The complex Langevin method: successes and open problems

Erhard Seiler

Max-Planck-Institut für Physik, München

(Werner-Heisenberg-Institut)

part of collective effort involving G. Aarts, F. Atanasio, L. Bongiovanni, P. Giudice,

B. Jäger, J. Pawlowski, D. Sexty, I.-O.Stamatescu,....

1. Sign problem

Functional measure $\rho \propto e^{-S}$ in Euclidean QFT not always positive:

- Real time Feynman integral
- Topological terms – nonzero vacuum angle θ
- Finite density - chemical potential
- $\ldots$

ρ Signed or Complex measure.

General Idea

(L. L. Salcedo 1993, 1997, 2007 Weingarten 2002):
Replace
complex (signed) measure ρ on $\mathcal{M}$ by
probability measure P on complexification $\mathcal{M}_c$
such that for holomorphic observables $\mathcal{O}$

$$\boxed{\langle\mathcal{O}\rangle \equiv \int_{\mathcal{M}} \mathcal{O}\rho d\mu = \int_{\mathcal{M}_c} \mathcal{O} dP\,.}$$

Note: P underdetermined.

Primitive example

$$\rho(x) = 1 + \kappa \cos x$$

For $\kappa \leq 1$: Real LE works. For $\kappa > 1$: Sign problem!

$$K(z) = -\frac{\sin z}{1+\kappa \cos z}$$

Claim:

$$P(x,y) \equiv (1 + \cos x)\frac{1}{\sqrt{2\pi\sigma}} \exp\left[\frac{-y^2}{2\sigma}\right] , \quad \sigma = 2 \log \kappa$$

solves problem for $\kappa > 1$:

Check elementary

More useful:

Complex Langevin (G. Parisi 1983, J. Klauder 1983):
Works 'in principle'.

Recent successes include:

- HDM approximation for QCD (β not too small) (E. S., D. Sexty, I.-O. Stamatescu 2012)
- Full QCD (β not too small) (D. Sexty 2013)

Important tool: Gauge cooling

More on this: talk by G. Aarts

2. Definition and justification

'Flat' case: defined on $\mathcal{M} = \mathbb{R}^n$ or $\mathcal{M} = U(1)^n$.
analytic extension of $\mathcal{M}$: $\mathcal{M}_c$.

Complex Langevin on $\mathcal{M}_c$

$$dz = K dt + dw, \quad K = -\nabla S$$

dw real Wiener increment $dw = \eta(t)dt$, η white noise).

$$\begin{aligned} dx =& K_x dt + dw, \quad & K_x = \mathrm{Re}\, K \\ dy =& K_y dt, \quad & K_y = \mathrm{Im}\, K \end{aligned}$$

real stochastic process on $\mathcal{M}_c$.

Evolution of observables $\mathcal{O}$: By Ito calculus

$$\langle\dot{\mathcal{O}}\rangle = \langle L\mathcal{O}\rangle$$

L real Langevin operator

$$L \equiv [\nabla_x + K_x]\,\nabla_x + K_y\nabla_y$$

$\mathcal{O}(z)$ holomorphic: $\nabla_y\mathcal{O} = i\nabla_x\mathcal{O}$ Cauchy-Riemann equation) $\Longrightarrow$

$$L\mathcal{O} = L_c\mathcal{O}$$

L_c complex Langevin operator

$$L_c \equiv [\nabla_x + K]\nabla_x = [\nabla_z + K]\nabla_z$$

Evolution of densities:

Positive density P:

$$\frac{\partial}{\partial t} P(x,y;t) = L^T P(x,y;t)\,; \quad P(x,y;0) = \delta(x-x_0)\delta(y)\,,$$

$L^T \equiv \nabla_x[\nabla_x - K_x] - \nabla_y K_y$ real Fokker-Planck operator.

Complex density ρ:

$$\frac{\partial}{\partial t} \rho(x;t) = L_c^T \rho(x;t)\,; \quad \rho(x;0) = \delta(x-x_0),$$

$L_c^T \equiv \nabla_x\left[\nabla_x + K\right]$ complex Fokker-Planck operator.

Relation of evolutions

$$\langle \mathcal{O} \rangle_{P(t)} \equiv \frac{\int \mathcal{O}(x) P(x,y;t) dx\, dy}{\int P(x,y;t) dx\, dy}, \quad \langle \mathcal{O} \rangle_{\rho(t)} \equiv \frac{\int \mathcal{O}(x) \rho(x;t) dx}{\int \rho(x;t) dx}.$$

Two time evolutions:

$$\partial_t \langle \mathcal{O} \rangle_{\rho(t)} = \int dx \mathcal{O}(x)\, L_c^T \rho(x;t)$$

$$\partial_t \langle \mathcal{O} \rangle_{P(t)} = \int dx dy \mathcal{O}(x+iy) L^T P(x,y;t)\,.$$

Consistent? Formally yes: use CRE and integration by parts.

Result (not fully rigorous)

$$\boxed{\langle \mathcal{O} \rangle_{\rho(t)} = \langle \mathcal{O} \rangle_{P(t)} \quad \forall\, t \geq 0}$$

Requirements:

- agreement of initial conditions
- holomorphy of drift $K \equiv K_x + iK_y$
- sufficient decay of $P\mathcal{O}$ at imaginary infinity

Idea of proof

Interpolate between evolutions of P and $\mathcal{O}$:

1. Initial conditions agree.

2. Let $\mathcal{O}(x+iy;t) \equiv \exp\left[tL\right]\mathcal{O}(x+iy)$ be unique solution of DE

$$\partial_t \mathcal{O}(x+iy;t) = L\mathcal{O}(x+iy;t) \quad (t \geq 0)\,;$$

3. Consider $F(t,\tau) \equiv \int P(x,y;t-\tau)\mathcal{O}(x+iy;\tau)$.
Interpolates between $\langle \mathcal{O} \rangle_{P(t)}$ and $\langle \mathcal{O} \rangle_{\rho(t)}$:

$$F(t,0) = \langle \mathcal{O} \rangle_{P(t)}; \quad F(t,t) = \langle \mathcal{O} \rangle_{\rho(t)}$$

Formally: $F(t,\tau)$ independent of τ:

$$\frac{\partial}{\partial\tau}F(t,\tau) = -\int L^T P(x,y;t-\tau)\mathcal{O}(x+iy;\tau)dxdy$$
$$+\int P(x,y;t-\tau)L\mathcal{O}(x+iy;\tau)dxdy$$

Integration by parts and holomorphy of $\mathcal{O}(z;t) \Rightarrow$

$$\boxed{\frac{\partial}{\partial\tau}F(t,\tau) = 0} \quad \Longrightarrow \quad \langle\mathcal{O}\rangle_{\rho(t)} = \langle\mathcal{O}\rangle_{P(t)}$$

Assumption: no boundary terms

3. Consistency conditions

Recall

$$0 = \frac{\partial}{\partial\tau} F(t,\tau) = -\int L^T P(x,y;t-\tau)\mathcal{O}(x+iy;\tau)dx\,dy$$
$$+\int P(x,y;t-\tau)\mathcal{O}(x+iy;\tau)dx\,dy\,.$$

Take $\tau = 0, t \to \infty$, assume convergence to equilibrium:

$$\exp(tL^T)P(x,y;t) \to P(x,y;\infty); \quad L^T P(x,y;\infty) = 0\,,$$

$$\boxed{\langle\dot{\mathcal{O}}\rangle = \langle L\mathcal{O}\rangle \equiv \int P(x,y;\infty)L\mathcal{O}(x+iy)dx\,dy = 0.} \quad \text{(CC)}$$

Expresses stationarity of noise averaged observables.

A "quasi-theorem"

Consider compact $\mathcal{M}$. If

- CC hold for a dense (in sup norm) set of observables and
- $\left|\int_{\mathcal{M}_c} P\mathcal{O}\right| \leq C \sup_{\mathcal{M}} |\mathcal{O}|$

Then

$$\int_{\mathcal{M}_c} P\mathcal{O} = \frac{1}{Z}\int_{\mathcal{M}} \mathrm{e}^{-S}\mathcal{O}\,.$$

i.e. Equilibrium measure correct.

"Proof":

For simplicity consider $\mathcal{M} = S_1^n$, Function space $\mathcal{C} = \mathcal{C}(\mathcal{M})$ (continuous functions)

$$||f|| = \sup_{x \in \mathcal{M}} |f|$$

Assume (CC) for $\mathcal{O} \in \mathcal{D} \subset \mathcal{C}$, $\mathcal{D}$ dense, $L\mathcal{O} \subset \mathcal{D}$, and $\exists C$ such that

$$|\langle \mathcal{O} \rangle| \leq C ||\mathcal{O}||$$

Riesz-Markov: $\langle \cdot \rangle$ given by complex measure σdx on $\mathcal{M}$.
(CC) $\Rightarrow \int_{\mathcal{M}} (L\mathcal{O}) d\sigma(x) = 0$.
Integration by parts on S_1^n (harmless): $L_c^T \sigma = 0$.
Usual spectral properties of $L_c^T \Longrightarrow \sigma \propto \exp(-S)$. $\square$

4. Problem #1: slow decay

Typical:
$\mathcal{M}$ compact, $\mathcal{M}_c$ noncompact
Example: $\mathcal{M} = SU(N)$, $\mathcal{M}_c = SL(N, \mathbb{C})$

Note:
Holomorphic functions grow $\Longrightarrow$
Drift K grows; observables $\mathcal{O}$ as well $\Longrightarrow$
Large excursions possible

“Skirts”, “tails” of distribution P on $\mathcal{M}_c$.

Integration by parts without boundary terms:
Questionable

Simple examples

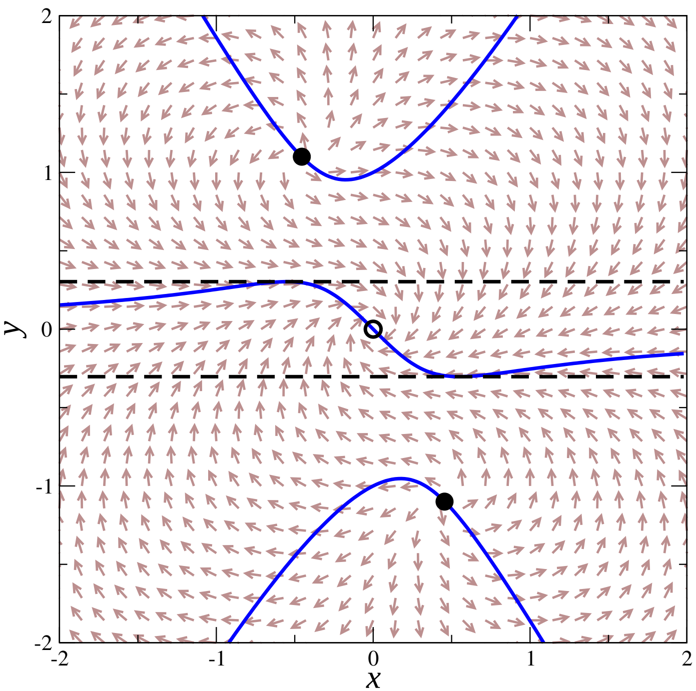
Quartic model: $\mathcal{M} = \mathbb{R}$, $\mathcal{M}_c = \mathbb{C}$:

$$S = \frac{1}{2}\sigma x^2 + \frac{1}{4}\lambda x^4, \quad \sigma = A + iB, \;\; \lambda = 1\,.$$

G. Aarts, P. Giudice, E. S. 2013

Lucky case

$3A^2 > B^2$: Process confined in strip.

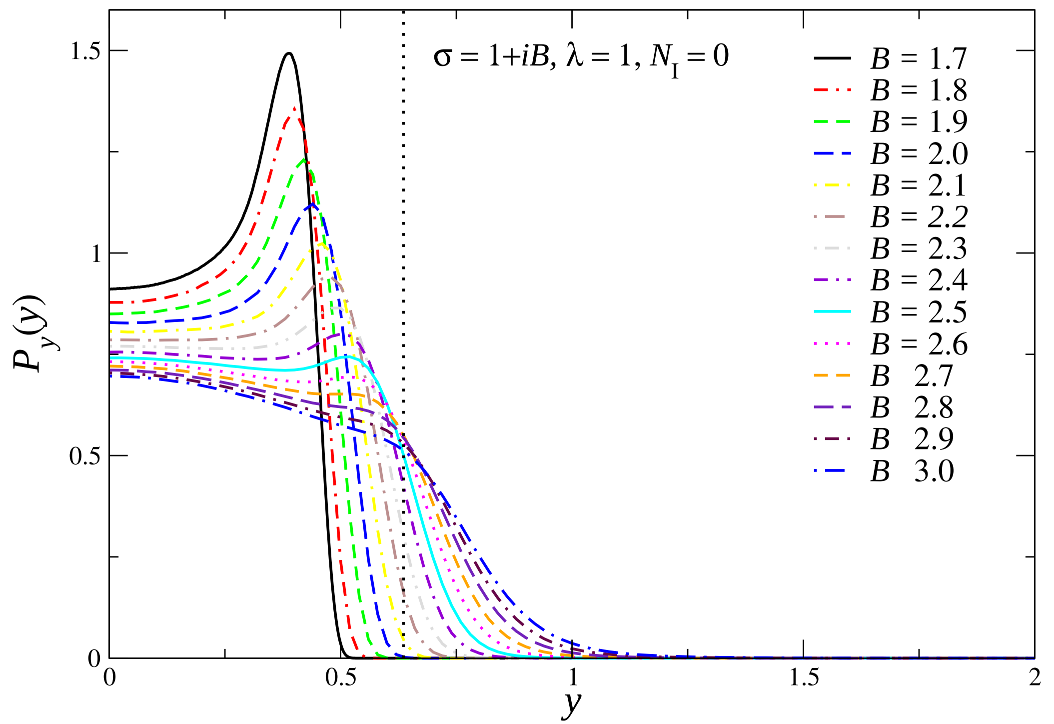


$\sigma = 1 + i, \lambda = 1$. Solid lines: $K_y = 0$.

CLE results correct

General case

Process not always confined in strip.



$\sigma = 1 + iB, \lambda = 1$. Tail $O((x^2 + y^2)^{-3})$.

CLE results deteriorate for $3A^2 < B^2$.

5. Gauge cooling

Eliminates some skirts in gauge models

Polyakov loop model:

Chain of N links, periodic bc. Analytically 1-link integral.

$$-S = \beta_1 \mathrm{tr}\,(U_1 \ldots U_N) + \beta_2 \mathrm{tr}\,(U_N^{-1} \ldots U_1^{-1})$$

$U_i \in SU(3), \;\; i = 1, \ldots N, \quad \beta_{1,2} \in \mathbb{C}$.

$$\beta_1 = \beta + \kappa e^{\mu}, \quad , \beta_1 = \beta^* + \kappa e^{-\mu}$$

For β complex, S complex.

For large N, simple CLE simulation fails.

Reason: process wanders very far from $SU(3)$.

Unitarity norm(s)

Quantify nonunitarity by

$$F(\{U\}) \equiv \sum_i \mathrm{tr}\left[U_i{}^\dagger U_i + (U_i{}^\dagger)^{-1} U_i{}^{-1} - 2 \right] \geq 0\,,$$

Idea: use $SL(3,\mathbb{C})$ gauge transformations to reduce F.

$$U_i \mapsto \exp(\alpha_i \lambda_a) U_i, \quad U_{i-1} \mapsto U_{i-1} \exp(-\alpha_i \lambda_a)$$

'Gauge gradient' at i

$$\begin{aligned} G_{a,i} \equiv & 2\,\mathrm{tr}\lambda_a \left[U_i U_i^\dagger - U_{i-1}^\dagger U_{i-1} \right] \\ & + 2\,\mathrm{tr}\lambda_a \left[-(U_i^\dagger)^{-1} U_i^{-1} + (U_{i-1}^\dagger)^{-1} U_{i-1}^{-1} \right]\,. \end{aligned}$$

Cooling step

$$U_i \mapsto \exp\left(-\sum_a \tilde{\alpha}\lambda_a G_{a,i}\right) U_i$$

$$U_{i-1} \mapsto U_{i-1} \exp\left(\sum_a \tilde{\alpha}\lambda_a G_{a,i}\right),$$

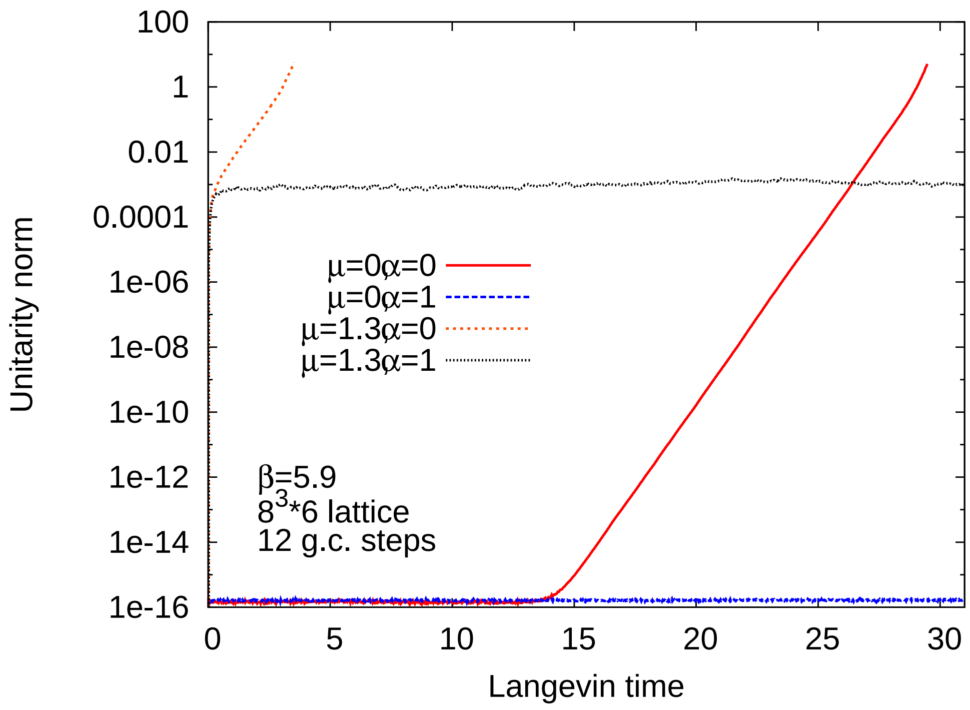
$\alpha = \epsilon\tilde{\alpha}$, ϵ discretization parameter

Two possibilities:

(1) Gauge cool appropriately between dynamical updates

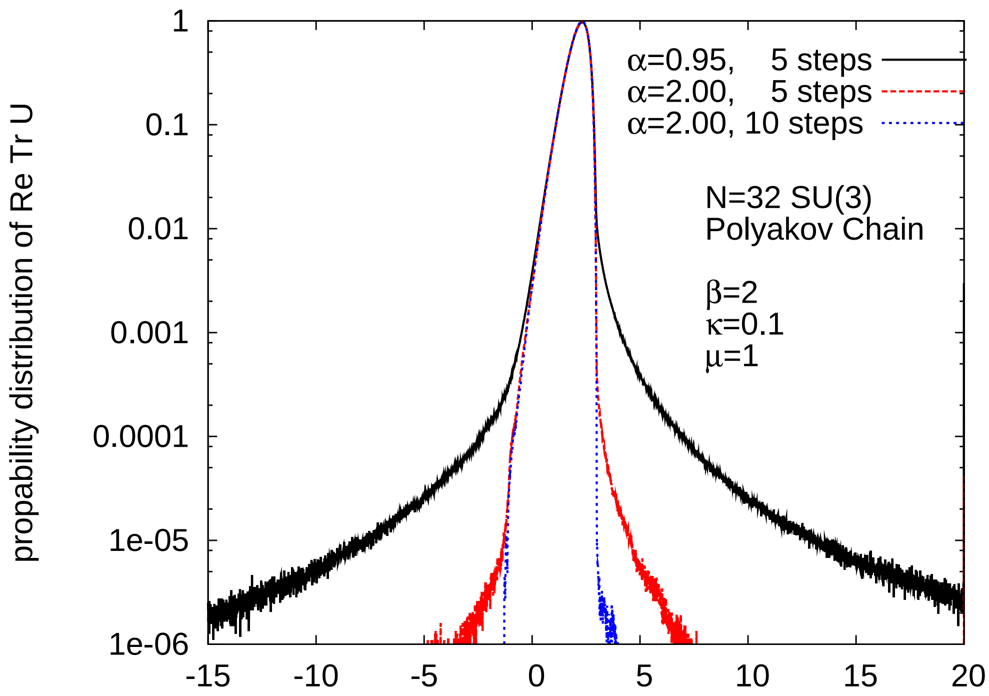
(2) Add 'cooling drift' to dynamical drift

Cooling keeps U norm in check

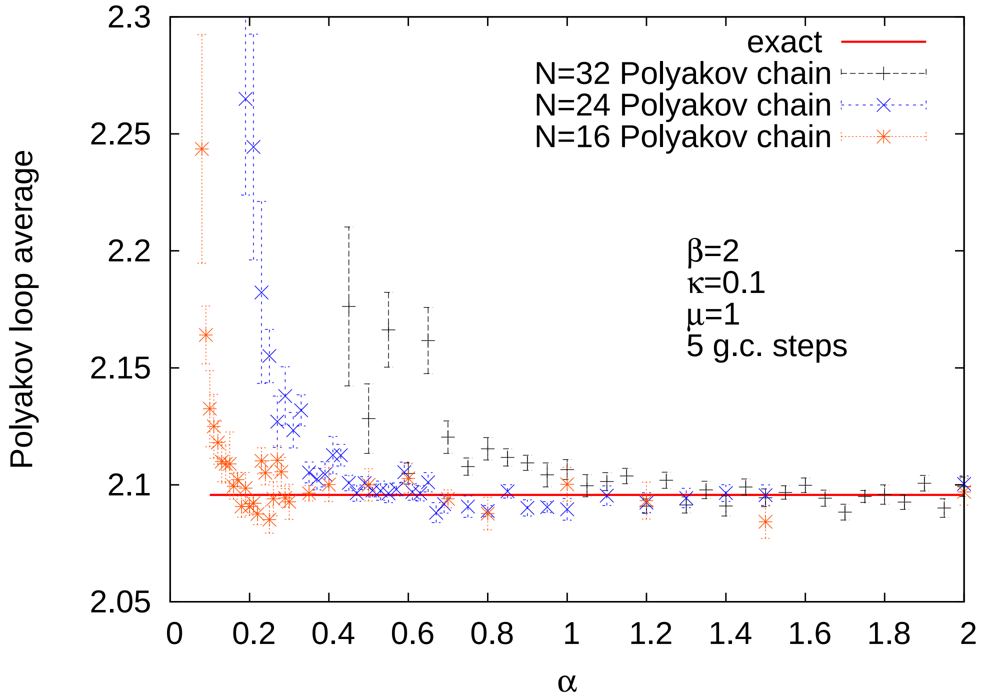


Cooling reduces skirts

Histograms of Polyakov loops

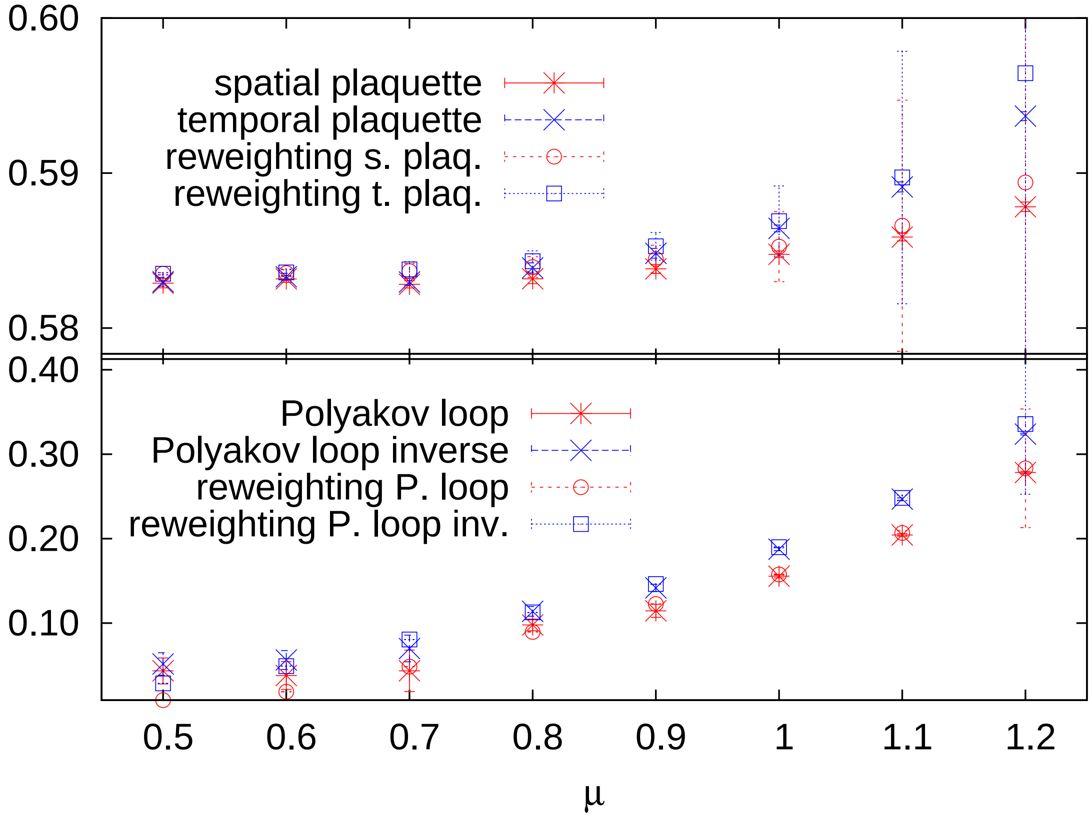


Cooling produces correct results



Gauge cooling in HD QCD

Validated by comparison with reweighting



Lattice 6^4, $\beta = 5.9, \kappa = 0.12$. Reweighting breaks down at $\mu \gtrsim 1.2$. Agreement deteriorates from smaller β.

More on successful applications of CL with gauge cooling in HDQCD and QCD:

G. Aarts's talk next week.

6. Problem #2: poles in drift

If ρ has zeroes in $\mathcal{M}_c$

$\Longrightarrow$ drift only meromorphic (positive integer residues)

$\Longrightarrow$ Problem:

$\dot{\mathcal{O}} = L\mathcal{O}$ does not preserve holomorphy of $\mathcal{O}$, justification of CLE destroyed.

Full QCD:

Fermion determinant

$$\det(\not{D}_U + M)$$

generically vanishes for some $U \in SL(3, \mathbb{C})$.

But: D. Sexty 2013 finds in QCD: eigenvalues avoid 0.

Real models

Ambjørn, Flensburg & Peterson (1986) studied CLE for

$$\rho(x) \equiv \exp(-S) = \cos(x)\exp[\beta\cos(x)]\,.$$

ρ real signed measure. AMP find “disaster”.

Unavoidable:

Real axis attractor $\Longrightarrow$ equilibrium density

$$P(x,y) = \delta(y)\sigma(x), \quad \sigma(x) \geq 0\,,$$

Incompatible with $\rho(x)$, CLE must fail!

Of course: RLE fails as well!

Flower, Otto&Callahan(1986):
'segregation phenomenon'
K. Fujimura et al (1994):
attempted cure adding $i\pi\delta(\cos x)$ to drift –
does not work in general.

We (2013) find:
Mathematically process not ergodic. Stationary FPE has two linearly independent solutions:

$$P_+(x) = \rho(x)\theta(\rho(x)) \quad \text{and} \quad P_-(x) = \rho(x)\theta(-\rho(x))\,;$$

Numerically: Get phase quenched result
$P \propto P_+(x) + P_+(x)$.

7. *Cures for real models*

Cure #1: “Sign reweighting”

$$\rho(x) = \theta(\rho(x))P_+(x) - \theta(-\rho(x))P_-(x)$$

whereas simulation yields

$$\rho(x) = \theta(\rho(x))P_+(x) + \theta(-\rho(x))P_-(x)\,,$$

$\Longrightarrow$ the following reweighting should work:

$$\langle\mathcal{O}\rangle_{rew} \equiv \frac{\mathcal{O}(x+iy)\mathrm{sgn}\rho(x)\rangle}{\langle\mathrm{sgn}\rho(x)\rangle}\,.$$

$\langle\cdot\rangle$: ordinary Langevin average.

Works for toy models.

Cure #2 for compact real models:
"Shifting poles"
Idea: Choose c s.t.

$$\sigma(x) \equiv \rho(x) + c \geq 0 \,,$$

Consider $\mathcal{O}$ with $\int \mathcal{O} = 0$. Then

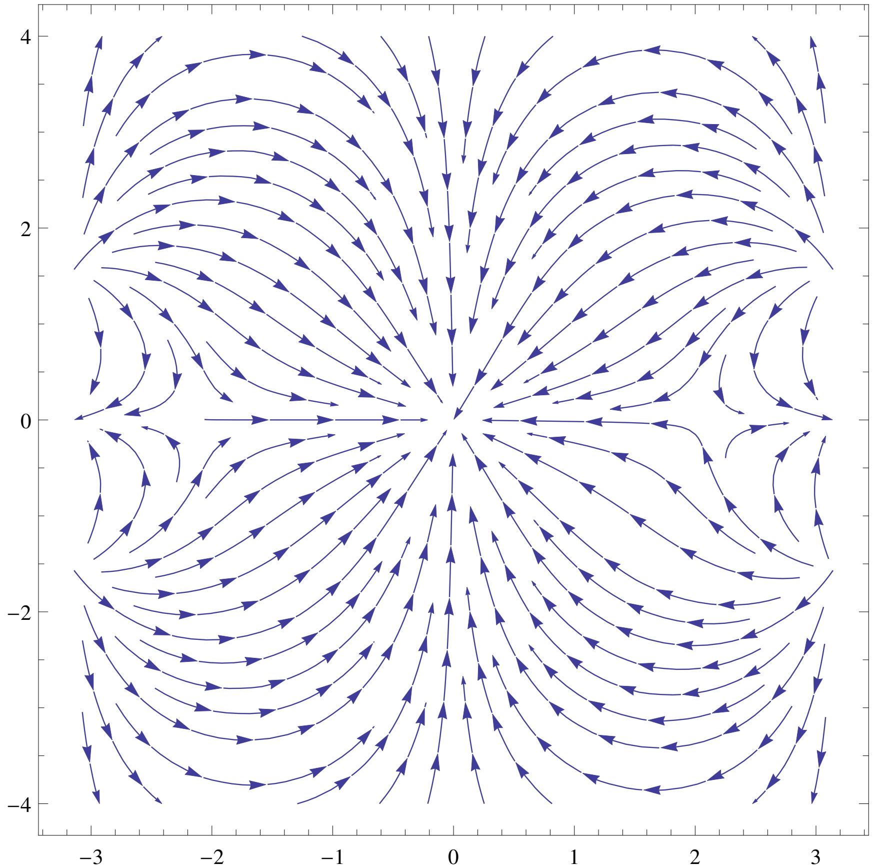
$$\langle \mathcal{O} \rangle_\rho = \frac{\langle \mathcal{O} \rangle_\sigma}{\langle \rho/\sigma \rangle_\sigma} \,,$$

Different kind of reweighting: Change drift from $K = \rho'/\rho$ to

$$K_\sigma = \frac{\sigma'}{\sigma} = \frac{\rho'}{\sigma} \,.$$

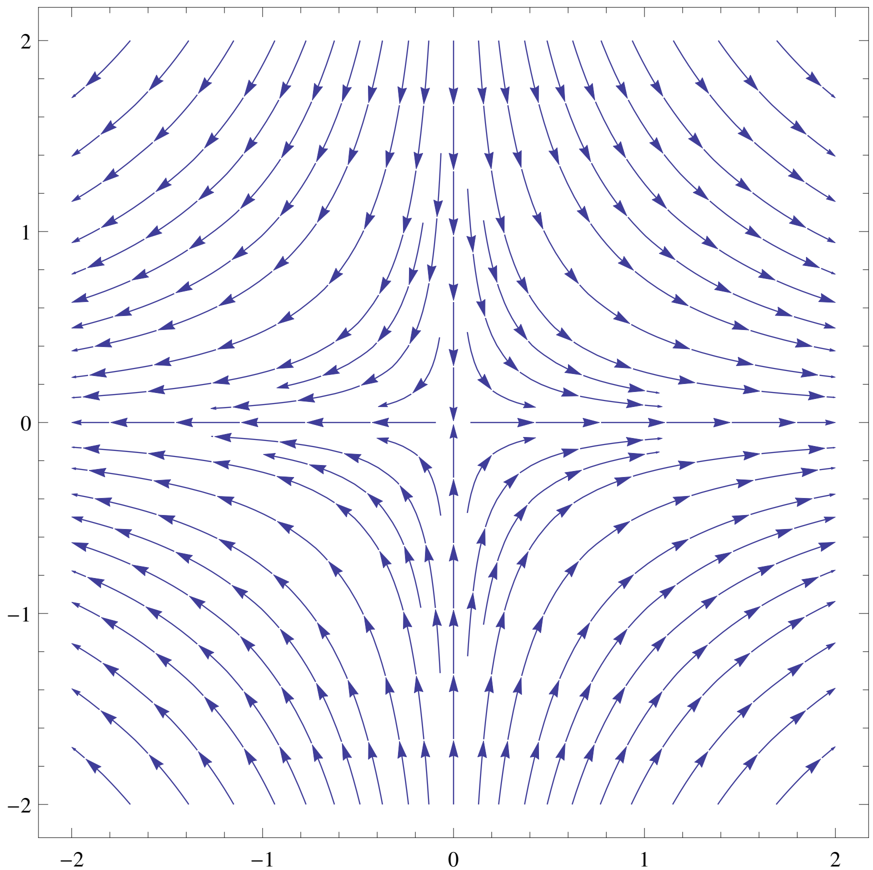
Poles pushed way from real axis!

Flow pattern for $\alpha = 0$, ($h = 0, \kappa = 2, \mu = 0, \beta = 0.5$):



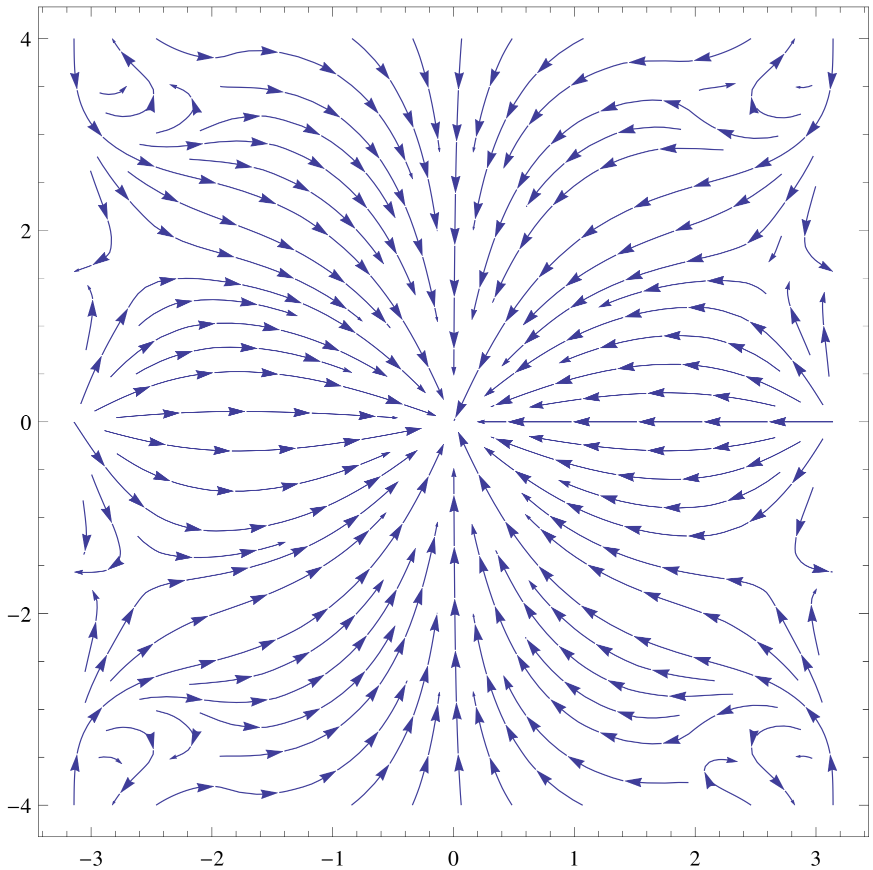
Poles at $x = \pm 2.0944, y = 0$

Flow near pole



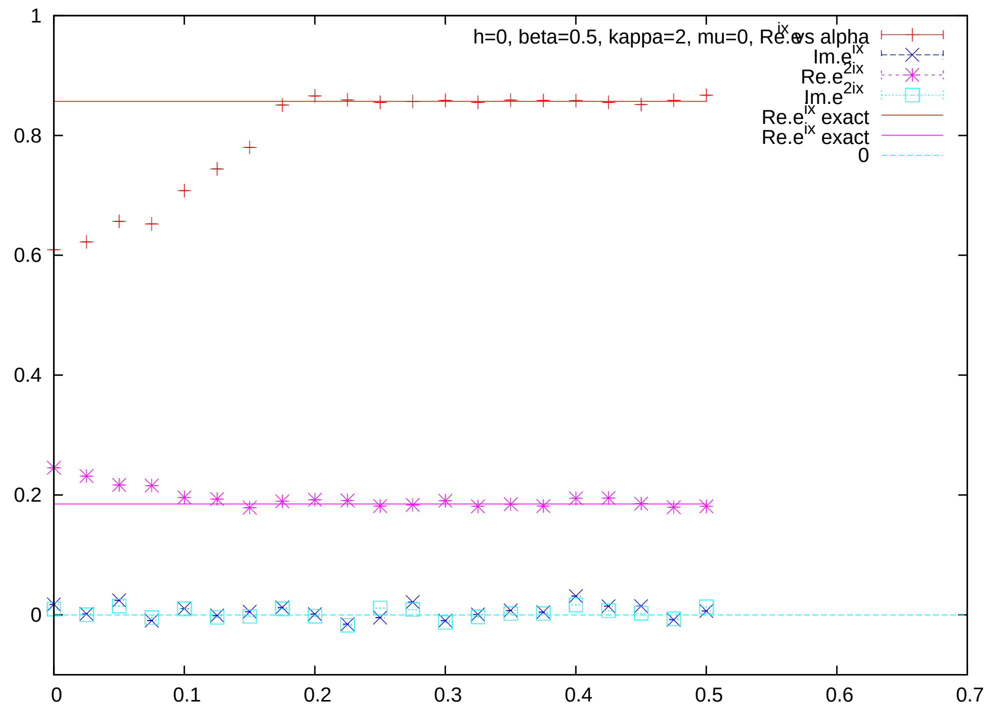
Pattern characteristic for any pole with positive residue.

Flow pattern for $\alpha = 0.4$, ($h = 0, \kappa = 2, \mu = 0, \beta = 0.5$):



Poles at $x = \pm 2.68297, y = \pm 1.60041$

Numerical example:



$\kappa = 2, \mu = 0, \beta = 0.5$;

data points: CLE with cure #2 vs $\alpha = ce^{-\beta}/\kappa$,

solid lines: exact results.

Upshot

Real models can be cured.

Ambjørn-Flensburg-Peterson 'quantum mechanical desasters' averted.

But: Reweighting painful for lattice models

8. Complex case:

How bad are poles?

Toy model

$$\rho(x) \equiv \exp(-S) = (1 + \kappa \cos(x - i\mu)) \exp[\beta \cos(x)] \, .$$

Poles:

(1) $\kappa \leq 1$: $z_P = \pm\pi + iy_P$

(2) $\kappa > 1$: $z_P = x_P + i\mu, \; x_p \neq \pm\pi$

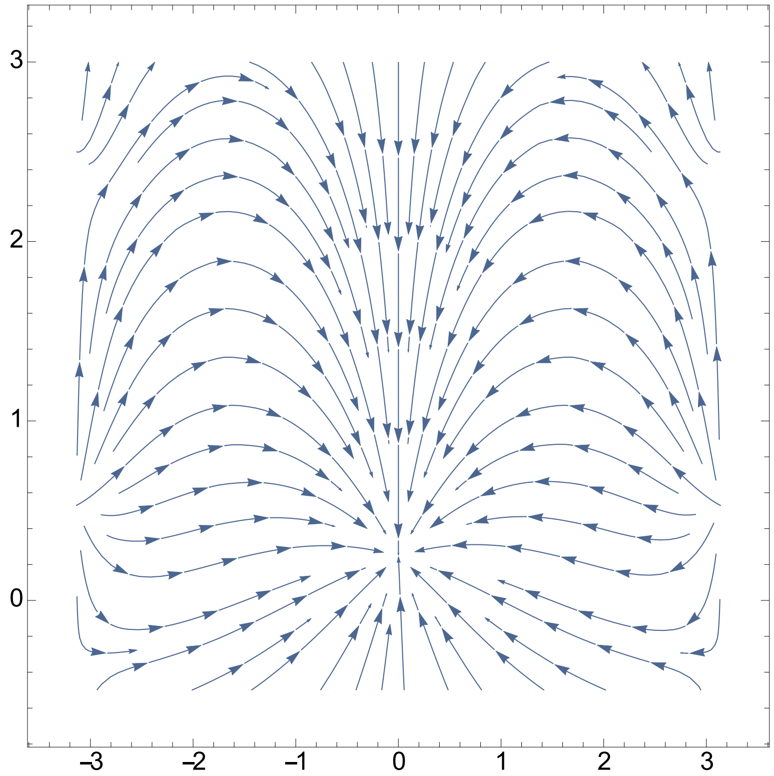
Three examples:

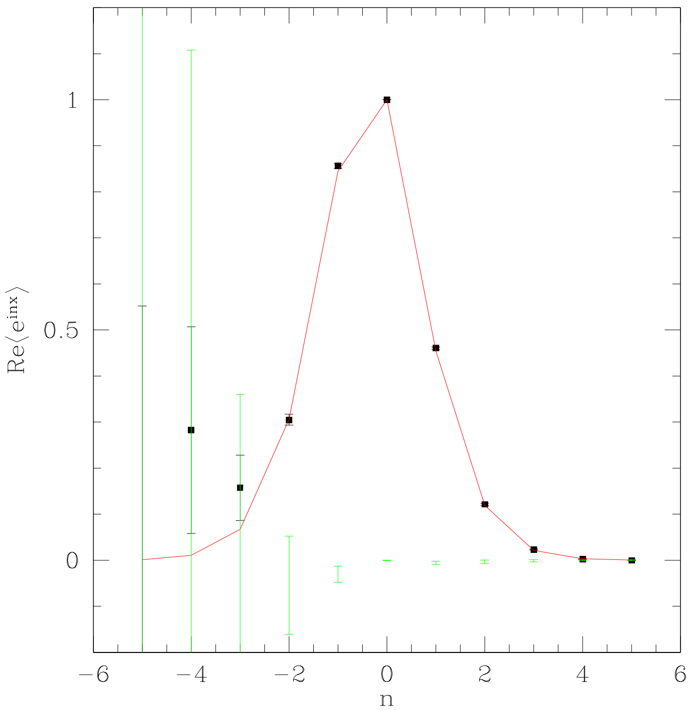
$\kappa = 0.5, \beta = 1, \mu = 1$

$\kappa = 2, \beta = 0.3, \mu = 1$: the worst case

$\kappa = 2, \beta = 5, \mu = 1$

(a) $\kappa = 0.5, \beta = 1, \mu = 1$

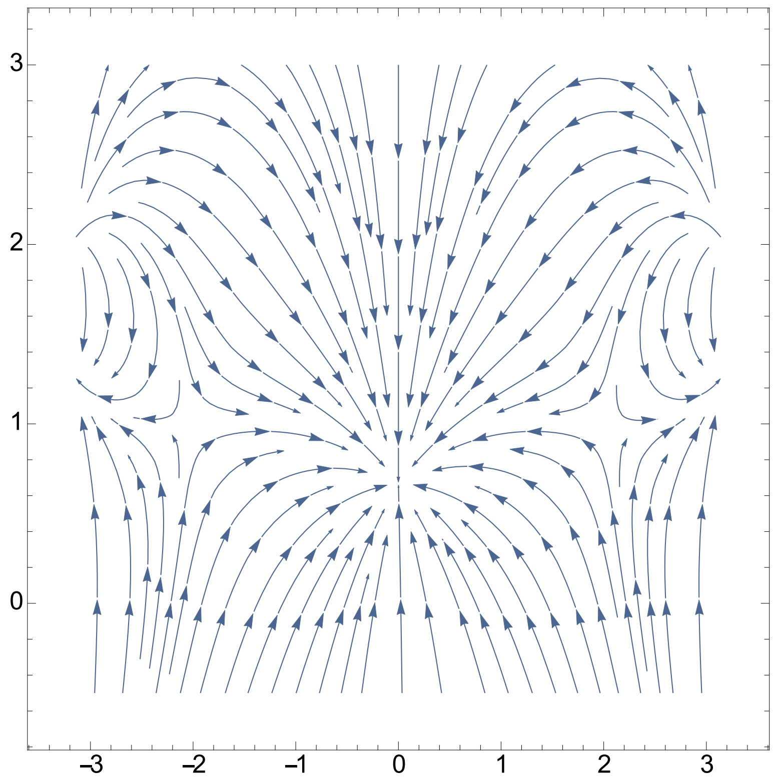


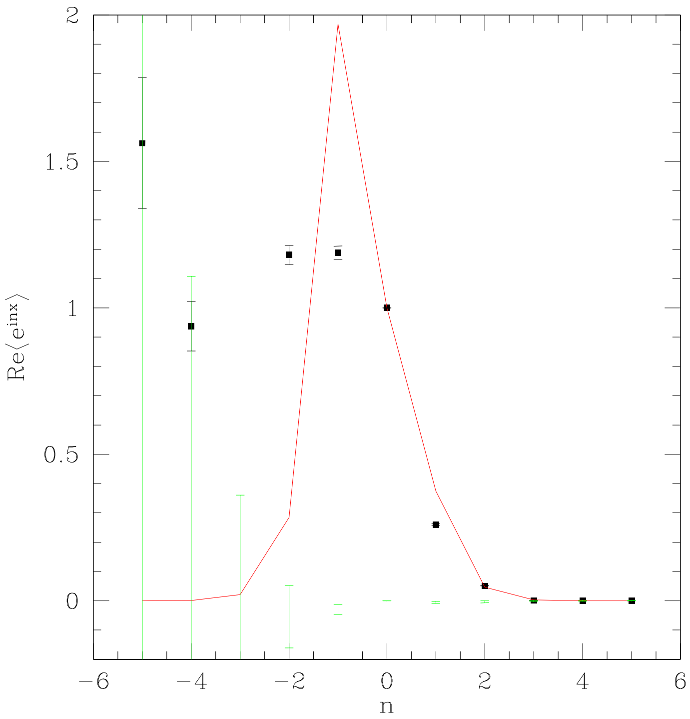


Left: flow pattern; only 1 attractive fixed point; large excursions upwards possible.
Right: $\langle e^{inz} \rangle$ (red) and CC (green)

Negative modes poor for $-n \geq 3$

(b) $\kappa = 2., \beta = 0.3, \mu = 1$

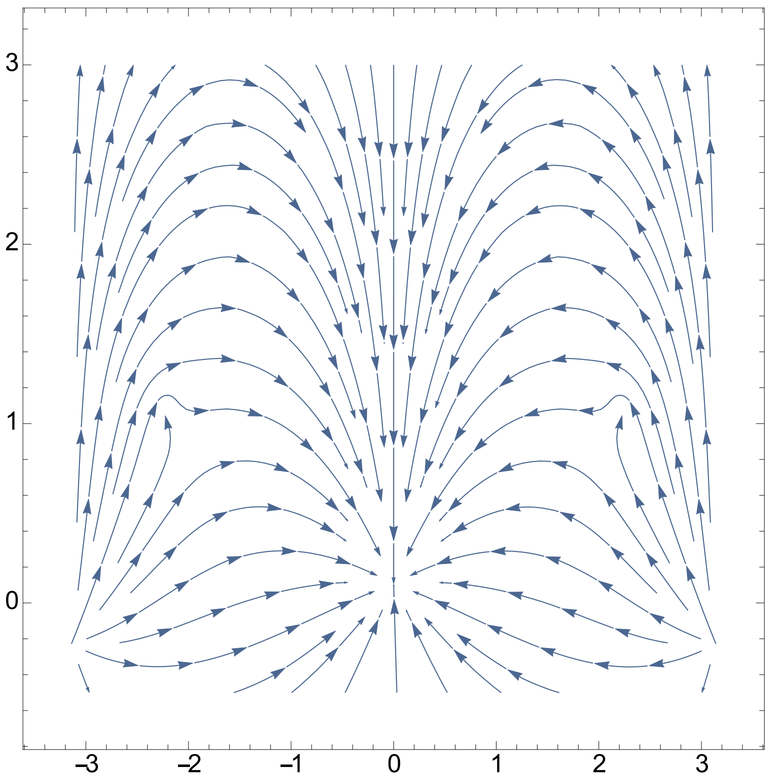


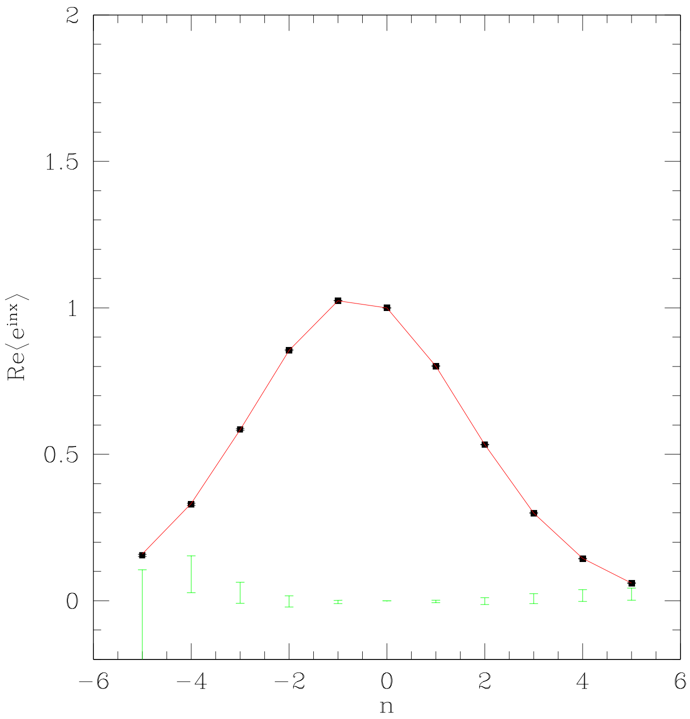


Left: flow pattern; secondary attractive fixed point at $\pm\pi + 1.2i$. Right: $\langle e^{inz}\rangle$ (red) and CC (green)

All negative modes deviate and grow, CC alright

(c) $\kappa = 2, \beta = 5, \mu = 1$



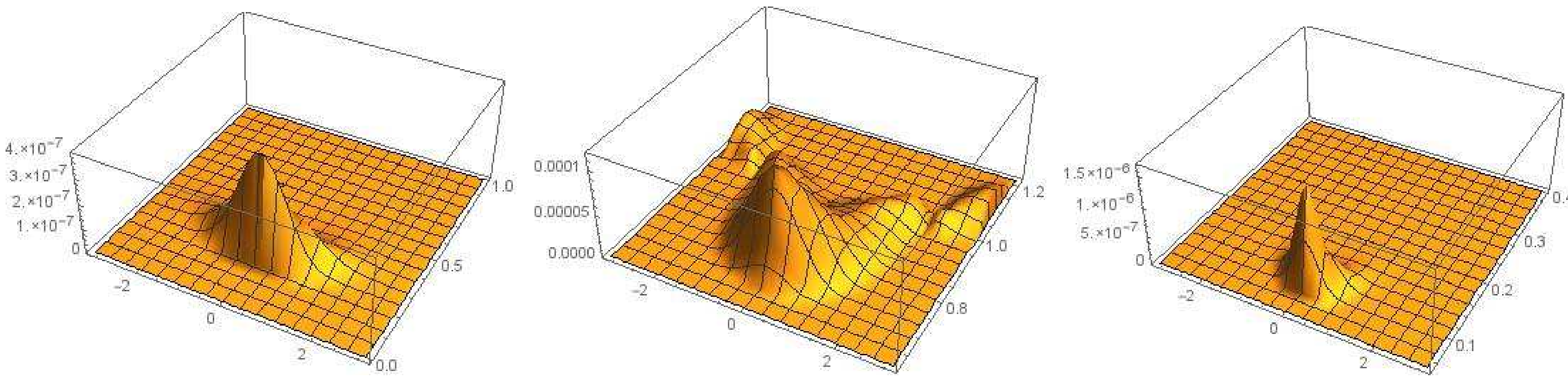


Left: flow pattern; only one attractive fixed point.

Right: $\langle e^{inz} \rangle$ (red) and CC (green)

Modes up to $|n| = 5$ good

Histograms



Left: $\kappa = 0.5, \beta = 1, \mu = 1$

Middle: $\kappa = 2, \beta = 0.3, \mu = 1$

Right: $\kappa = 2, \beta = 5, \mu = 1$

A paradox?

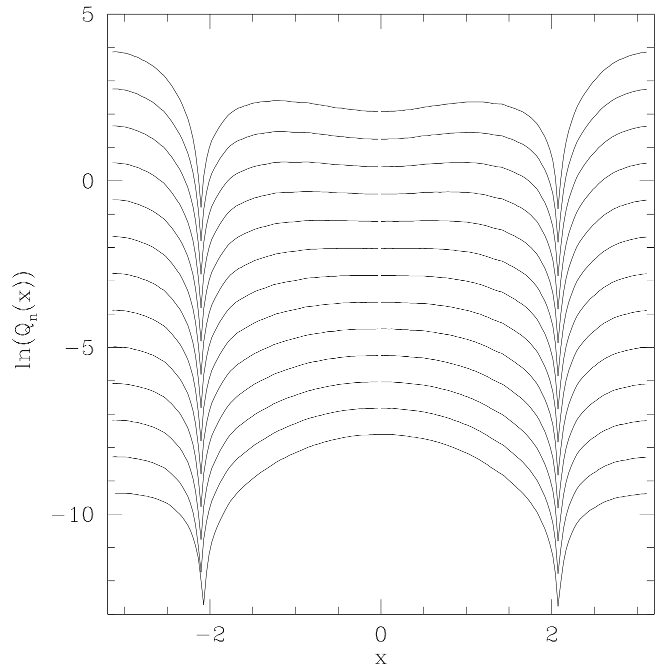
CC always seem to be True (but noisy),
even if results wrong.
Quasi-theorem: bounds must fail.
Claim: $\langle e^{inx} \rangle$ will grow with $|n|$.

$$Q_n(x) \equiv \int dy P(x,y) e^{-ny} .$$

$$\langle e^{inx} \rangle = \int dx Q_n(x) e^{inx} .$$

Worst case $\kappa = 2, \beta = 0.3, \mu = 1$:
Q_n grow exponentially in n.

Plotting Q_n



$\ln Q_n(x)$ from $n = 0$ to $n = -12$

Q_n grows exponentially and has kinks $\Longrightarrow$
Fourier transform decay only powerlike $\Longrightarrow$ Bounds fail

What have we learned?

- Poles can lead to wrong convergence
- CC not sufficient to exclude wrong convergence
- Poles harmless if process stays away from them (Møllgaard & Splittorff 2013)
- Poles harmless if obervables small near them
- Small κ helps
- Large β helps

9. Cures for complex case?

Cure #1 generalized to complex situations:
Reweight

$$\mathcal{O}(z) \mapsto \mathcal{O}(z)\,\mathrm{sgn}\,c(z)$$

with (for instance)

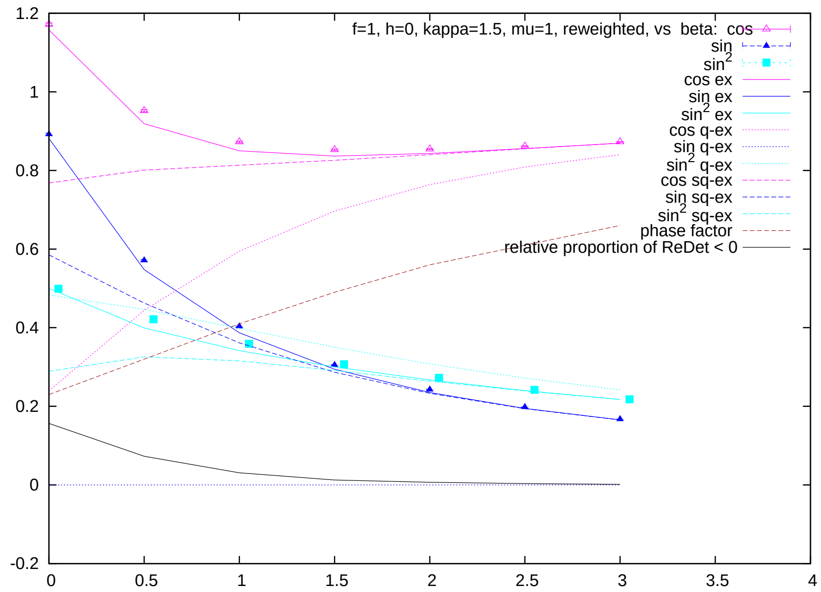
$$c(z) = \mathrm{sgn}\,\mathrm{Re}\,\det(x + iy; \kappa, \mu)$$

and compute by CLE

$$\langle \mathcal{O} \rangle_{corr} \equiv \frac{\langle \mathcal{O}\, c \rangle}{\langle c \rangle}\,.$$

$\approx$ works in toy models. Overlap problem for lattices.

Numerical example



Data points: CLE with sign reweighting,
solid lines: unquenched exact results,
dotted lines: quenched results,
dashed lines: quenched results on the line $y = \mu$)

Cure #2:

Consider $\mathcal{O}$ with

$$\int \mathcal{O} = 0\,.$$

Define as before

$$\sigma \equiv \rho + c\,,$$

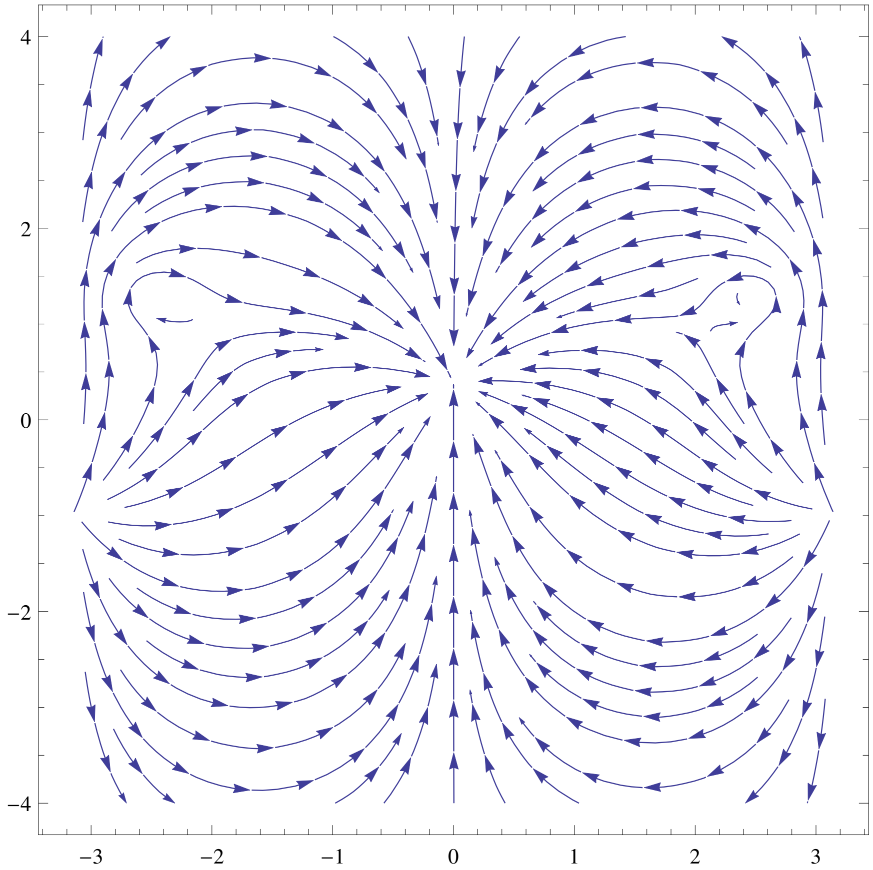
$$\langle \mathcal{O} \rangle_\rho = \frac{\langle \mathcal{O} \rangle_\sigma}{\langle \rho/\sigma \rangle_\sigma}\,.$$

Fixed points not moved; poles shifted.

Møllgaard&Splittorff 2013 (random matrix model):

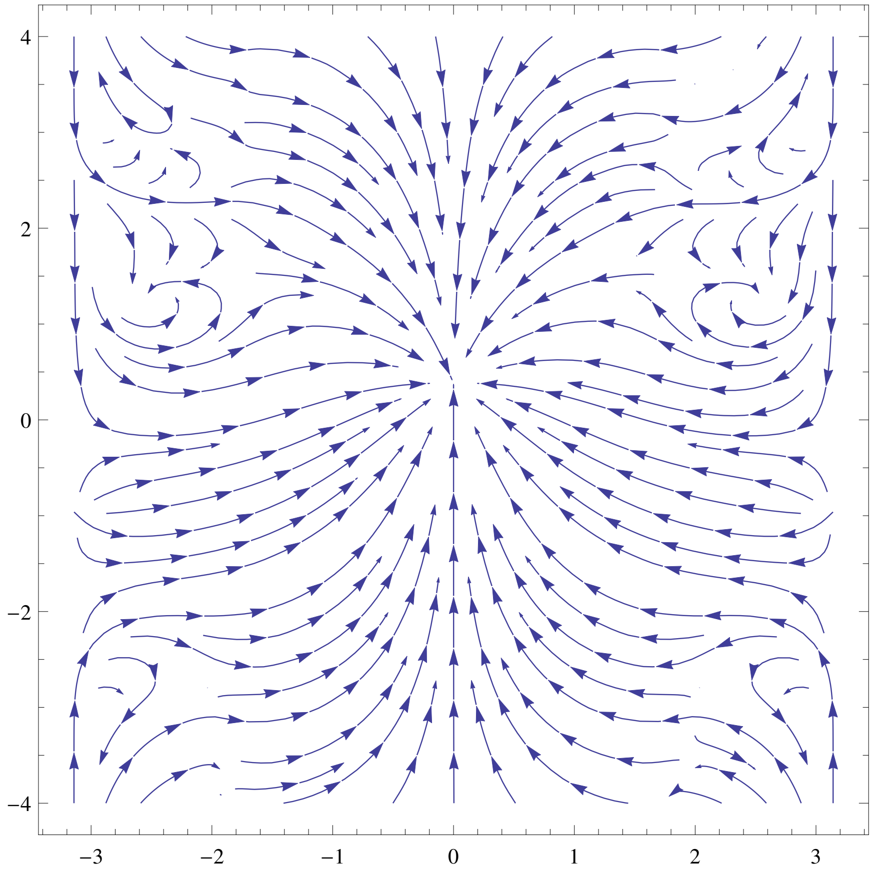
If paths ‘do not wind around’ poles CLE ok.

Flow without shift



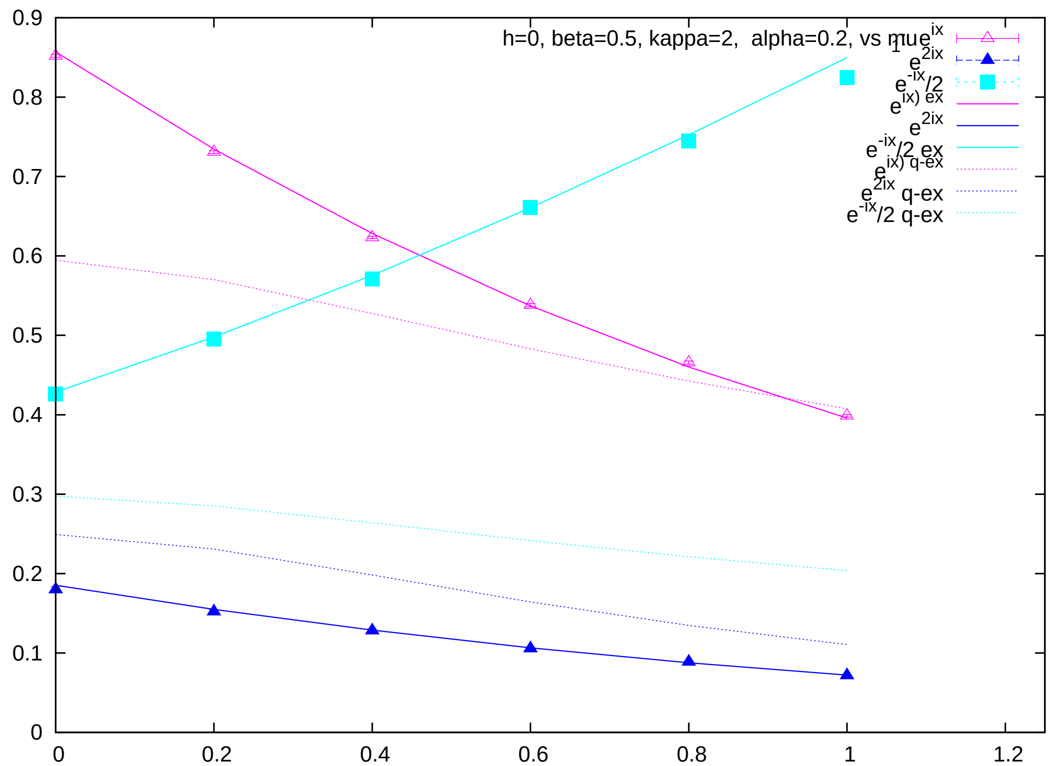
Poles at $x = \pm 2.094395, y = 1$

Flow with shift $c = 1$



Poles at $z = \pm 1.684981 + 1.52266i$ and $x = \pm\pi - 0.261275i$

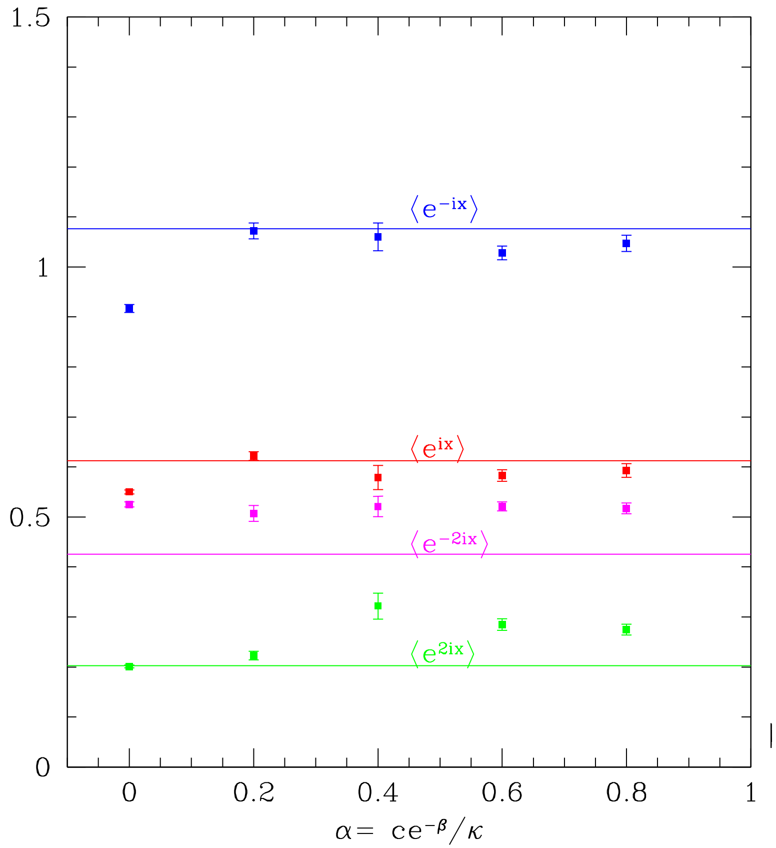
$h = 0,\ \beta = 0.5,\ \kappa = 2.0$, $\mu = 0 \,\mathrm{to}\, 1$



Data points: CLE with shift by $c = \alpha\kappa \exp(\beta) \approx 0.66$, solid lines: exact results.

Looks good

$h = 0,\ \beta = 1.0,\ \kappa = 2.0,\ \mu = 0.5$



Data points: CLE with shift by $c = \alpha\kappa\exp(\beta)$,
solid lines: exact results.
Cure #2 $\approx$ doesn't work very well in complex models.

10. Outlook

- Slow decay partly tamed by gauge cooling

10. Outlook

- Slow decay partly tamed by gauge cooling
- Skirts still have to be monitored

10. Outlook

- Slow decay partly tamed by gauge cooling
- Skirts still have to be monitored
- Poles harmless if process ‘stays away from them’; monitoring costly in QCD

10. Outlook

- Slow decay partly tamed by gauge cooling
- Skirts still have to be monitored
- Poles harmless if process ‘stays away from them’; monitoring costly in QCD
- Cures not perfect for complex case

10. Outlook

- Slow decay partly tamed by gauge cooling
- Skirts still have to be monitored
- Poles harmless if process ‘stays away from them’; monitoring costly in QCD
- Cures not perfect for complex case
- Cures involve reweighting: Bad for QCD

10. Outlook

- Slow decay partly tamed by gauge cooling
- Skirts still have to be monitored
- Poles harmless if process ‘stays away from them’; monitoring costly in QCD
- Cures not perfect for complex case
- Cures involve reweighting: Bad for QCD
- Large β helps: Gaussian FP dominates

10. Outlook

- Slow decay partly tamed by gauge cooling
- Skirts still have to be monitored
- Poles harmless if process ‘stays away from them’; monitoring costly in QCD
- Cures not perfect for complex case
- Cures involve reweighting: Bad for QCD
- Large β helps: Gaussian FP dominates
- Validation important


La
lucha
SIGUE

General U(1) toy models

$$\rho(x) = \sum a_n \mathrm{e}^{inx}, \quad a_n = A_n + iB_n$$

with technical condition on growth of a_n's for $n \to \infty$.
Then $\exists$ solution of the form

$$P(y) = \sum_{n=-\infty}^{\infty} \frac{\lambda_n \mathrm{e}^{inx}}{2\sqrt{2\pi\sigma}} \exp\left(-\frac{(y-1)^2}{2\sigma}\right) + \sum_{n=-\infty}^{\infty} \frac{\mu_n \mathrm{e}^{inx}}{2\sqrt{2\pi\sigma}} \exp\left(-\frac{(y+1)^2}{2\sigma}\right)$$

with $\lambda_0 = \mu_0 = 1/2$ and for $n \neq 0$

$$\lambda_n = \exp\left(-\frac{n^2\sigma}{2}\right)\left\{A_n \frac{\cosh(ny_0)}{\cosh(2ny_0)} - iB_n \sinh(ny_0)\right\},$$

$$\mu_n = \exp\left(-\frac{n^2\sigma}{2}\right)\left\{-A_n \frac{\sinh(ny_0)}{\cosh(2ny_0)} + iB_n \cosh(ny_0)\right\}.$$