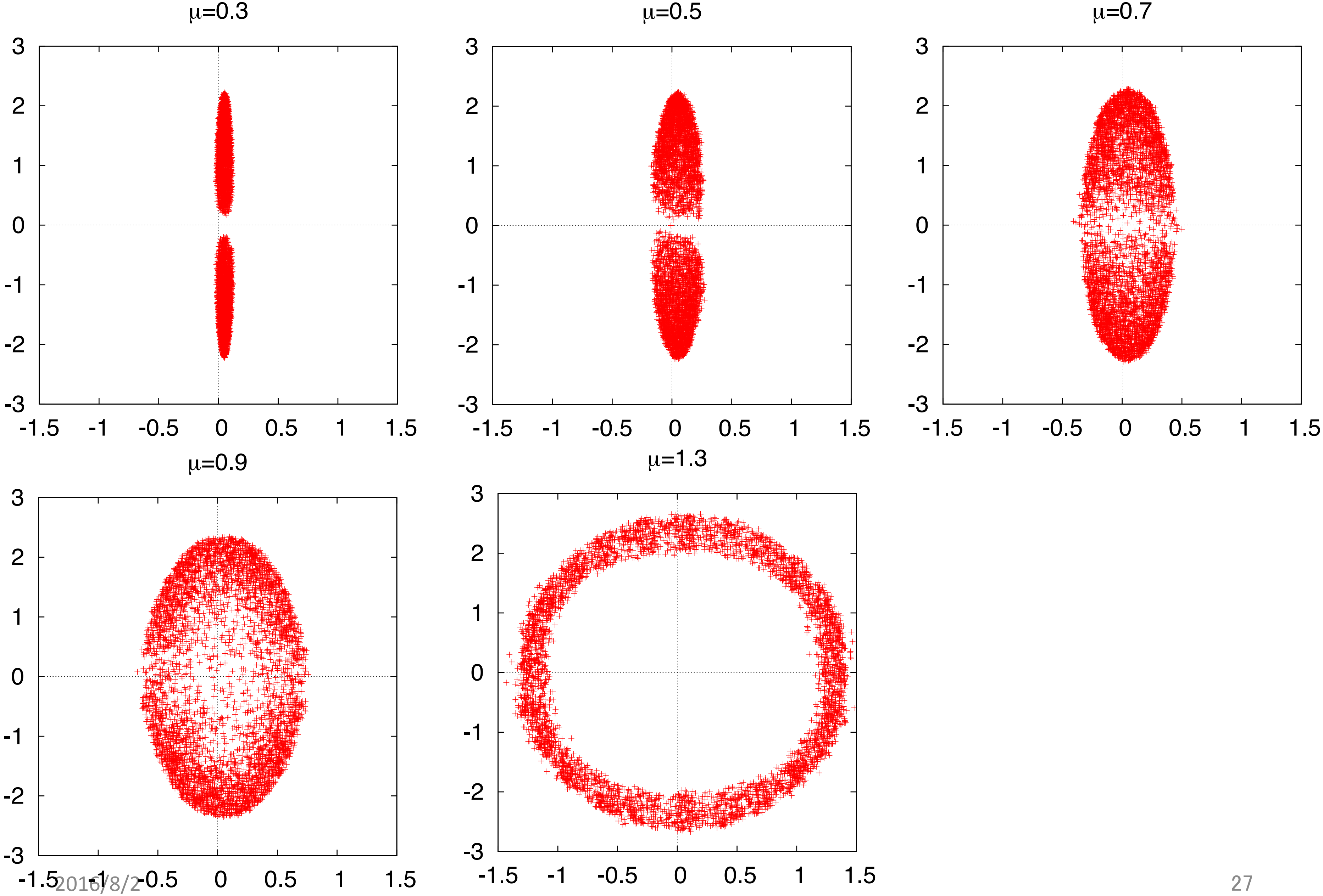
- Results in CLM agree with those in PQ, but deviation found at μ=0.5.
- Are the results reliable ?

Unitarity norm

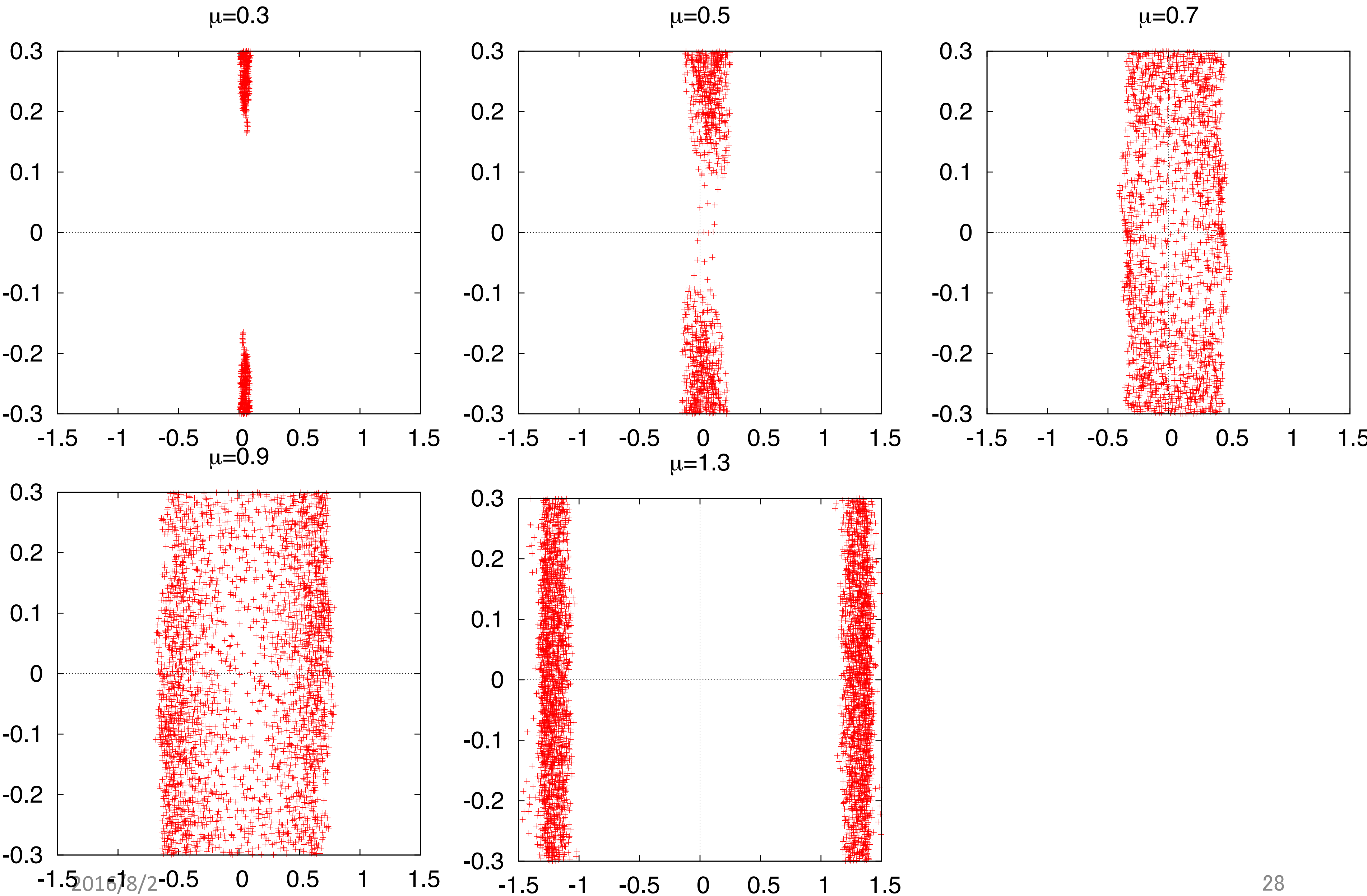


- Unitarity norm is almost under control for all the cases.

Dirac Eigenvalues (ev(D+m))

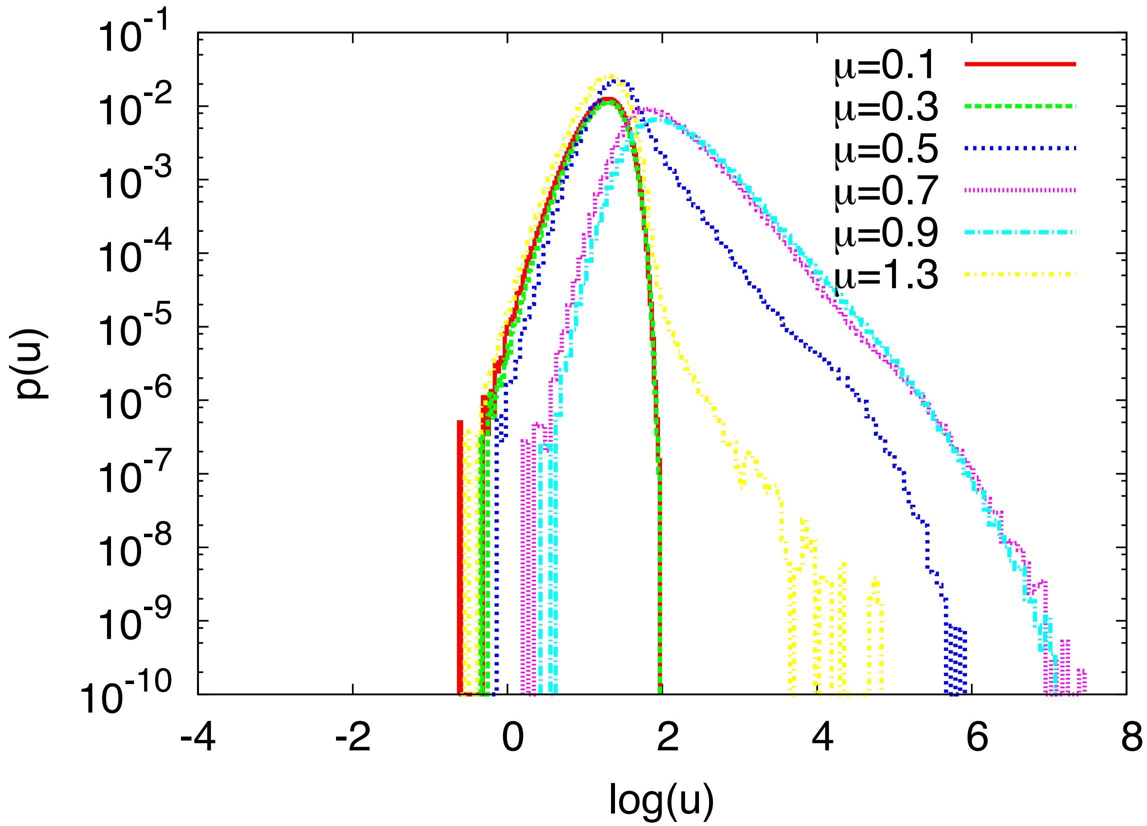


Dirac Eigenvalues (near the origin)



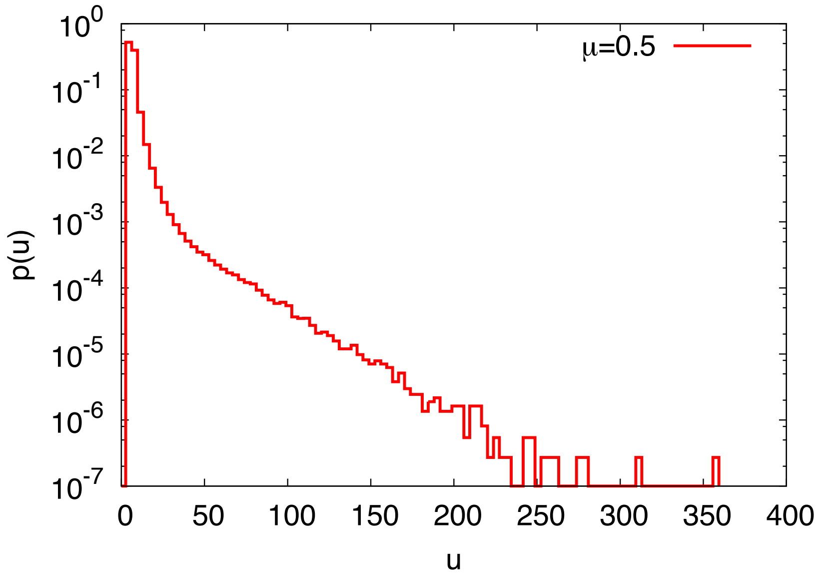
It is difficult to judge correctness of the results from the unitarity norm and Dirac evs.

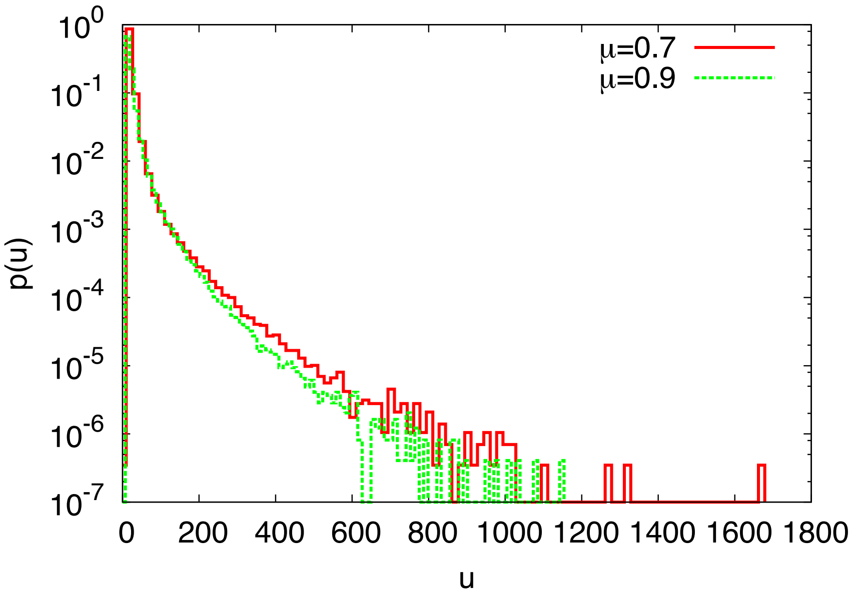
Probability of drift terms

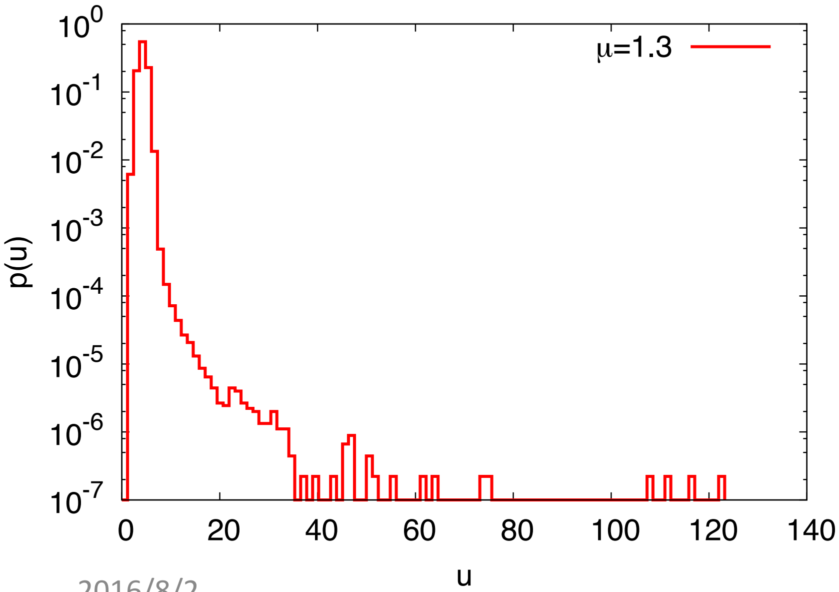


- μ≦0.3 : fall-off exponentially or faster => reliable
- μ = 0.5 : fall-off exponentially => reliable
- 0.7 ≦ μ : power law

Data in semi-log plot

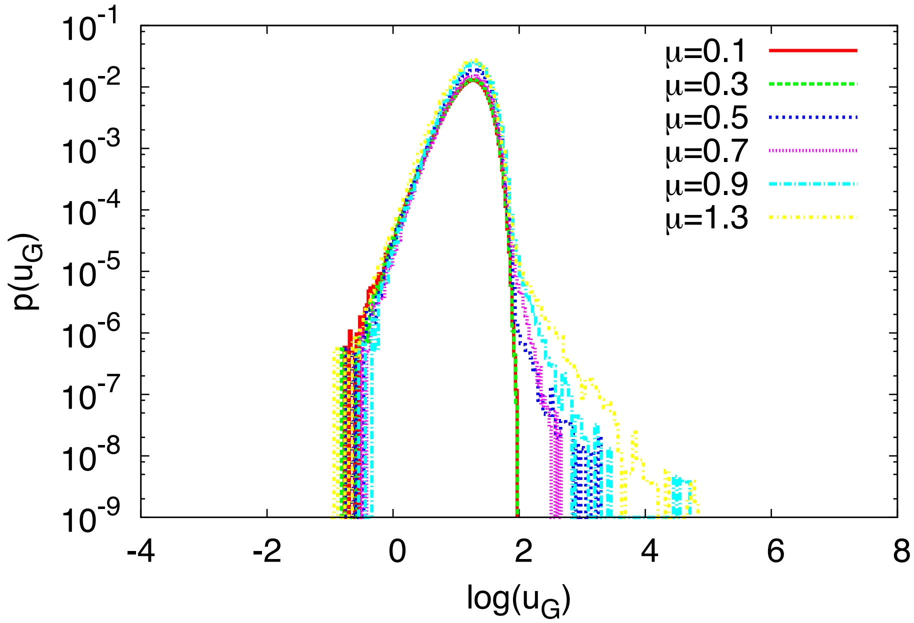




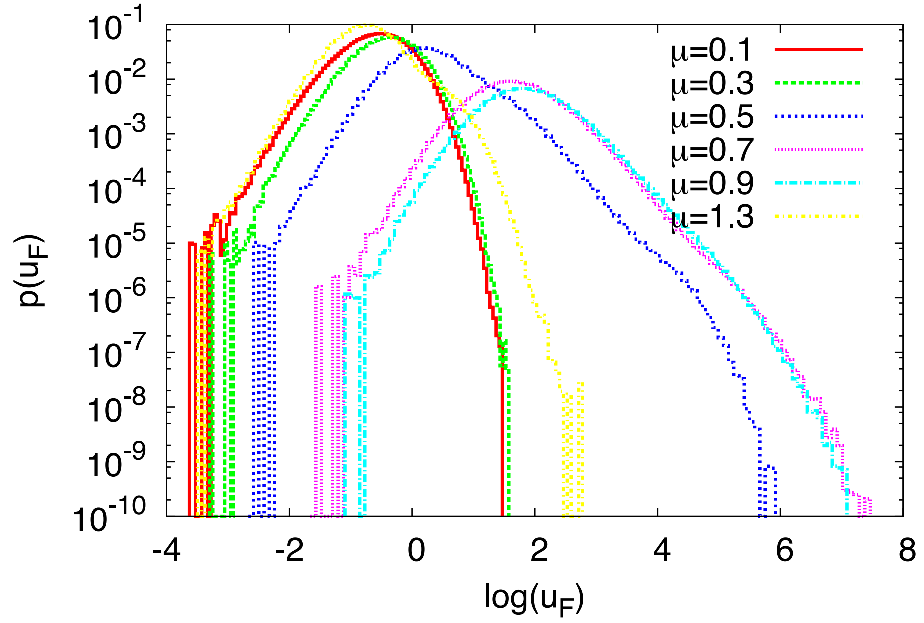


Probability of drift: (L) gauge part, (R) fermion part

$v = \mathcal{D}S_G$



$v = \mathcal{D}S_F$



- μ = 0.7, 0.9 => singular drift problem
- μ = 1.3 => excursion problem
- From the criterion, we can identify the origin of problems

Conclusion

- We proposed a new criterion of correctness of results in CLM by revisiting the argument for justification.
  - The criterion works well for 2d SU(2) (and also cRMT [Shimasaki'talk]).

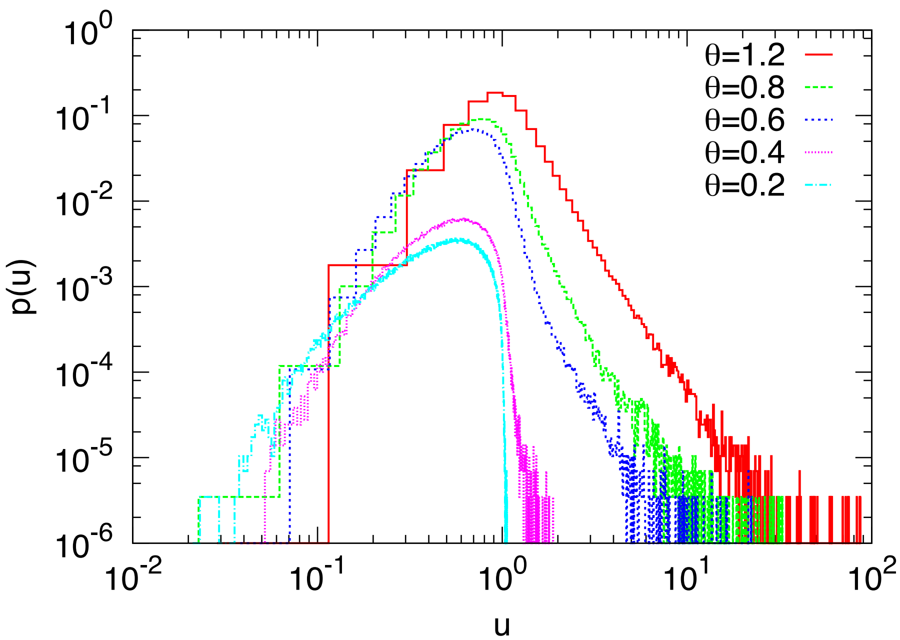
- We study QCD at low T with light quakrs using CLM, we determine the reliable range of chemical potential.
  - singular drift problem occurs at intermediate values of μ.

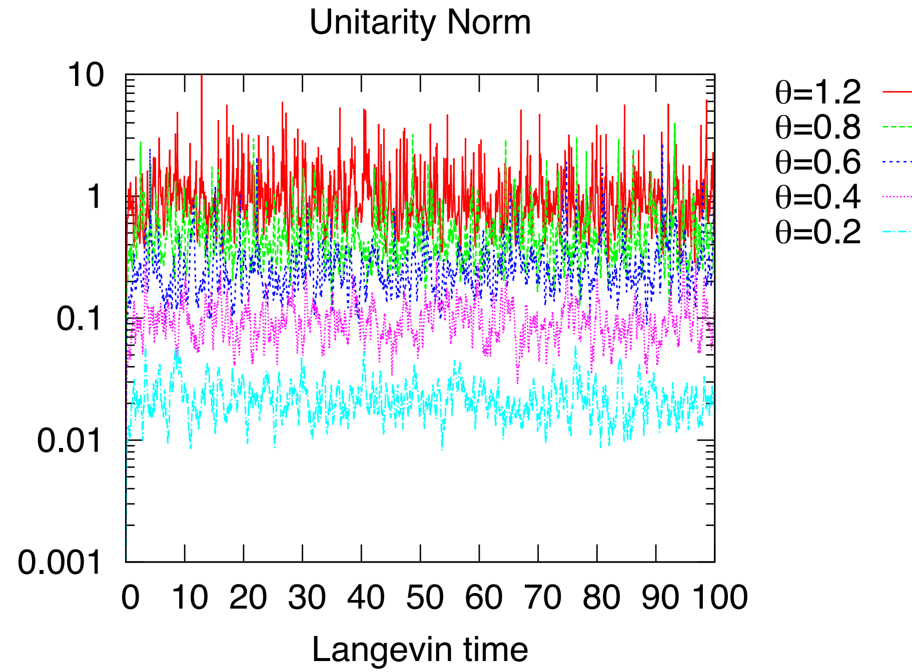
- [work in progress]
  - application of new norms with Dirac operators to avoid the singular drift problem.

Conclusion

- ✓ **The probability of the drift terms is a reliable criterion of correctness**
- ✓ **It is a legitimate, cheap and clear way of judging correctness.**

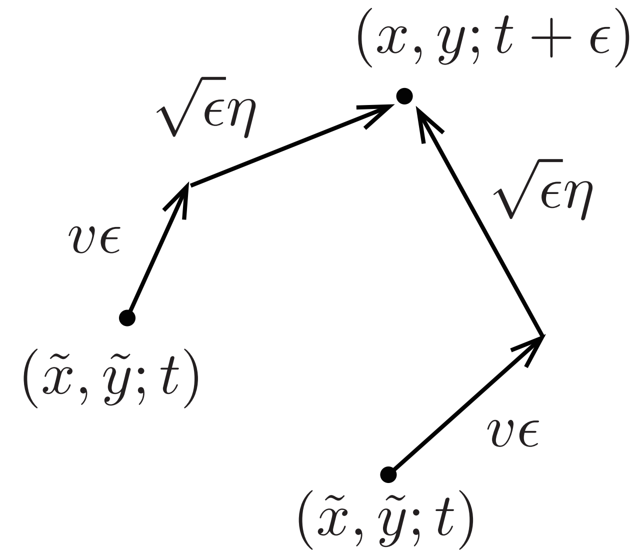
Remark : large drift and spikes in norms





- Large drift is correlated with spikes of unitarity norm.
- If the unitarity norm has spilkes frequently, results may be wrong.
  => The probability of drift tell you if the result is reliable !
- Langevin time should be sufficiently larger than the auto-correlation of norms to use the criterion.

stochastic process with ε



- probability at t+ε

$$P(x,y;t+\epsilon) = \frac{1}{\mathcal{N}} \int d\eta \, e^{-\frac{1}{4}\{\frac{1}{N_R}\eta_k^{(R)2}+\frac{1}{N_I}\eta_k^{(I)2}\}} \int d\tilde{x} d\tilde{y}$$

$$\times \delta\Big(x-\tilde{x}-\epsilon \mathrm{Re} v(\tilde{z}) - \sqrt{\epsilon}\eta^{(R)}\Big) \delta\Big(y-\tilde{y}-\epsilon \mathrm{Im} v(\tilde{z}) - \sqrt{\epsilon}\eta^{(I)}\Big) P(\tilde{x},\tilde{y};t)$$

$$= \frac{1}{\epsilon\mathcal{N}} \int d\tilde{x} d\tilde{y} \exp\left[-\left\{ \frac{\Big(x-\tilde{x}-\epsilon \mathrm{Re} v(\tilde{z})\Big)^2}{4\epsilon N_R} + \frac{\Big(y-\tilde{y}-\epsilon \mathrm{Im} v(\tilde{z})\Big)^2}{4\epsilon N_I} \right\} \right]$$

$$\times P(\tilde{x},\tilde{y};t)$$

ε-evolution (suppl.)

- Taking into accout the Gaussian factor as a part of observable

$$\Phi(t+\epsilon) = \int dx\, dy\, \mathcal{O}_\epsilon(x+iy)\, P(x,y;t)$$

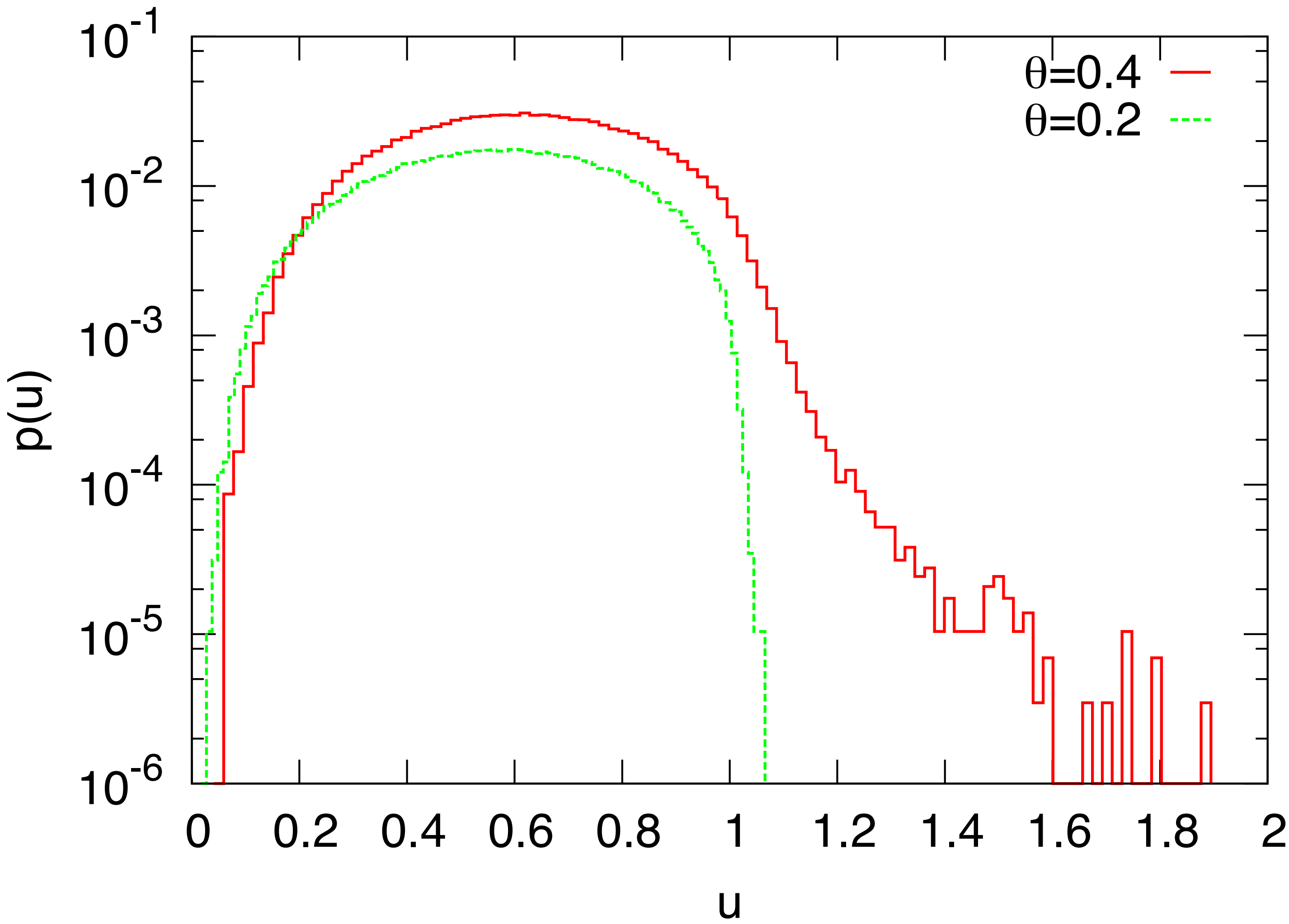
$$\mathcal{O}_\epsilon(z) = \frac{1}{\mathcal{N}} \int d\eta\, e^{-\frac{1}{4}\left\{\frac{1}{N_{\mathrm{R}}}\eta_k^{(\mathrm{R})2} + \frac{1}{N_{\mathrm{I}}}\eta_k^{(\mathrm{I})2}\right\}} O\Big(z + \epsilon\, v(z) + \sqrt{\epsilon}\,\eta\Big) \ .$$

- ε-expansion (for holomorphic O(z))

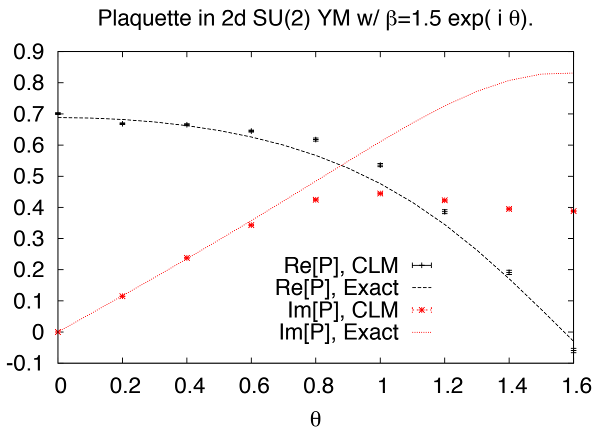
$$\mathcal{O}_\epsilon(z) = \sum_{n=0}^{\infty} \frac{1}{n!}\epsilon^n : \tilde{L}^n : \mathcal{O}(z)$$

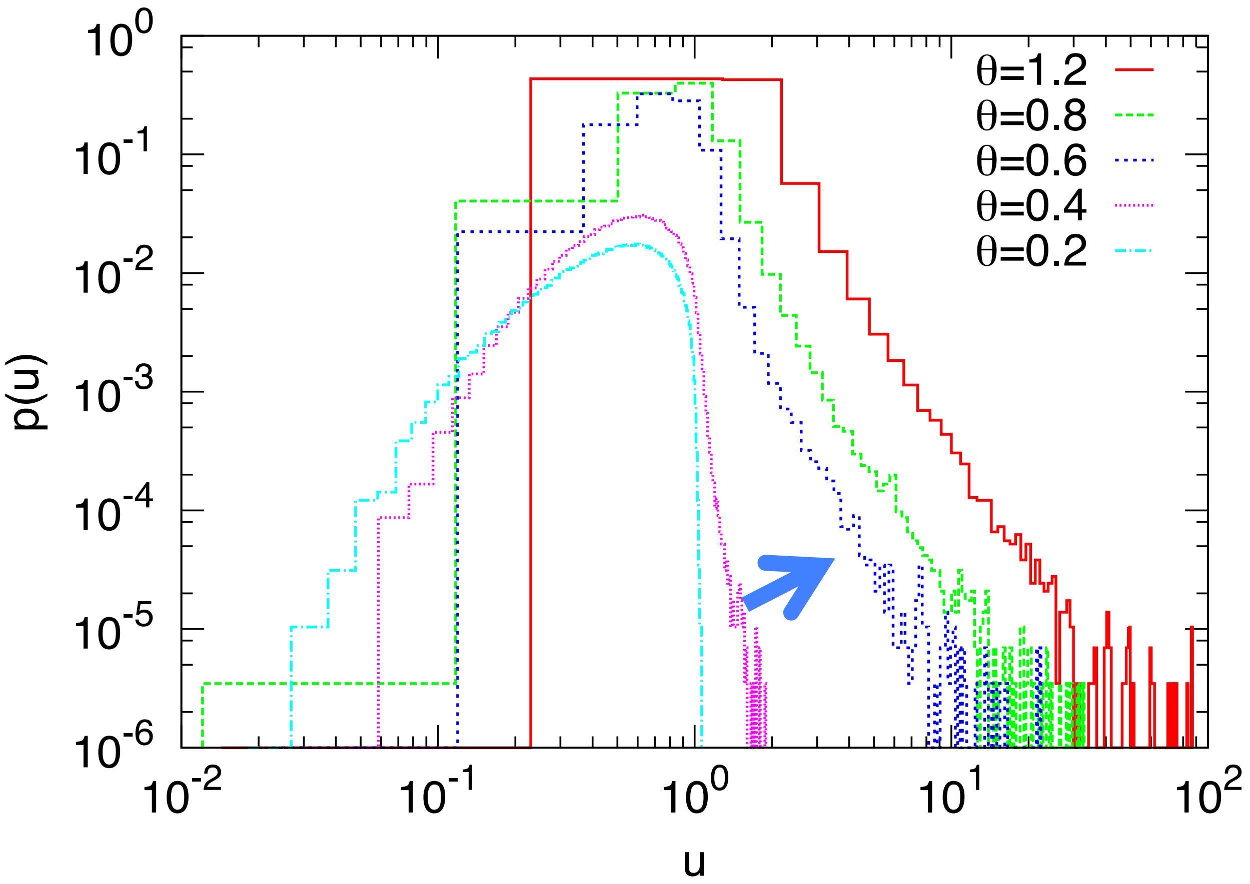
Back up for 2d SU(2)

Probability of drift –semi-log plot



intermediate value of θ

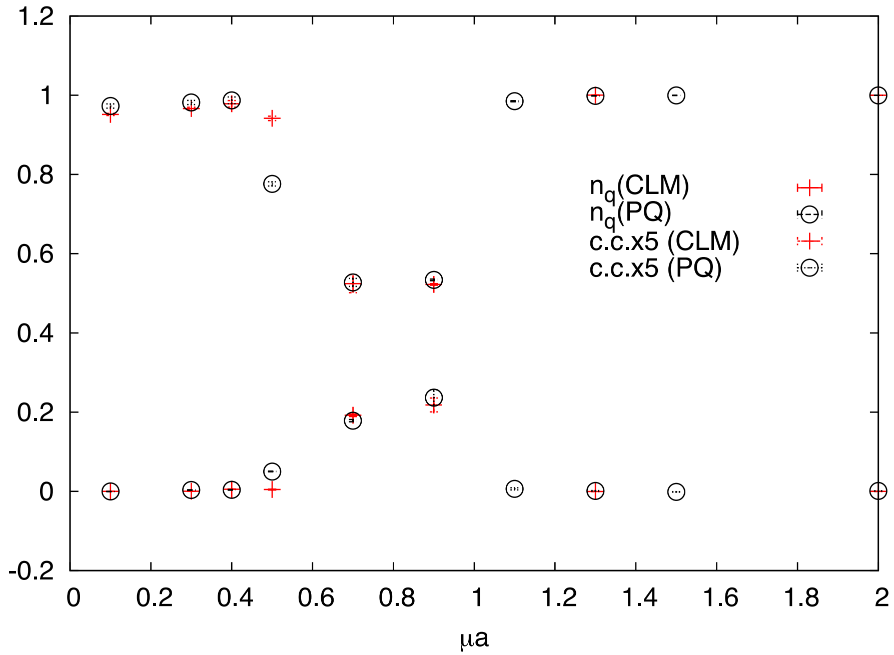


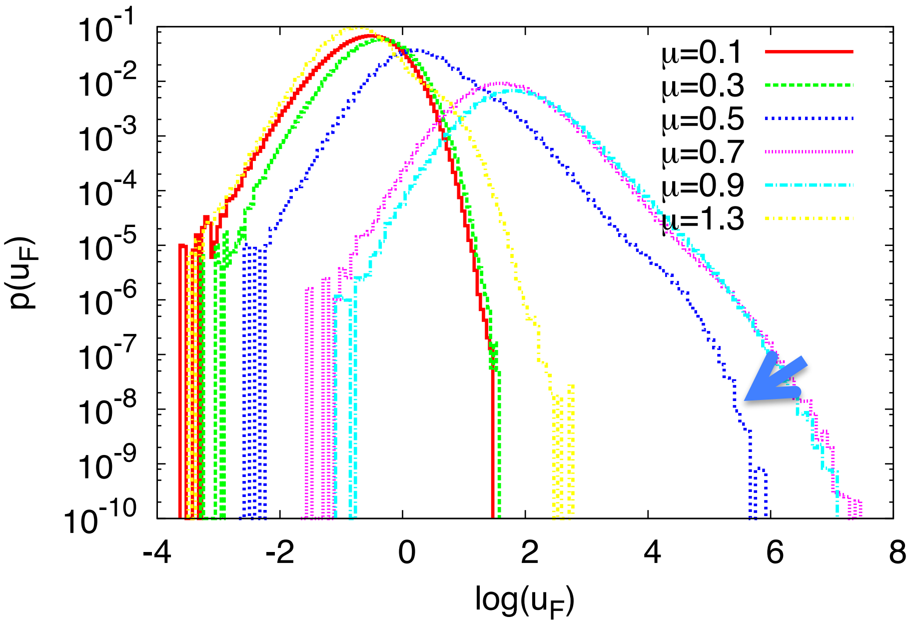


- Our criterion: CLM is correct for the fall off exponentially or faster.
- It is possible that the CLM is correct even if it has power law.

Back up slide
data at μ=0.5

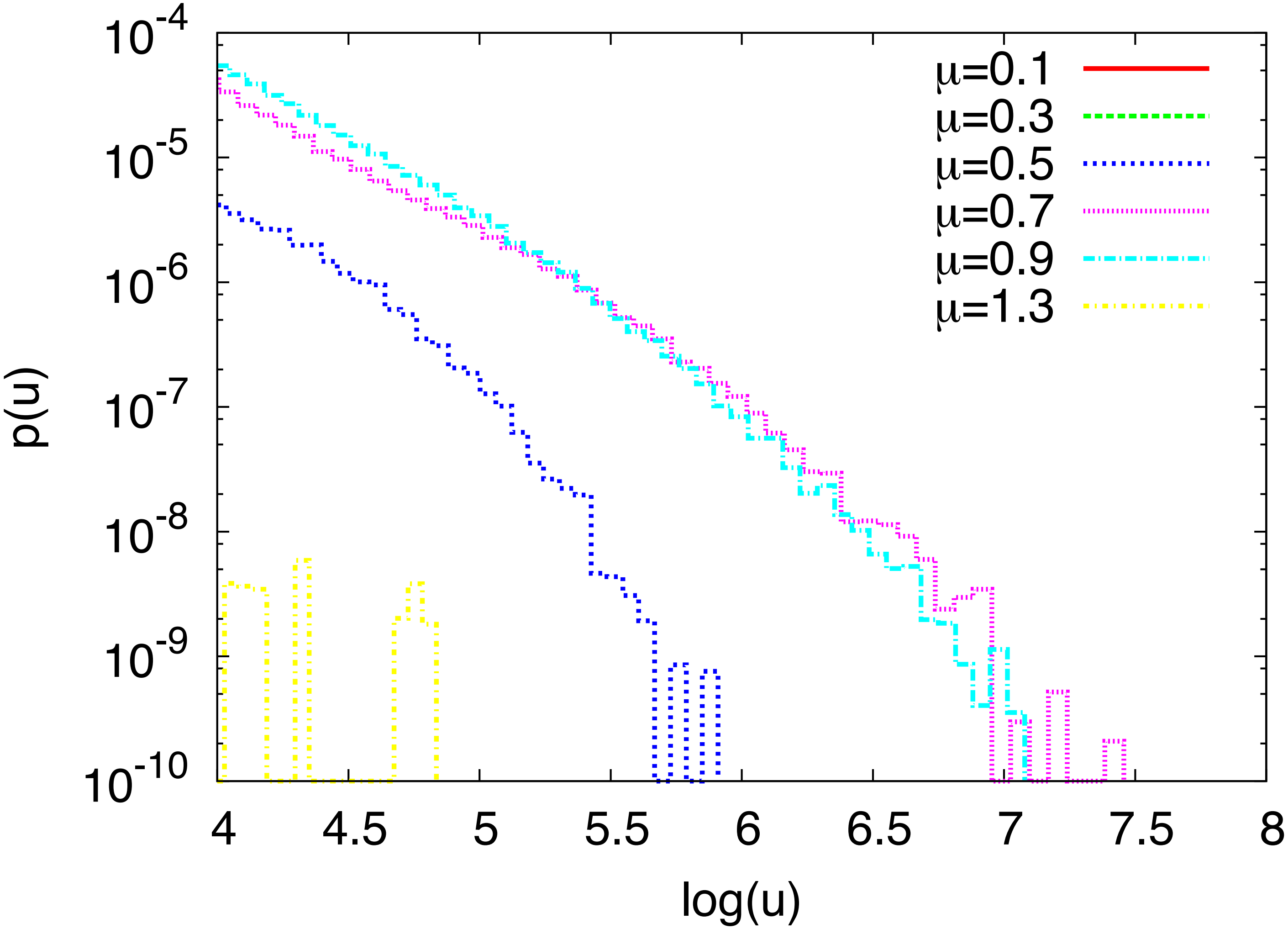
Distribution



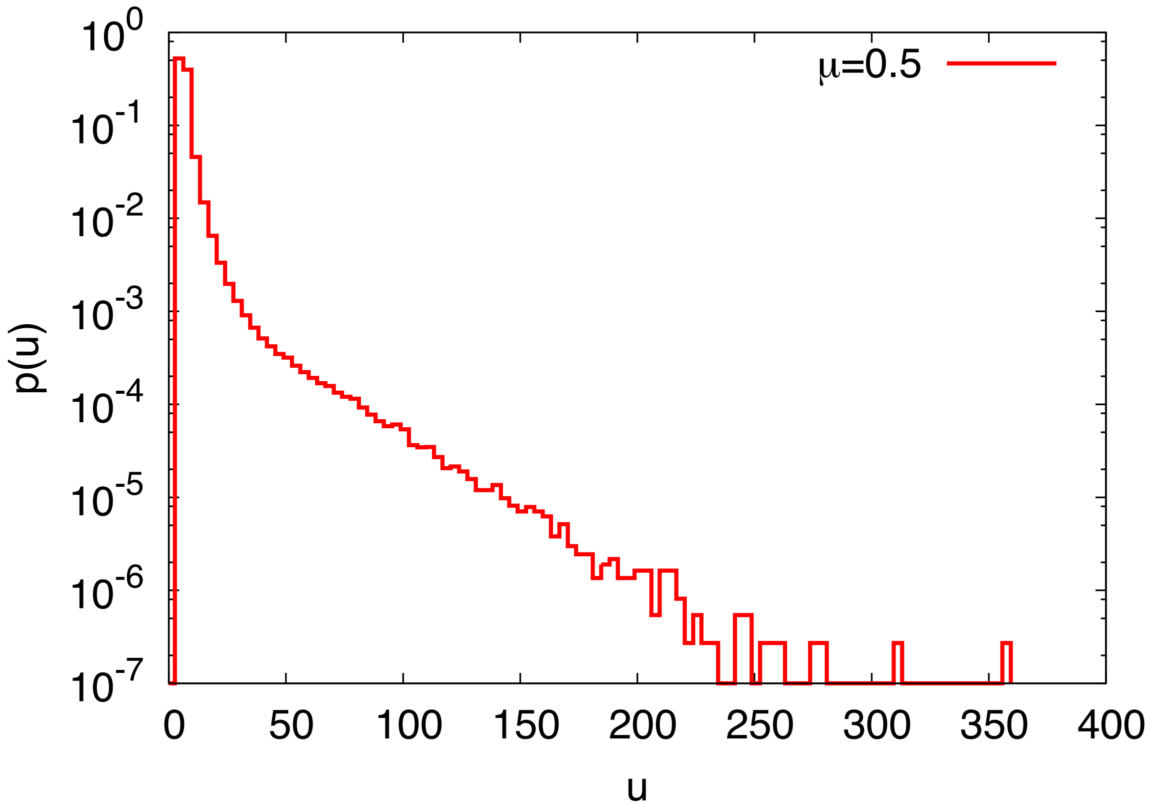


- The new criterion suggests the data at μ = 0.5 is reliable.
  - At μ=0.5: the probability falls off faster than power law.

Tail part of the distribution

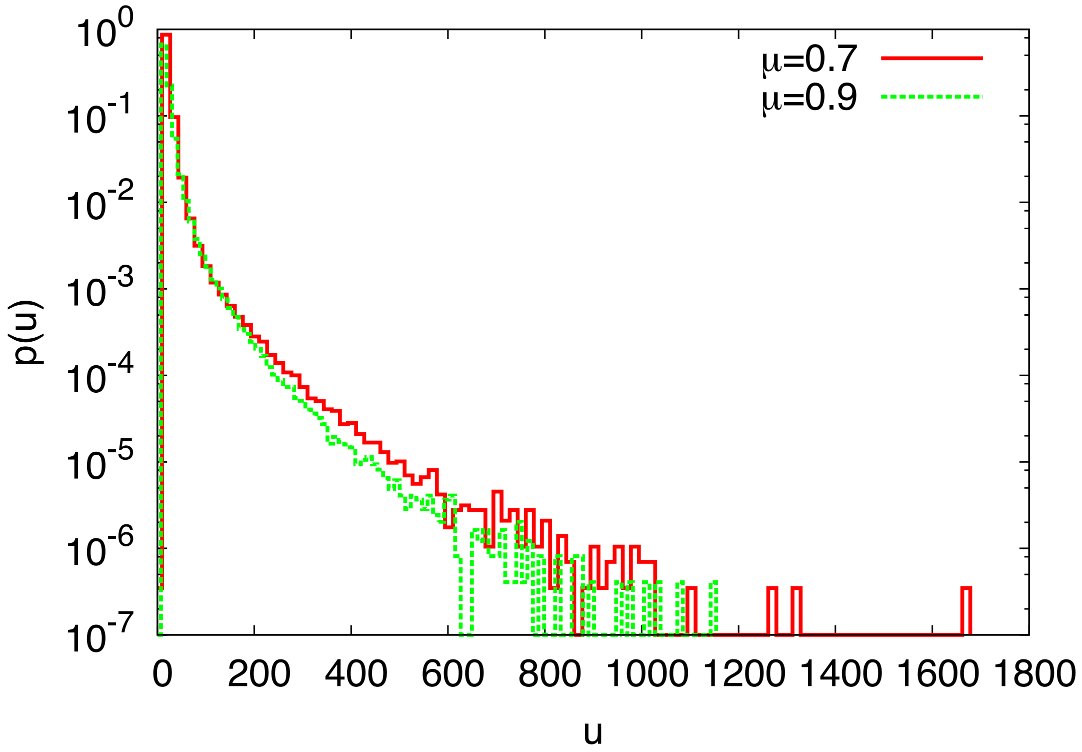


Data at μ = 0.5 in semi-log plot



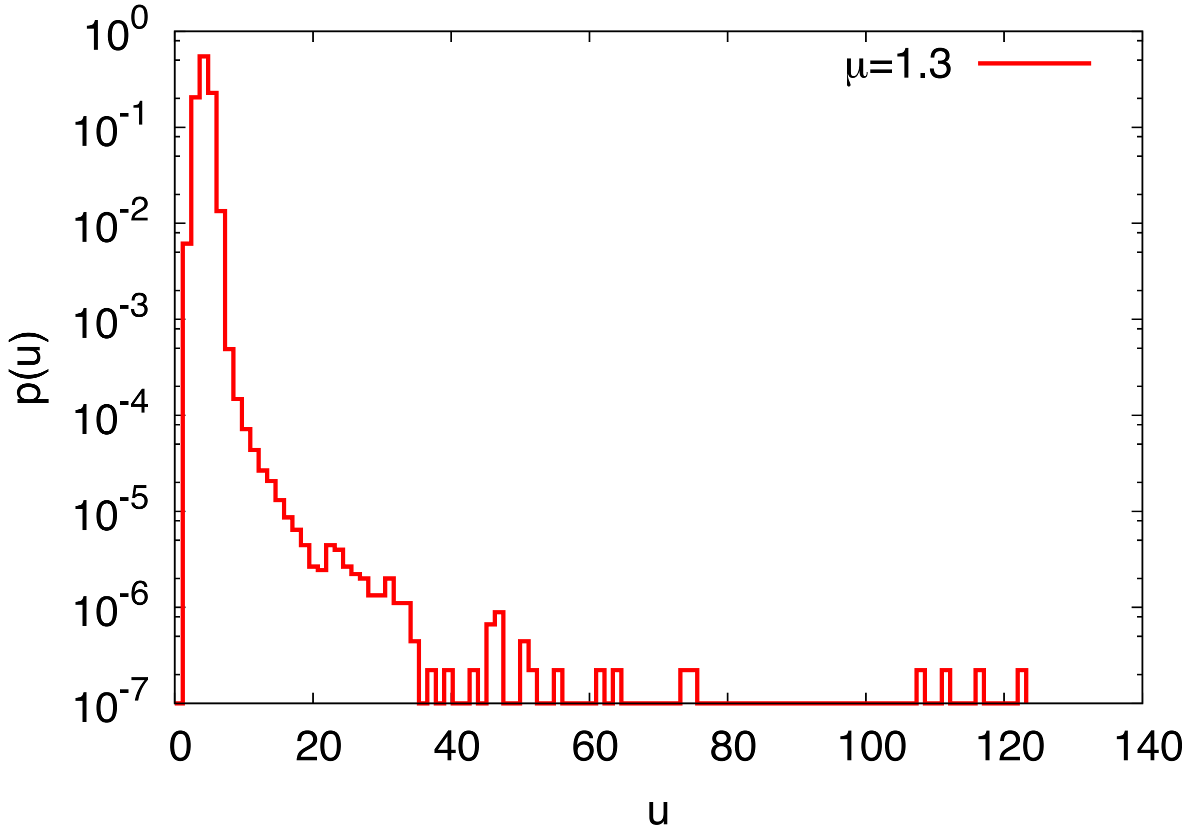
- The data in semi-log plot seems to fit with linear
- The probability falls off exponentially.

Data at μ = 0.7 and 0.9 in semi-log plot



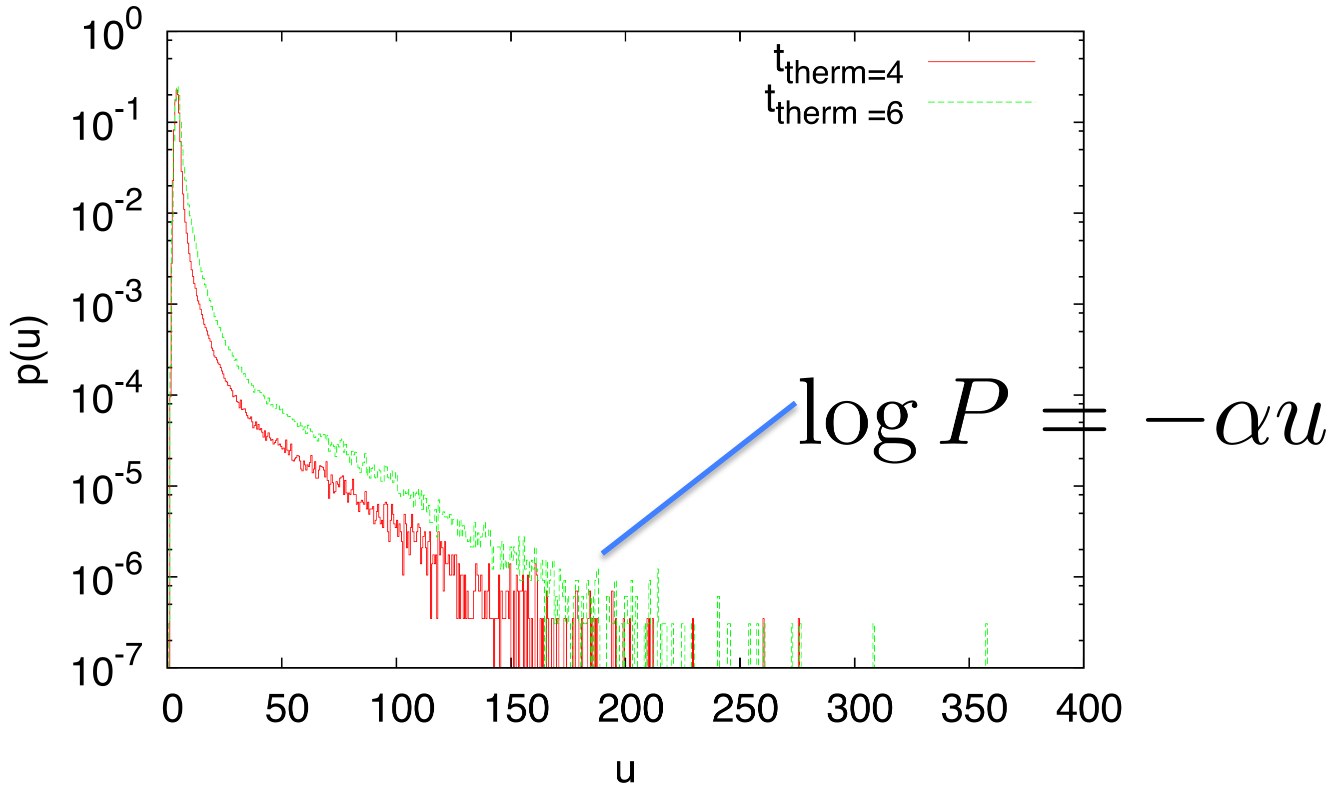
- The data in semi-log plot seems to fall off slower than linear.
- The probability falls off slower than exponentially.

Data at μ = 1.3 in semi-log plot



- The data in semi-log plot seems to fall off slower than linear.
- The probability falls off slower than exponentially.

T_{therm} dep-Data at μ = 0.5 in semi-log-plot



- Data for μ = 0.5 with t_{therm} = 4 and 6
- The plot implies the exponential damp.
  - the data is reliable.

Unitariry norm and anti-hermiticity

- We found the similarity between unitarity norm and anti-hermiticity norm

$$N_{\rm u} \equiv \frac{1}{4N_V}\sum_{x,\nu} {\rm tr}[(\mathcal{U}_{x\nu})^\dagger \mathcal{U}_{x\nu} + (\mathcal{U}_{x\nu}^{-1})^\dagger \mathcal{U}_{x\nu}^{-1} - 2].$$

$$\mathcal{N}_{\rm a.h.} = \frac{1}{4N_V}{\rm tr}\,(D+D^\dagger)(D+D^\dagger)^\dagger$$

$$= \frac{2}{4N_V}\sum_{x,\sigma}{\rm tr}_{\rm c}\left[e^{2\mu a\delta_{4,\sigma}}U_\sigma^\dagger(x)U_\sigma(x) + e^{-2\mu a\delta_{4,\sigma}}\left(U_\sigma^{-1}(x-\hat\sigma)\right)^\dagger U_\sigma^{-1}(x-\hat\sigma) - 2\mathbb{1}_{3\times 3}\right]$$

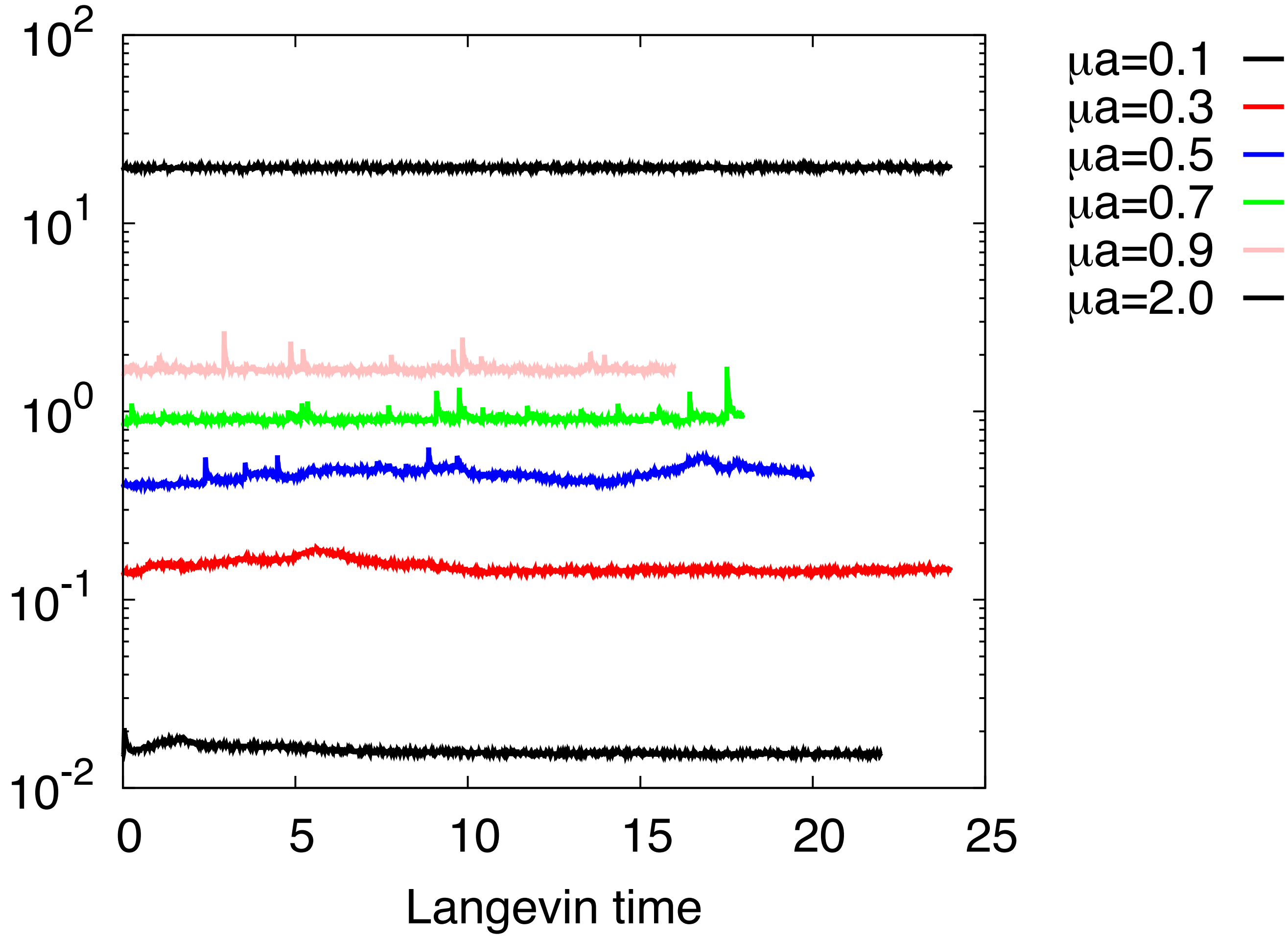
At μ=0, they are equivalent.
Cooling for unitarity norm has an effect to reduce the anti-hermiticity norm.

Cooling for new norm may extend the applicable range further, work in progress.

Back up for 4^3x8

Anti-Hermiticity norm

- 4^3x8



Back up – framework

Sign problem

$$Z = \int \prod_k dx_k \, e^{-S(x)}$$

- Monte Carlo method with importance sampling
  - powerful tool to solve path-integrals non-perturbatively

- Importance sampling breaks down if S is complex
  - QCD at finite density or with theta-term
  - Chern-Simons gauge theories
  - Hubbard model away from half-filling

(real) Langevin Method (LM)

$$Z = \int \prod_k dx_k \, e^{-S(x)}, \; (x_k, S \in \mathbb{R})$$

- Generation of ensemble using Langevin eq. (stochastic quantization)

$$\frac{dx_k^{(\eta)}}{dt} = -\frac{\partial S}{\partial x_k^{(\eta)}} + \eta_k(t)$$

t: parameter(Langevin time)
η: Gaussian white noise

Ensemble is generated by the stochastic equation rather than importance sampling

Parisi-Wu 1981

proof of LM

- Average of an observable in LM is given by

$$\langle O(x^{(\eta)}(t))\rangle_\eta = \int dx\, O(x) P(x;t)$$

$$P(x;t) = \left\langle \prod_k \delta(x_k - x_k^{(\eta)}(t)) \right\rangle_\eta$$

$$\langle \cdots \rangle_\eta = \frac{\int \mathcal{D}\eta \cdots e^{-\frac{1}{4}\int d\tau \eta^2}}{\int \mathcal{D}\eta e^{-\frac{1}{4}\int d\tau \eta^2}}$$

- According to the Fokker-Planck equation, P converges to

$$\lim_{t\to\infty} P(x;t) \propto e^{-S(x)}$$

- Average of the observable converges to

$$\lim_{t\to\infty} \langle O(x(t))\rangle_\eta = \lim_{t\to\infty} \int \prod dx_k\, O(x) P(x;t),$$

$$\propto \int \prod_k dx_k\, O(x) e^{-S(x)}$$

average in LM

physical expectation value

Complex Langevin method(CLM)

- Stochastic quantization is available for complex action
  - LM is free from the probability interpretation of exp(-S)
  - however, complexification is inevitable

[Parisi('83), Klauder('83)]

- CLM
  - extend originally real variables to complex

    $$x \in \mathbb{R} \to z = x + iy \in \mathbb{C}$$

  - extend also action and observables in a holomorphic manner

    $$S(x) \to S(z) = S(x + iy)$$

  - Langevin equation

    $$\frac{\partial z}{\partial t} = -\frac{\partial S}{\partial z} + \eta(t)$$

(noise term can be complex. However, we prefer to use real noise throughout this talk)

Problem of convergence

- In CLM, P(x;t) in equilibrium is not ensured to converge to correct limit

$$\int dxdy\, O(x+iy)P(x,y;t) \overset{?}{=} \int dx\, O(x)\rho(x;t)$$

$$\lim_{t\to\infty} \rho(x;t) = e^{-S(x)}$$

  - CLM works well for some cases, but fails for other.
  - there had been no criteria to distinguish if results in CLM are correct or not.

Justification of CLM / criteria of correctness

- CLM is justified if some conditions are satisfied
  [Aarts, et. al. PRD81, 054508('10), EPJC71,1756('11)].

$$\int dxdy\, O(x+iy)P(x,y;t) = \int dx\, O(x)\rho(x;t)$$

$$\lim_{t\to\infty} \rho(x;t) = e^{-S(x)}$$

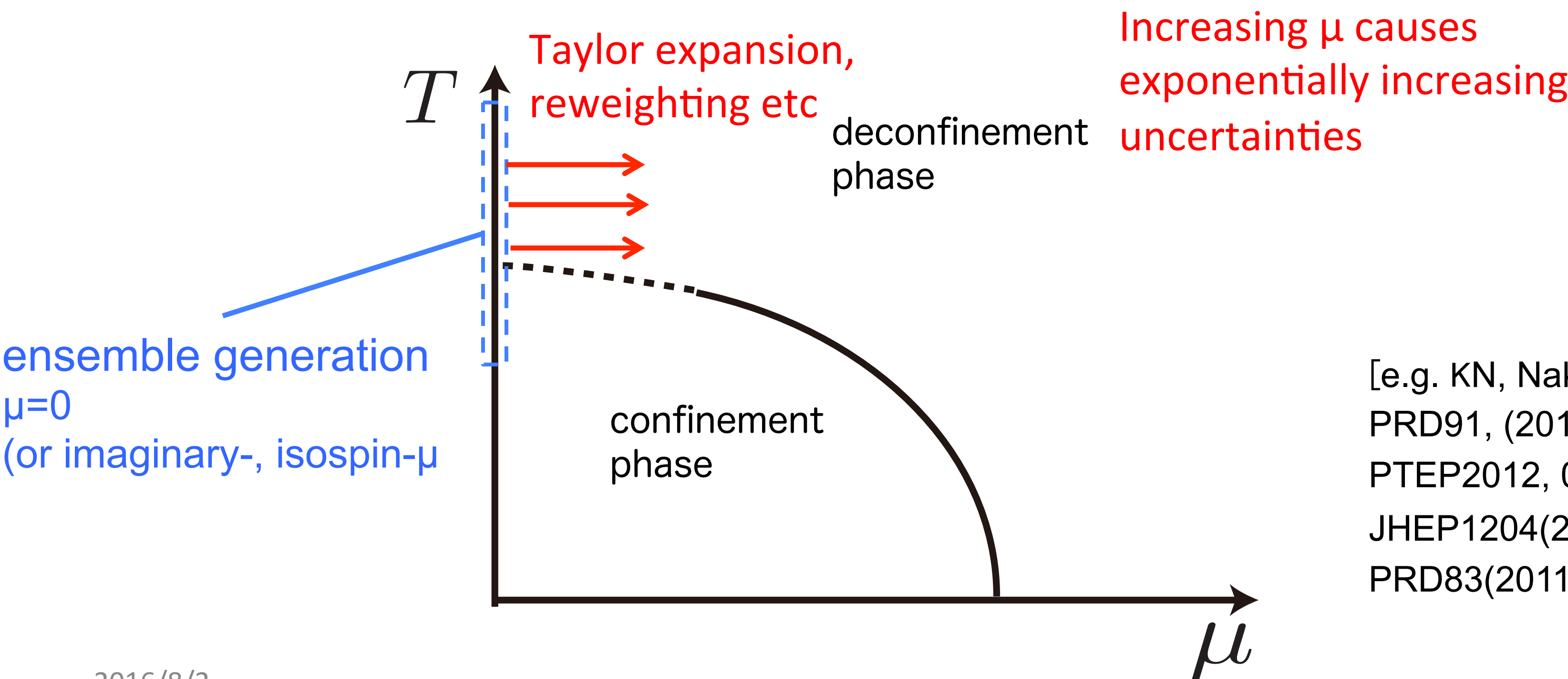
  – fast fall-off of the probability distribution in the imaginary direction
  – holomorphy of action and observables

This argument also tells what causes the failure of the CLM.

("Revisit the argument of justification", KN, Nishimura, Shimasaki in preparation.)

Advantages of CLM

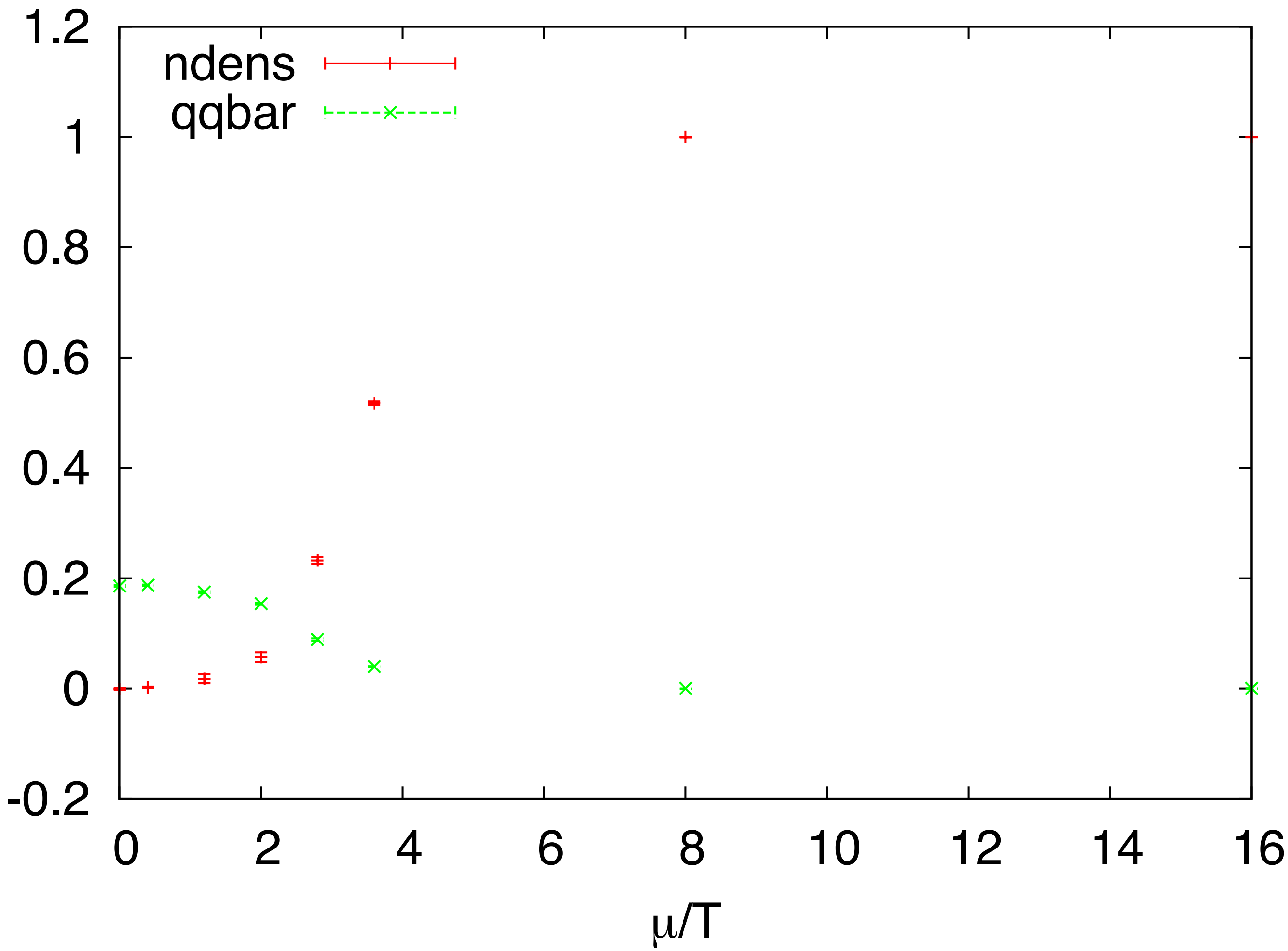
- CLM overcomes several points which are serious difficulties for approaches based on importance sampling
  - CLM is possible even if the phase fluctuation is very large
  - exponential increase of numerical cost does not occur



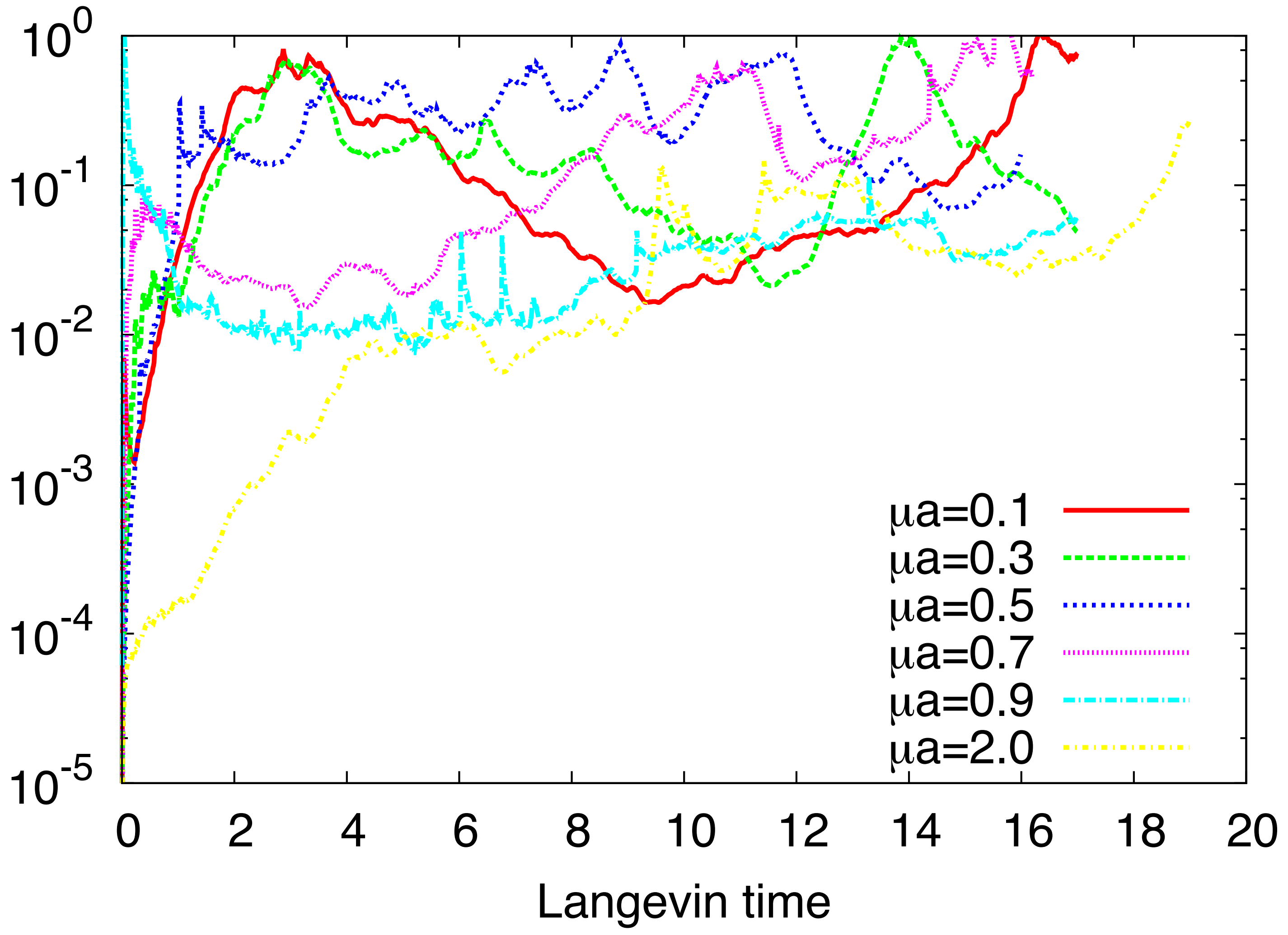
[e.g. KN, Nakamura,
PRD91, (2015) no9. 094507
PTEP2012, 01A103,
JHEP1204(2012)092,
PRD83(2011)114507]

Back up for 4^4

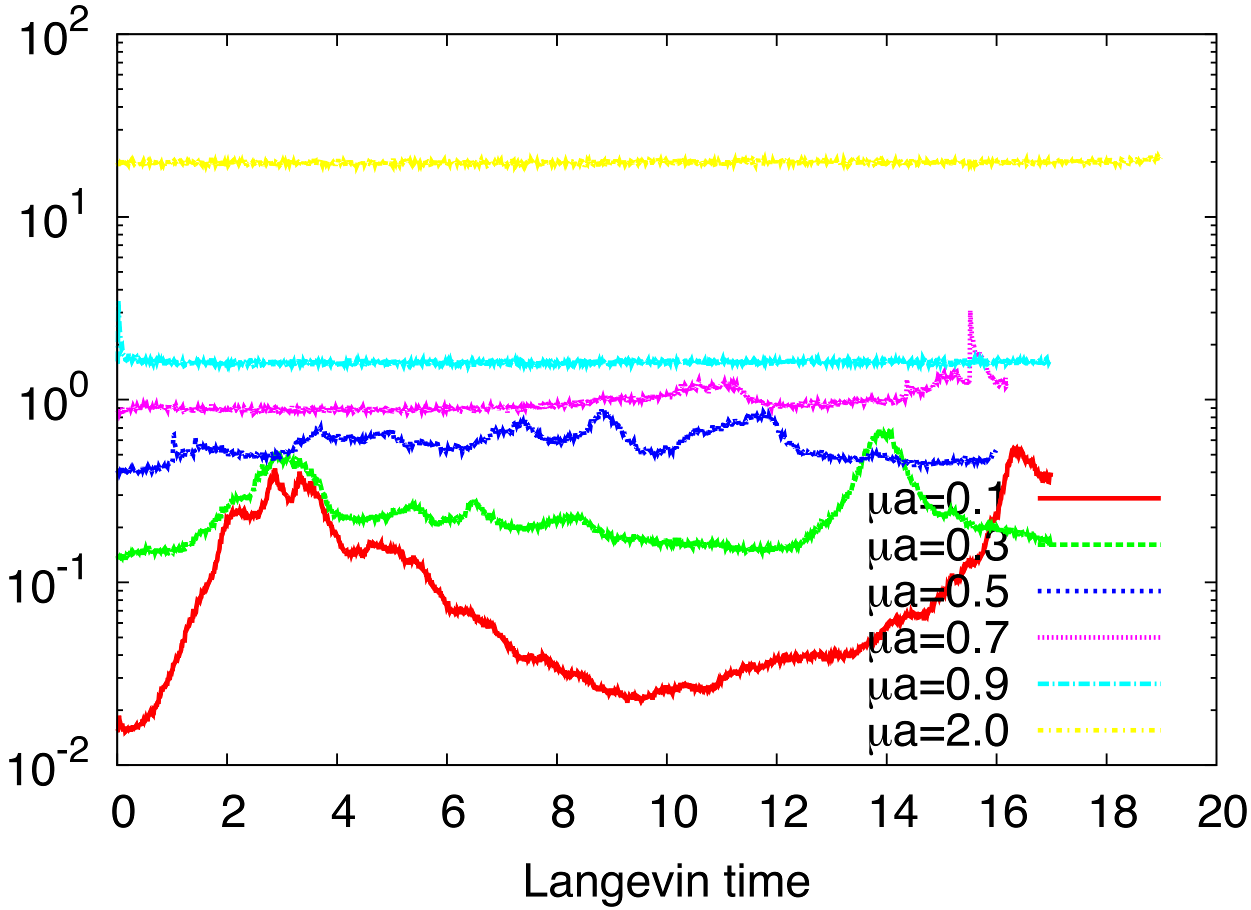
Chiral condensate and number density



Unitarity norm



anti-hermiticity norm



New types of norm for singular drift problem

- CLM fails when such Dirac eigenvalues appear that the fermion drift becomes singular [Mollgaard & Splittorff, Greensite]

- We showed that the singular drift problem can be avoided by chosing suitable norm in RMT
  - norms including Dirac operator
  - e.g. anti-hermiticity norm

$$\mathcal{N}_{\rm a.h.} = \frac{1}{N_V} {\rm tr}[(D + D^\dagger)^2]$$