On the condition for correct convergence in the complex Langevin method

Shinji Shimasaki (Keio U.)

based on the work done in collaboration with
Keitaro Nagata (KEK), Jun Nishimura (KEK)

arXiv:1606.07620 [hep-lat] and more

Complex Langevin Method (CLM)

[Parisi 83][Klauder 84]

- Path integral for a theory with complex S

$$Z = \int_{-\infty}^{\infty} dx e^{-S(x)} \qquad \langle \mathcal{O}(x) \rangle = \frac{1}{Z} \int dx \mathcal{O}(x) e^{-S(x)}$$

- Langevin eq. for complexified variables $x \to z = x + iy \in \mathbf{C}$

$$\frac{dz(t)}{dt} = v(z) + \eta(t)$$

- drift term $v(z) = -\frac{dS(z)}{dz}$, gaussian noise η (t)

$$\lim_{t \to \infty} \langle \mathcal{O}(z_\eta(t)) \rangle_\eta = \langle \mathcal{O}(x) \rangle \qquad \langle \cdots \rangle_\eta = \int \mathcal{D}\eta \cdots e^{-\frac{1}{4}\int \eta(t')^2 dt'}$$

Key identity

[Aarts, Seiler, Stamatescu 09][Aarts, James, Seiler, Stamatescu 11]

$$\boxed{\begin{gathered}\int dx dy \mathcal{O}(x+iy) P(x,y;t) = \int dx \mathcal{O}(x) \rho(x;t) \\ P(x,y;0) = \rho(x)\delta(y), \qquad \rho(x;0) = \rho(x)\end{gathered}}$$

- Probability distribution $P(x,y;t) = \langle \delta(x - x_\eta(t)) \delta(y - y_\eta(t)) \rangle_\eta$
- complex weight ρ (x;t) $\dfrac{\partial \rho(x;t)}{\partial t} = \dfrac{\partial}{\partial x}\left(\dfrac{dS}{dx} + \dfrac{\partial}{\partial x}\right)\rho(x;t)$ (Fokker-Planck eq)

  → e^{-S} is a stationary solution

- If the key identity holds and $\lim_{t\to\infty} P(x,y;t)$ converges to a unique function, then $\lim_{t\to\infty} \rho(x;t) = e^{-S}/Z$ and

$$\lim_{t\to\infty} \int dx dy \mathcal{O}(x+iy) P(x,y;t) = \frac{1}{Z}\int dx \mathcal{O}(x) e^{-S}$$

Argument on the conditions for the CLM to work

- The previous argument for the justification of the CLM uses the continuous Langevin equation from the outset.
- Starting from a discretized Langevin equation, the $\varepsilon \to 0$ limit turns out to be subtle.
- The expectation value of the time-evolved observable $\mathcal{O}(z;t)$, which plays an important role in the argument, can be ill-defined.
- ☆ We fix these subtleties and establish the argument for the justification. This leads to a simple condition, which tells us whether the results are reliable or not.

Discretized Langevin equation

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

- $\eta(t) = \eta^{(R)}(t) + i\eta^{(I)}(t)$ is a gaussian noise with probability

$$\langle \cdots \rangle_\eta \sim \int d\eta^{(R)}(t) d\eta^{(I)}(t) \cdots e^{-\frac{1}{4}\sum_t \left(\frac{1}{N_R}\eta^{(R)}(t)^2 + \frac{1}{N_I}\eta^{(I)}(t)^2\right)}$$

$$N_R - N_I = 1$$

- Expectation value

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

Infinitesimal (ε) time-evolution of the expectation value

$$\langle \mathcal{O}(z^{(\eta)}(t+\epsilon))\rangle_\eta = \int dxdy \mathcal{O}_\epsilon(x+iy)P(x,y;t)$$

$$\mathcal{O}_\epsilon(z) = \int d\eta_R d\eta_I \, e^{-\frac{1}{4N_\mathrm{R}}\eta_R^2 - \frac{1}{4N_\mathrm{I}}\eta_I^2} \mathcal{O}\Big(z + \epsilon\, v(z) + \sqrt{\epsilon}\,(\eta_R + i\eta_I)\Big)$$

ε -expansion and η -integration

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

$$L = \left\{ \mathrm{Re}v(z) + N_R \frac{\partial}{\partial x} \right\} \frac{\partial}{\partial x} + \left\{ \mathrm{Im}v(z) + N_I \frac{\partial}{\partial y} \right\} \frac{\partial}{\partial y}$$

$$: (f(x) + \partial_x)^2 := f(x)^2 + 2f(x)\partial_x + \partial_x^2$$

- For holomorphic O(z)

$$L = \left\{ \mathrm{Re} v(z) + N_R \frac{\partial}{\partial x} \right\} \frac{\partial}{\partial x} + \left\{ \mathrm{Im} v(z) + N_I \frac{\partial}{\partial y} \right\} \frac{\partial}{\partial y}$$

$$\frac{\partial}{\partial x} \to \frac{\partial}{\partial z}, \quad \frac{\partial}{\partial y} \to i \frac{\partial}{\partial z} \qquad N_R - N_I = 1$$

$$\tilde{L} = \left(v(z) + \frac{\partial}{\partial z} \right) \frac{\partial}{\partial z}$$

$$LO(z) \to \tilde{L}O(z)$$ holomorphic

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

Subtlety in the ε time-evolution

$$\tilde{L} = \left(v(z) + \frac{\partial}{\partial z} \right) \frac{\partial}{\partial z}$$

$$\begin{aligned}\langle \mathcal{O}(z^{(\eta)}(t+\epsilon)) \rangle_\eta &= \int dx dy \mathcal{O}_\epsilon(x+iy) P(x,y;t) \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} \epsilon^n \underline{\int dx dy \left\{ : \tilde{L}^n : \mathcal{O}(x+iy) \right\} P(x,y;t)}\end{aligned}$$

- The integrand involves the n-th power of the drift term
  $\to$ The prob. dist. of the drift should fall off faster than any power law at large magnitude

- If the integrals are convergent, we can take the ε →0 limit

$$\frac{d}{dt} \langle \mathcal{O}(z^{(\eta)}(t)) \rangle_\eta = \int dx dy \left\{ \tilde{L} \mathcal{O}(x+iy) \right\} P(x,y;t)$$

finite (τ) time-evolution and its subtlety

- Repeat the ε time-evolution

$$\tilde{L} = \left(v(z) + \frac{\partial}{\partial z}\right)\frac{\partial}{\partial z}$$

$$\langle \mathcal{O}(z^{(\eta)}(t+\tau))\rangle_\eta = \sum_{n=0}^{\infty}\frac{1}{n!}\tau^n \int dxdy \left\{\tilde{L}^n \mathcal{O}(x+iy)\right\} P(x,y;t)$$

- The infinite series should have a finite convergence radius
- The prob. dist. of the drift : $p(u;t) = \int dxdy\delta(u - |v(z)|)P(x,y;t)$
- i.e. $p(u;t) \sim e^{-\kappa u}$ ➡ $\sum_{n=0}^{\infty}\frac{1}{n!}\tau^n\int_0^\infty duu^n p(u;t) = \sum_{n=0}^{\infty}\frac{1}{\kappa}\left(\frac{\tau}{\kappa}\right)^n$

convergence radius ~ κ

- For the finite t-evolution of O(x) to be well-defined, the prob. dist. of the magnitude of the drift should fall off exponentially or faster

Proof of the key identity

- We prove the following identity for any k by induction wrt t

$$\boxed{\begin{gathered}\int dxdy\left\{\tilde{L}^k\mathcal{O}(x+iy)\right\}P(x,y;t)=\int dx\left\{L_0^k\mathcal{O}(x)\right\}\rho(x;t)\\ \tilde{L}=\left(v(z)+\frac{\partial}{\partial z}\right)\frac{\partial}{\partial z}\qquad L_0=\left(v(x)+\frac{\partial}{\partial x}\right)\frac{\partial}{\partial x}\end{gathered}}$$

- At t=0, it is trivially satisfied because $P(x,y;0)=\rho(x)\delta(y),\quad \rho(x;0)=\rho(x)$

- Assume this holds at some t, then for $\tau < \tau_{\text{conv}}$

$$\sum_{n=0}^{\infty}\frac{1}{n!}\tau^n\int dxdy\left\{\tilde{L}^{n+k}\mathcal{O}(x+iy)\right\}P(x,y;t)=\sum_{n=0}^{\infty}\frac{1}{n!}\tau^n\int dx\left\{(L_0)^{n+k}\mathcal{O}(x)\right\}\rho(x;t)$$

integration by parts
& FP eq for ρ (x;t)

$$\int dxdy\left\{\tilde{L}^k\mathcal{O}(x+iy)\right\}P(x,y;t+\tau)\qquad\qquad\int dx\left\{(L_0)^k\mathcal{O}(x)\right\}\rho(x;t+\tau)$$

Demonstration of our condition in simple examples

- A model with a singular drift [Nishimura-SS 15]

$$Z = \int dx (x + i\alpha)^p e^{-\frac{1}{2}x^2} \qquad v(z) = \frac{p}{z + i\alpha} - z$$

- A model with a possibility of excursions

[Aarts-Giudice-Seiler 13]

$$Z = \int dx e^{-\frac{1}{2}(A+iB)x^2 - \frac{1}{4}x^4} \qquad v(z) = (A + iB)z - z^3$$

A model with a singular drift

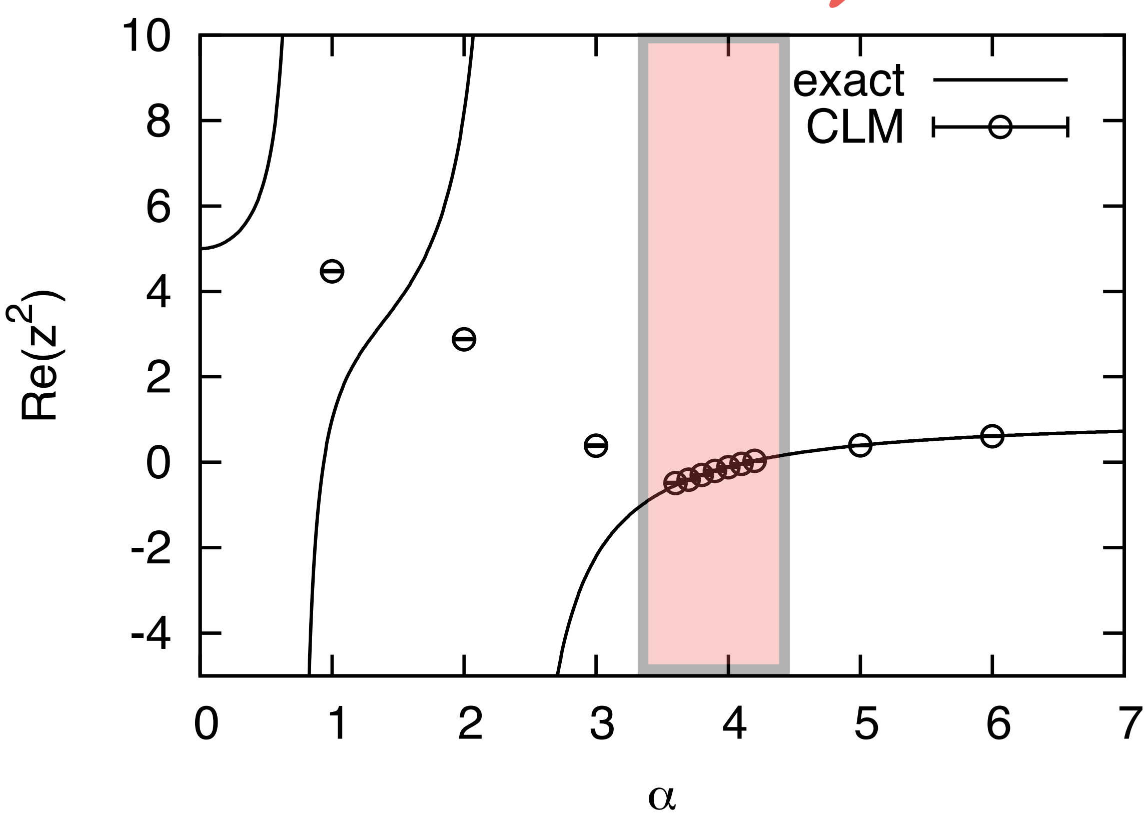
[Nishimura-SS 15]

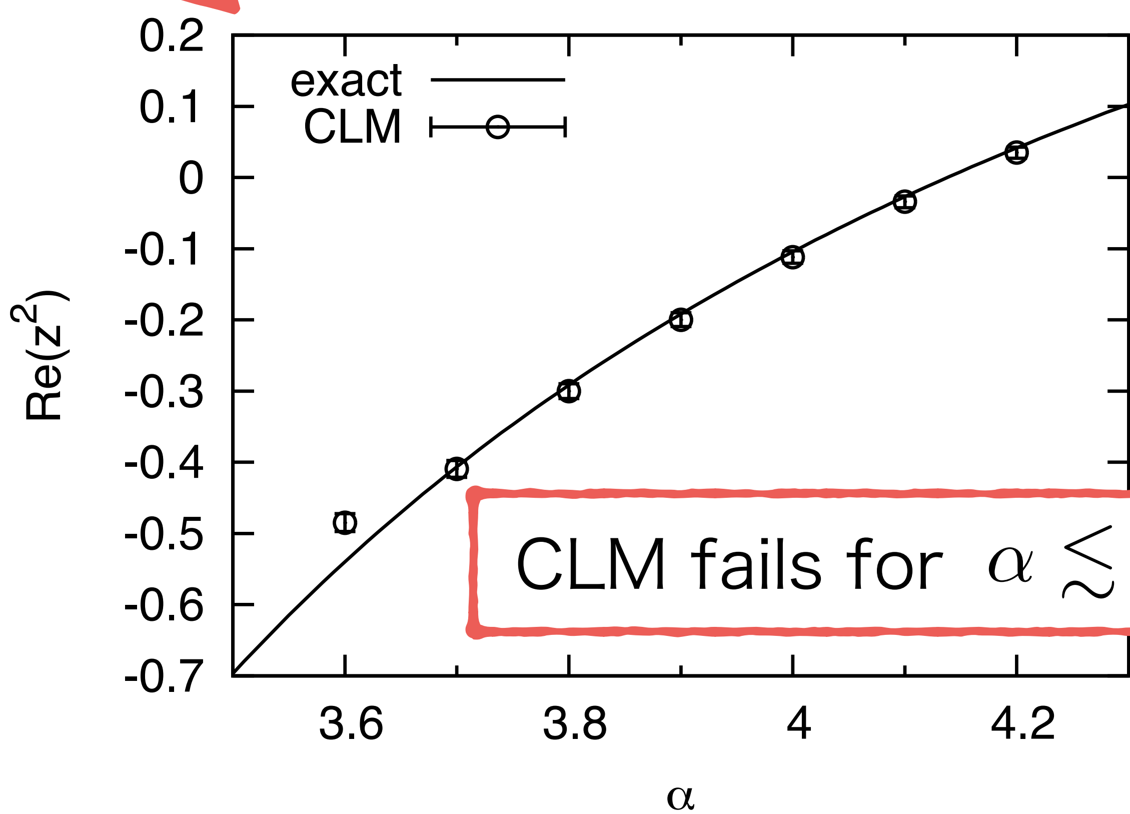
$$Z = \int dx (x + i\alpha)^p e^{-\frac{1}{2}x^2}$$

drift : $v(z) = \dfrac{p}{z + i\alpha} - z$ singularity at z=-iα

- p=4

[Nagata-Nishimura-SS 16]

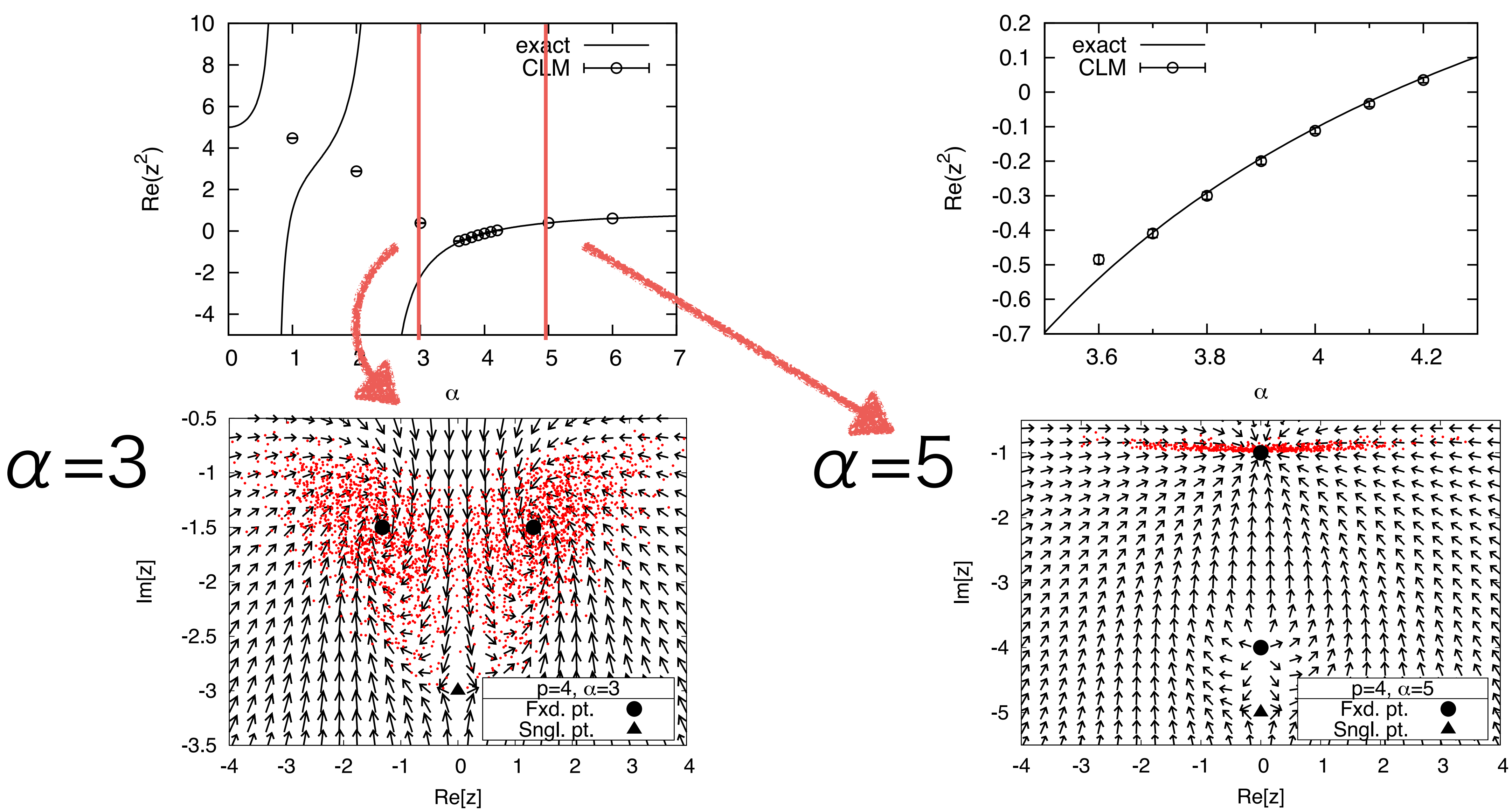




$\boxed{\text{CLM fails for } \alpha \lesssim 3.6}$

Probability distribution

[Nagata-Nishimura-SS 16]



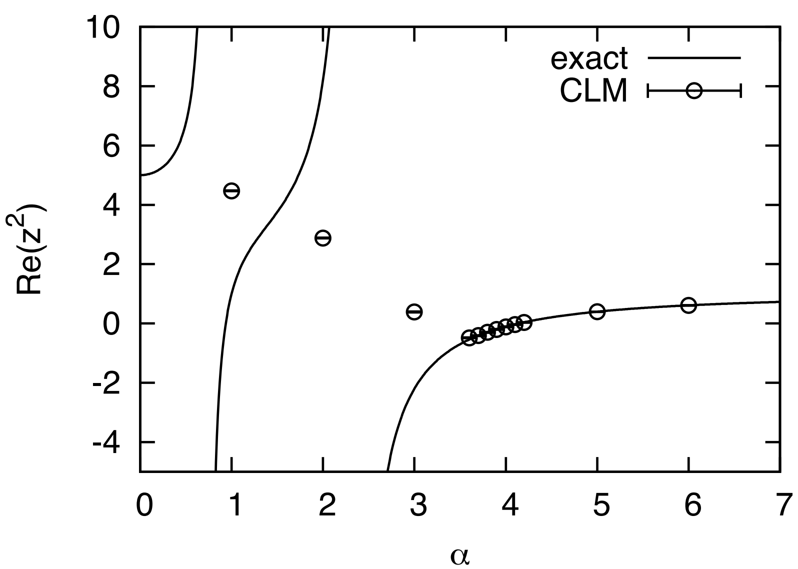
many configurations
near the singularity

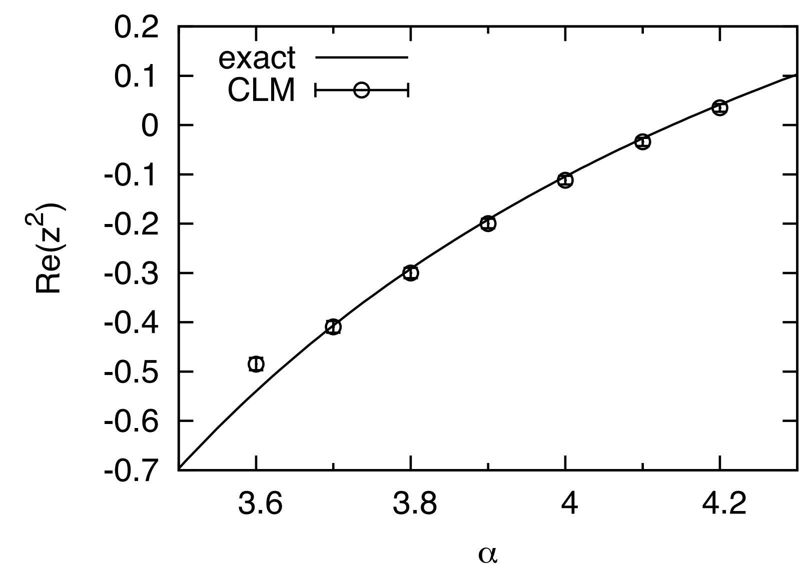
→ Large drifts appear

no configurations
near the singularity

Demonstration of our condition

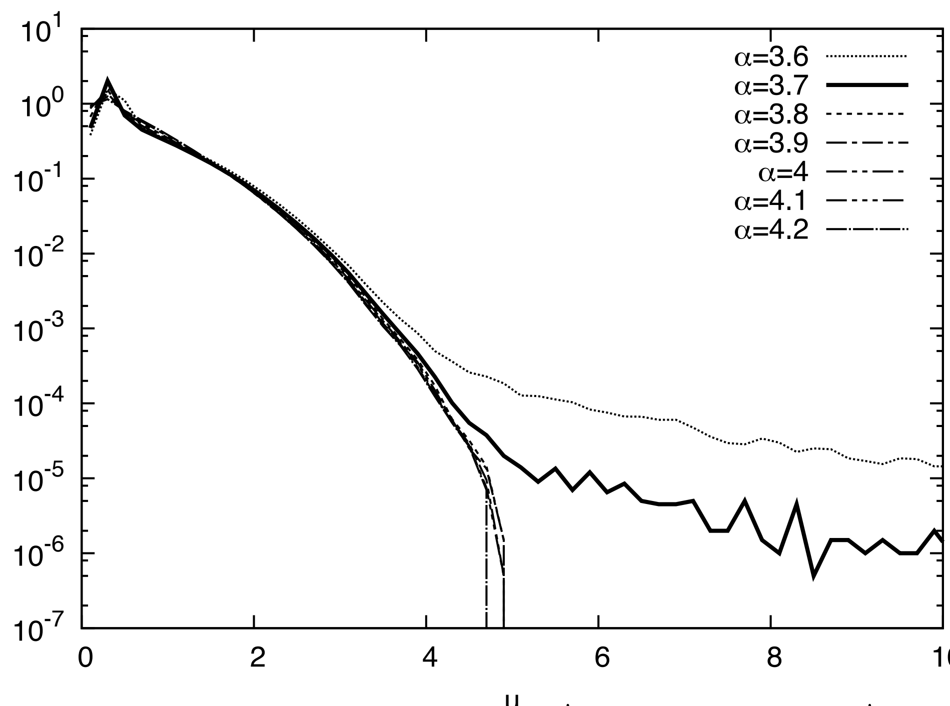
[Nagata-Nishimura-SS 16]





semi-log plot

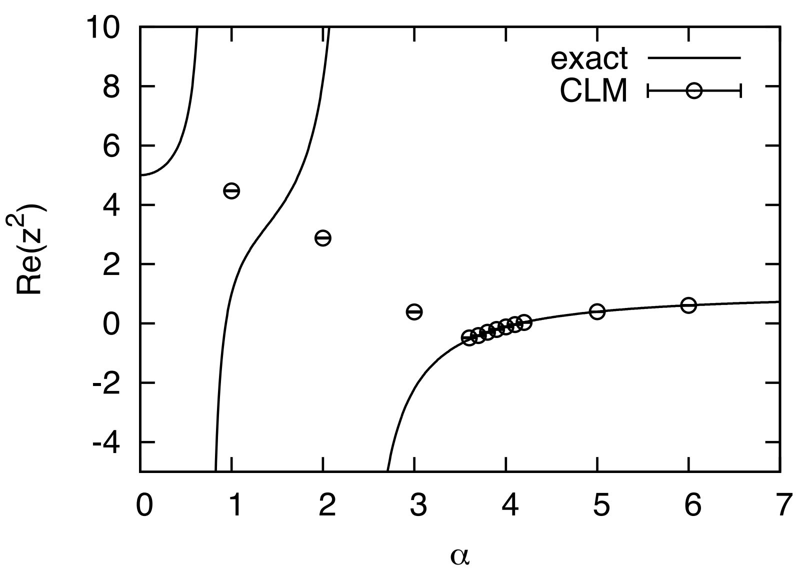
prob. dist. of the drift

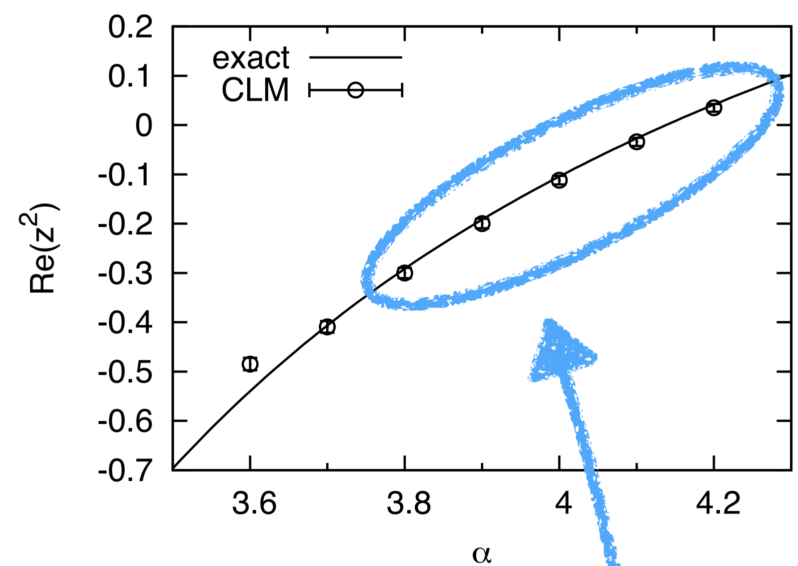


$$u = |v(z)| = \left| \frac{p}{z + i\alpha} - z \right|$$

Demonstration of our condition

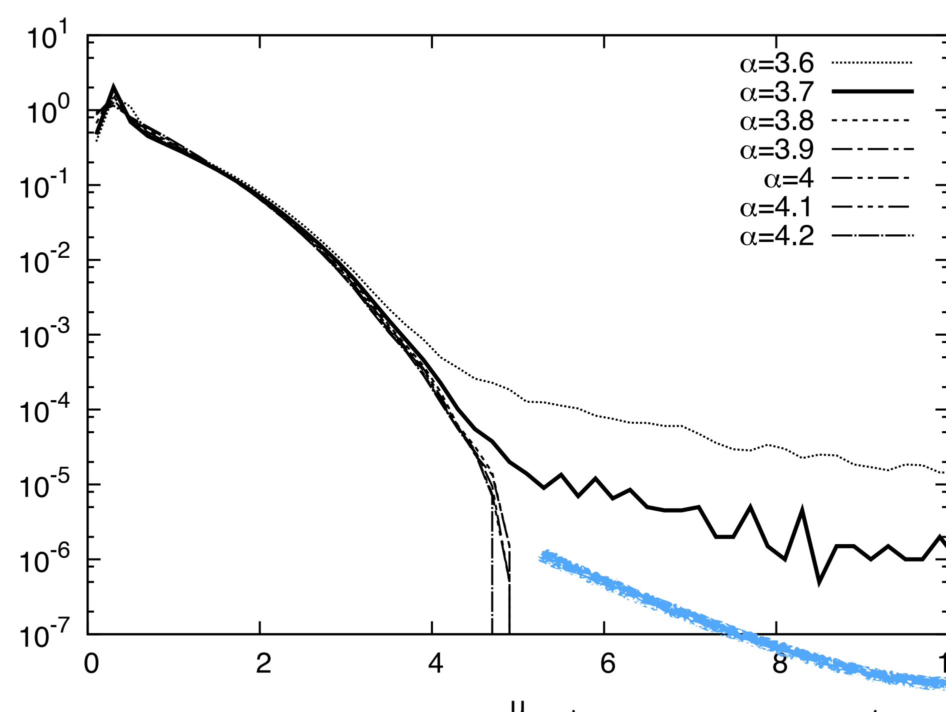
[Nagata-Nishimura-SS 16]





semi-log plot

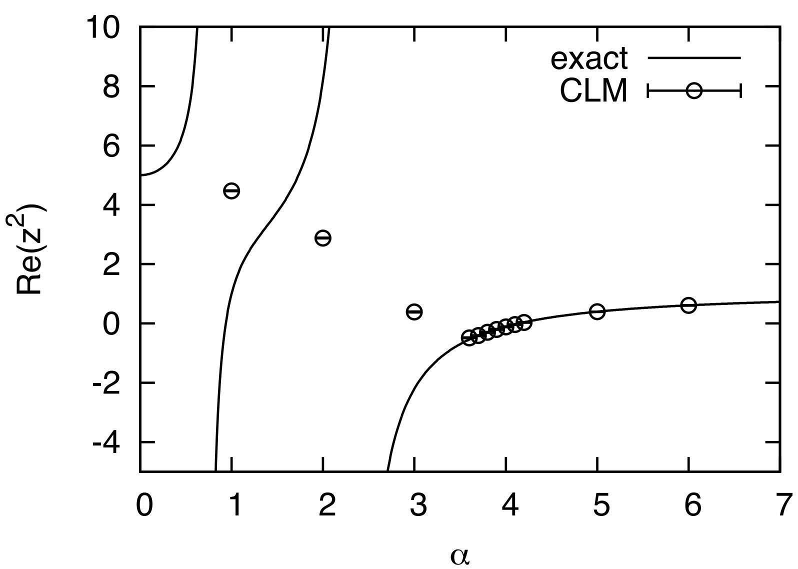
prob. dist. of the drift

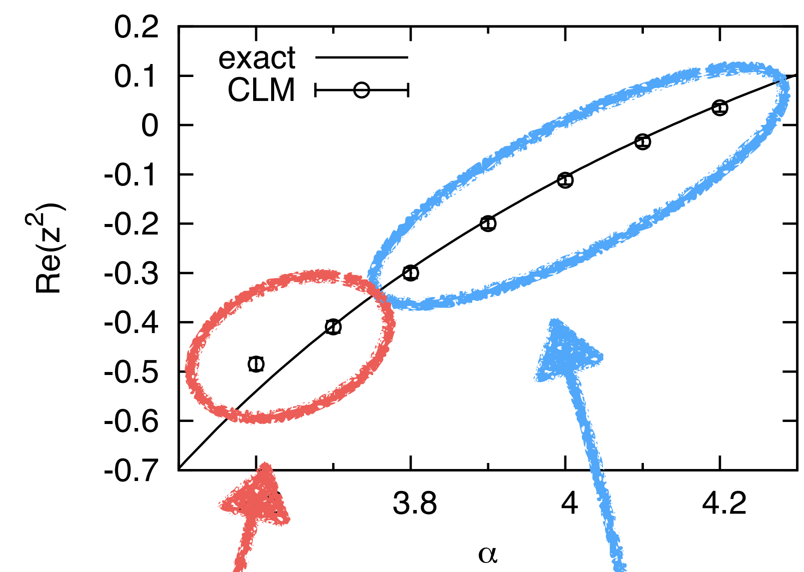


$$u = |v(z)| = \left| \frac{p}{z + i\alpha} - z \right|$$

Demonstration of our condition

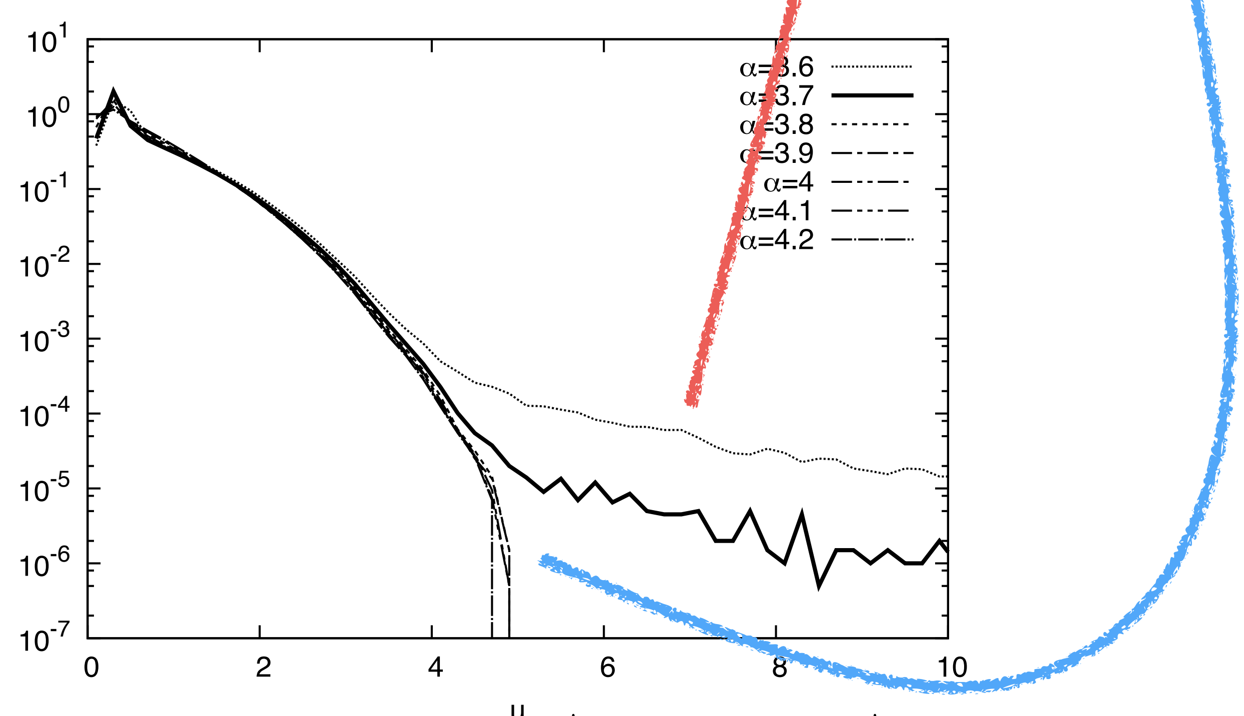
[Nagata-Nishimura-SS 16]





semi-log plot

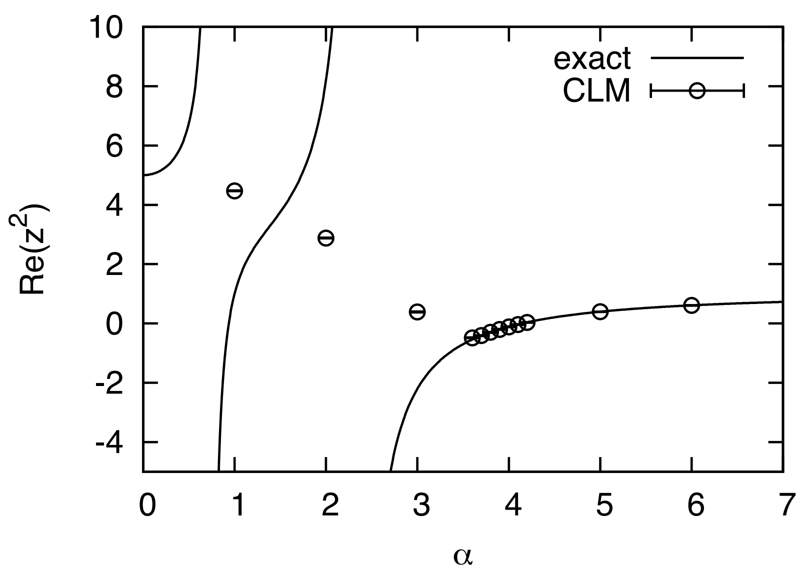
prob. dist. of the drift

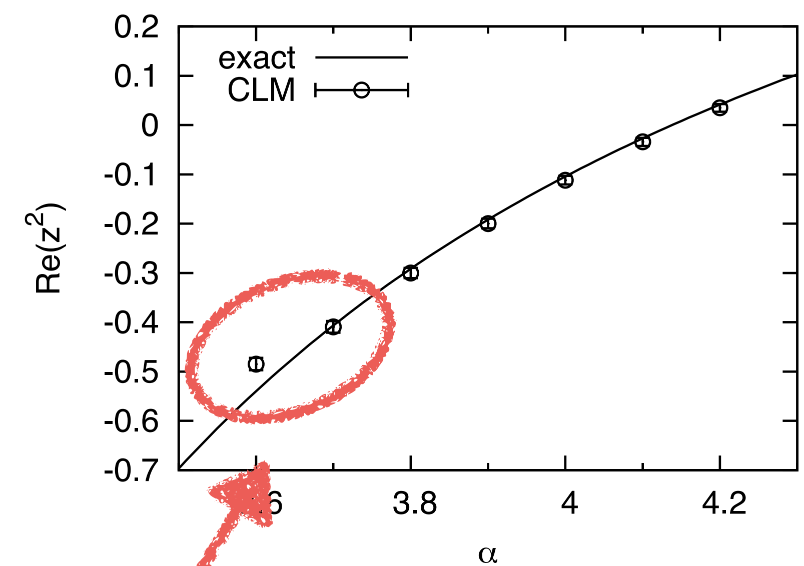


$$u = |v(z)| = \left| \frac{p}{z + i\alpha} - z \right|$$

Demonstration of our condition

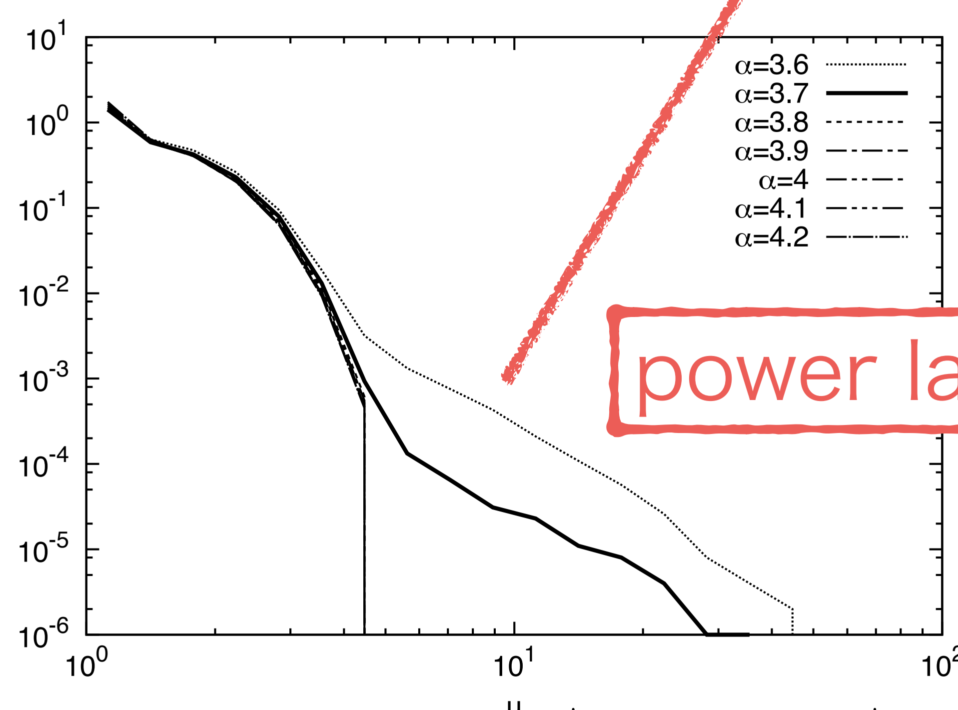
[Nagata-Nishimura-SS 16]





log-log plot

prob. dist. of the drift



power law fall off

$$u = |v(z)| = \left| \frac{p}{z + i\alpha} - z \right|$$

A model with a possibility of excursions

[Aarts-Giudice-Seiler 13]

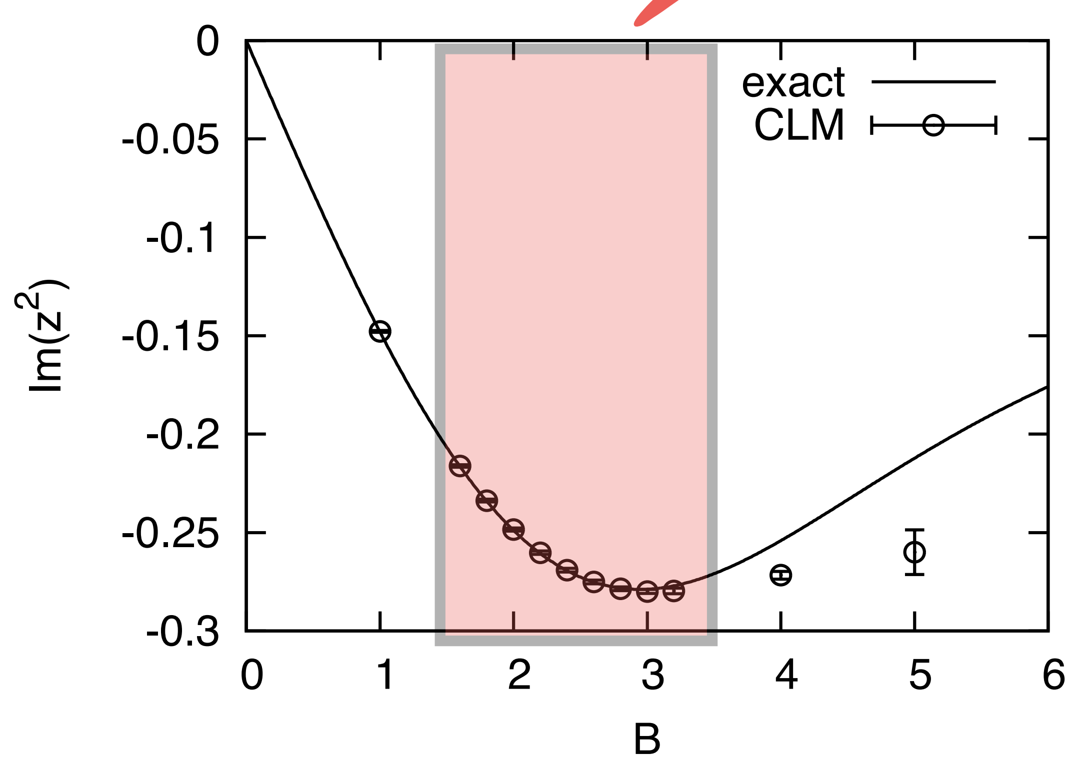
$$Z = \int dx e^{-\frac{1}{2}(A+iB)x^2 - \frac{1}{4}x^4}$$

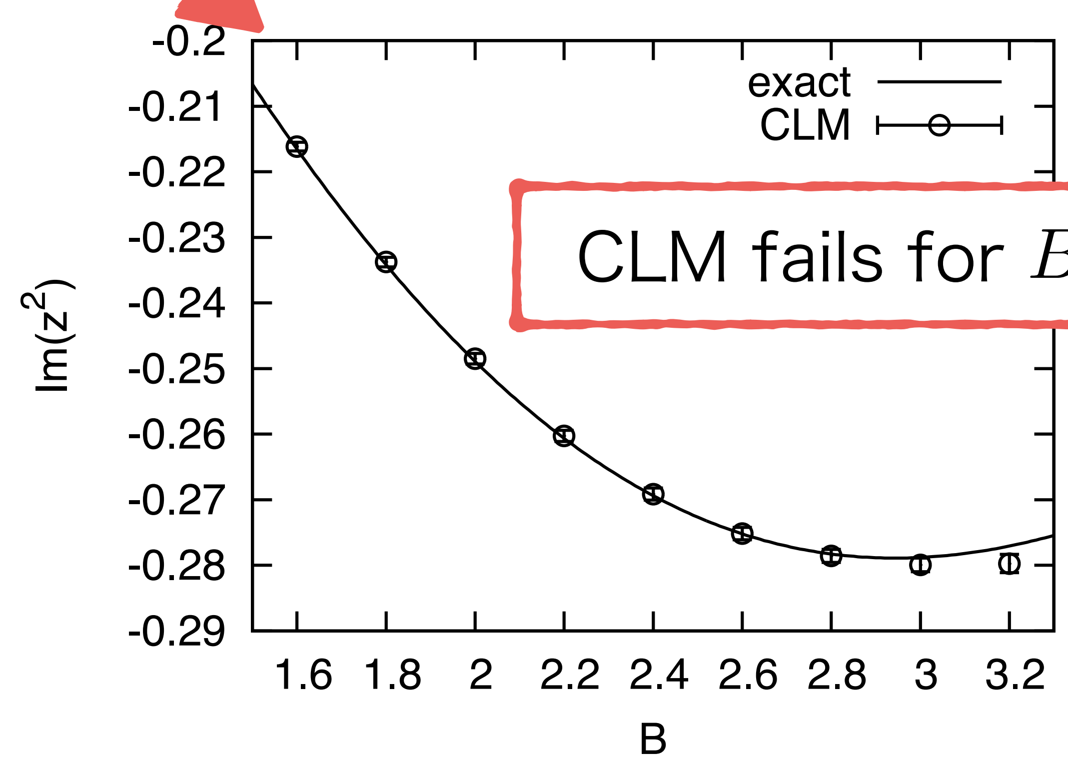
drift : $v(z) = -(A+iB)z - z^3$

potential danger of excursions in large |z|

- A=1

[Nagata-Nishimura-SS 16]

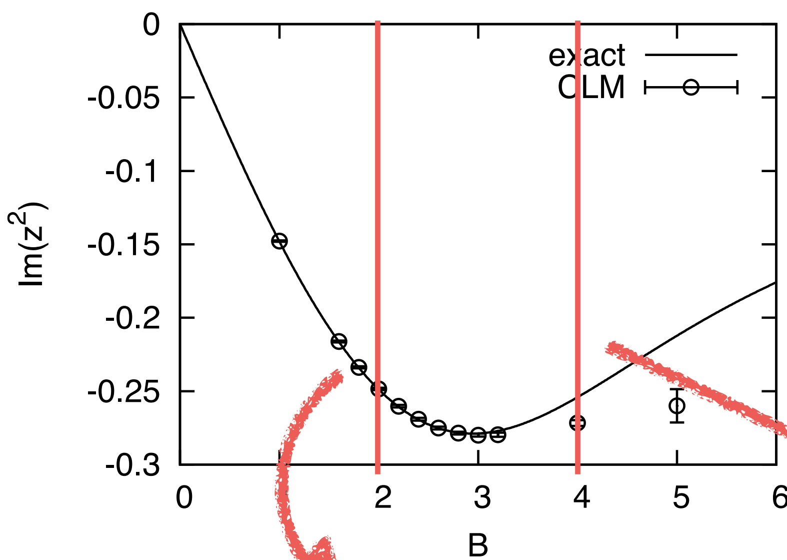


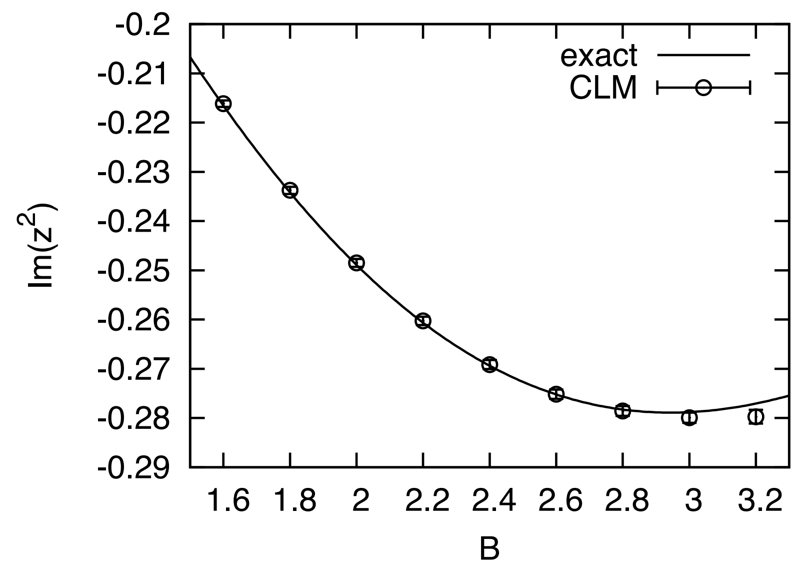


CLM fails for $B \gtrsim 3.0$

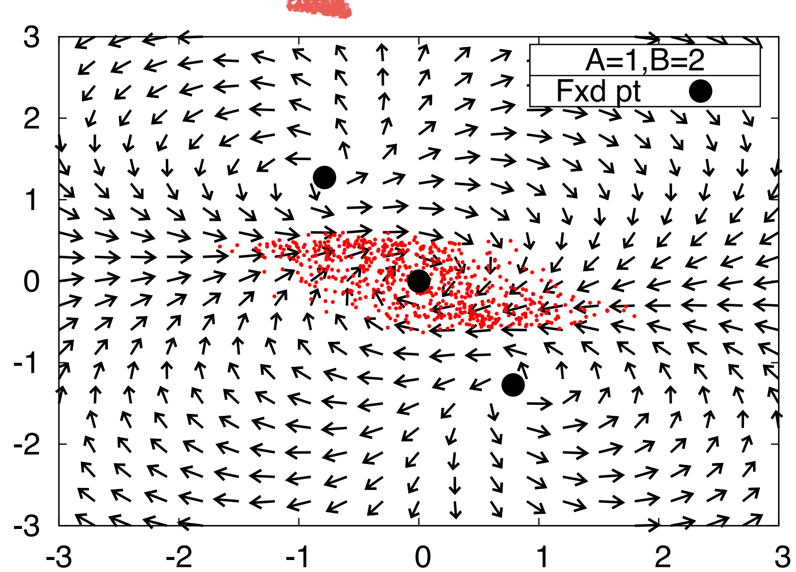
Probability Distribution

[Nagata-Nishimura-SS 16]

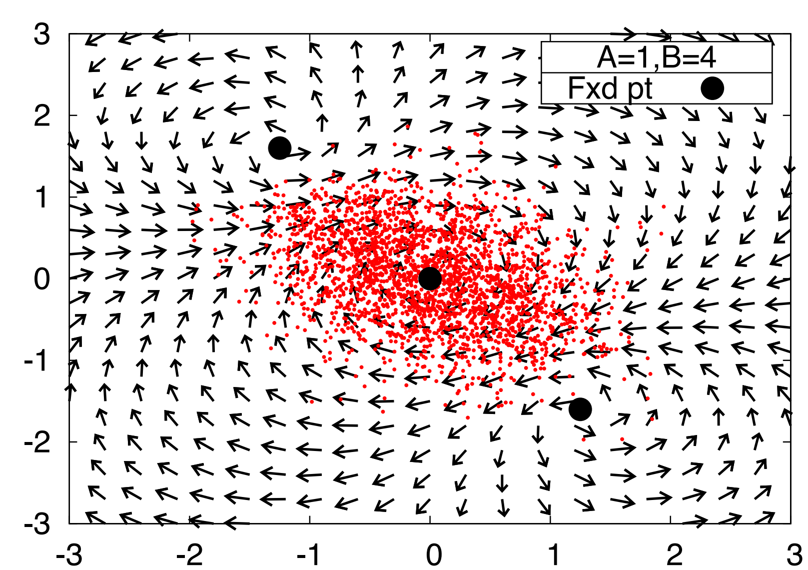




B=2



B=4



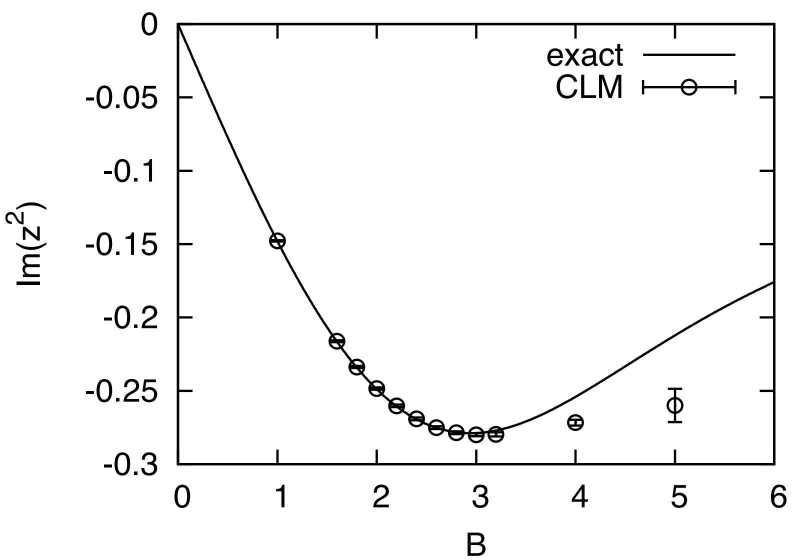
configurations are restricted in |y|<C

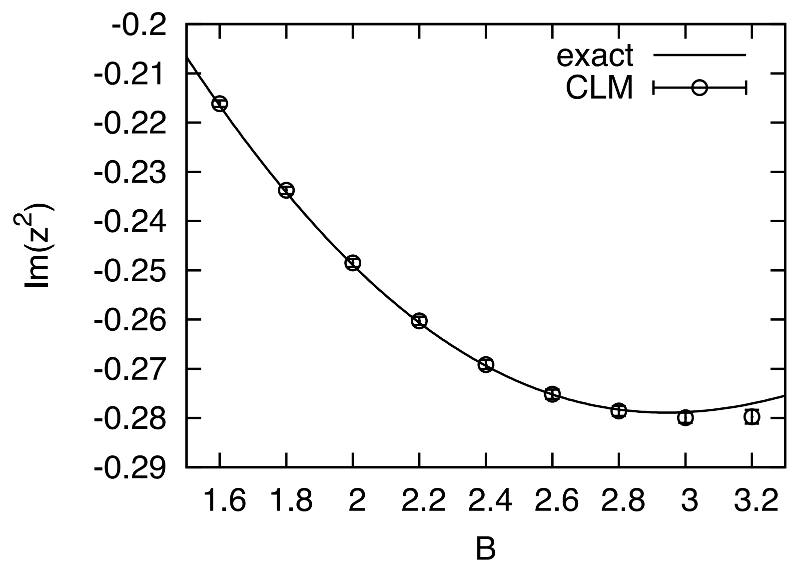
configurations spread out into large |y| region

→ **Large drifts appear**

Demonstration of our condition

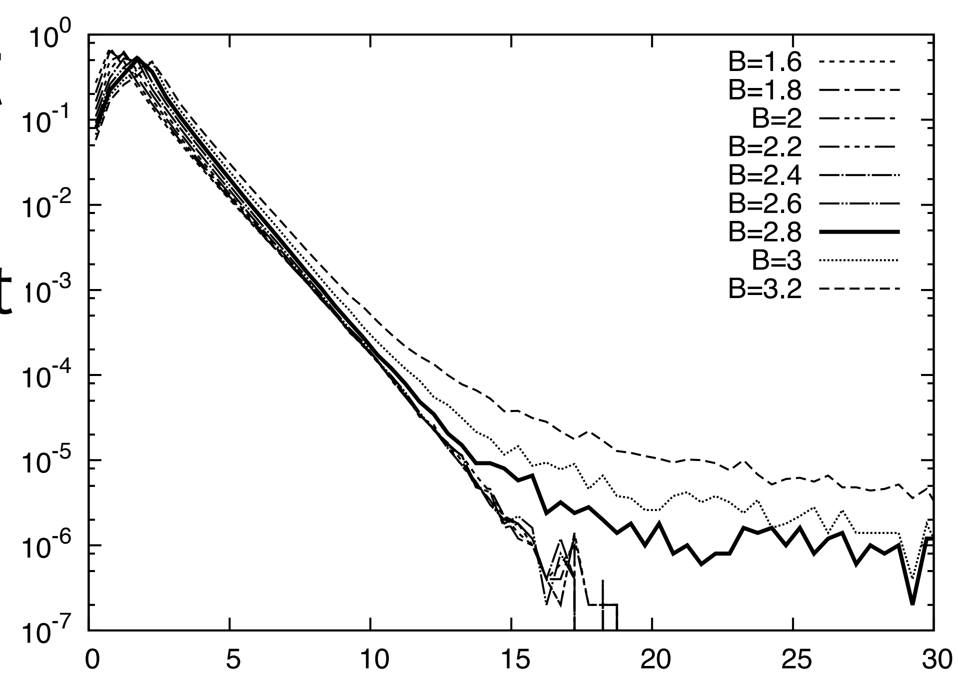
[Nagata-Nishimura-SS 16]





semi-log plot

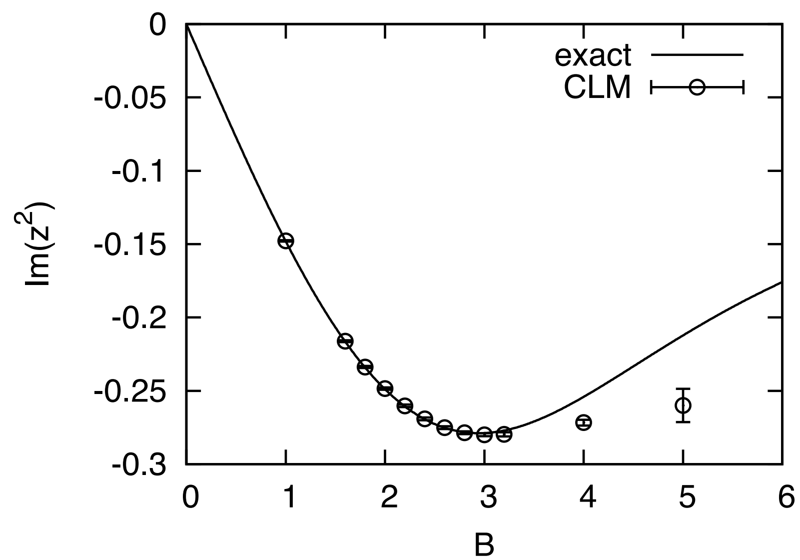
prob. dist. of the drift

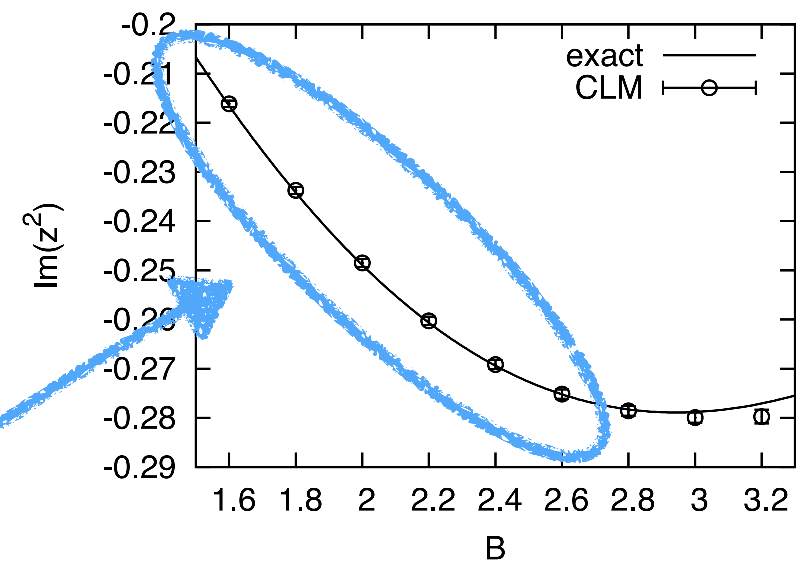


$$u = |v(z)| = |(1+iB)z - z^3|$$

Demonstration of our condition

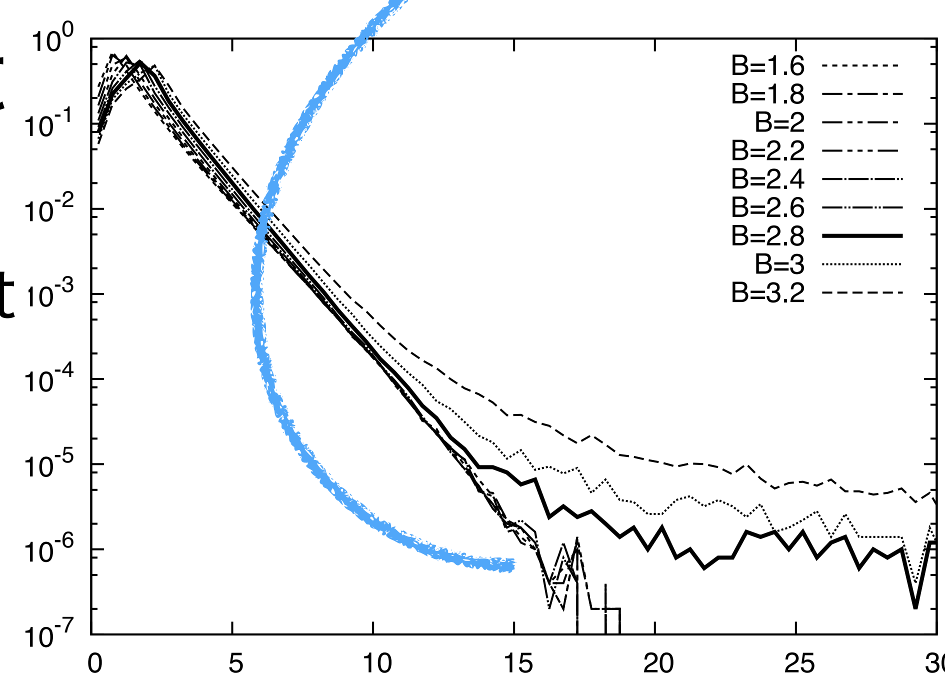
[Nagata-Nishimura-SS 16]





semi-log plot

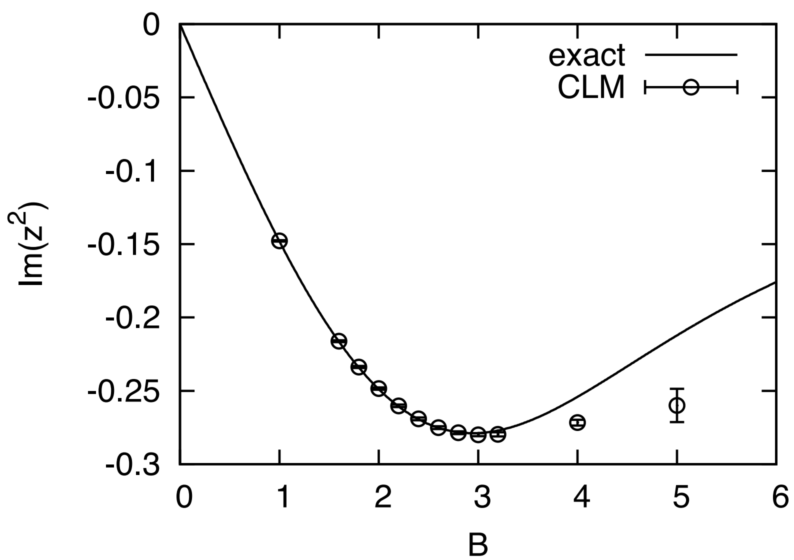
prob. dist. of the drift

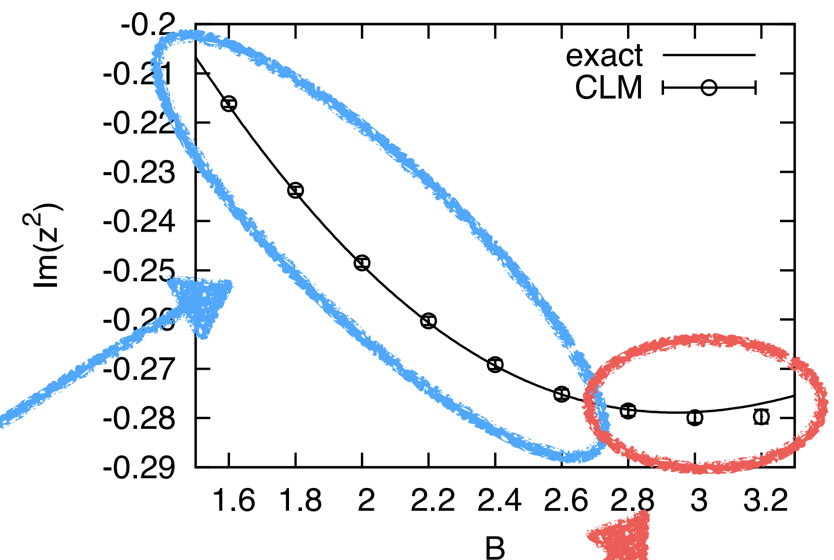


$$u = |v(z)| = |(1+iB)z - z^3|$$

Demonstration of our condition

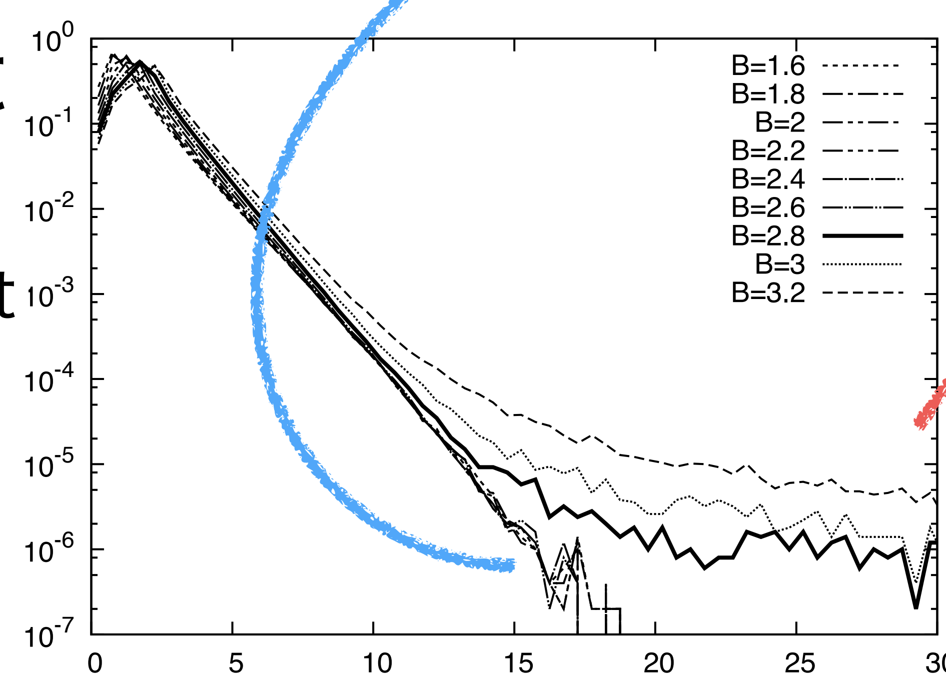
[Nagata-Nishimura-SS 16]





semi-log plot

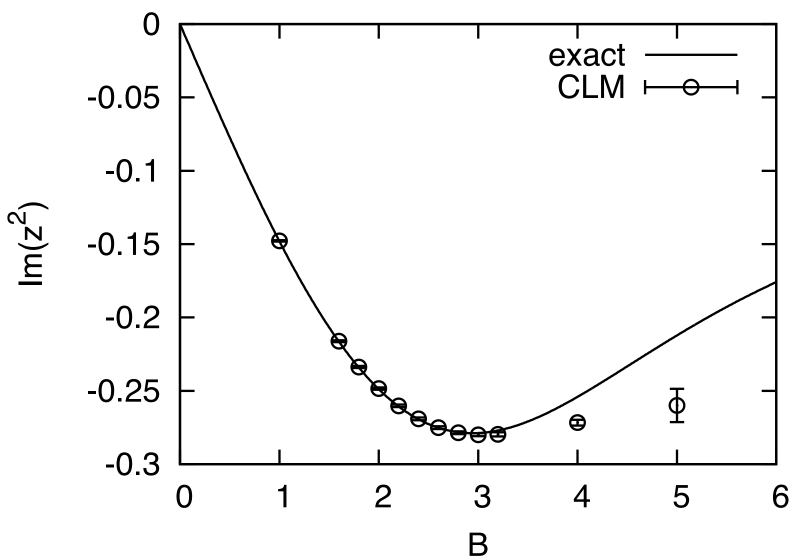
prob. dist. of the drift

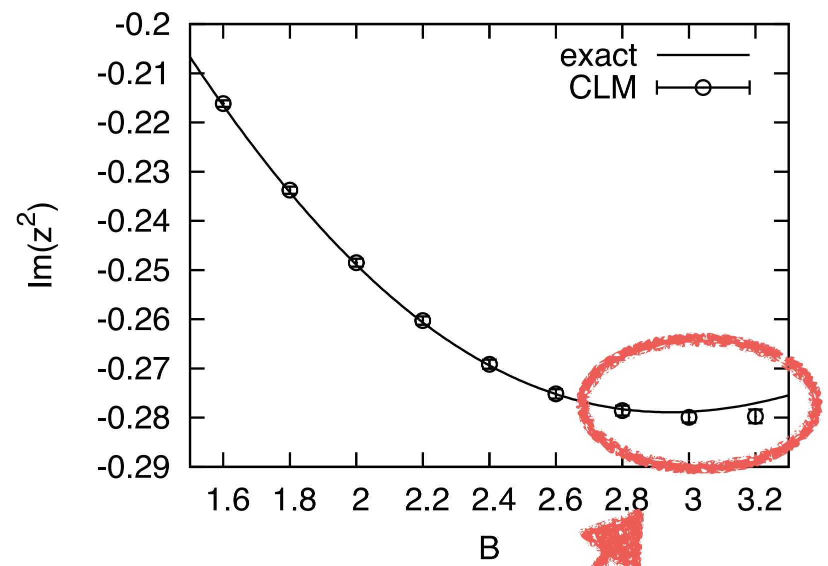


$$u = |v(z)| = |(1+iB)z - z^3|$$

Demonstration of our condition

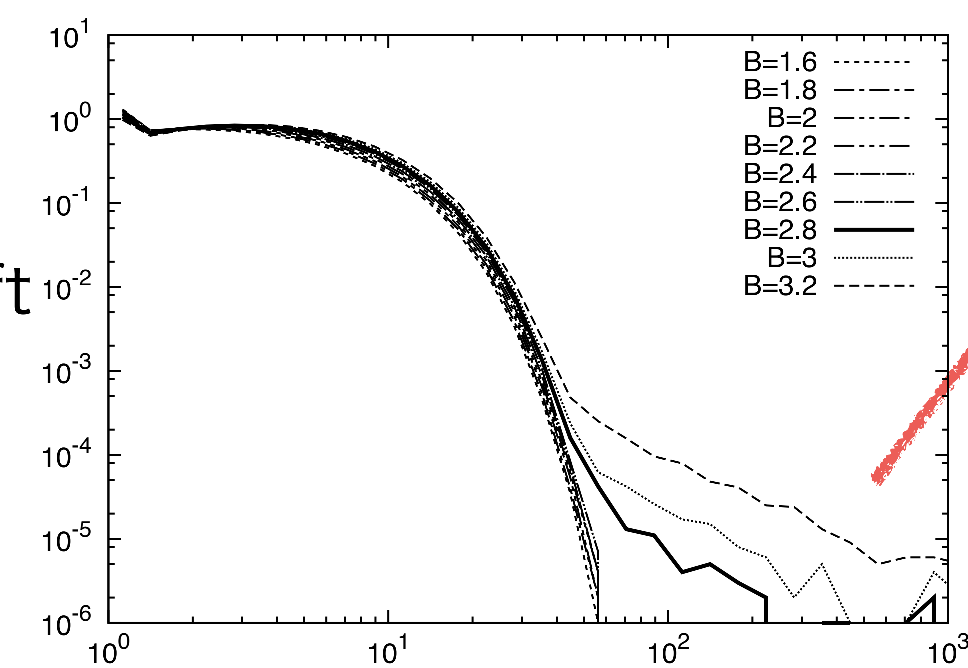
[Nagata-Nishimura-SS 16]





log-log plot

prob. dist. of the drift



$$u = |v(z)| = |(1+iB)z - z^3|$$

power law fall off

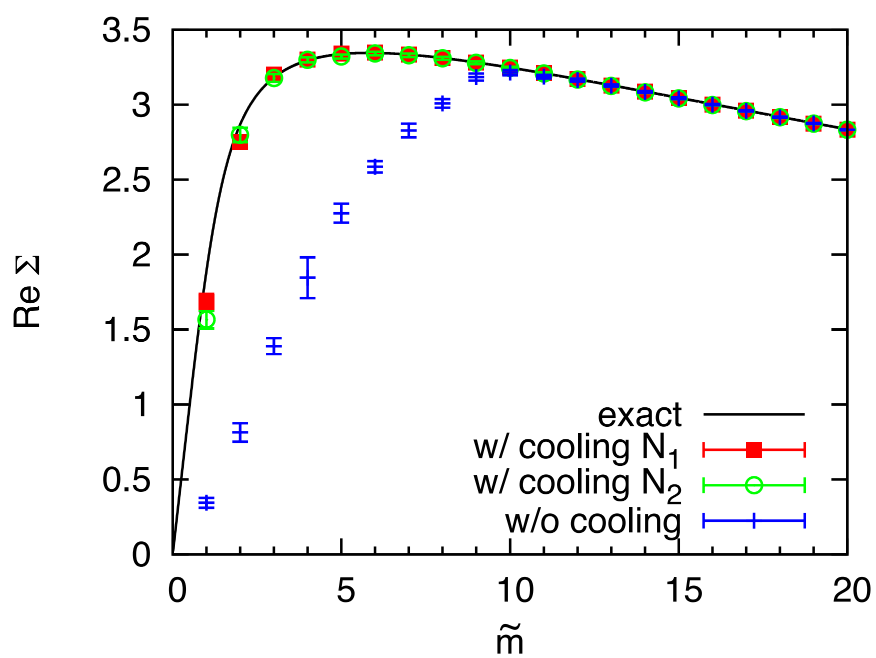
Demonstration of our condition in Chiral Random Matrix Theory

CLM for Chiral Random Matrix Theory

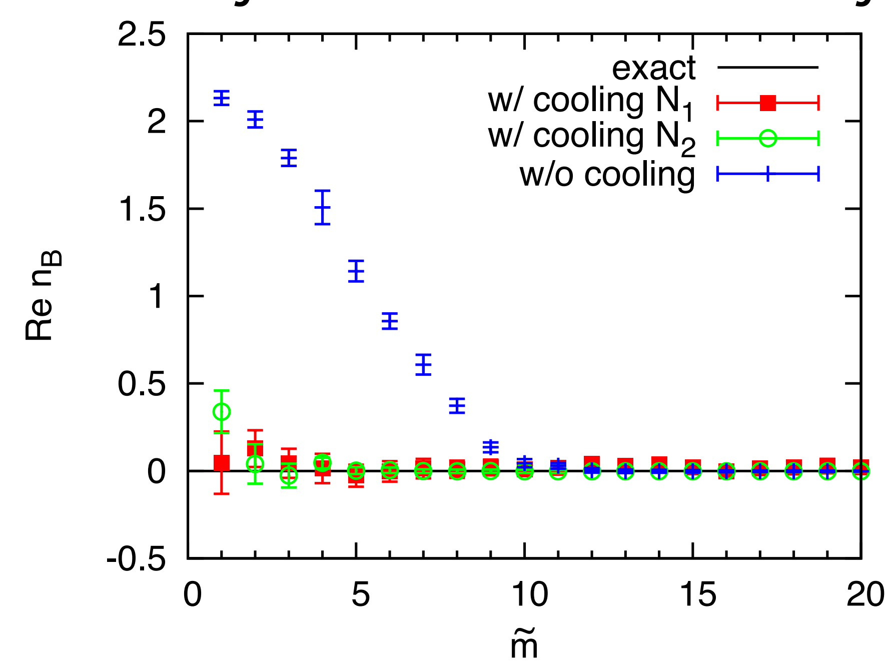
[Bloch-Bruckmann-Kieburg-Splitorff-Verbaarschot 13]

- $$Z = \int d\Phi_1 d\Phi_2 d\Psi_1 d\Phi_2 [\det(D+m)]^{N_f} e^{-S}$$ [Mollgaard-Splittorff 13, 15] [Nagata-Nishimura-SS 16]
- For small masses, naive application of the CLM fails due to the singular drift problem (large drift) [Mollgaard-Splittorff 13]
- By using a gauge cooling, we can avoid the problem and make the CLM work even at small masses [Nagata-Nishimura-SS 16]

chiral condensate



baryon number density

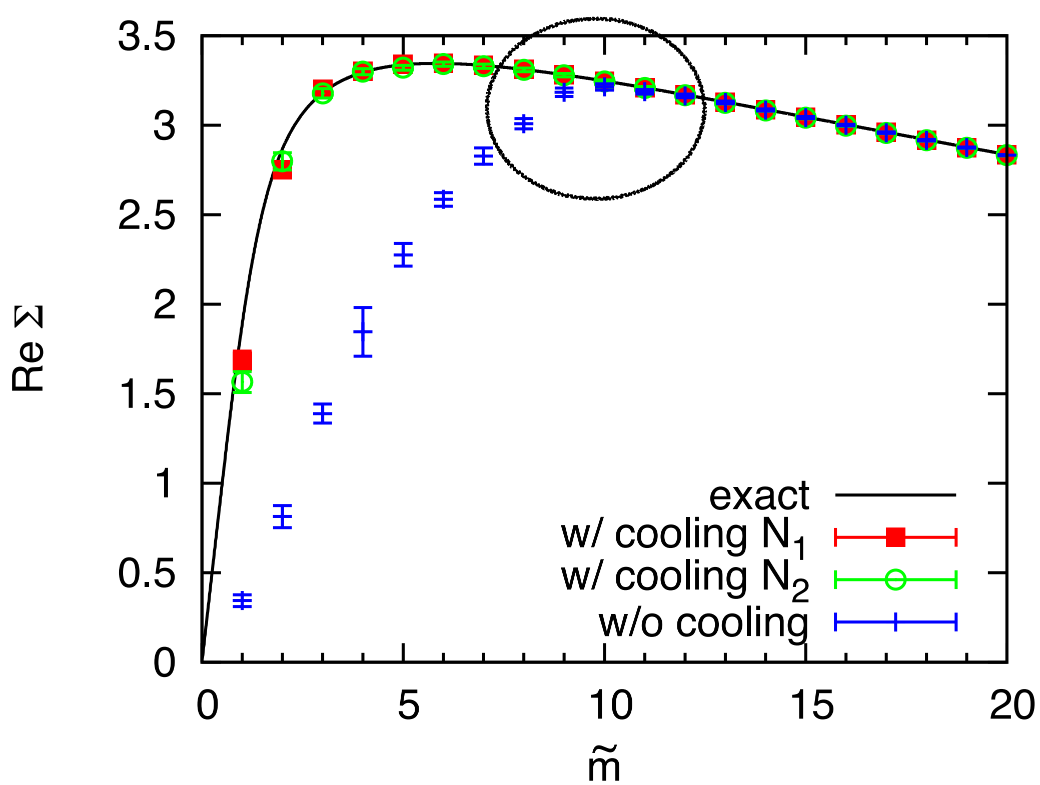


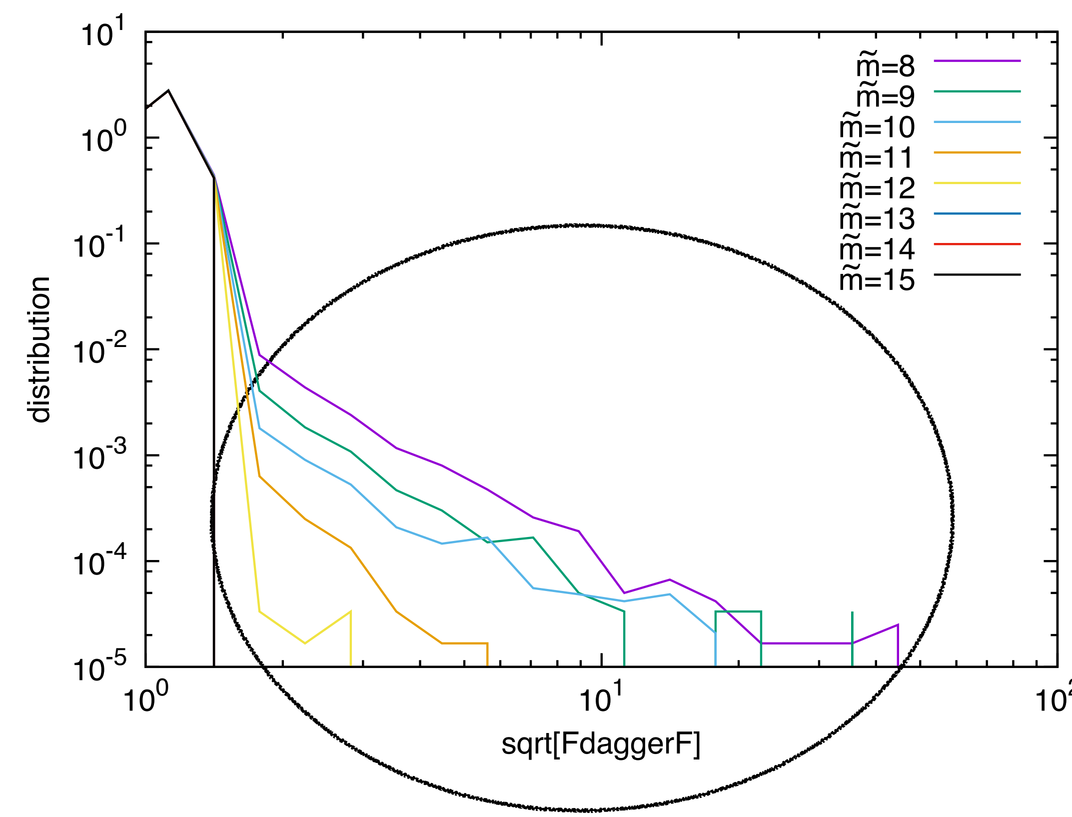
Demonstration of our condition in CRMT

$\Phi_1, \ \Phi_2, \ \Psi_1, \ \Psi_2 \quad \rightarrow$ drift F_i (i=1,2,3,4)

eigenvalues of $\sqrt{F_i^\dagger F_i}$ (i=1,2,3,4)

without cooling (blue points)



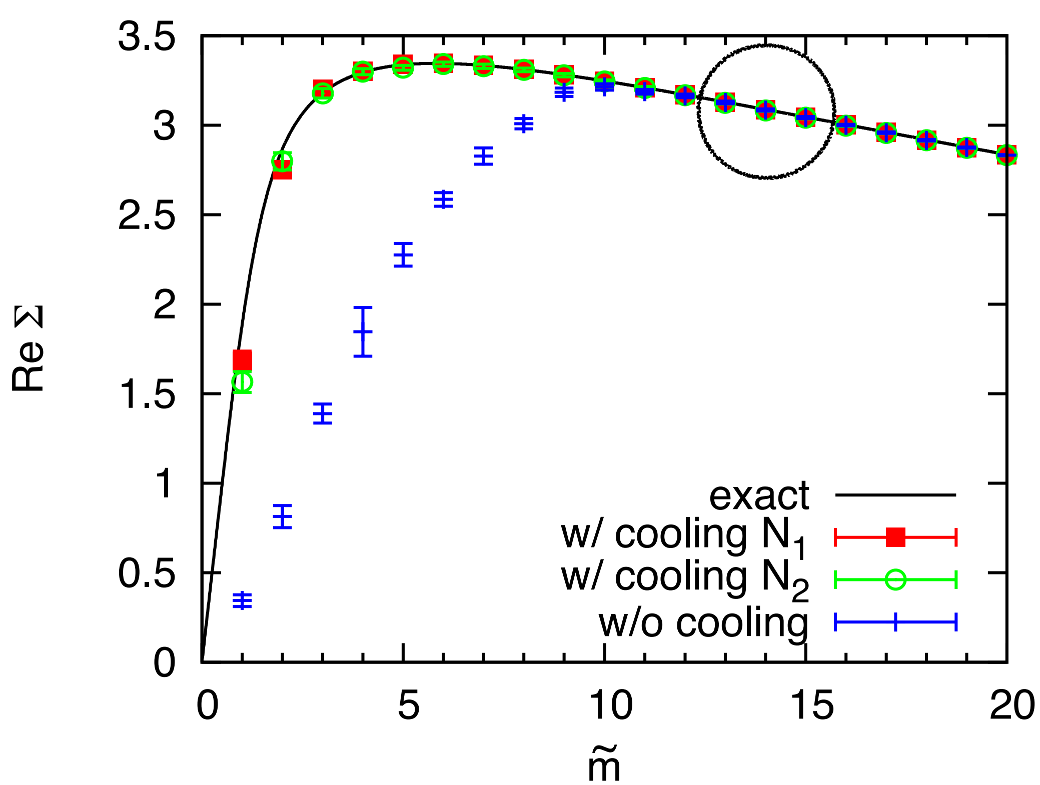


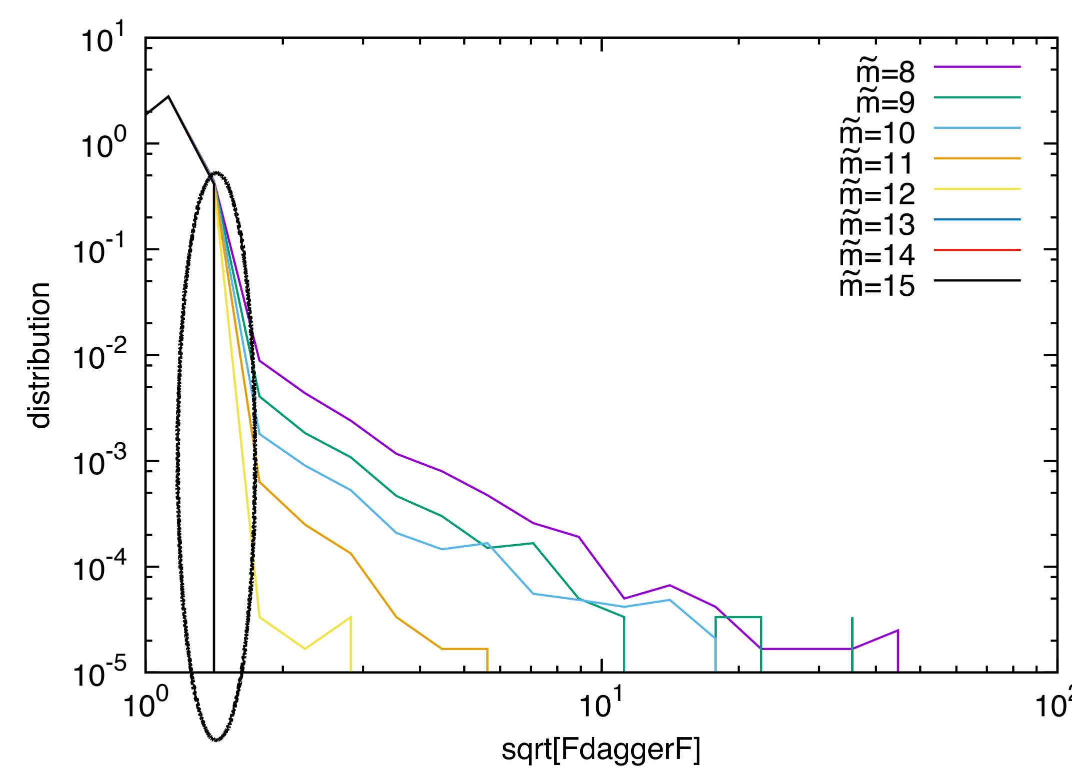
Demonstration of our condition in CRMT

$\Phi_1,\ \Phi_2,\ \Psi_1,\ \Psi_2 \quad \rightarrow$ drift F_i (i=1,2,3,4)

eigenvalues of $\sqrt{F_i^\dagger F_i}$ (i=1,2,3,4)

without cooling (blue points)



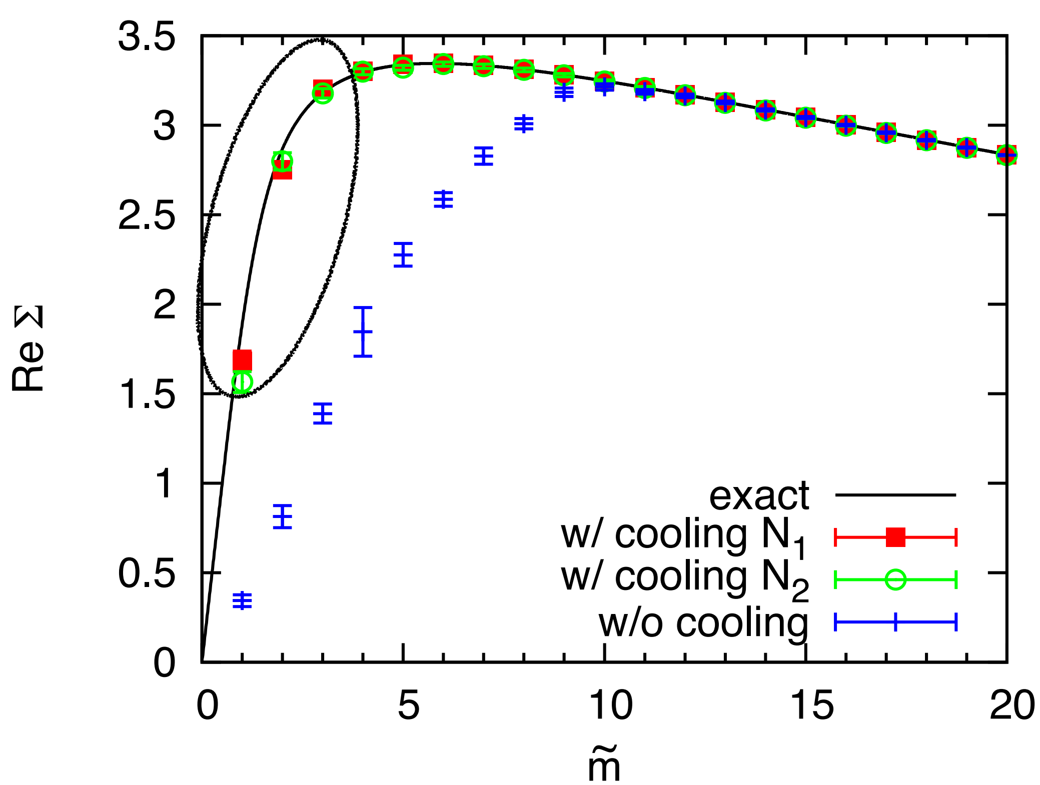


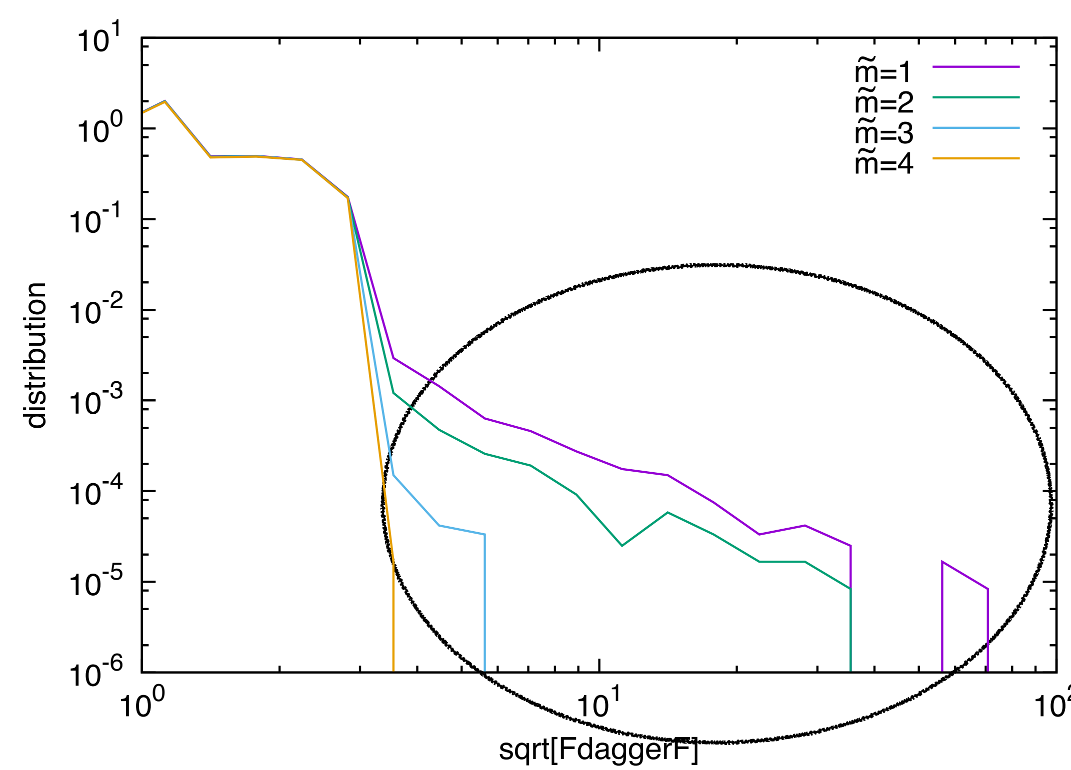
Demonstration of our condition in CRMT

$\Phi_1,\ \Phi_2,\ \Psi_1,\ \Psi_2\quad \to$ drift F_i (i=1,2,3,4)

eigenvalues of $\sqrt{F_i^\dagger F_i}$ (i=1,2,3,4)

with a cooling (red points)



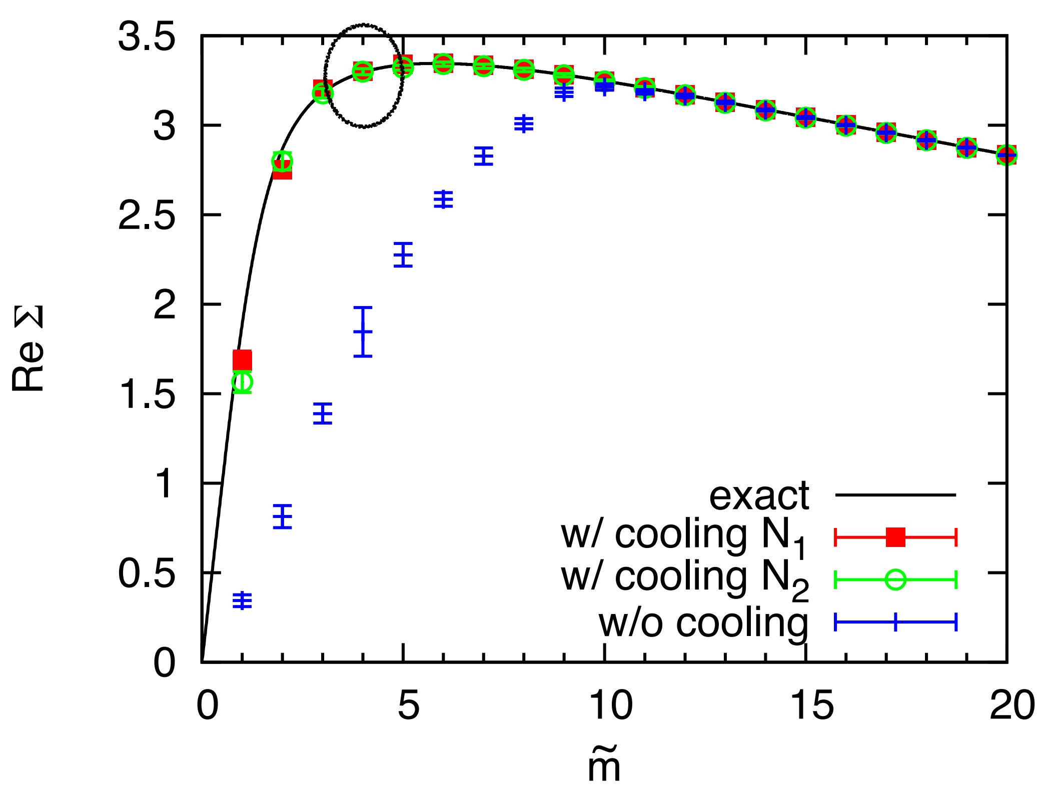


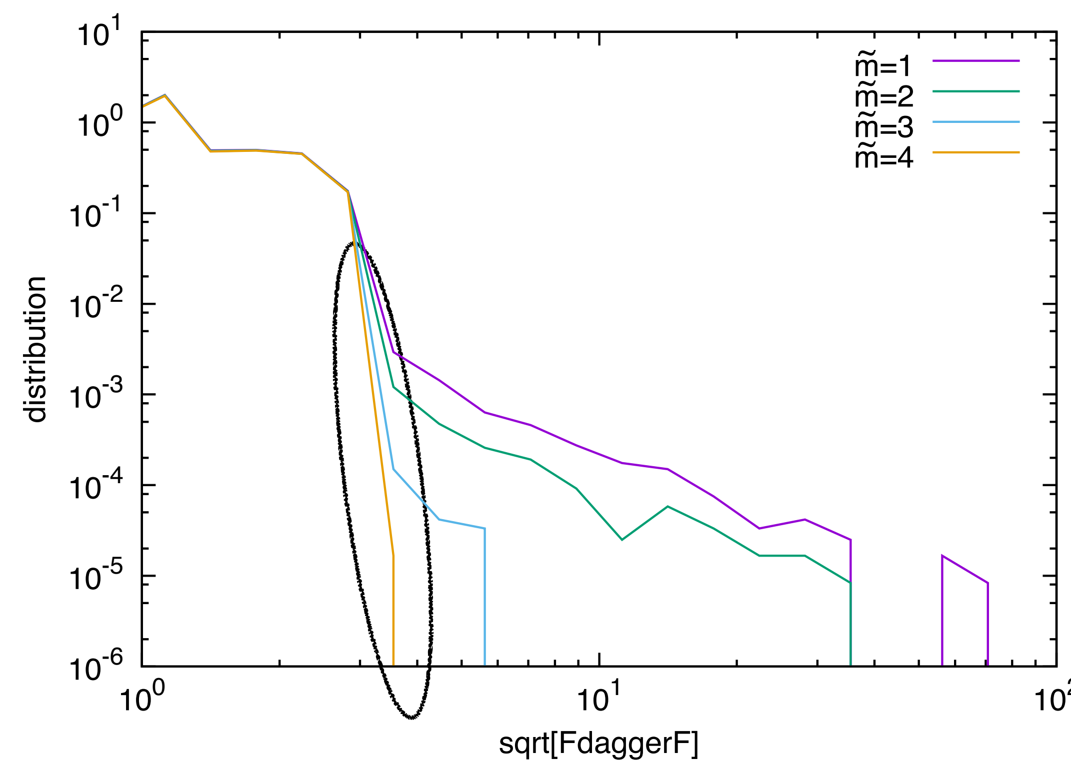
Demonstration of our condition in CRMT

$\Phi_1,\ \Phi_2,\ \Psi_1,\ \Psi_2$ $\rightarrow$ drift F_i (i=1,2,3,4)

eigenvalues of $\sqrt{F_i^\dagger F_i}$ (i=1,2,3,4)

with a cooling (red points)





Summary

- We have established the argument for justification of the CLM by starting with a finite step-size ε
- We find that in order for the CLM to work, the probability distribution of the drift should fall off exponentially or faster.
- We demonstrate the validity of our condition in simple models involving the singular drift problem or the excursion problem and Chiral Random Matrix Theory.
- Importantly, our condition turns out to be practically useful in the CLM for lattice QCD → Nagata's talk (next)