Status of the Complex Langevin Method

Erhard Seiler

Max-Planck-Institut für Physik, München

(Werner-Heisenberg-Institut)

With many thanks to my collaborators over many years

Gert Aarts, Dénes Sexty, Ion-Olimpiu Stamatescu

1. "Sign problem"

Functional measure (density) $\rho \propto e^{-S}$ 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 density.

Principle:

Want complex measure $\longrightarrow$ positive measure:

Go Complex!

Concretely:

ρ complex (signed) density on manifold $\mathcal{M}$

search for

$P \geq 0$ probability density on complexification $\mathcal{M}_c$ s.t.

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

for holomorphic observables $\mathcal{O}$.
($d\mu, d\mu_c$ positive a priori measures)

Possibilities:

- Solve underdetermined problem directly
  Salcedo 1993–2016; ES&Wosiek, arXiv:1702.06012; Wosiek, Ruba: this conference
- Saddle point method $\longrightarrow$ Lefschetz thimble
  Cristoforetti et al 2014; Alexandru, Basar, Bedaque, Ridgway and Warrington 2016
- Path optimization
  Mori, Kashiwa, Ohnishi 2017
- Complex Langevin
  Parisi 1983, Klauder 1983

2. Complex Langevin

First proposed:
Parisi 1983; Klauder 1983

Many studies in 1980's and 1990's:
Sometimes success, sometimes no convergence, sometimes worse:

Convergence to wrong limit!
(Ambjørn, Flensburg&Peterson 1986)

Interest rekindled:
Berges&Stamatescu 2005:
Simulation of Minkowski space QFT

Recall: Real Langevin

aka Stochastic Quantization:
(Nelson 1966 (different story);
Parisi & Wu 1981; ES 1984; Batrouni et al 1985;
Damgaard&Hüffel 1987)

$$d\vec{x} = \vec{K}dt + d\vec{w}, \quad \vec{K} = -\vec{\nabla}S$$

Convergence of P_t to invariant measure $P_\infty \propto \exp(-S)$
provided

- $\exp(-S)$ integrable,
- process ergodic.

Justification of Real Langevin

Fokker-Planck eqation for probability density $P \equiv \rho$:

$$\partial_t P(\vec{x};t) = L^T P(\vec{x};t)$$

where

$$L^T \equiv \vec{\nabla}\left(\vec{\nabla} + (\vec{\nabla} S)\right)$$

Similarity transformation

$$H_{FP} \equiv \exp(-S/2)(-L^T)\exp(S/2) =$$

$$-\left(\vec{\nabla} + \tfrac{1}{2}(\vec{\nabla} S)\right)\left(\vec{\nabla} + \tfrac{1}{2}(\vec{\nabla} S)\right) \geq 0\,.$$

Consequence

$\text{spec}(-L^T) = \text{spec}(H_{FP}) \in \mathbb{R}^+$

H_{FP} has unique ground state $|0\rangle \iff$ process ergodic:

$$\lim_{t\to\infty} \exp(-H_{FP}t) = |0\rangle\langle 0|;$$

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

But there are counterexamples to ergodicity!

3. Go complex: CL

Klauder, Parisi 1983:

$\mathcal{M} = \mathbb{R}^N$, $\vec{K}$ complex: postulate as before

$$d\vec{z} = \vec{K}dt + d\vec{w}, \quad \vec{K} = -\nabla \vec{S}$$

$\vec{dw}$ real Wiener increment $\vec{dw} = \vec{\eta}(t)dt$, $\vec{\eta}$ white noise).
Process wanders into complexification $\mathcal{M}_c$

$$d\vec{x} = \vec{K}_x dt + d\vec{w}, \quad \vec{K}_x = \mathrm{Re}\ \vec{K}$$

$$d\vec{y} = \vec{K}_y dt, \quad \vec{K}_y = \mathrm{Im}\ \vec{K}$$

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

Why should this be right?

Justification of CL

Early attempt:

Nakazato 1987, Okano, Schülke&Zheng 1993 argue:

Relate P to ρ by

$$\int d\vec{y} \exp(i\vec{y}\partial_{\vec{x}}) P(\vec{x}, \vec{y}; t) = \rho(\vec{x}; t)$$

Problem:

$$\exp(i\vec{y}\partial_{\vec{x}}) P(\vec{x}, \vec{y}; t) \equiv P(\vec{x} - i\vec{y}, \vec{y}; t)$$

meaning? Requires analyticity of P.

But: want to start with $P(\vec{x}, \vec{y}; 0) = \delta(\vec{x})\delta(\vec{y})$.

Our approach

(AS^3 $\equiv$ Aaarts, Seiler, Sexty, Stamatescu)

$\mathcal{O}$ holomorphic, $\vec{K}$ at least meromorphic:

Evolution of averaged observables $\mathcal{O}(\vec{z};t) \equiv \langle \mathcal{O}(\vec{z}(t)) \rangle$:

Ito calculus $\Longrightarrow$

$$\partial_t \mathcal{O}(\vec{z};t) = L \mathcal{O}(\vec{z};t)\,, \quad L \equiv \left[\vec{\nabla}_x + \vec{K}_x\right] \cdot \vec{\nabla}_x + \vec{K}_y \cdot \vec{\nabla}_y\,;\ (*)$$

$(*)$ solved formally by $\mathcal{O}(\vec{z};t) \equiv \exp\left[tL\right] \mathcal{O}(\vec{z})$.

$\mathcal{O}(\vec{z};t)$ holomorphic where $\vec{K}$ is:

$\vec{\nabla}_y \mathcal{O}(\vec{z};t) = i \vec{\nabla}_x \mathcal{O}(\vec{z}:t)$ (Cauchy-Riemann) $\Longrightarrow$

$$L \mathcal{O}(\vec{z};t) = L_c \mathcal{O}(\vec{z};t)\,, \quad L_c \equiv \left[\vec{\nabla}_x + \vec{K}\right] \cdot \nabla_x$$

Evolutions 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:

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

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

Relation of evolutions?

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

Evolutions of expectations:

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

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

Consistent?

Formally yes: use CRE and integration by parts.

Result

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

Assumptions:

- agreement of initial conditions (irrelevant for $t \to \infty$ if process ergodic)
- meromorphy of drift $\vec{K} \equiv \vec{K}_x + i\vec{K}_y$
- sufficient decay of $|\vec{K} P \mathcal{O}|$ at imaginary infinity and near poles of $\vec{K}$: to be checked.

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 PDE

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

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

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

(last eq. requires integration by parts)

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 at ∞ and at poles of $\vec{K}$.

Consistency conditions

In equilibrium

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

Expresses stationarity of noise averaged observables
Equivalent to Schwinger-Dyson equations.

Equivalent to $L^T P = 0$ if integration by parts without boundary terms.

A theorem

Consider compact $\mathcal{M}$. If

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

Then equilibrium measure correct: Then equilibrium measure correct:

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

(Aarts, James, Seiler, Stamatescu 2011)

Not easy to check. In some cases: negative evidence

4. Problems

Questions for mathematicians:

- Existence and uniquenss of process: no theorems, numerically ok
- Convergence of measures: ditto (need at least an attractive fixed point)

Would need: $\mathrm{spec}(L), \mathrm{spec}(L^T)$ in the left half of $\mathbb{C}$, unique eigenvalue at 0, but:
"... conspicuous absence of general spectral theorems"
(for non-self-adjoint and non-normal operators).
(Klauder&Petersen 1985)

Still true in 2017, but: let's be pragmatic!

Questions for practitioners:

- Boundary terms
  1. at ∞
  2. at poles of $\vec{K}$
- Lack of ergodicity

Both can lead to failure. Examples below.

Other criteria for failure: Hayata et al 2016, Salcedo 2016
Covered by criteria above.

5. Boundary terms at ∞

Typical:

$\mathcal{M}$ compact, $\mathcal{M}_c$ noncompact

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

Note:

Holomorphic functions grow at $\infty \Longrightarrow$

Drift $\vec{K}$ grows; observables $\mathcal{O}$ as well $\Longrightarrow$

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

Integration by parts without boundary terms:

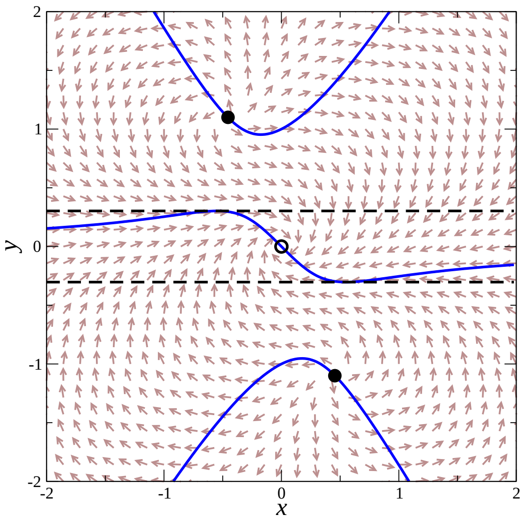
depends on behavior of $|\vec{K}P\mathcal{O}|$ at ∞.

Lucky case

G. Aarts, P. Giudice, E. S. 2013: $\mathcal{M} = \mathbb{R}$, $\mathcal{M}_c = \mathbb{C}$:

$$S = \frac{1}{2}(1 + iB)x^2 + \frac{1}{4}x^4 \, .$$

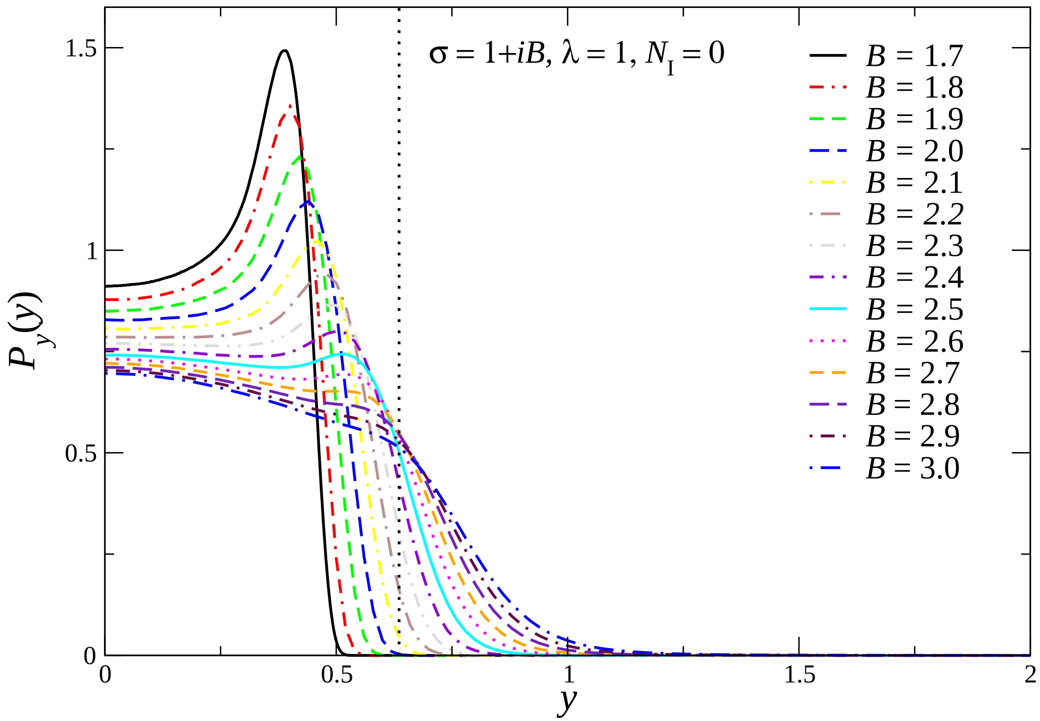
$B < \sqrt{3}$: Process confined in strip.



(Solid lines: $K_y = 0$).

CLE results correct!

More generally



Tail $O((x^2 + y^2)^{-3})$.

Tails present: CLE results deteriorate for $B > \sqrt{3}$.

6. Boundary terms at poles

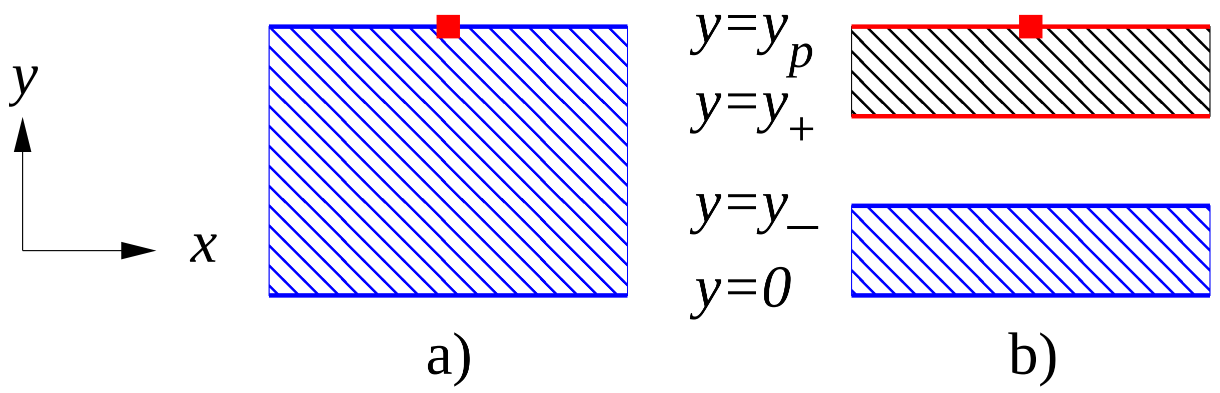
If ρ has zeroes in $\mathcal{M}_c$, $\vec{K}$ has poles there.
$\dot{\mathcal{O}} = L_c \mathcal{O}$ generically produces essential singularities at poles of $\vec{K}$.
In equilibrium distribution poles may be

- outside: harmless
- at the edge: depends on behavior of $|\vec{K}P|$
- inside: depends on behavior of $|\vec{K}P|$ and ergodicity

Nagata, Nishimura, Shimasaki 2016:
monitor $|\vec{K}P|$ at poles and ∞.

One-pole model

$$\rho(x) = (x - iy_p)^{n_p} \exp(-\beta x^2),$$



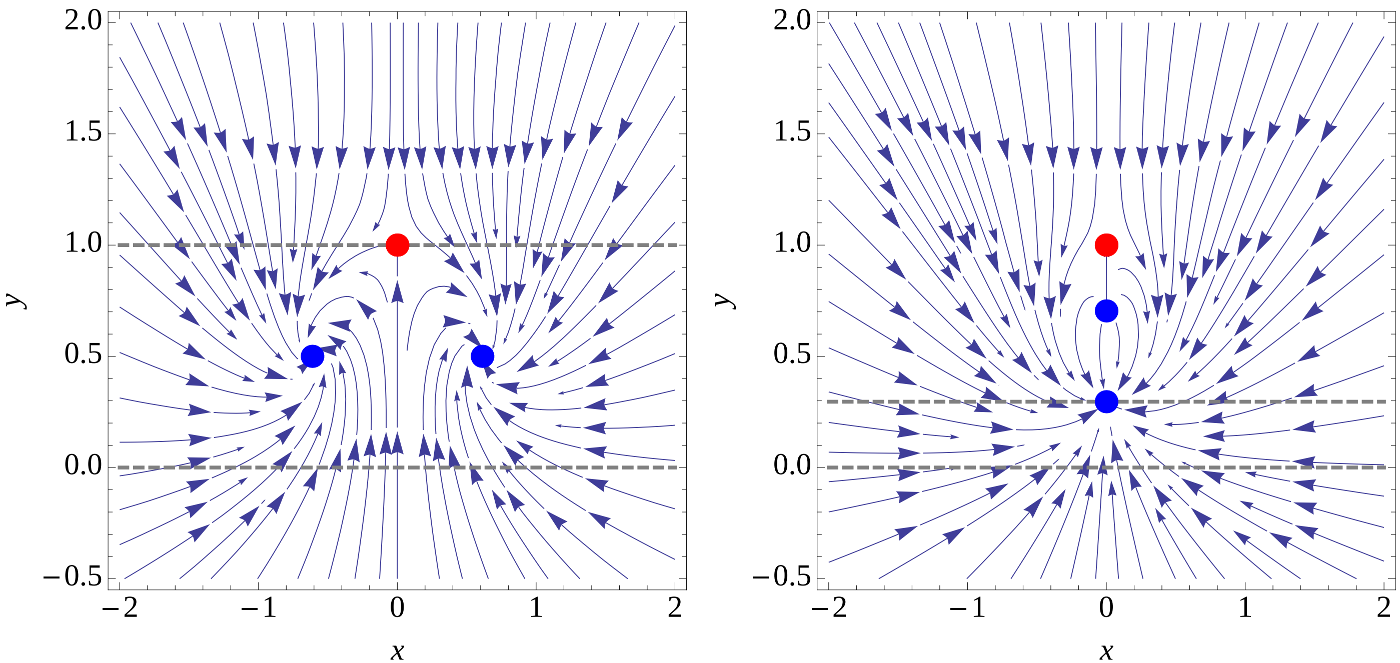
a) $\beta < 2n_p$: $P(x, y) > 0$ when $0 < y < y_p$; pole on the edge

b) $\beta > 2n_p$: $P(x, y) > 0$ when $0 < y < y_-$; pole outside

equilibrium distribution.

Upper strip transient.

Why? Classical flow diagrams



$y_p = 1$, $n_p = 2$

Left: $\beta = 1.6$, Right: $\beta = 4.8$.

Blue (red) circles fixed points (pole);

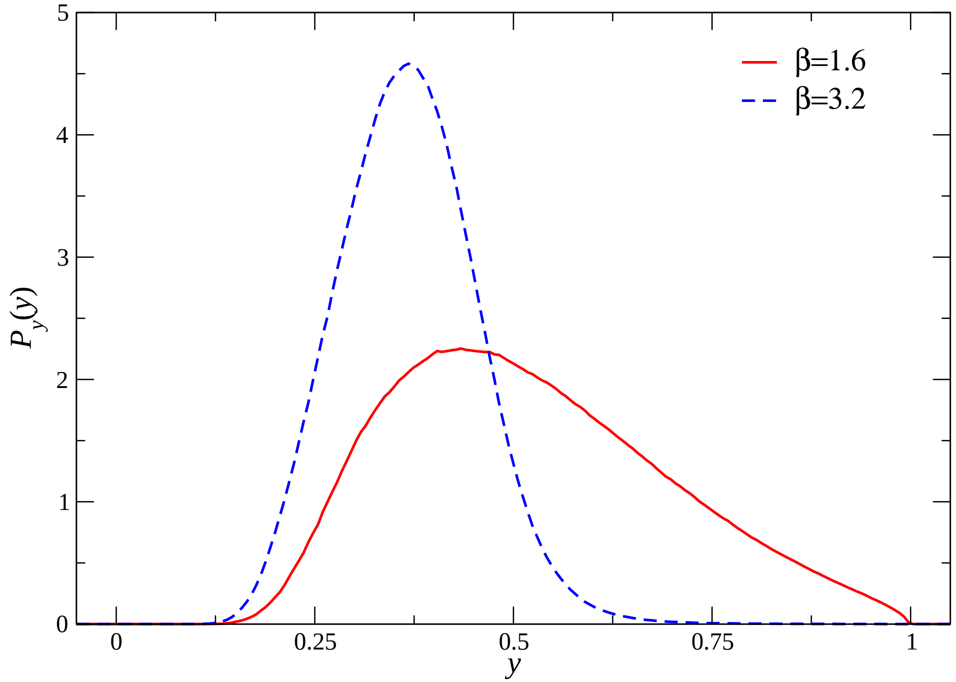
$P = 0$ outside dashed horizontal lines.

Pole on edge

Failure: $\beta = 1.6$: boundary term

Success: $\beta = 3.2$ no (or incredibly tiny) boundary term

Reason: Slow (fast) decay of P near pole:

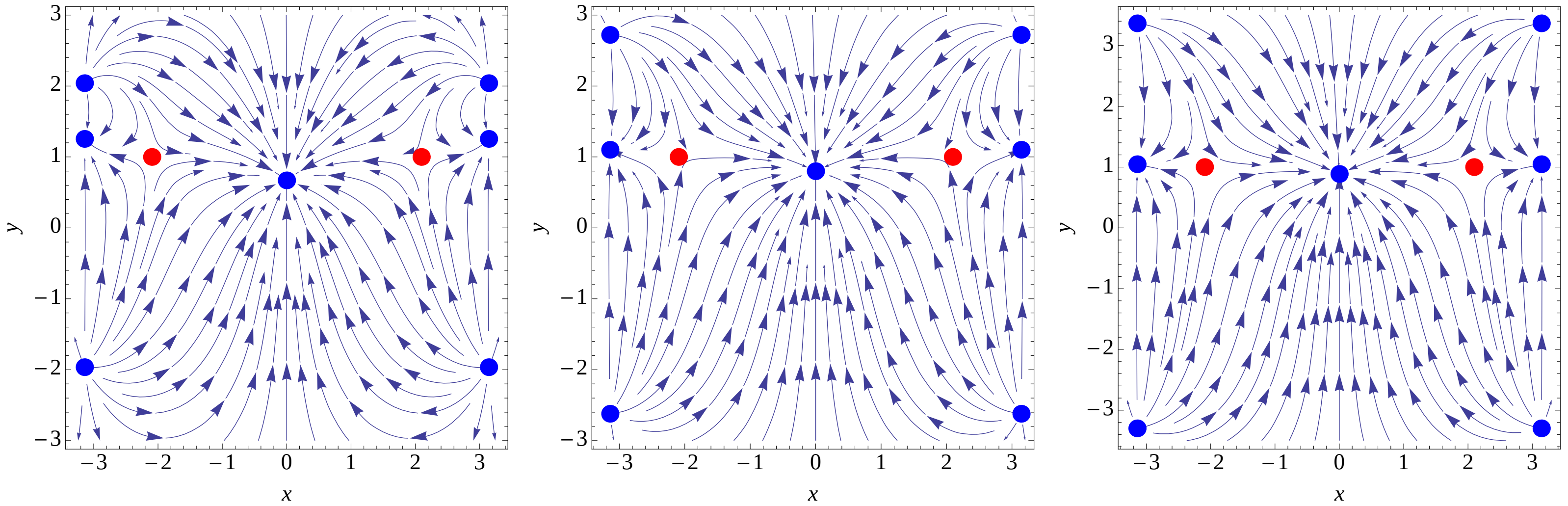


Partially integrated distributions $P_y(y) = \int dx P(x,y)$.

Poles inside distribution

One-link U(1) model:

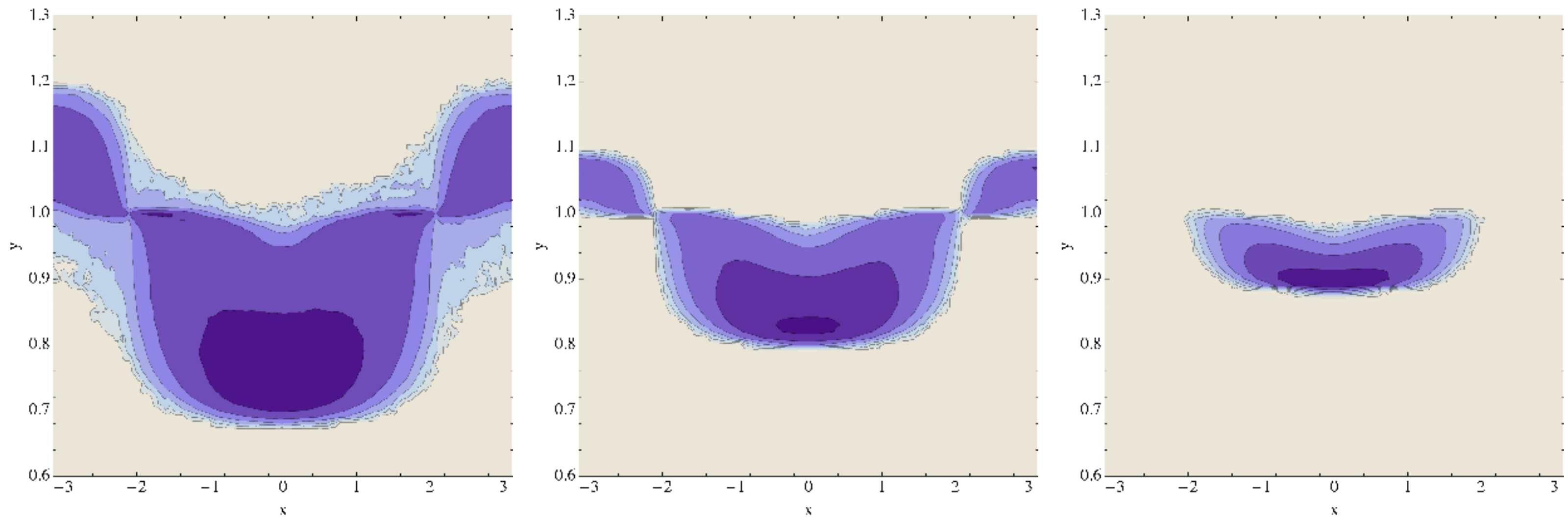
$$\rho(x) = [1 + \kappa \cos(x - i\mu)]^{n_p} \exp[\beta \cos(x)]; \quad \kappa = 2, \beta = 0.3, \mu = 1$$



$n_p = 1$ $\qquad$ $n_p = 2$ $\qquad$ $n_p = 4$

Blue dots: fixed points, red dots: poles.

Equilibrium distributions



$n_p = 1$ $n_p = 2$ $n_p = 4$

Logarithmic contour plots of $P(x, y)$:

Process tends to avoid poles,

but: Convergence to wrong limit

What to do in lattice models?

- Lucky strips not to be expected
- Always monitor tails
- Better: monitor distribution of drift (Nagata, Nishimura, Shimasaki 2016)

Not sufficient! Why?

7. Failure of ergodicity

Recall real one-pole model.

$$\rho = x^4 \exp\left(-\frac{x^2}{2\sigma}\right) \;; \quad K = -\frac{x}{\sigma} + \frac{4}{x}$$

Two ground states of $H_{FP} = -\frac{d^2}{dx^2} + \frac{2}{x^2} + \frac{x^2}{2\sigma} - \frac{5}{2\sigma}$:

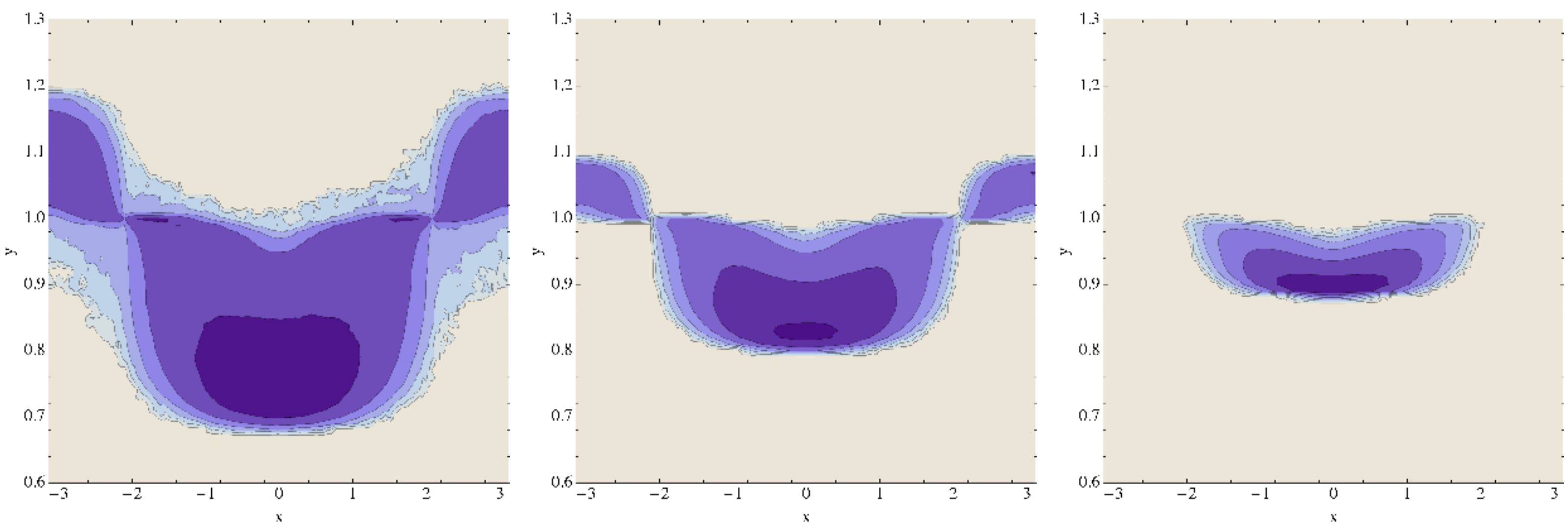
$$|\pm\rangle \propto \theta(\pm x) x^2 \exp\left(-\frac{x^2}{4\sigma}\right)$$

Two invariant probability densities:

$$P_\pm(x) \propto \theta(\pm x) x^4 \exp\left(-\frac{x^2}{2\sigma}\right) .$$

Poles are bottlenecks

Recall one-link U(1) model:



$n_p = 1$ $n_p = 2$ $n_p = 4$

Restrict process to “head/ears” region:

Simulates integral along 2 inequivalent paths between poles.

Non-ergodic!

Poles in QCD

Lattice models:

Fermion determinant

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

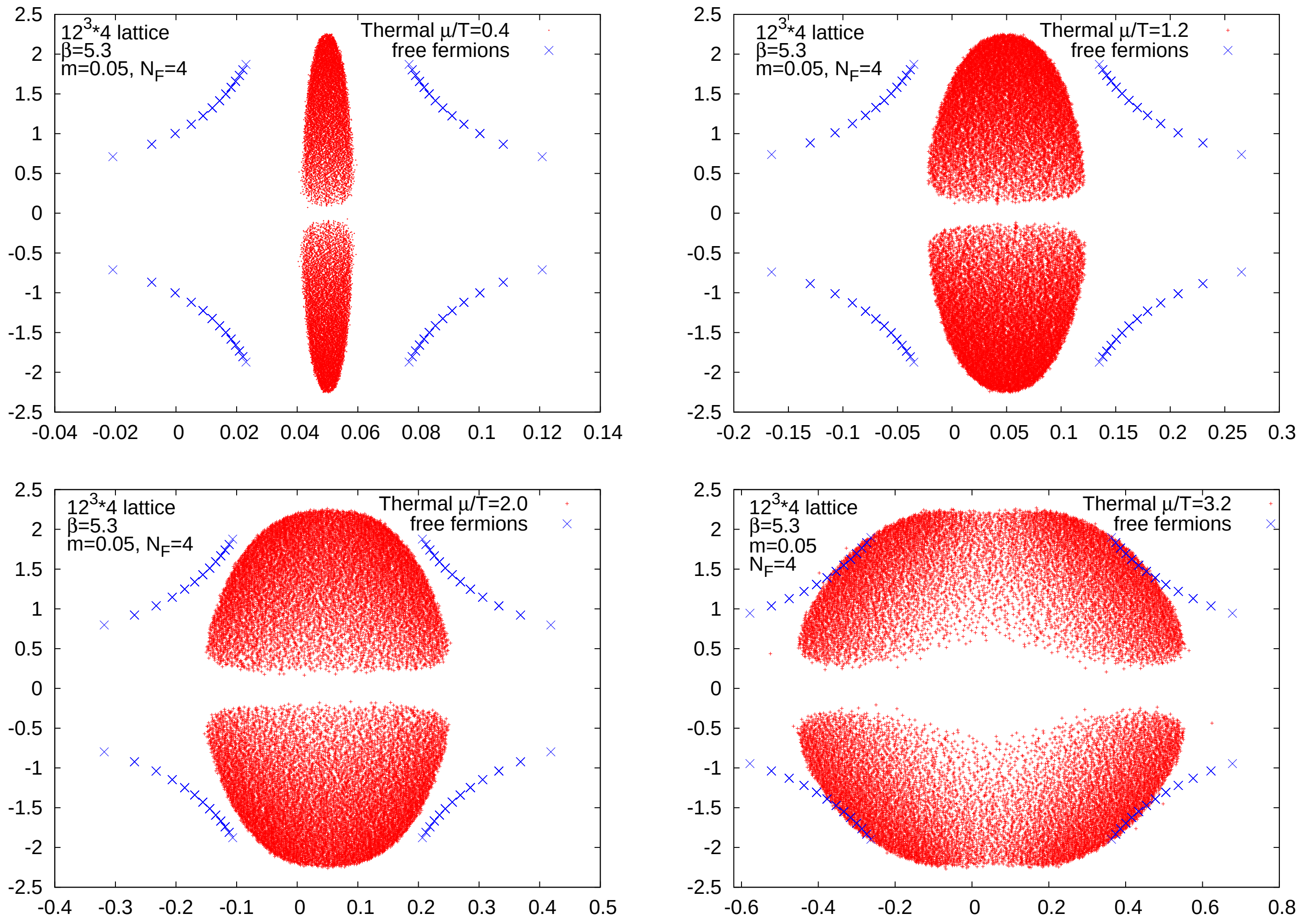
always has zeroes somewhere in $U \in SL(3, \mathbb{C})$.

Zeroes form manifolds of codimension 2:

$\Longrightarrow$ non-ergodicity?

But good news:

Sexty 2013, AS^3 2017 in QCD: eigenvalues avoid 0.



Spectrum of staggered Dirac op., $12^3 \times 4$ lattice, $\mu/T = 0.4, 1.2, 2.0, 3.2$.

8. Gauge cooling

(S^3 2013) 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

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

Update and cooling

$$U_{x,\mu} \mapsto \exp\left\{ -\sum_a i\lambda_a(\epsilon K_{a,x,\mu} + \sqrt{\epsilon}\eta_{a,x,\mu}) \right\} U_{x,\mu} \,,$$

$K_{a,x,\mu} = D_{a,x,\mu} S$ drift; η independent white noises. Gauge cooling step:

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

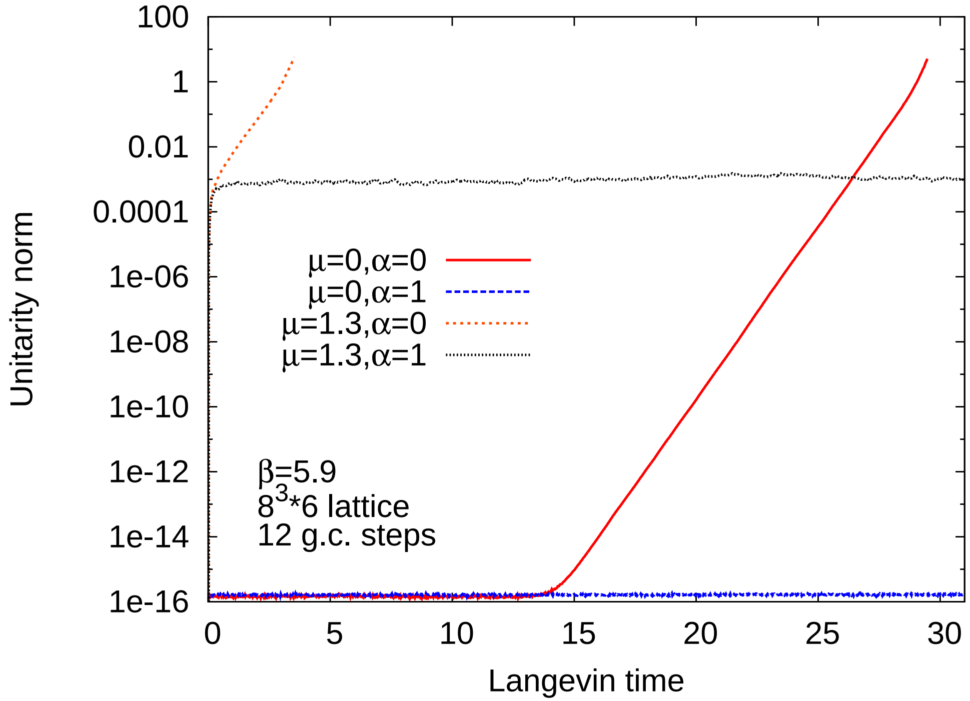
$\alpha = \epsilon\tilde{\alpha}$, ϵ discretization parameter

Two possibilities:

(1) Gauge cool appropriately between dynamical updates

(2) Additional 'cooling drift' (Nagata et al 2016)

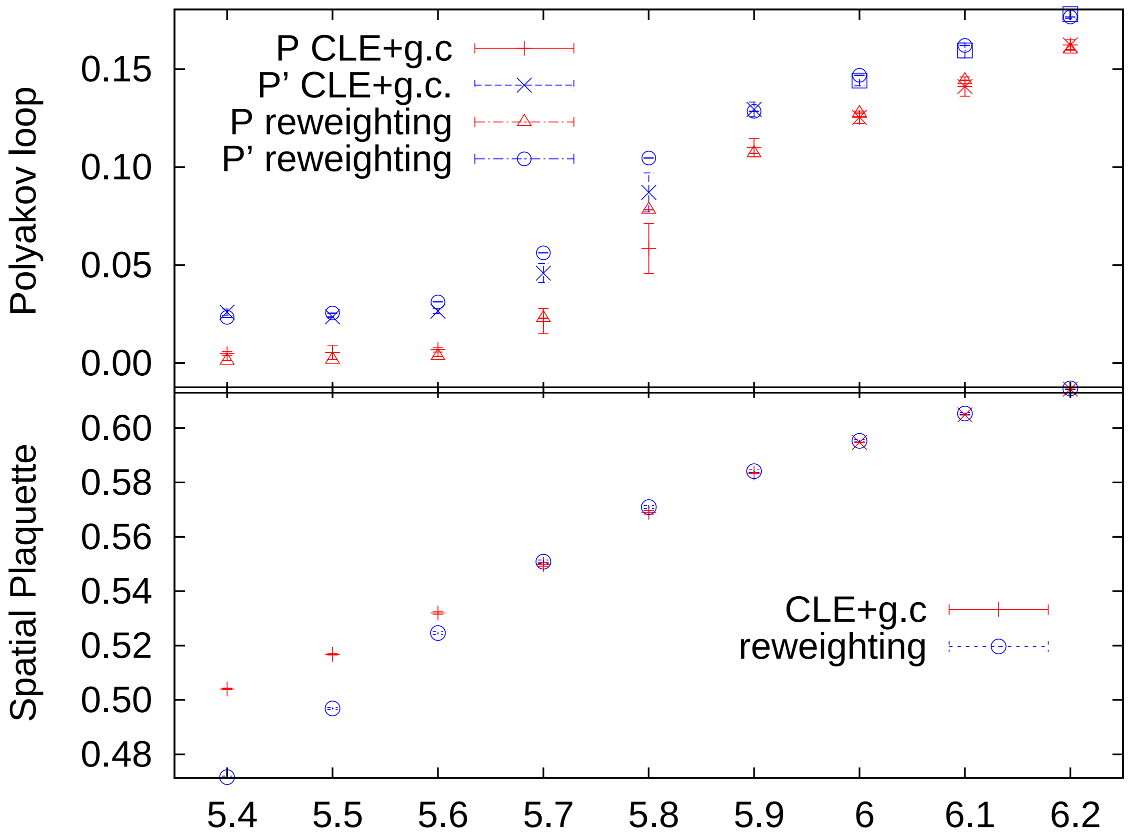
GC keeps U norm in check



GC eliminates skirts

GC produces correct results

GC in High Density QCD



$\mu = 0.85$, 6^4 lattice; adaptive step size.

Good for $\beta \geq 5.7$ but deviations for smaller β.

Full QCD: some results

Sexty 2014, AS^3 2017: Staggered quarks, lattices up to $12^3 \times 4$: deviations at small β.

AS^3 2014: Hopping expansion gives good results, lattices up to 8^4.

Sinclair & Kogut 2016: Staggered quarks, lattices up to 16^4; results at $\mu = 0$ disagree with known results.

Some not so good news

Long runs: U-norm has pleateau, then starts growing again.

Reason: Numerical instability, “true” instability?

Attempted cures:

Nagata & Nishimura & Shimasaki 2016: Different unitarity norms.

Schmalzbauer& Bloch 2016: Variations of gauge cooling.

Problem at small β remains.

Bloch 2016: Reweighting CL; works well in simple models, but expect overlap problems on larger lattices.

Dynamic stabilization

Aarts & Attanasio & Jäger &Sexty 2016:

Add extra stabilizing drift X

Invariant under $SU(3)$, not $SL(3,\mathbb{C})$.

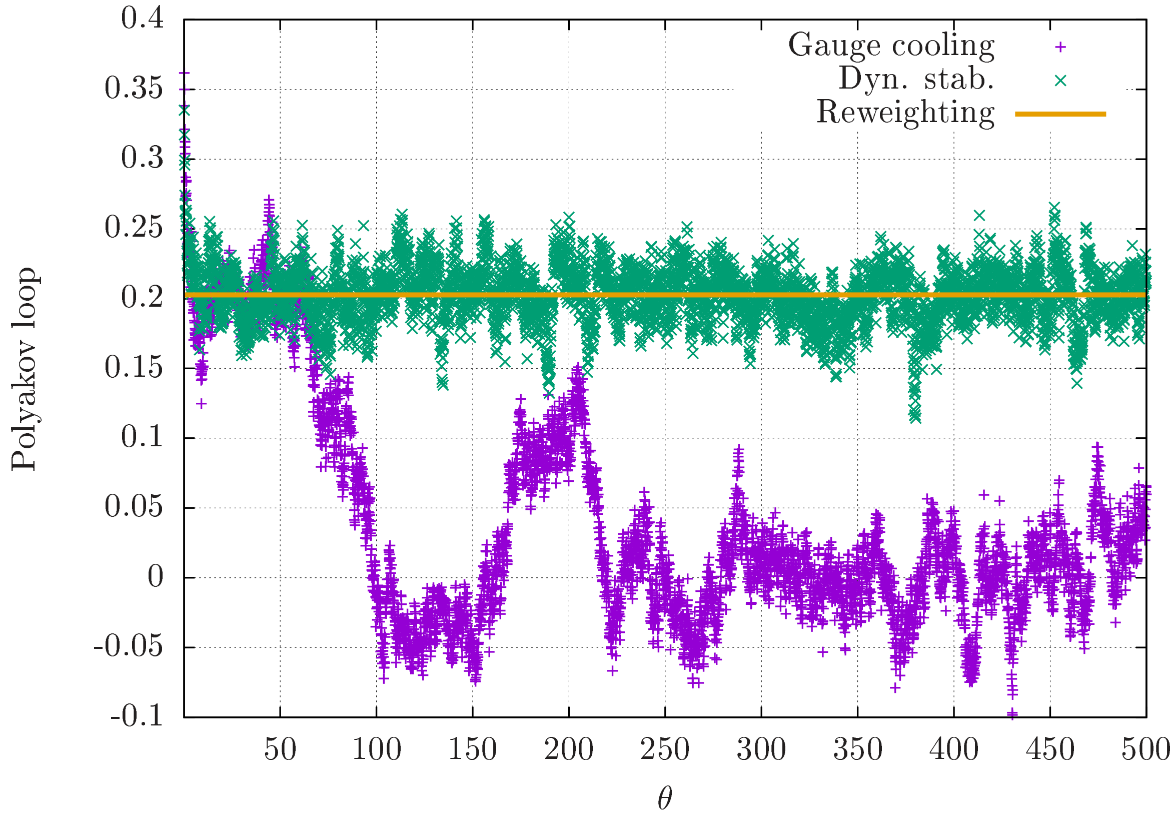
$$M_x^a[U] = i b_x^a (b_x^a)^3 \qquad b_x^a = \sum_\nu \mathrm{tr}[\lambda^a U_{x\nu} U_{x\nu}^\dagger]$$

$$X_{x\mu}^a = i\epsilon\alpha_{DS} M_x^a[U]\,. \qquad (1)$$

Argued to add only lattice artifact. Seems to work impressively.

Justification?

Example (F. Attanasio 2017)



HDQCD: average Polyakov loop vs Langevin time; lattice $10^3 \times 4, \kappa = 0.04, \beta = 5.8, \mu = 0.7$.

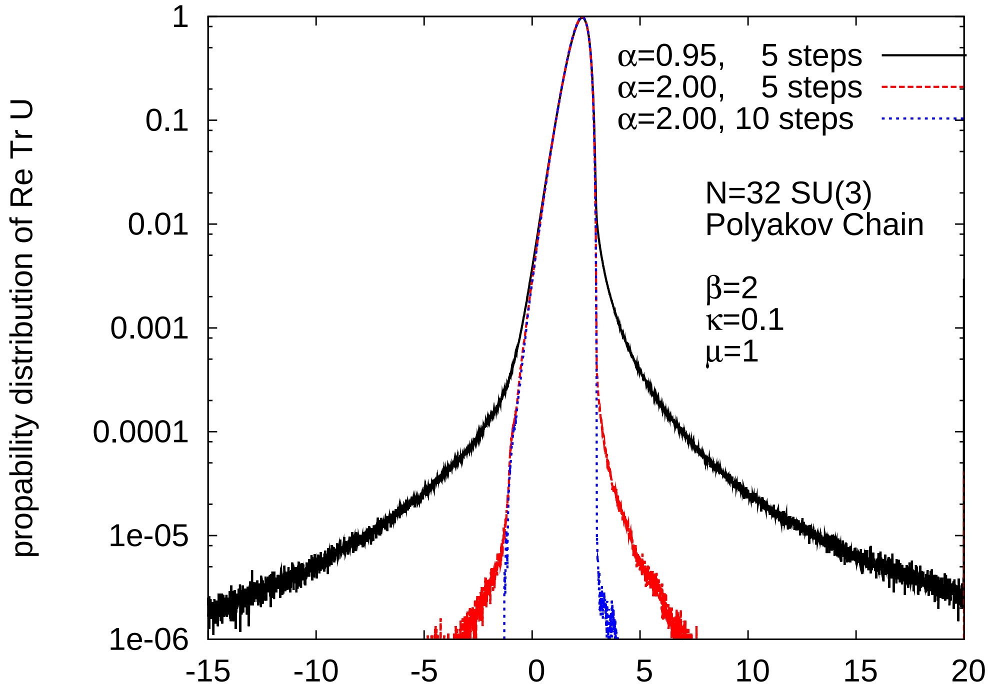
9. Summary

- 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
- Expansion avoids poles (AS^3 2014)
- Best hope: dynamic stabilization (DS)
- Task: Understand DS better

Apologies for contributions not mentioned.

10. Backup

Cooling can reduce skirts

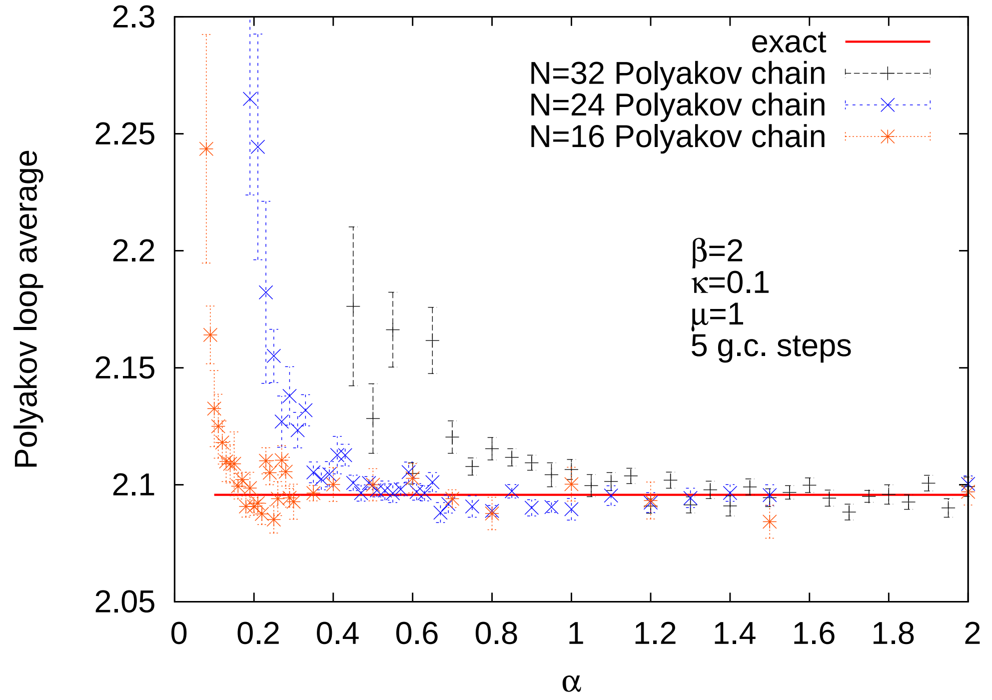


Histograms of Polyakov loops

Motto:

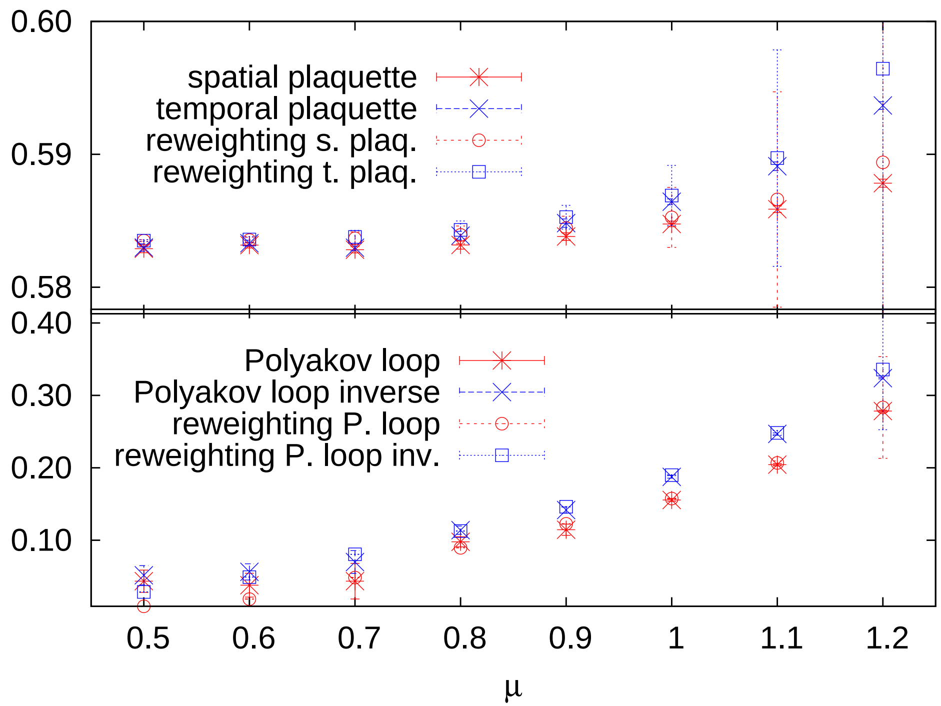
"Let 100 flowers bloom, let 100 schools of thought compete"

Cooling can produce correct results



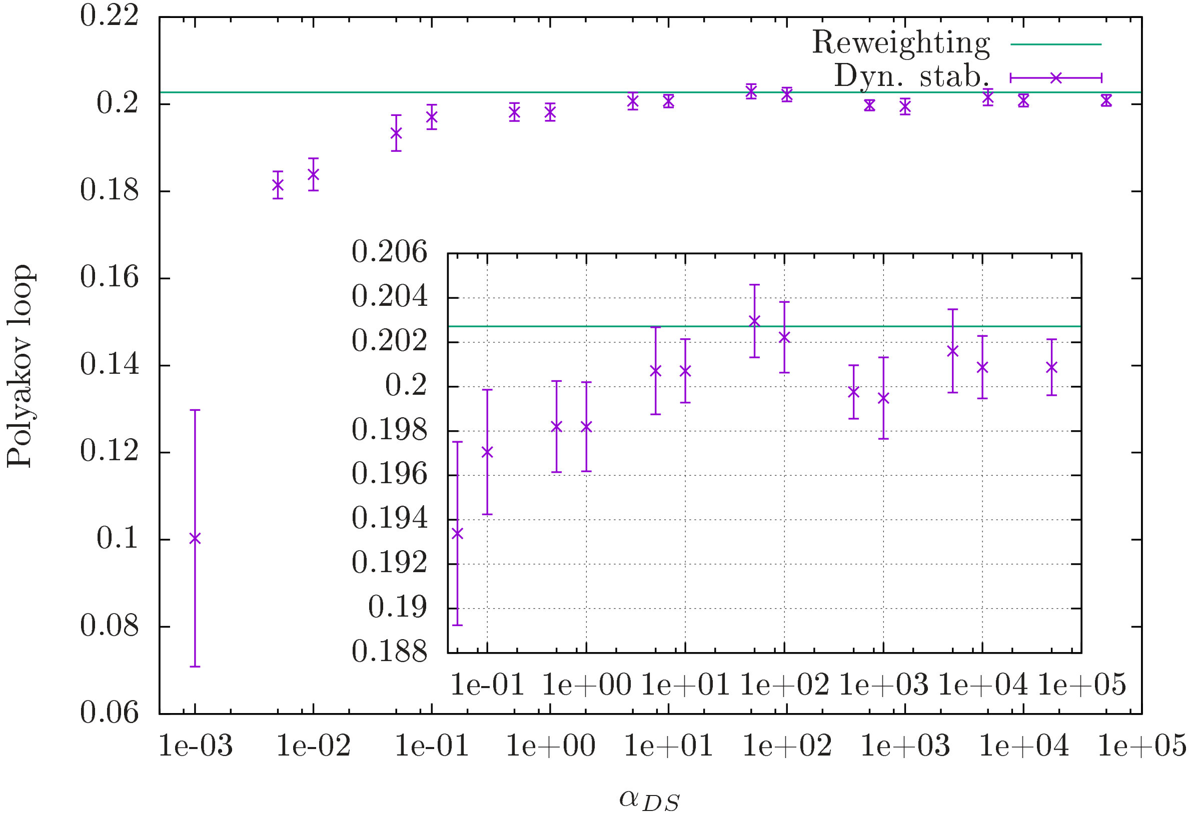
Gauge cooling in HD QCD

Compare with reweighting to validate:



Lattice 6^4, $\beta = 5.9, \kappa = 0.12$. Reweighting breaks down at $\mu \gtrsim 1.2$. Agreement deteriorates at $\beta \lesssim 5.3$.

Dynamical Stabilization



Average Polyakov loop vs α_{DS} compared with reweighting for HDQCD; lattice $10^3 \times 4$, $\kappa = 0.04, \beta = 5.8, \mu = 0.7$.