Parallel Tempered Metadynamics

an algorithm with potential

Christian Hoelbling

with

Timo Eichhorn

Gianluca Fuwa

Lukas Varnhorst

(Wuppertal U.)

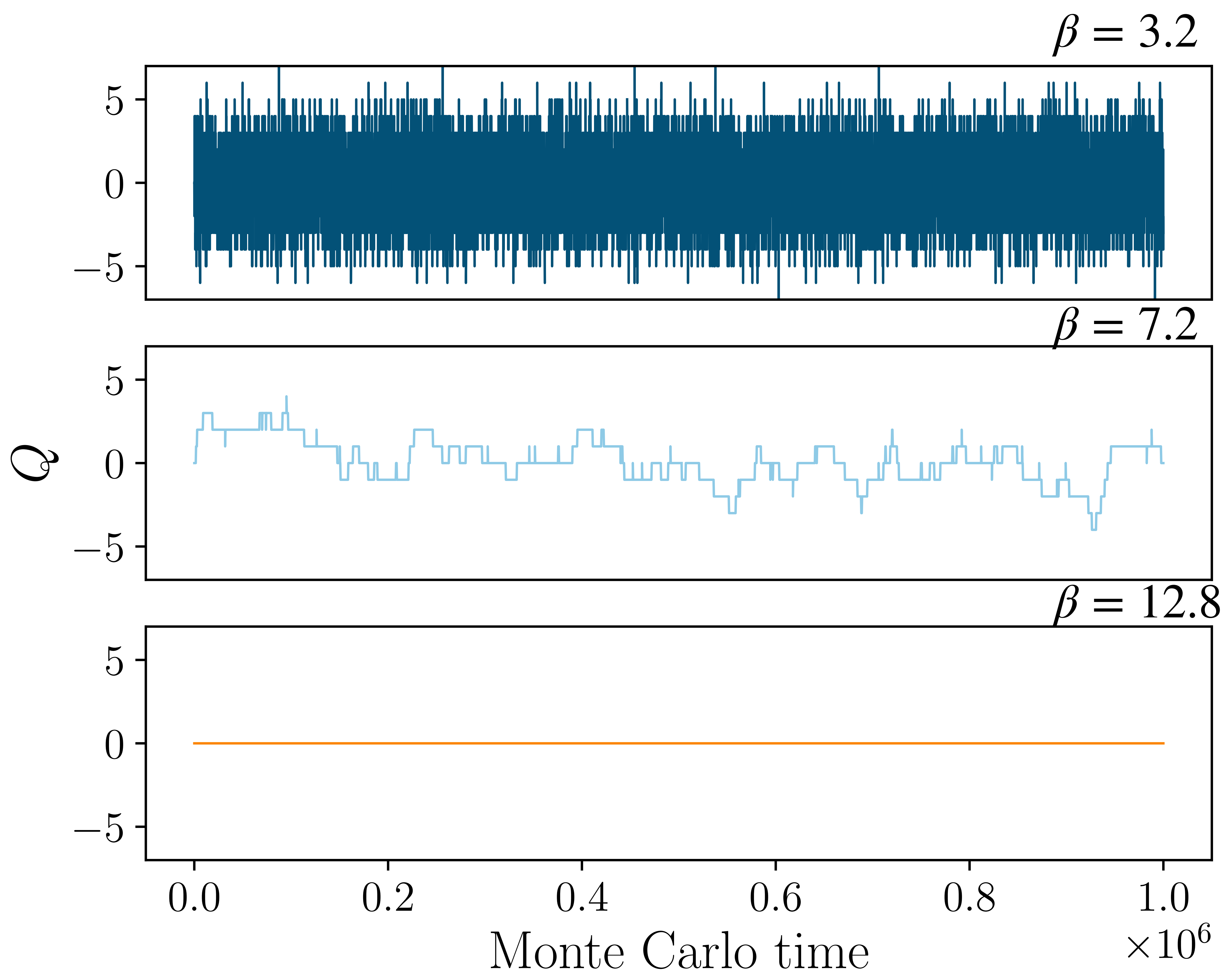
BERGISCHE UNIVERSITÄT WUPPERTAL

Japanese-German Workshop, Mainz, 27.9.2024

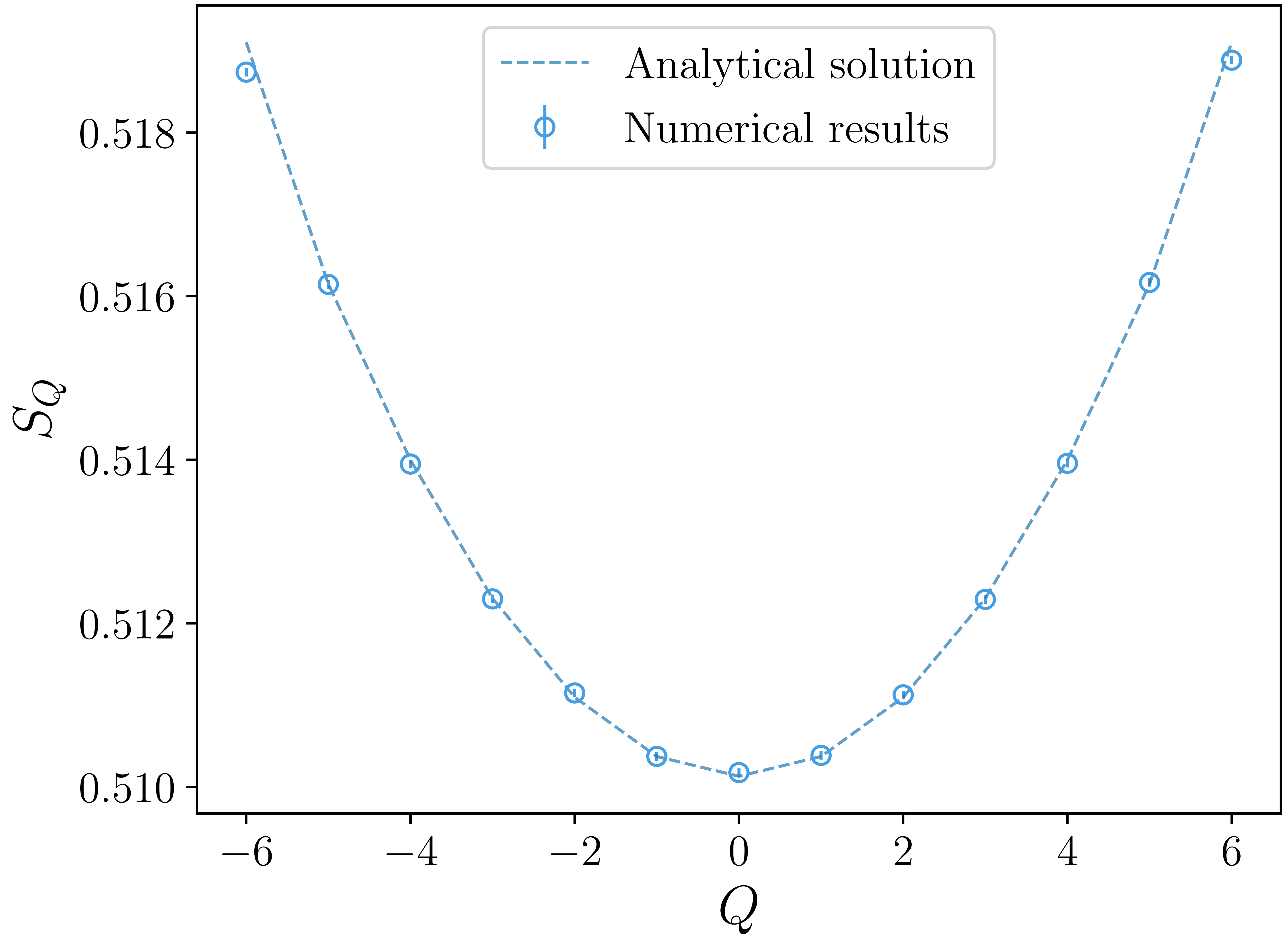
Funded by DFG Deutsche Forschungsgemeinschaft German Research Foundation

The problem

Topological freezing:
- Occurs even in the simplest theories with topology, e.g. 2D U(1)
- Affects even the most basic observables, e.g. the plaquette



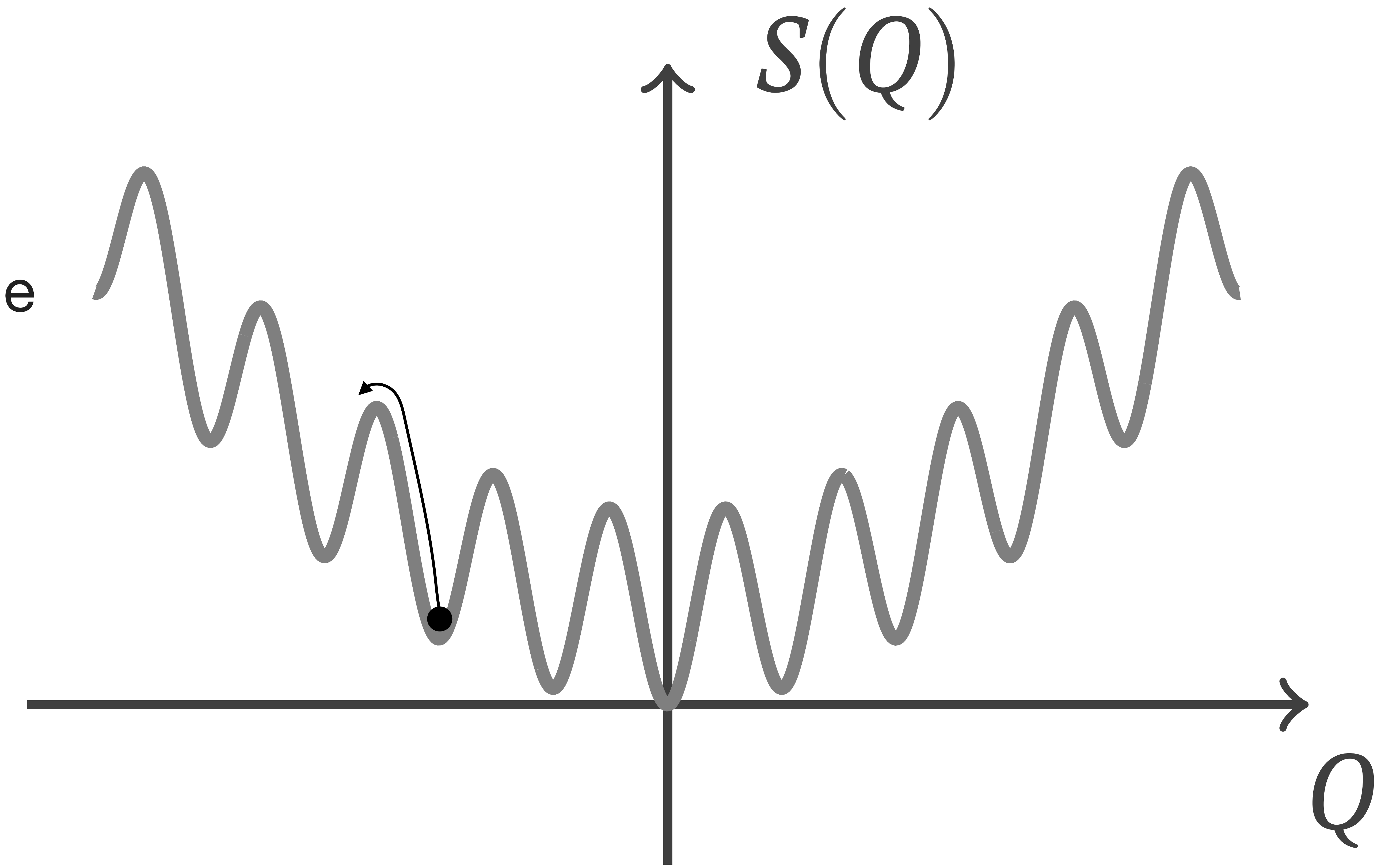
Plaquette



Why topological freezing?

- Continuum: Integer topological charge Q,
  disjoint topological sectors

- Lattice: Non-unique topological charge,
  sector boundaries smeared, become
  sharp towards the continuum



Conventional update algorithms need to transition between sectors, which can be ineffective for two reasons:

- Small corridor in configuration space
- Large action (small acceptance ratio) for intermediate configurations

→Critical slowing down $\propto a^{-5}$ or even exponential (Schaefer, Sommer, Virotta, 2010)

What to do?

- Circumvent the problem: Open boundary conditions →no topological sectors

(Lüscher, Schaefer, 2010)

Treat frozen topology as finite volume effect

(Brower et. al., 2003)

- Hop between topological sectors: Explicitly add an instanton to change topological charge

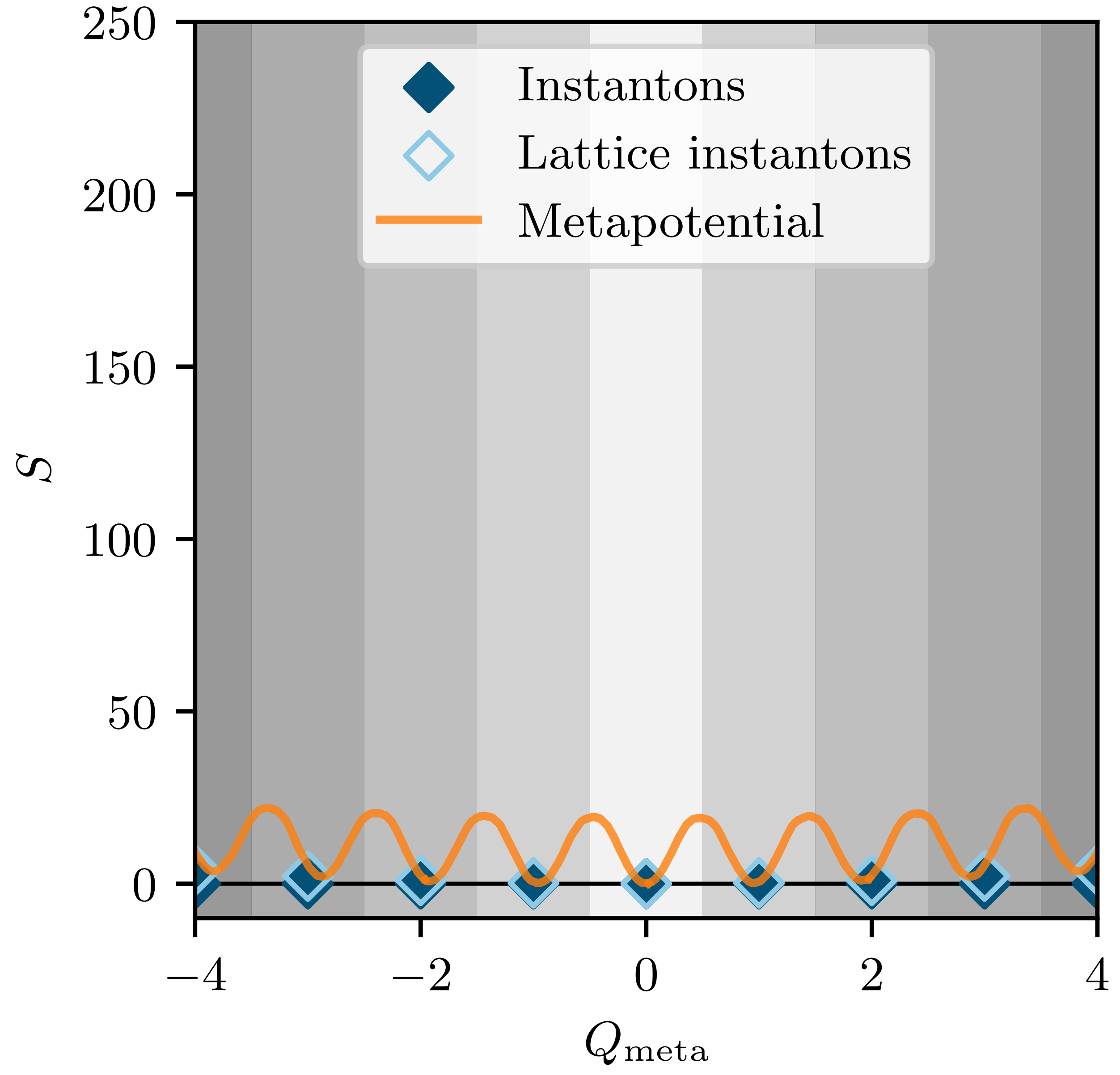
(Fucito, Solomon, 1984)

Works in 2D U(1):

$$Q \propto \int F_{01}$$

$$S \propto \int F_{01}^2$$

$$\lim_{V\to\infty} S_{\text{instanton}} = 0$$



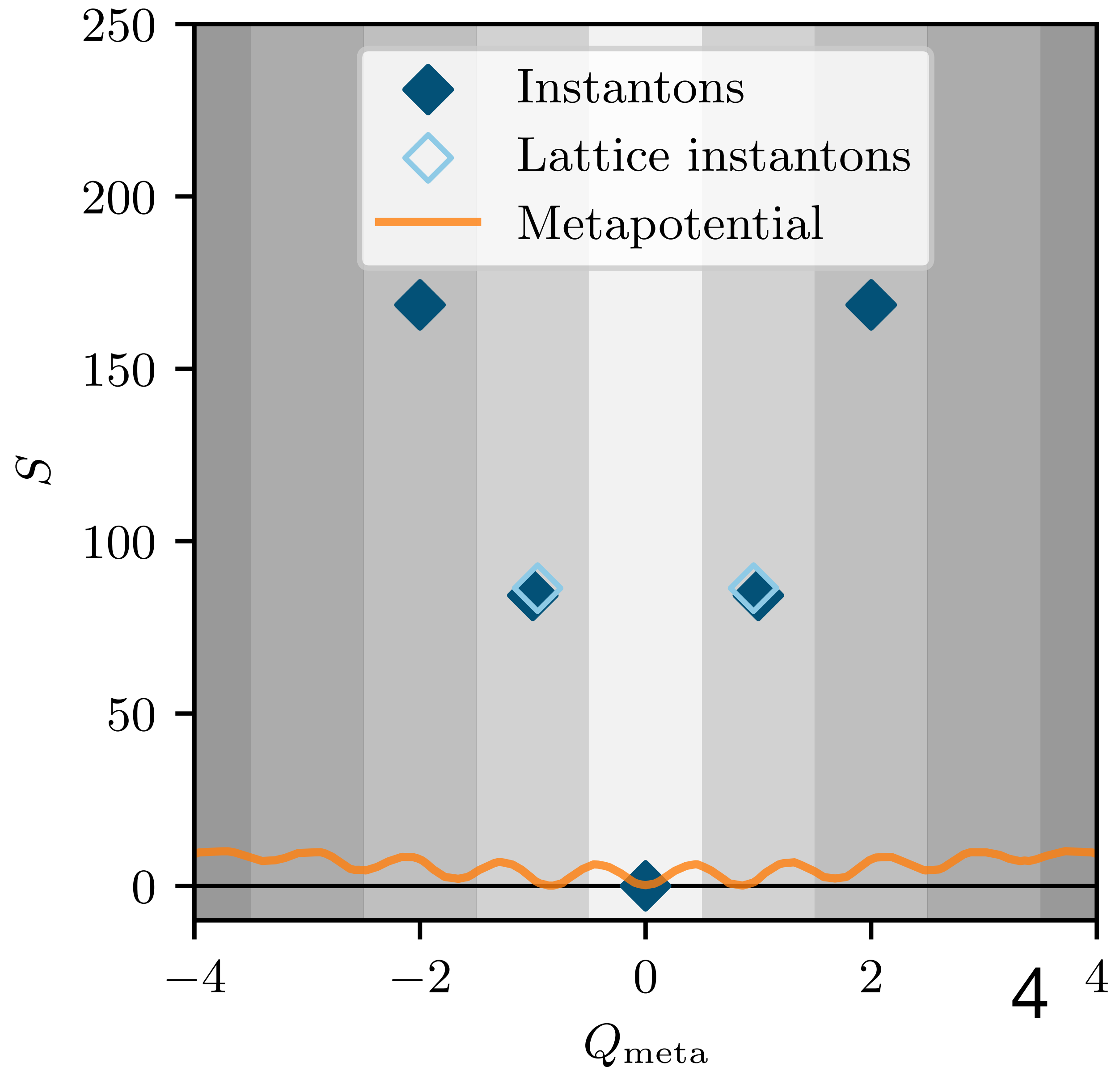
Does not work 4D SU(3):

$$Q \propto \int F\tilde{F}$$

$$S \propto \int F^2$$

$$S_{\substack{\text{instanton} \\ F=\tilde{F}}} = \frac{4}{3}\pi^2\beta\,|\,Q\,|$$

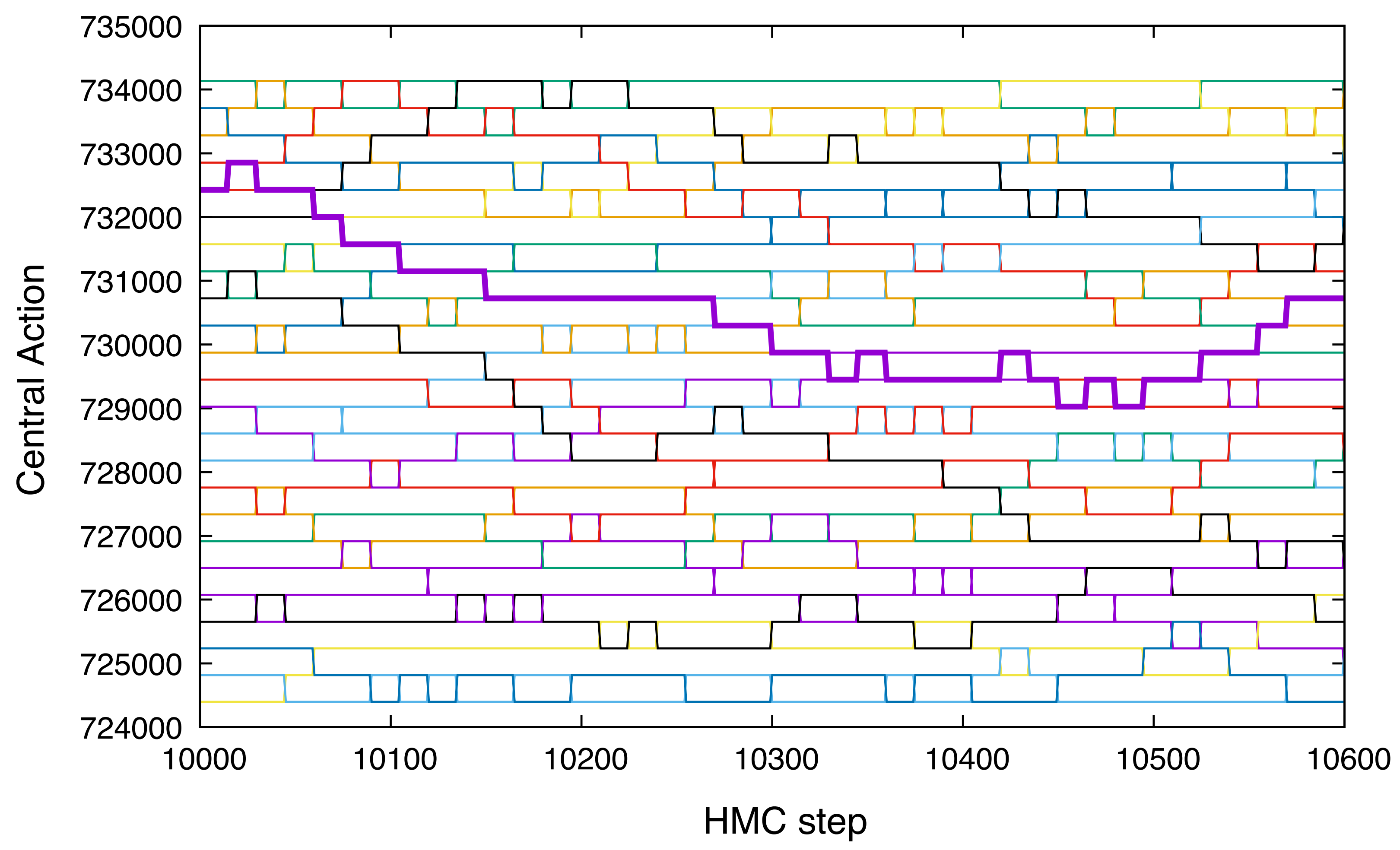
(Bogomolnyi, 1976)



What to do?

- Parallel tempering:
  - Simulate system simultaneously at different couplings
  - Periodically propose swaps between streams, Metropolis step

(Hukushima, Nemoto, 1996)
(UKQCD, 1998)

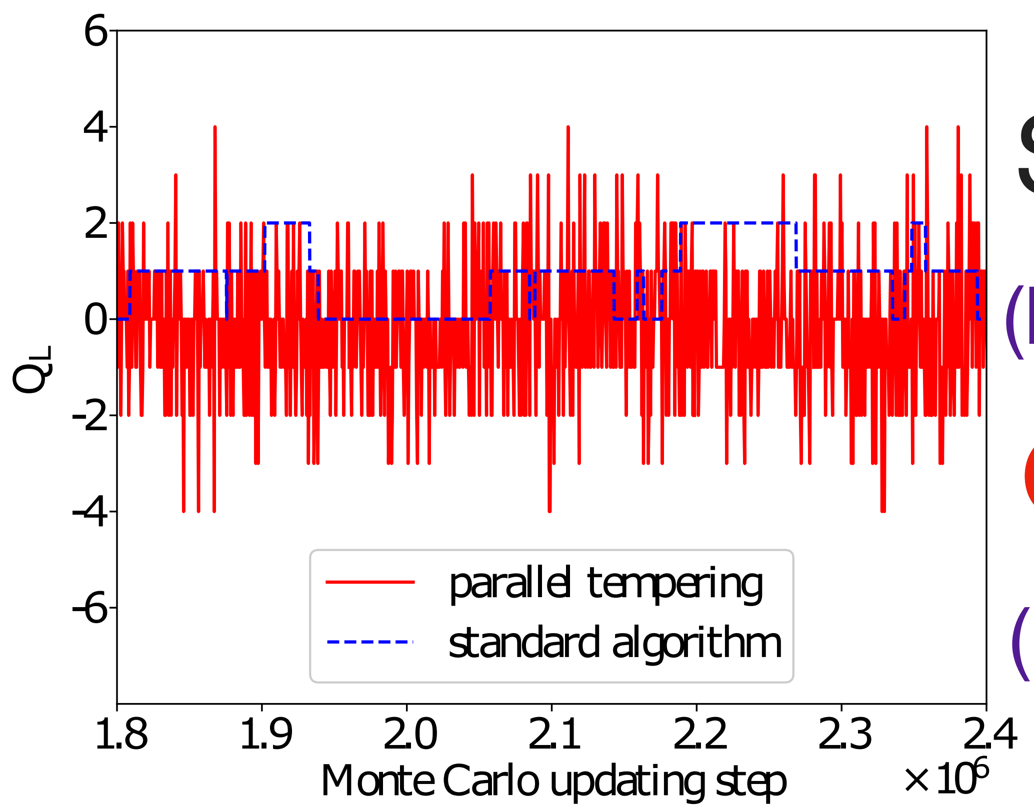


(Cossu et.al., 2021)

Idea is more general: Temper between systems with different boundary conditions

(Hasenbusch, 2017)

- Open defect volume, where topological charge can “vanish”
- Considerable number of intermediate streams



SU(6), 10-17 replica

(Bonnano, Bonati, d’Elia, 2021)

QCD, 10-24 replica

(Bonnano et.al., 2024)

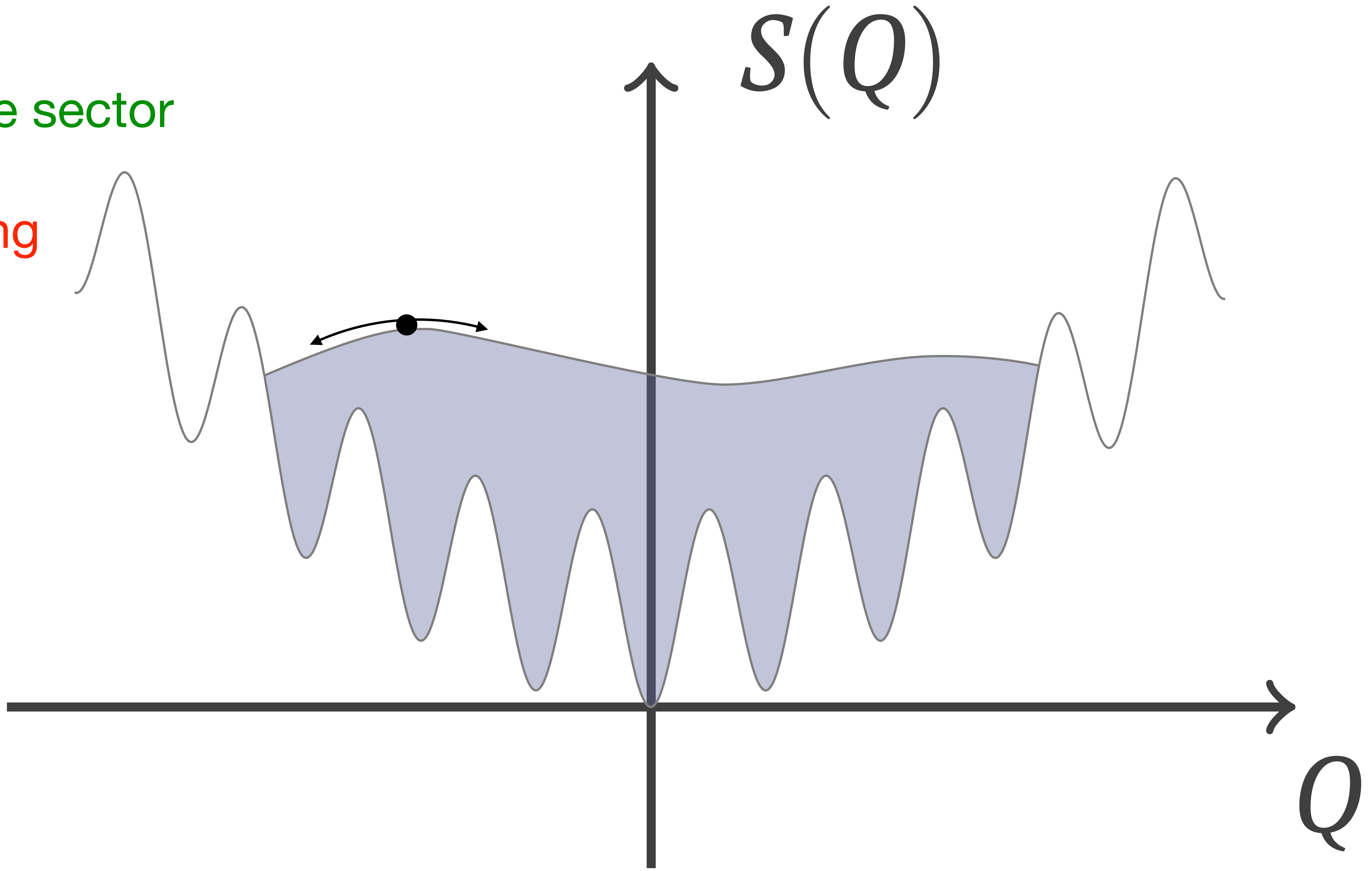
What to do?

- Flatten the action barriers: Introduce additional bias potential $V(Q)$

✓ Simulation no more stuck in one sector

✗ Relative weight of configs wrong

Reweighting factor $e^{-V(Q)}$

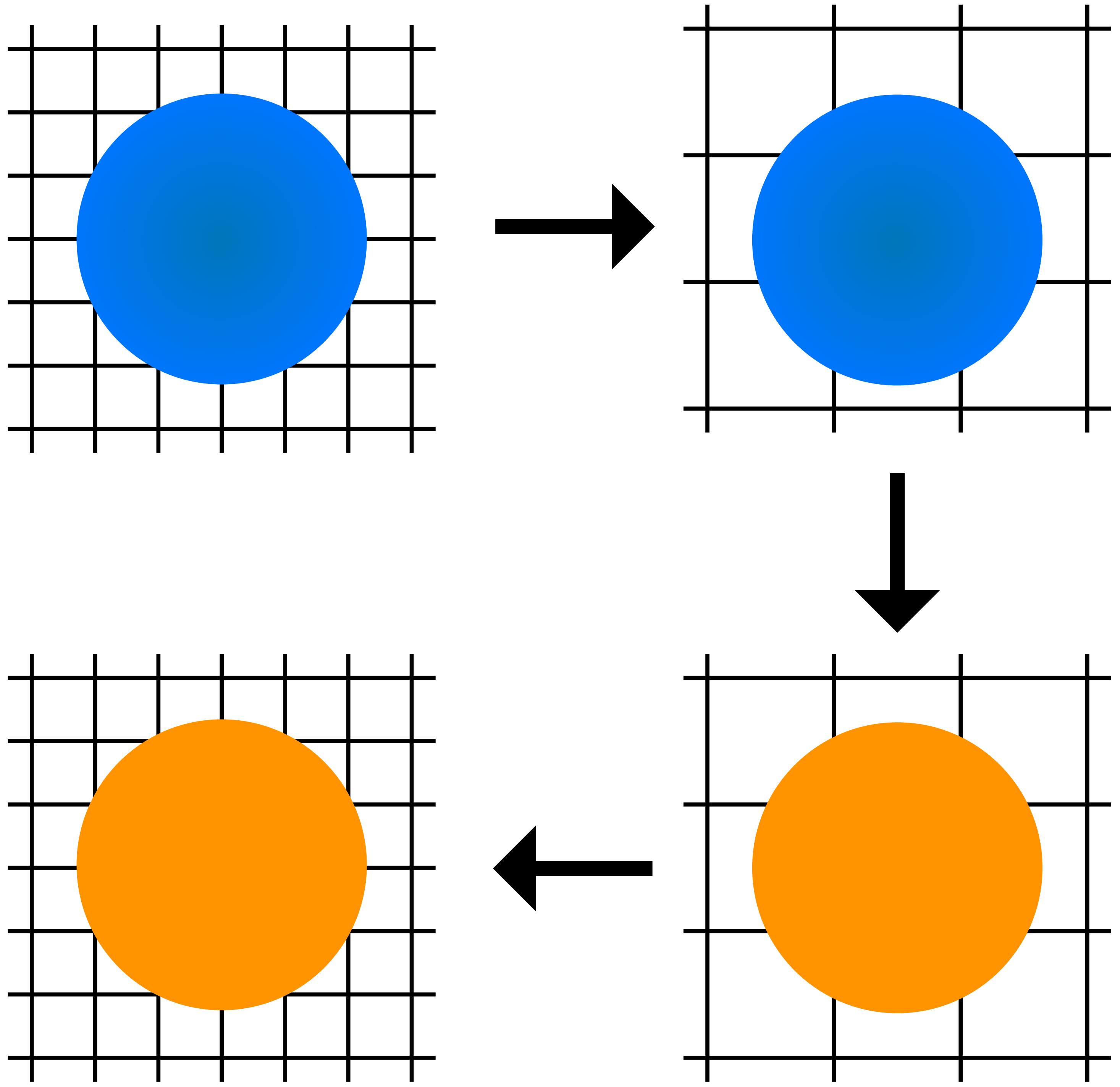


- Static bias potential: Multicanonical ensemble
  (Berg, Neuhaus, 1992)
- Dynamic buildup: Metapotential
  (Laio, Martinelli, Sanfilippo, 2015)

What to do?

Many other ideas (will not be covered)

- Update on coarse lattices via Wilson flow
  (Lüscher, 2009)
- Multi-scale thermalization
  (Endres et. al., 2015)
- Non-equilibrium ensembles (abandon detailed balance)
  (Pinto Barros, Marinkovic, 2022)
- Generate configs via ML (direct sampling or hybrid)



Metadynamics

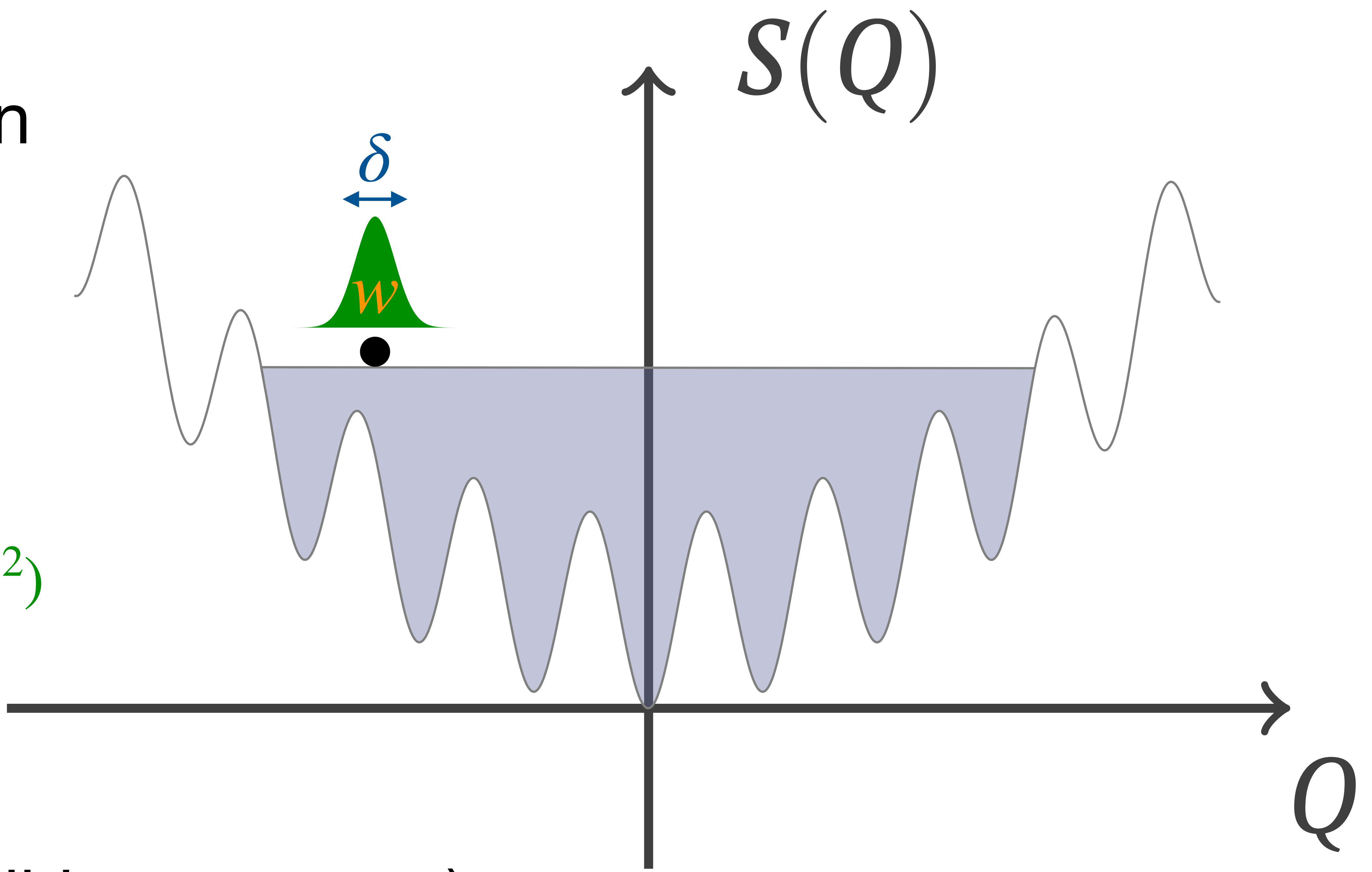
General idea:

- Pick collective variable(s): $Q_{\rm meta}$
- Add time dependent bias potential $V(Q_{\rm meta})$ to action
- Build up $V(Q_{\rm meta})$: -Start with $V = 0$

  (simplest case) -When at Q, increase $V(Q)$:

  $$V(q) \to V(q) + we^{-(Q-q)^2/(2\delta^2)}$$

  Gaussian, weight w, width δ



- $V(Q_{\rm meta})$ approaches negative free energy (up to additive constant)
- Probability density flattened over collective variable(s)


$V(Q_{\mathrm{meta}})$
0
−20
−40
−60
−80
−100
−120
−140
Metropolis update
Metropolis + Instanton update
−15
−10
−5
0
5
10
15
Q_{meta}

Topological charge definitions

2D U(1): Integer definition:

$$Q = \frac{1}{2\pi}\mathrm{Im}\left(\sum_x \ln P_{tx}\right)$$

Continuous definition:

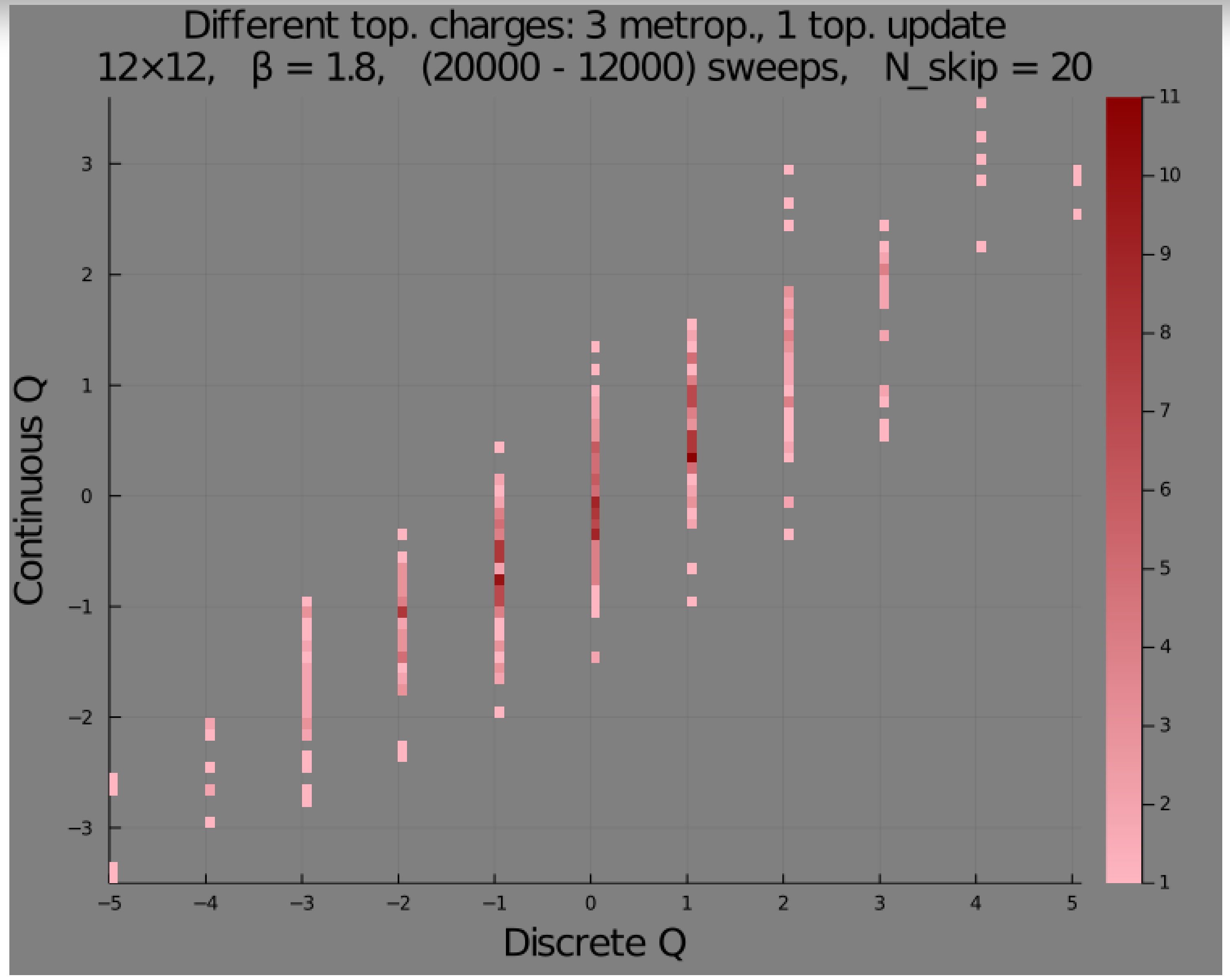
$$Q_{\mathrm{cont}} = Q_{\mathrm{meta}} = \frac{1}{2\pi}\mathrm{Im}\left(\sum_x P_{tx}\right)$$

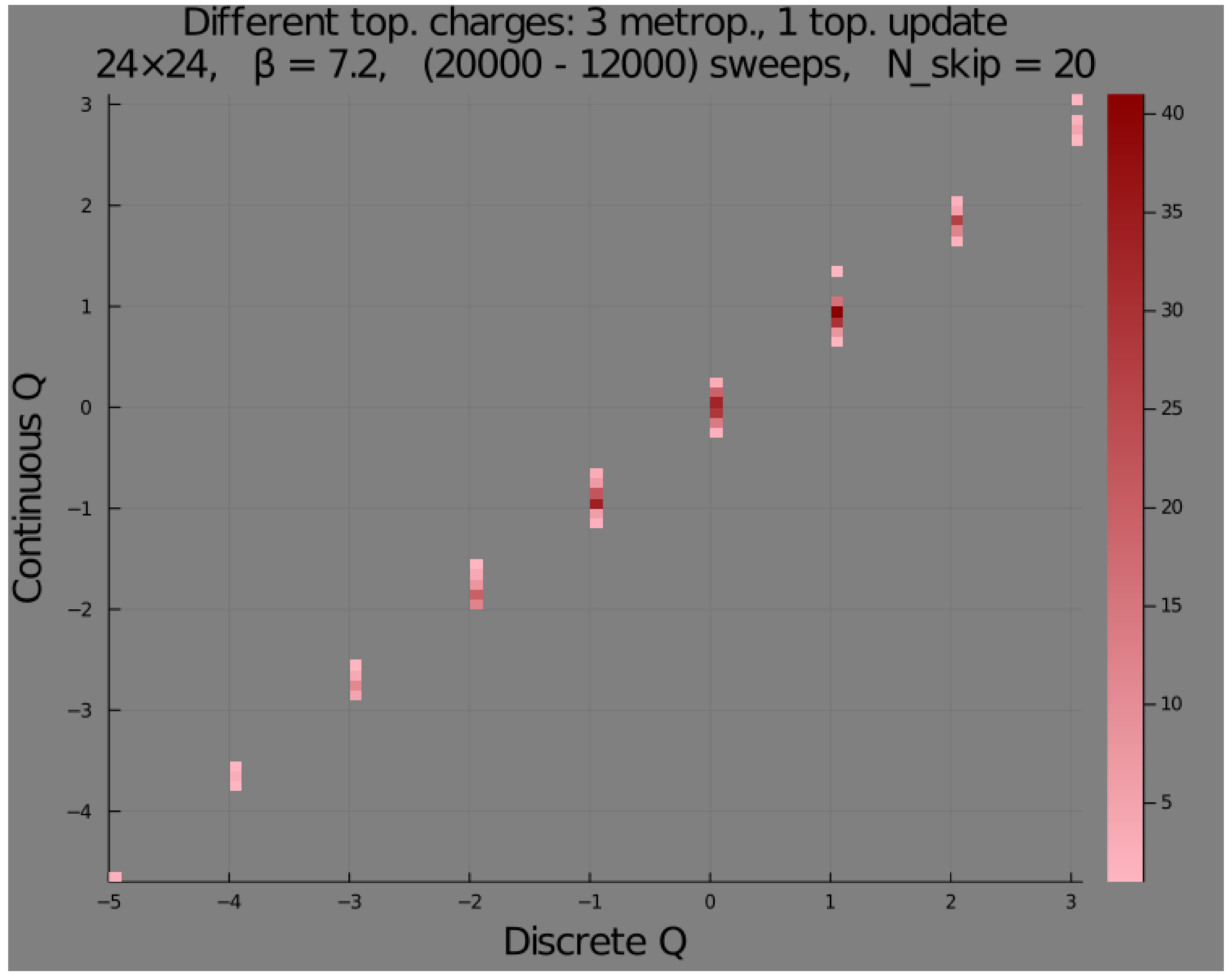
4D SU(3): $Q_c = \frac{1}{32\pi^2}\sum_x \mathrm{Tr}\left(F\tilde{F}\right)$

F, $\tilde{F}$ from clover averages with

~30 step stout smearing, $\rho = 0.12$

Q_{meta} with less smearing





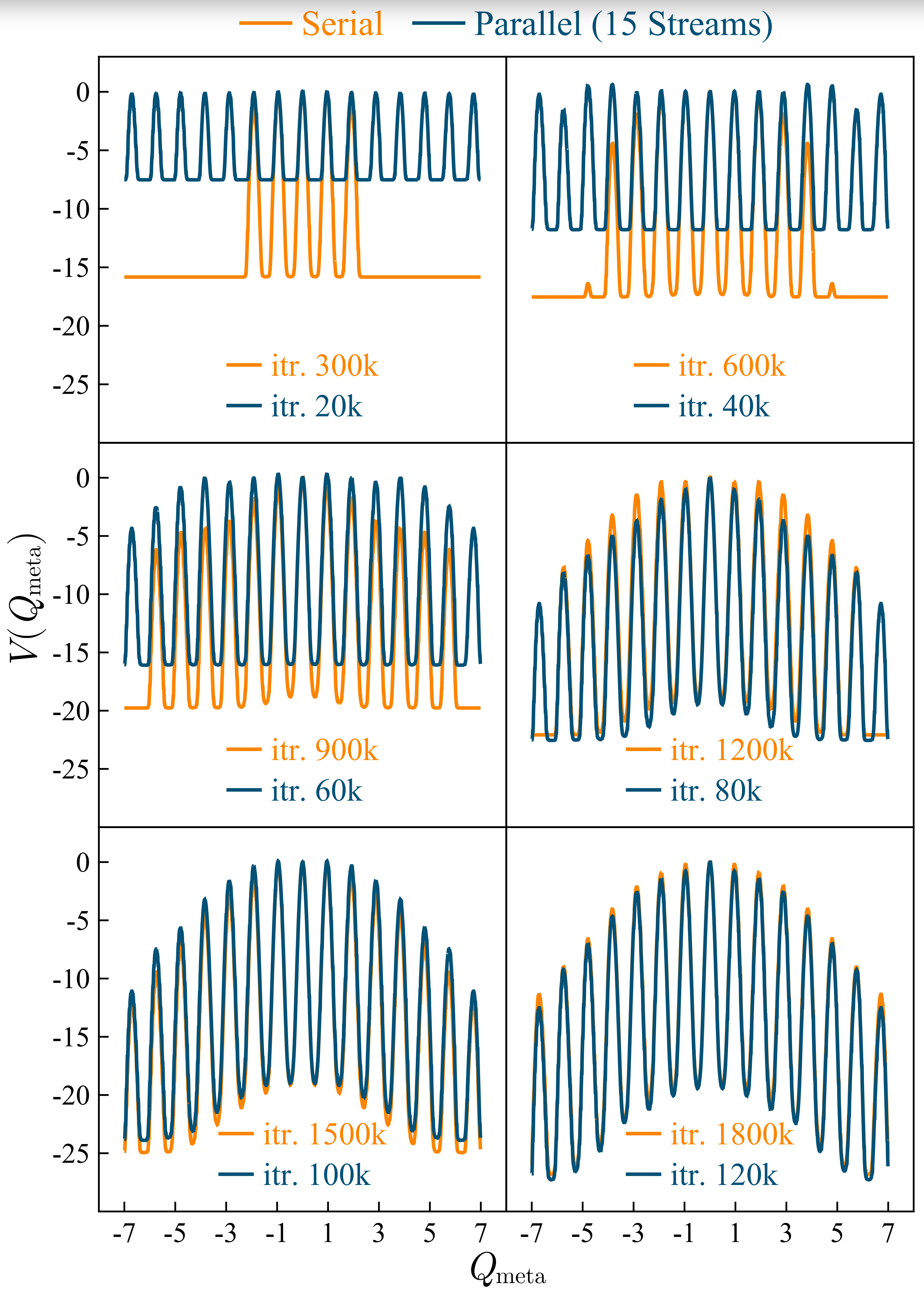
Parallel buildup

Bias potential buildup may be efficiently parallelized:

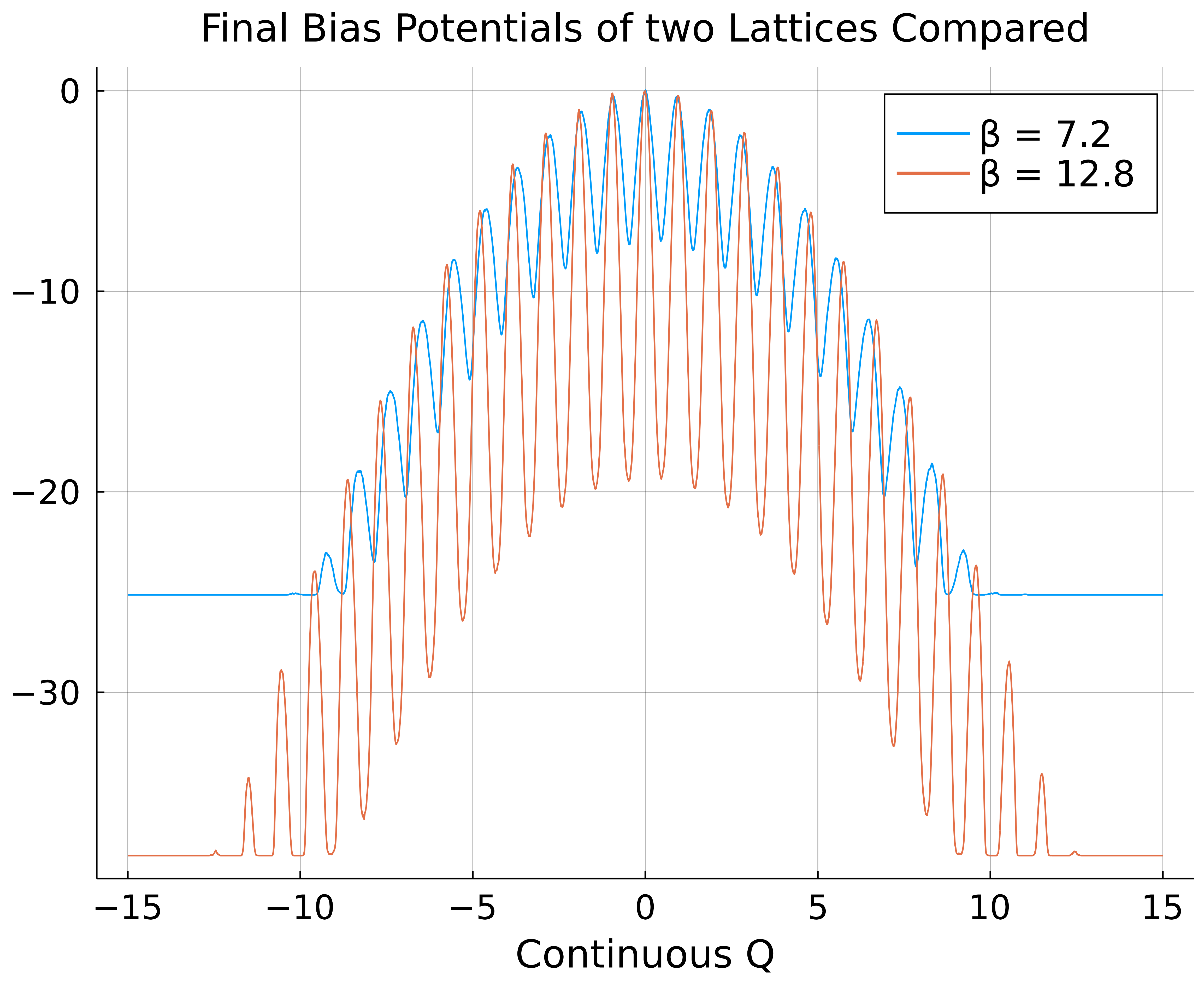
✓ Start in different sectors

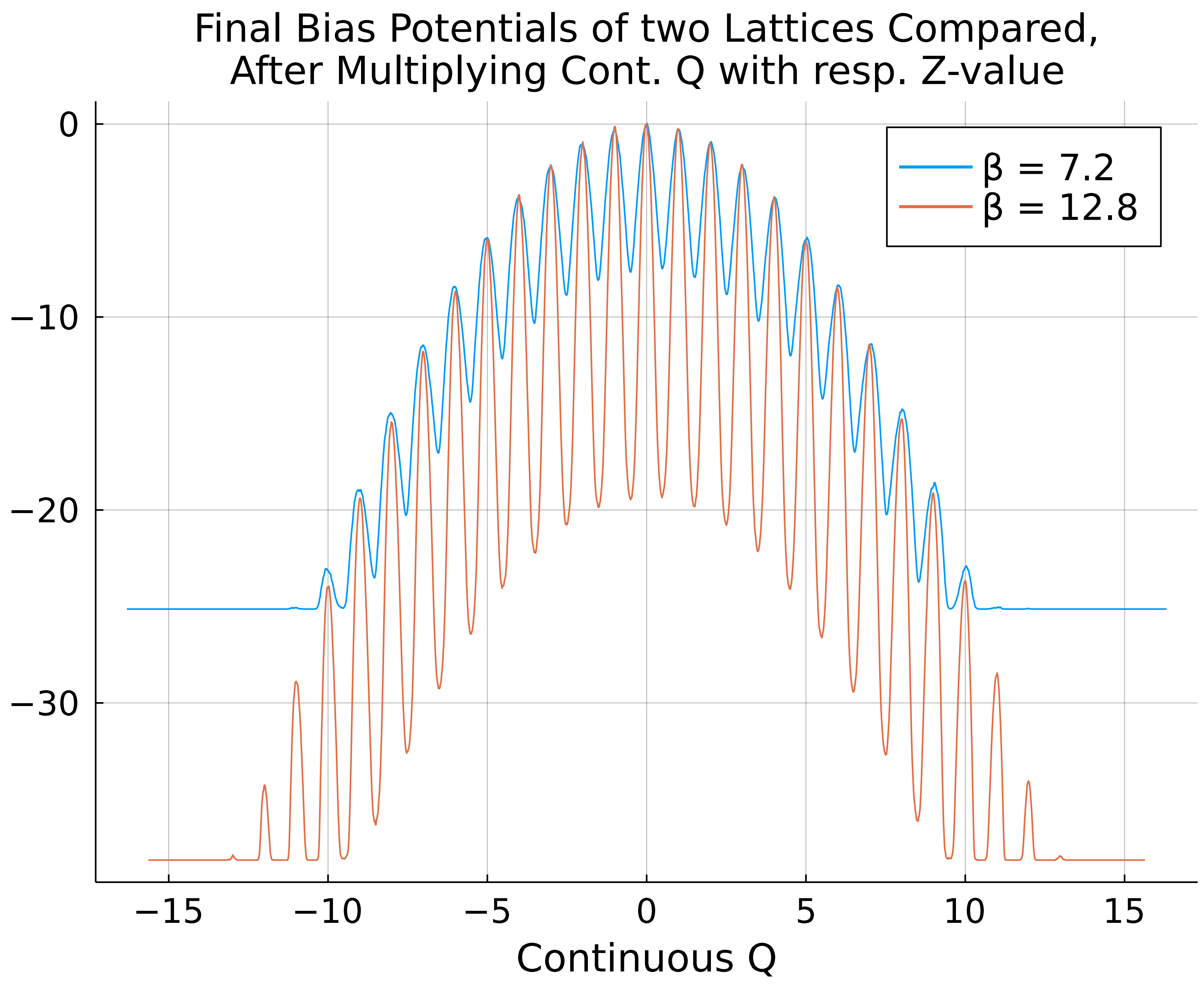
✓ Only bias potential needs to be communicated

✓ Convergence at least as good as single stream



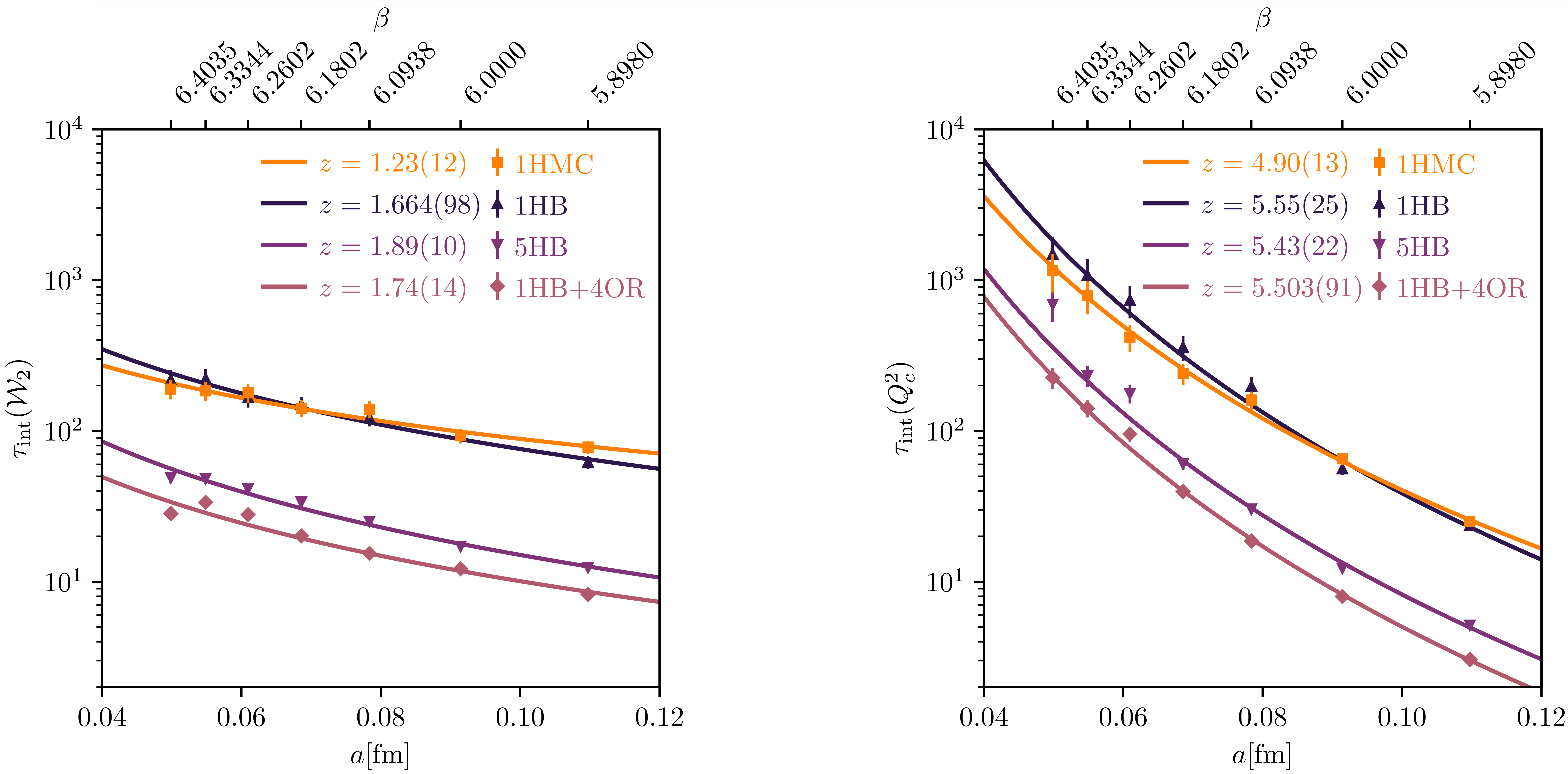
Bias potential in 2D U(1)





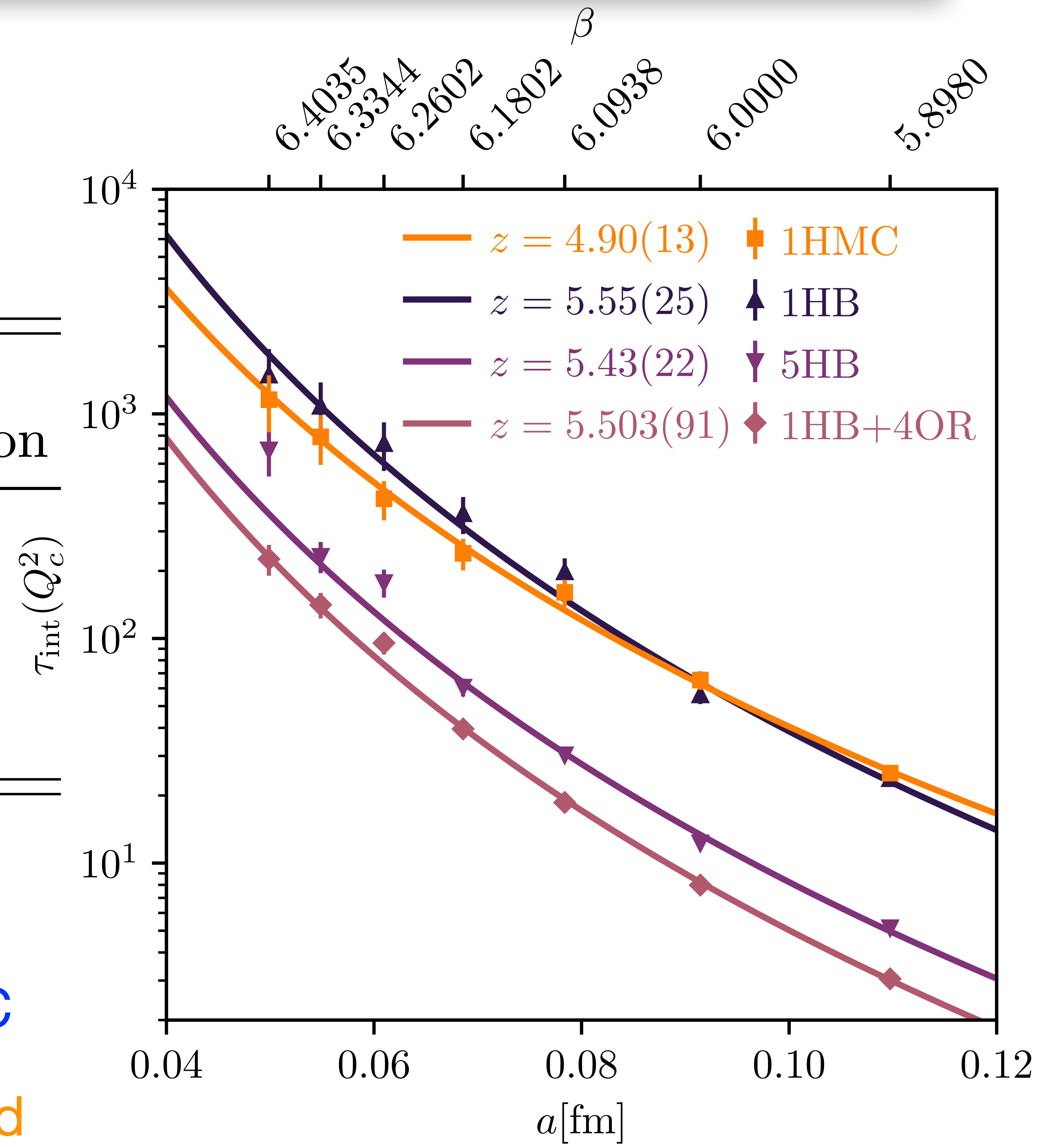
Strong dependence of barrier height on lattice spacing

Standard algorithm continuum scaling in 4D SU(3)



Standard algorithm continuum scaling in 4D SU(3)

Update scheme	Relative time	
	Wilson action	DBW2 action
1HB+4OR	1	1
1HMC	3.56	4.46
MetaD-HMC (`4stout`)	95.48	31.03
MetaD-HMC (`5stout`)	114.02	36.37



Q_{meta} expensive to compute (smearing):

local updates not feasible → HMC

HMC requires $\partial Q_{\mathrm{meta}} / \partial U_\mu$ → stout force recursion needed

Metadynamics in 4D SU(3)

Wilson plaquette action, $\beta = 6.4035$, $V = 22^4$

Bias potential buildup

DBW action, $\beta = 1.25$, $V = 16^4$

Large Q excursions as bias potential builds up

Metadynamics in 4D SU(3)

DBW action, $\beta = 1.25$, $V = 16^4$

Problem solved? Not quite!

Effective sample size

✓ Autocorrelation time of Q^2 dramatically improved:

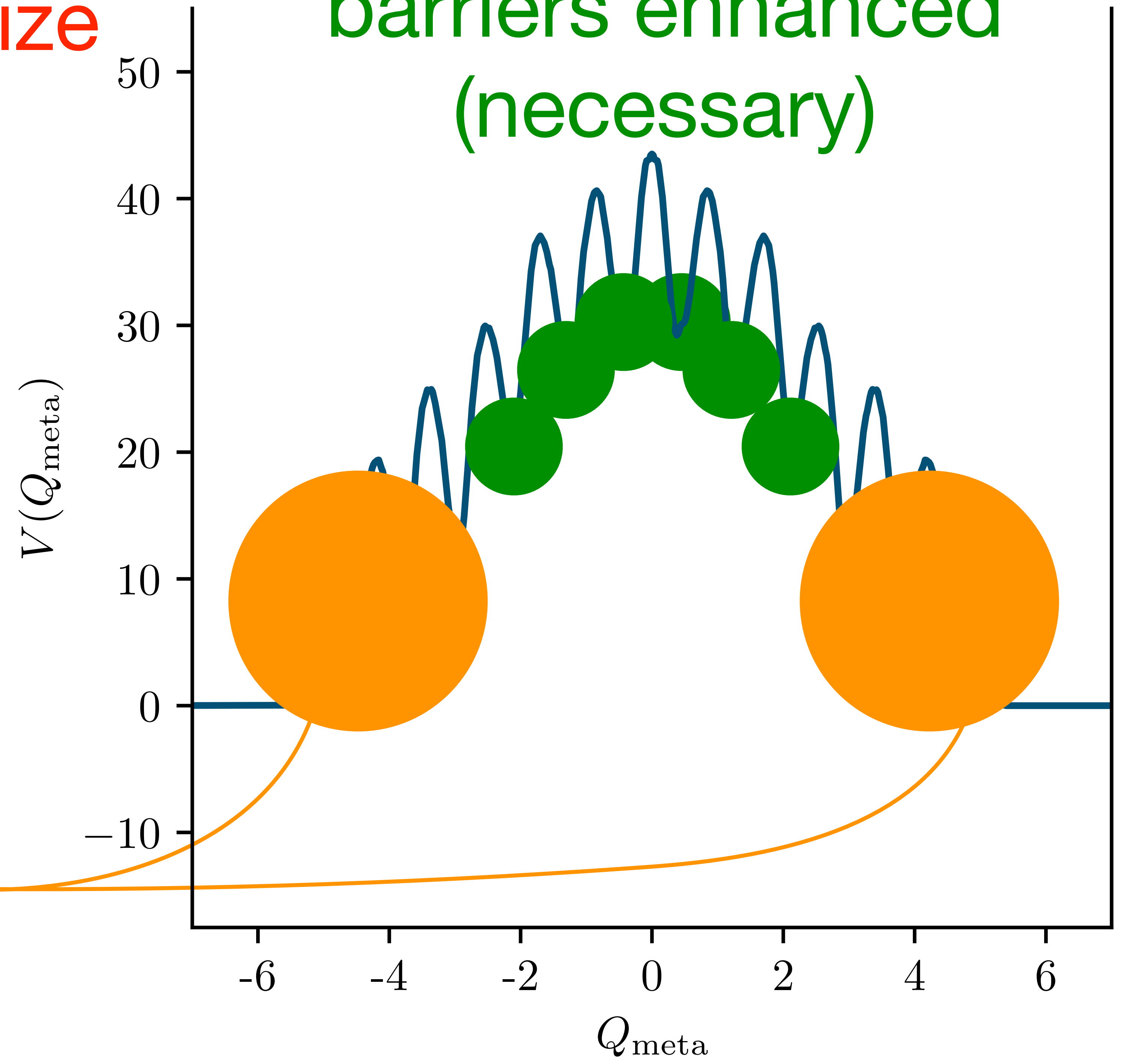
$\tau_{\text{int}}(Q^2)$: from $\gtrsim 4 \times 10^5$ (HMC, 1HB+4OR)

to ~ 2000 (MetaD-HMC)

✗ Relative weight of configs decreases with effective sample size

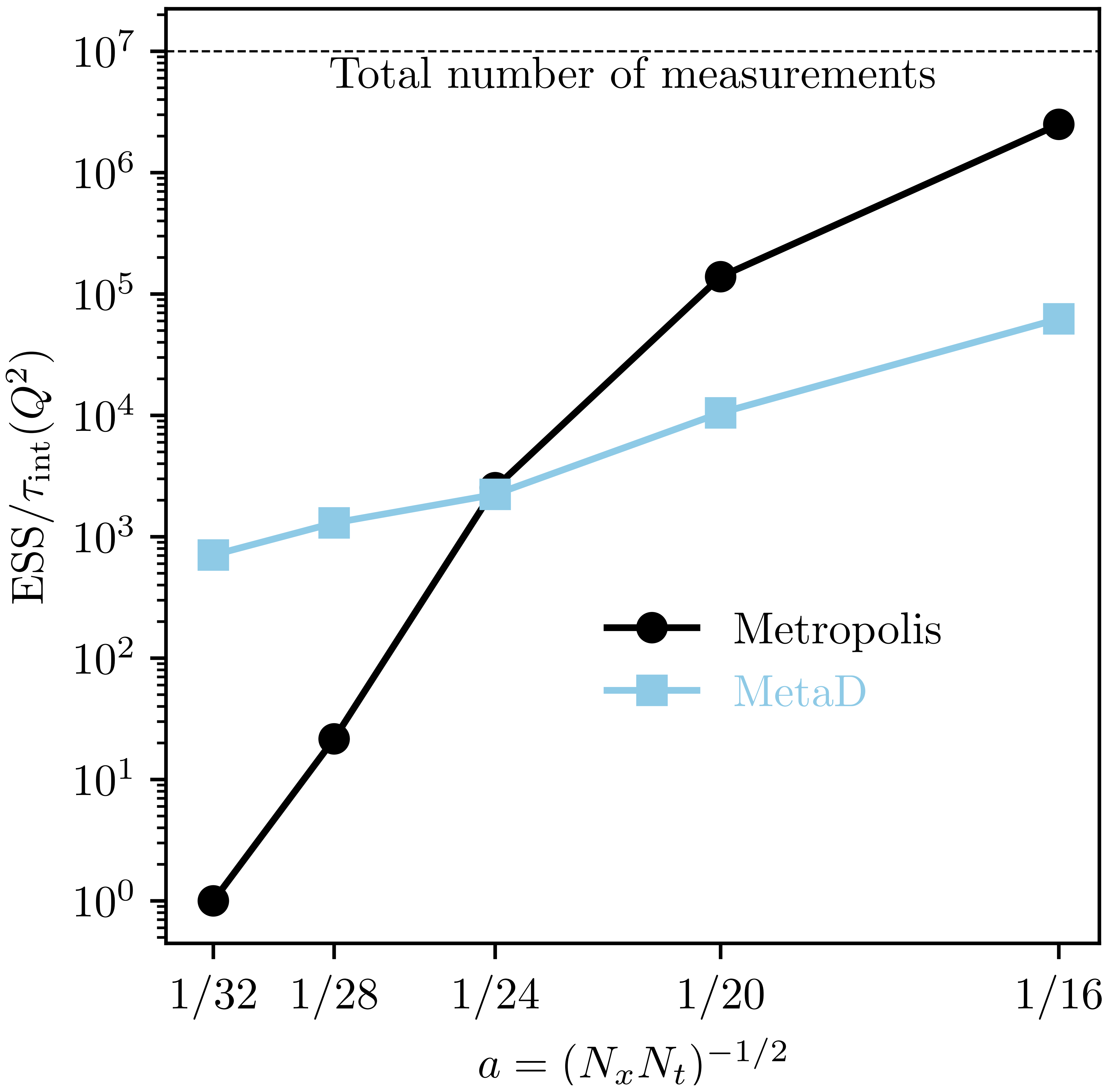
$$\text{ESS} = \frac{\left(\sum_i w_i\right)^2}{\sum_i w_i^2}$$

barriers enhanced (necessary)



large $|Q|$ enhanced (unwanted)

continuum scaling in 2D U(1)

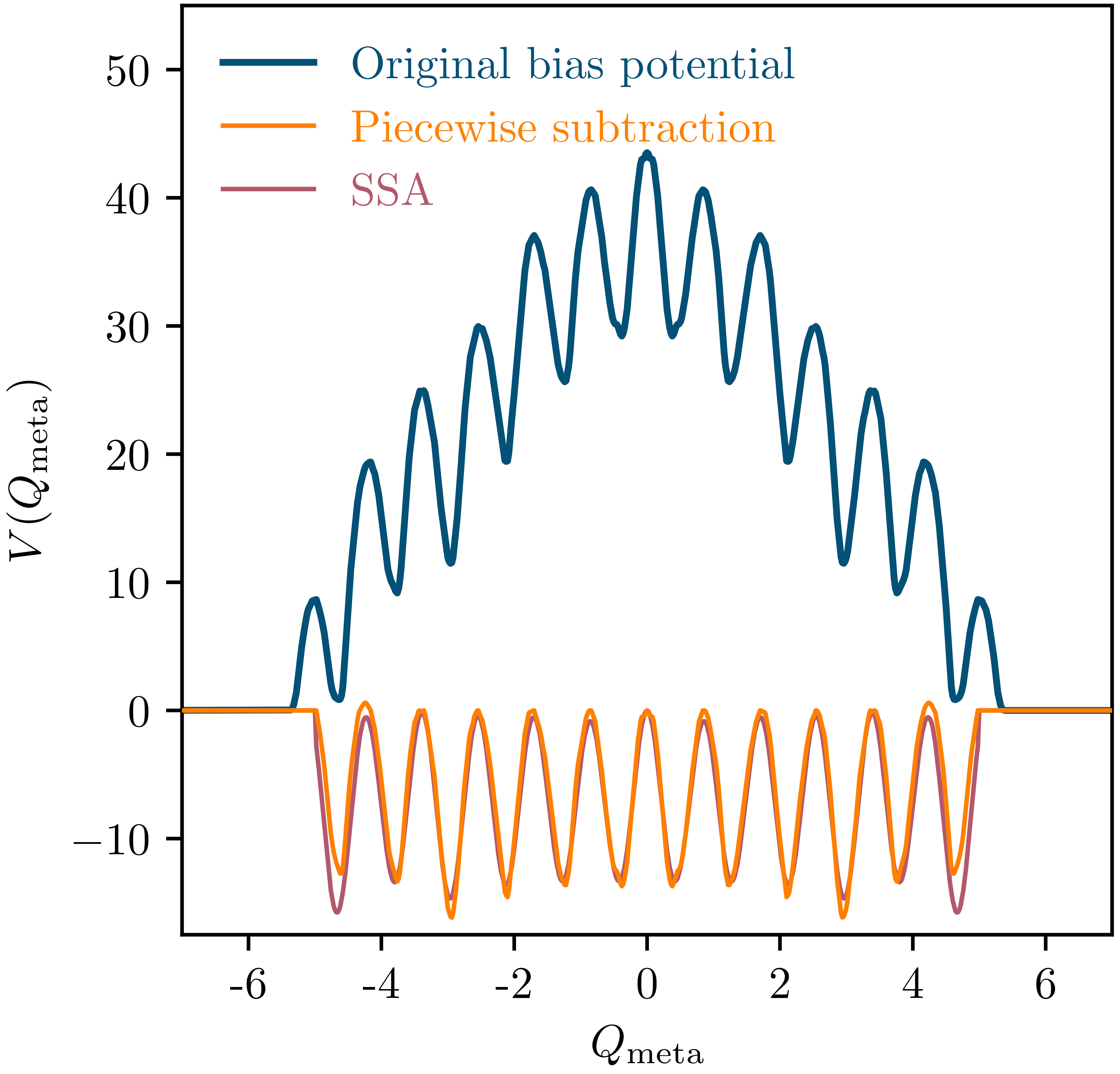


Modifying the bias potential

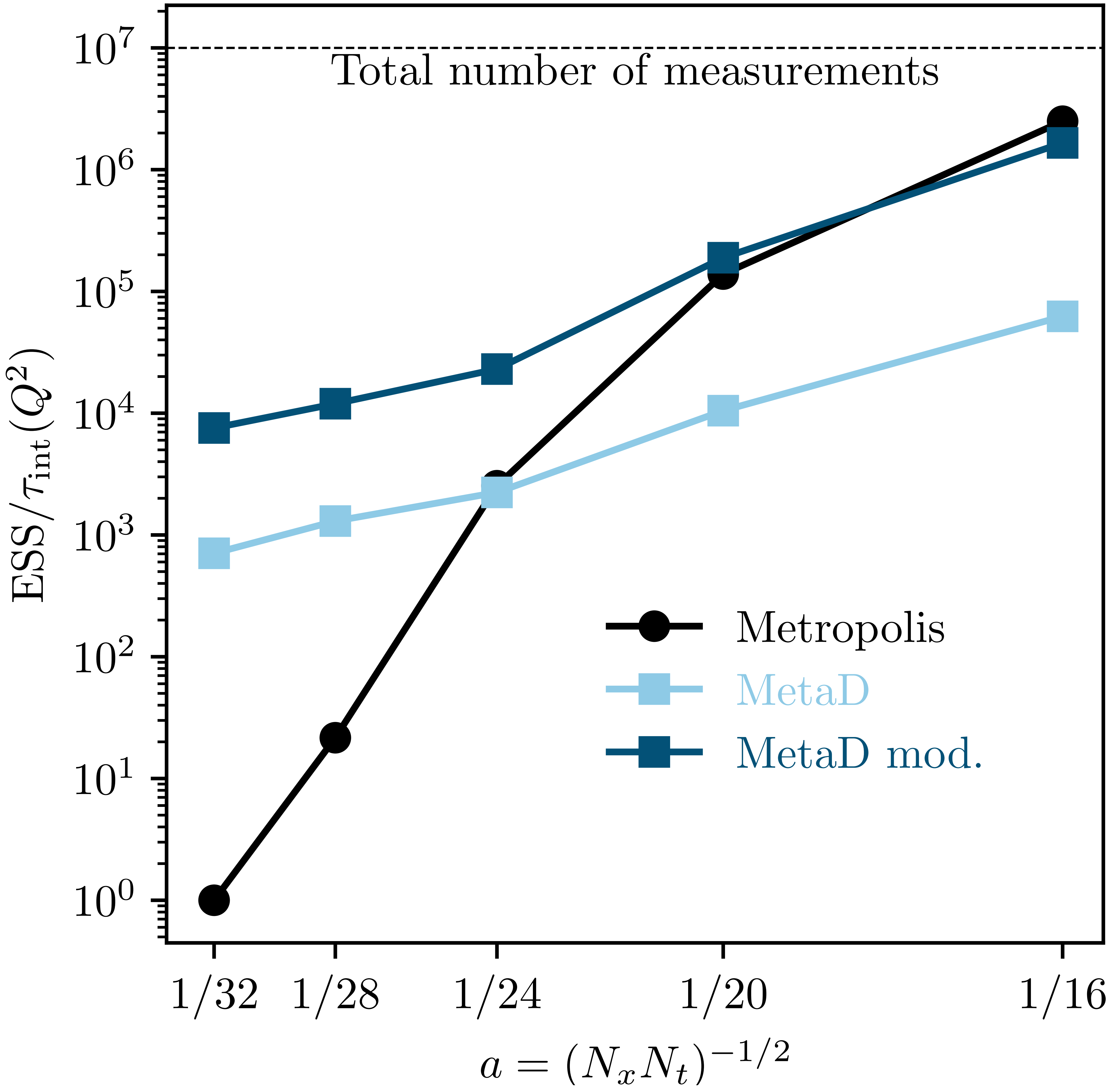
Remove long range (quadratic) term:

✓ Curbs large $|Q|$ excursions

✗ Small weight inter-barrier configs remain



continuum scaling in 2D U(1)



Technical details:

2D U(1): parametric description (parabolic + Fourier)

4D SU(3): piecewise subtraction, singular spectrum analysis, …
result insensitive to details of procedure

Parallel Tempered Metadynamics

Fully eliminate ESS problem:

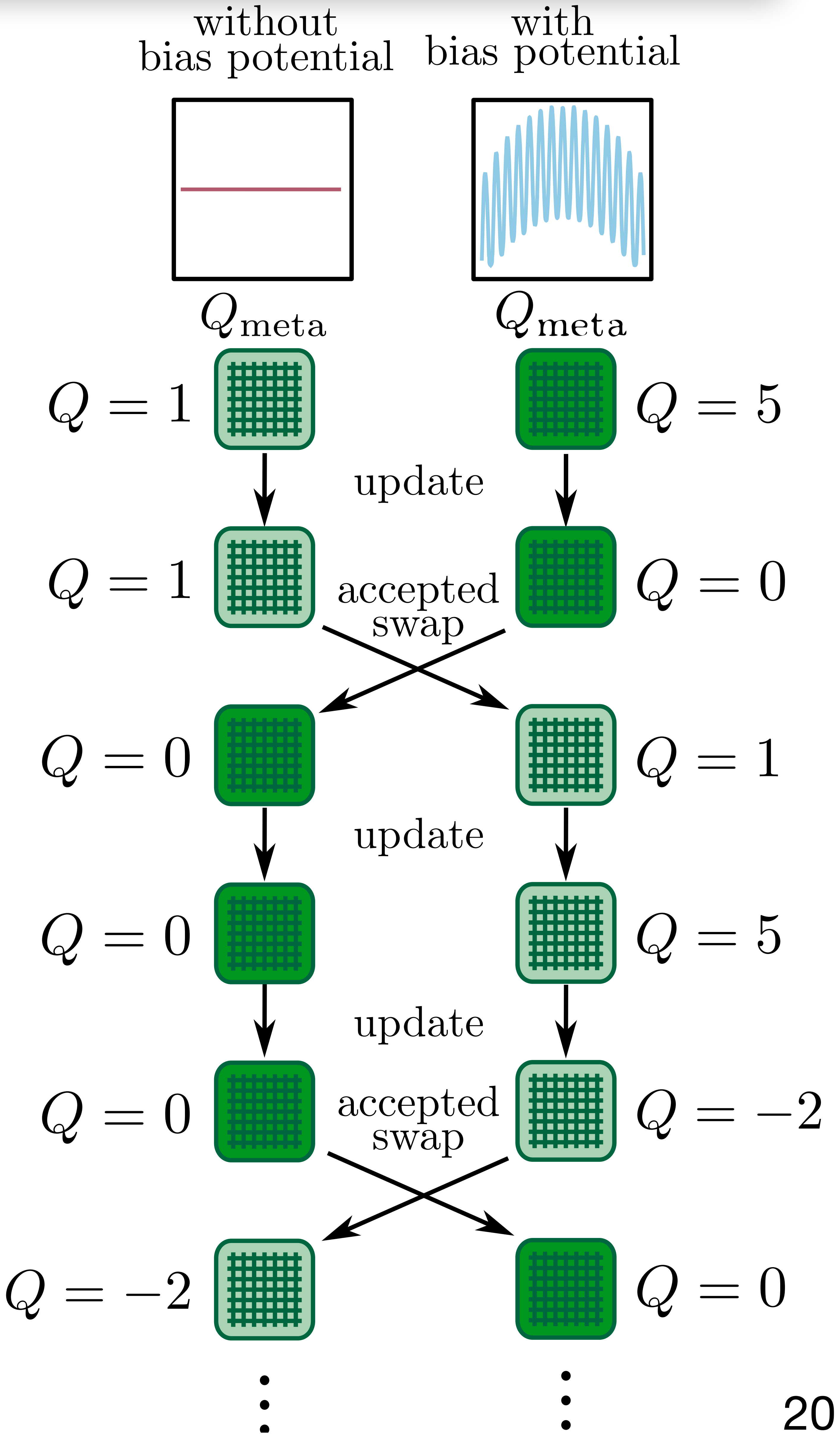
- One conventional, efficient update stream (frozen, no ESS problem) (measurement)
- One stream with a fixed bias potential (unfrozen, ESS problem) (tunneling)
- Periodically propose swaps, accept with Metropolis step

✓Same physical system → no physical action difference

✓Swap acceptance $p = \min(e^{V(Q_{\text{meta},1})-V(Q_{\text{meta},2})}, 1)$

✓No restriction on update algorithms in streams

✓Trivially expandable to multiple streams

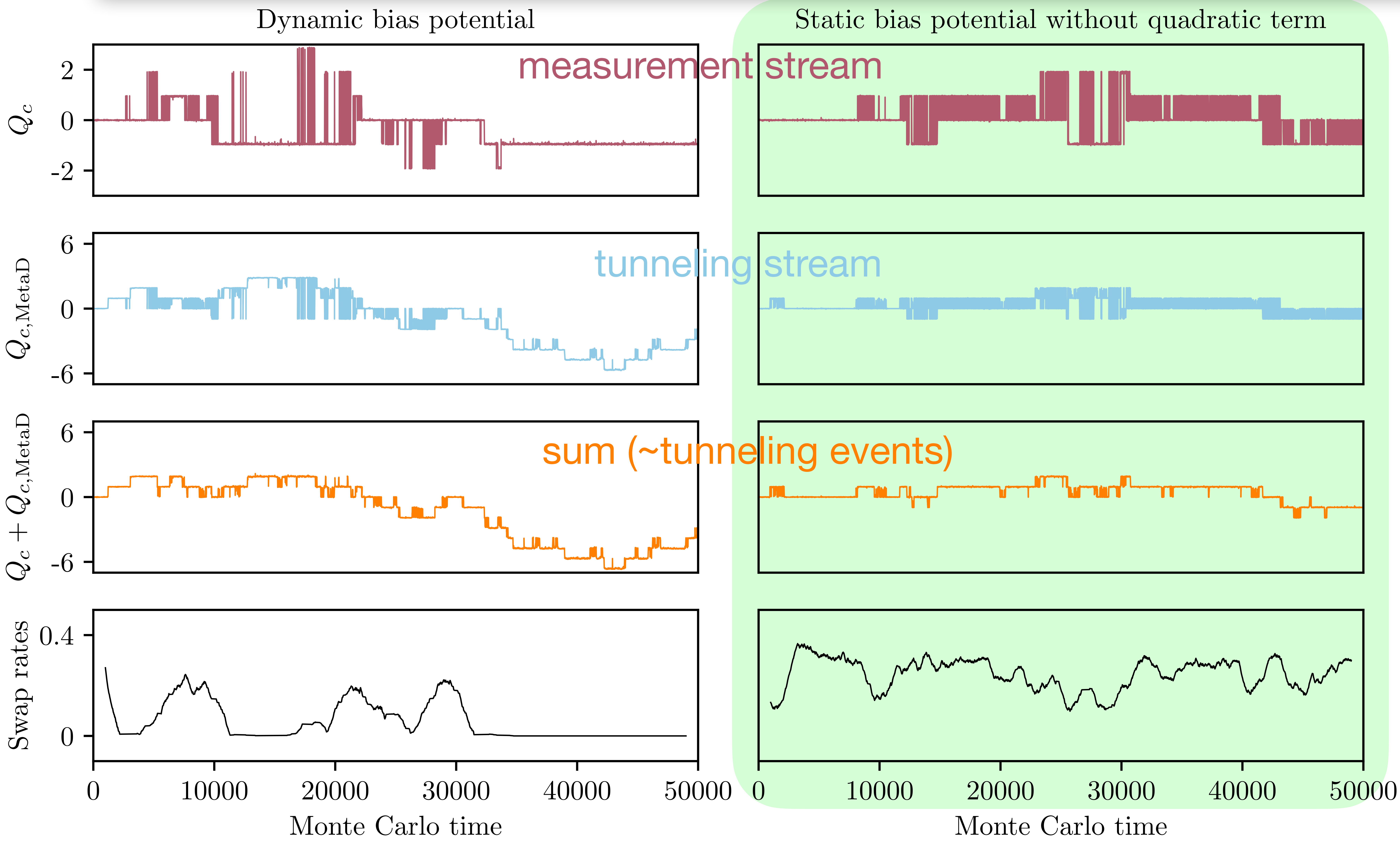


How does it perform?

Test in 4D SU(3)

✔ Tunneling in measurement stream ✔ ESS=#configs ✔Large $|Q|$ excursions curbed

Test in 4D SU(3)

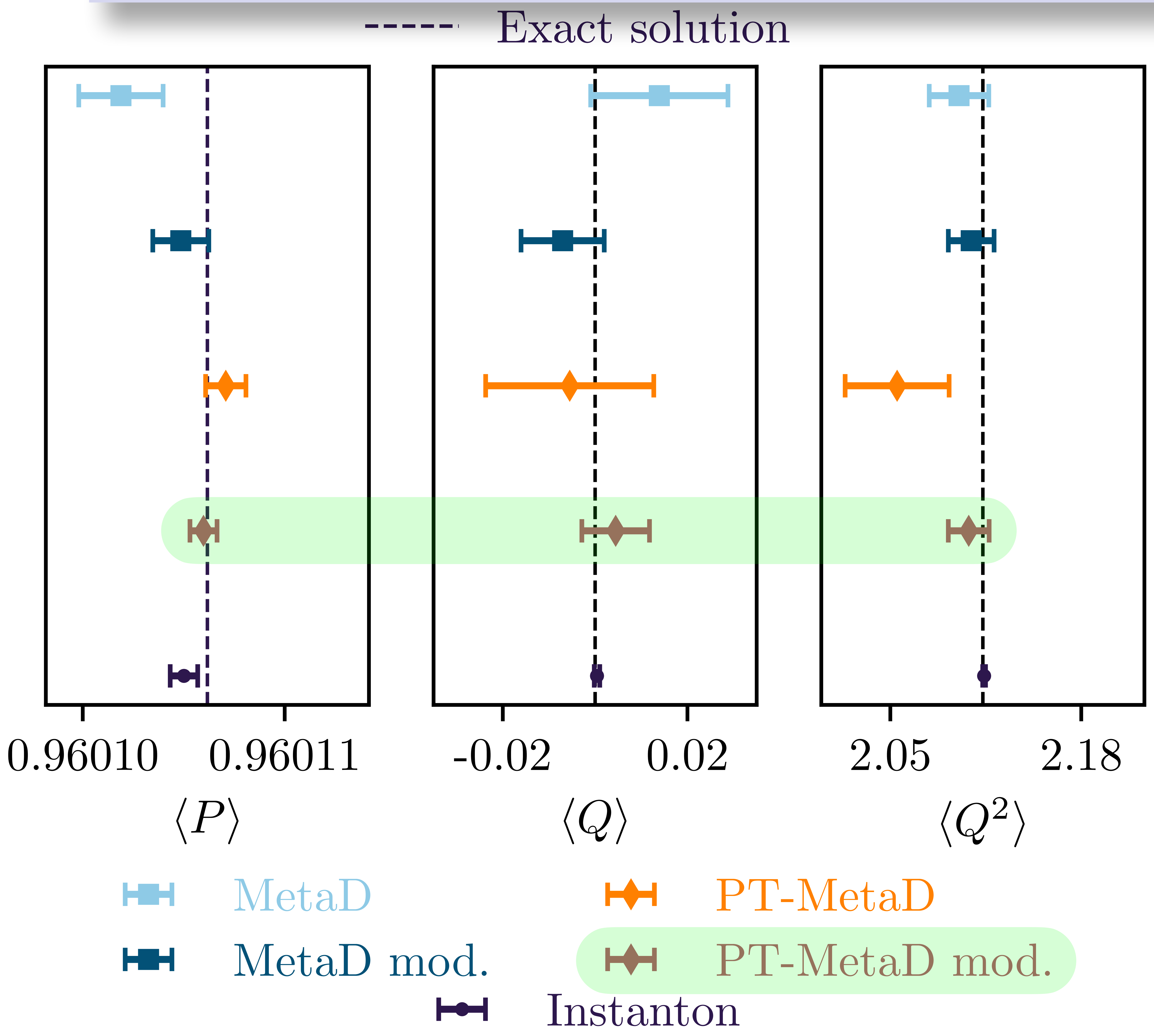


best results:

- static bias potential
- quadratic term removed

swap acceptance: ~0.25

2D U(1) comparison



✓Topological observables as good as MetaD

swaps with MetaD stream

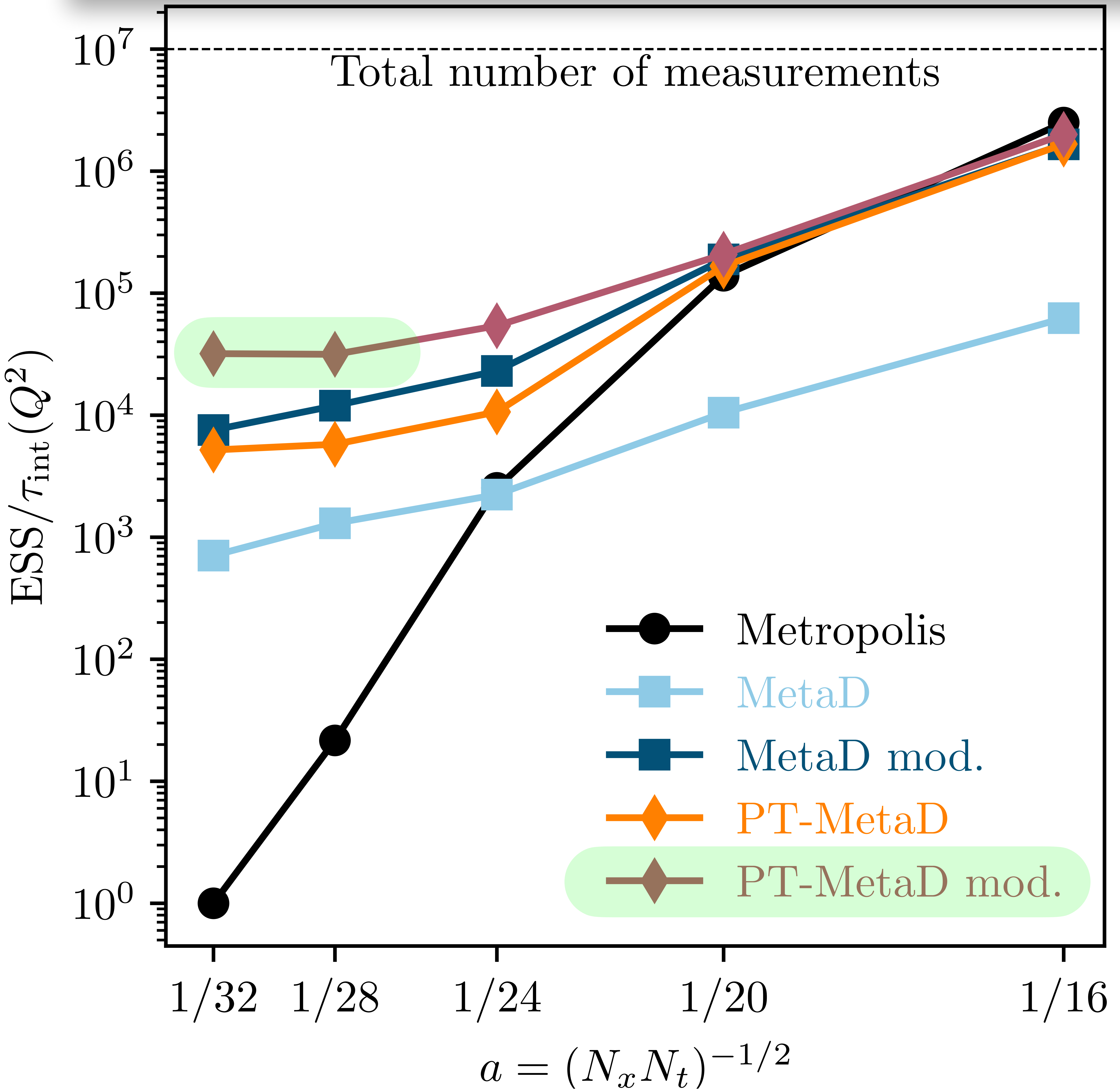
✓Plaquette improved

no reweighting (ESS=#configs)

✗Instanton hit still better

only 2D U(1) as we saw

Scaling in 2D U(1)



✓Indication for constant continuum scaling

How about QCD?

✓Tempering swaps totally independent of update in streams

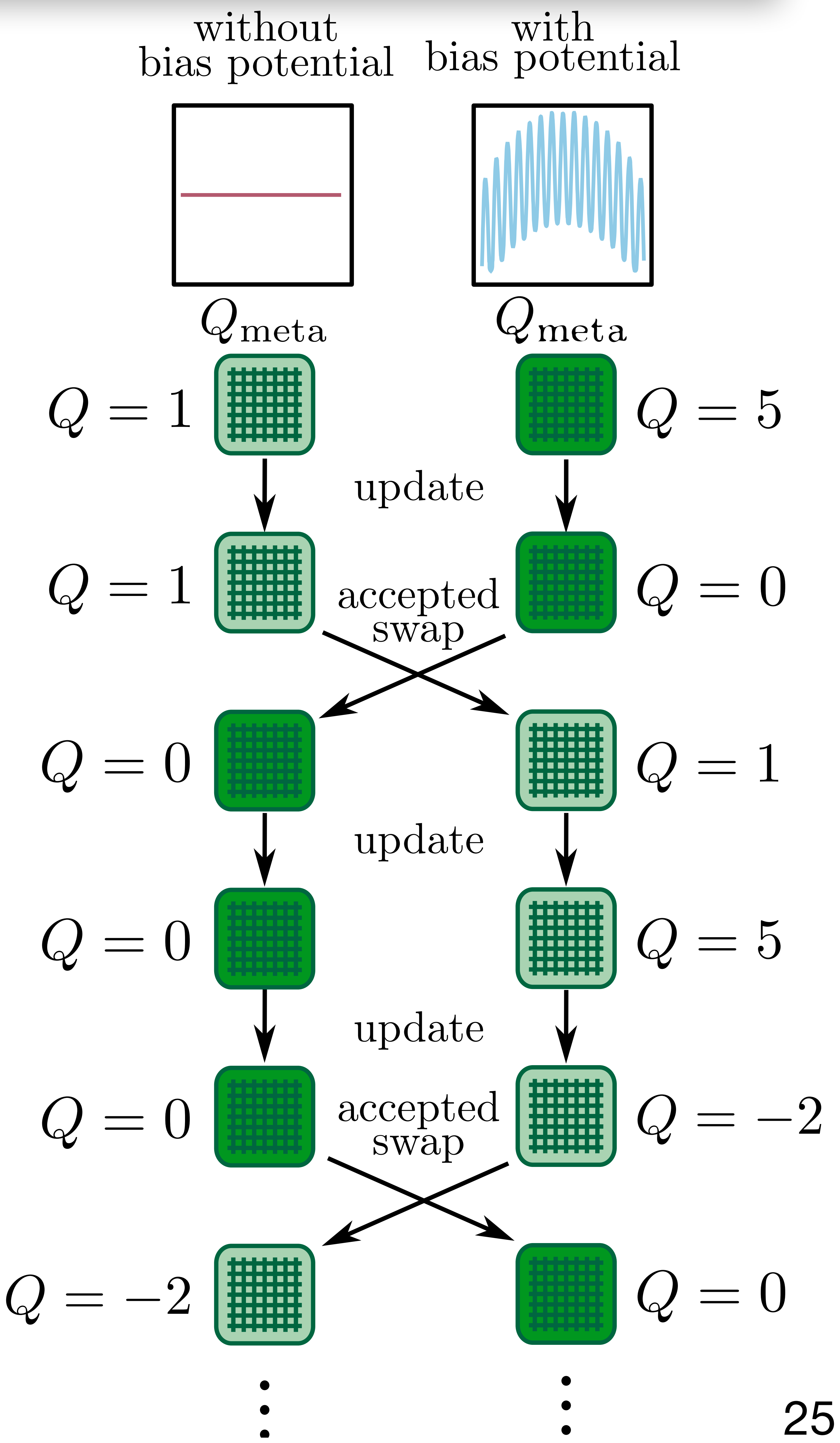
✓No need to compute physical action, as it is the same in all streams

✓Swaps only need $Q_{\rm meta}$, same as in pure gauge theory

✓Streams are trivially parallelizable

✓Relative overhead of $Q_{\rm meta}$ computation much smaller

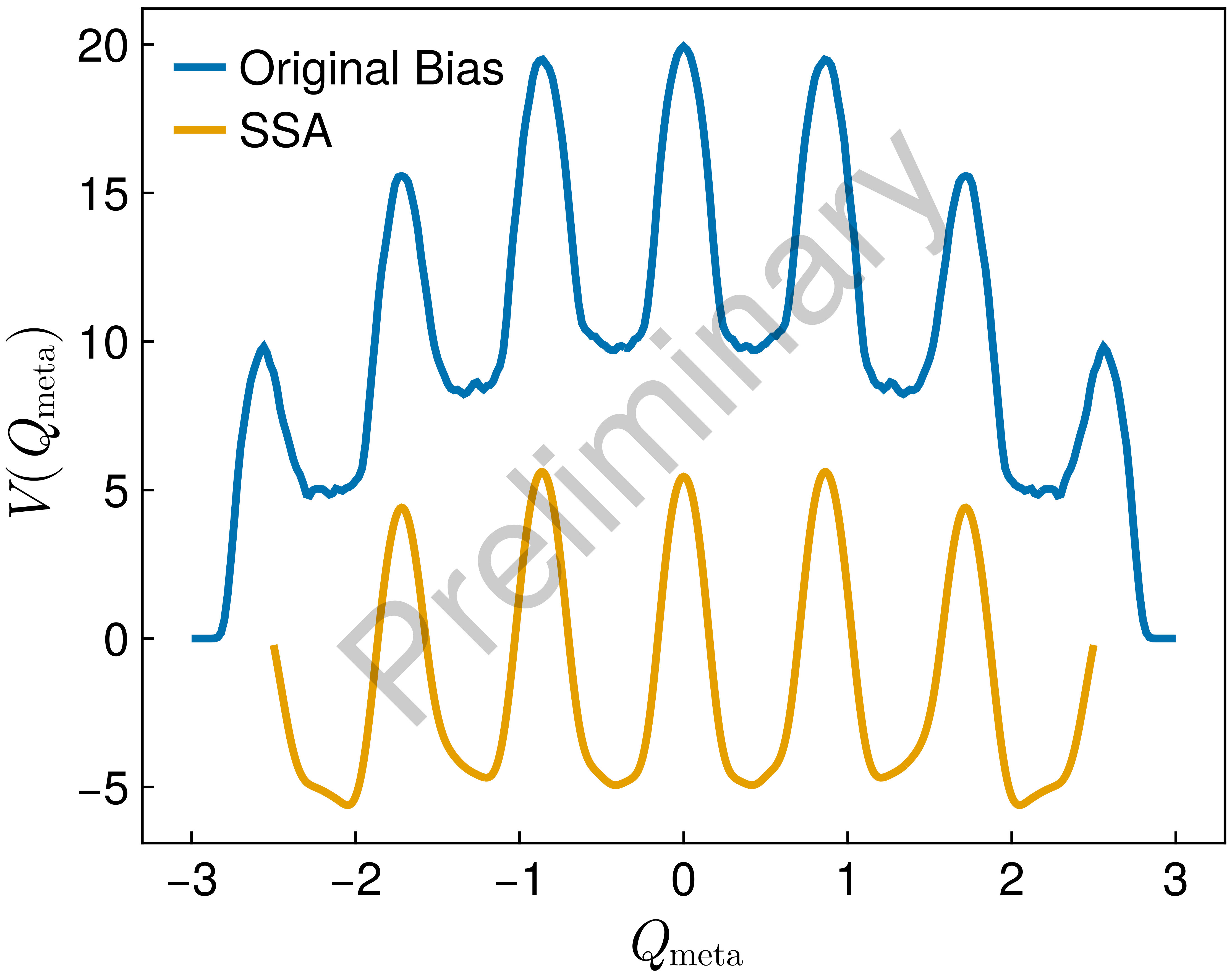
full QCD should just work



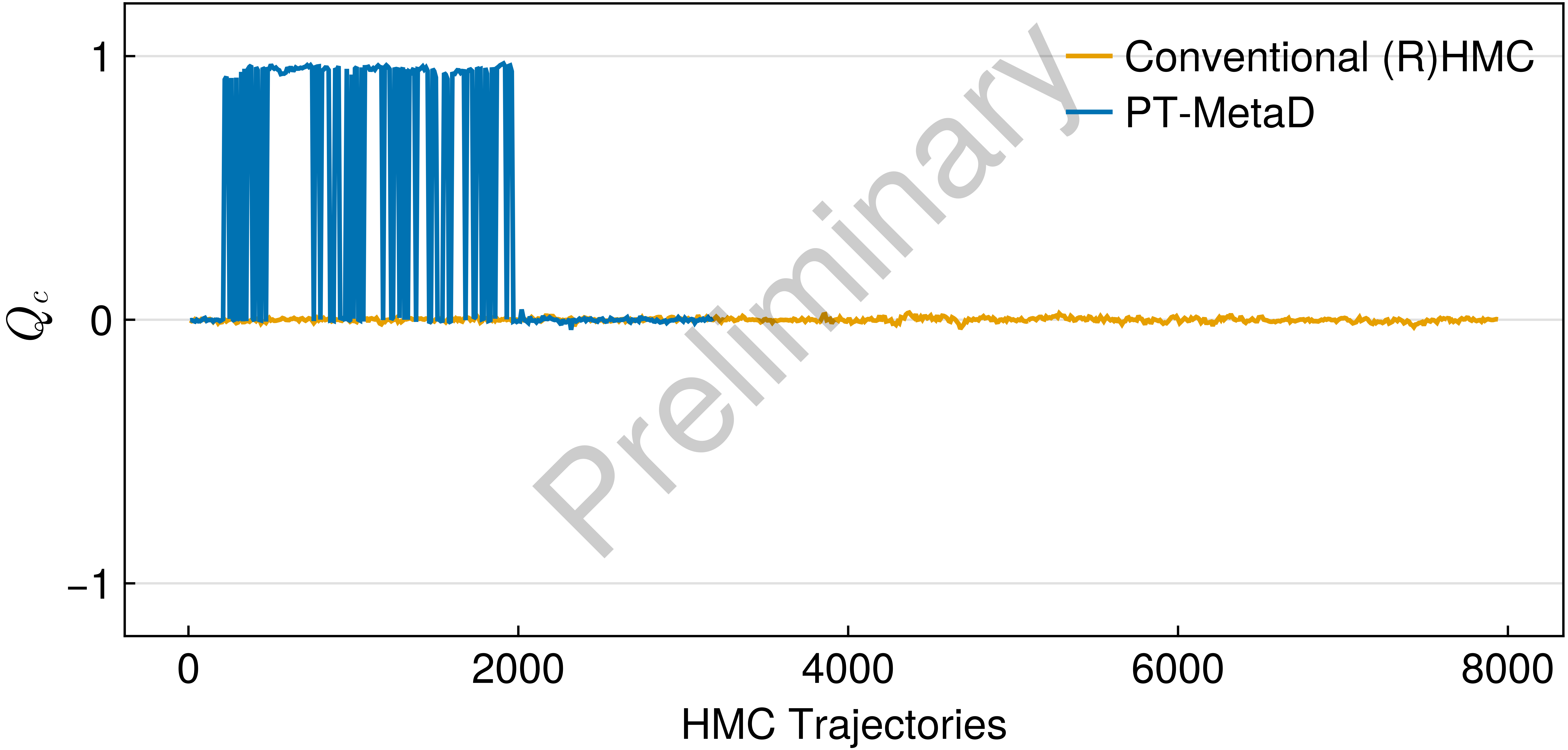
Full QCD: first look

- DBW2 gauge action, $\beta = 1.25 \quad \rightarrow \quad a^{-1} \sim 3.5 - 4$ GeV
- $N_f = 2$ staggered fermions $ma = 0.02$, 4-stout $\rho = 0.125$
- $V = 16^4 \;\rightarrow\; L \sim 0.8$ fm (close to T_c)
- Q_{meta}: 6-stout $\rho = 0.12$ clover charge
- 6 walkers, $\delta = w = 0.2$

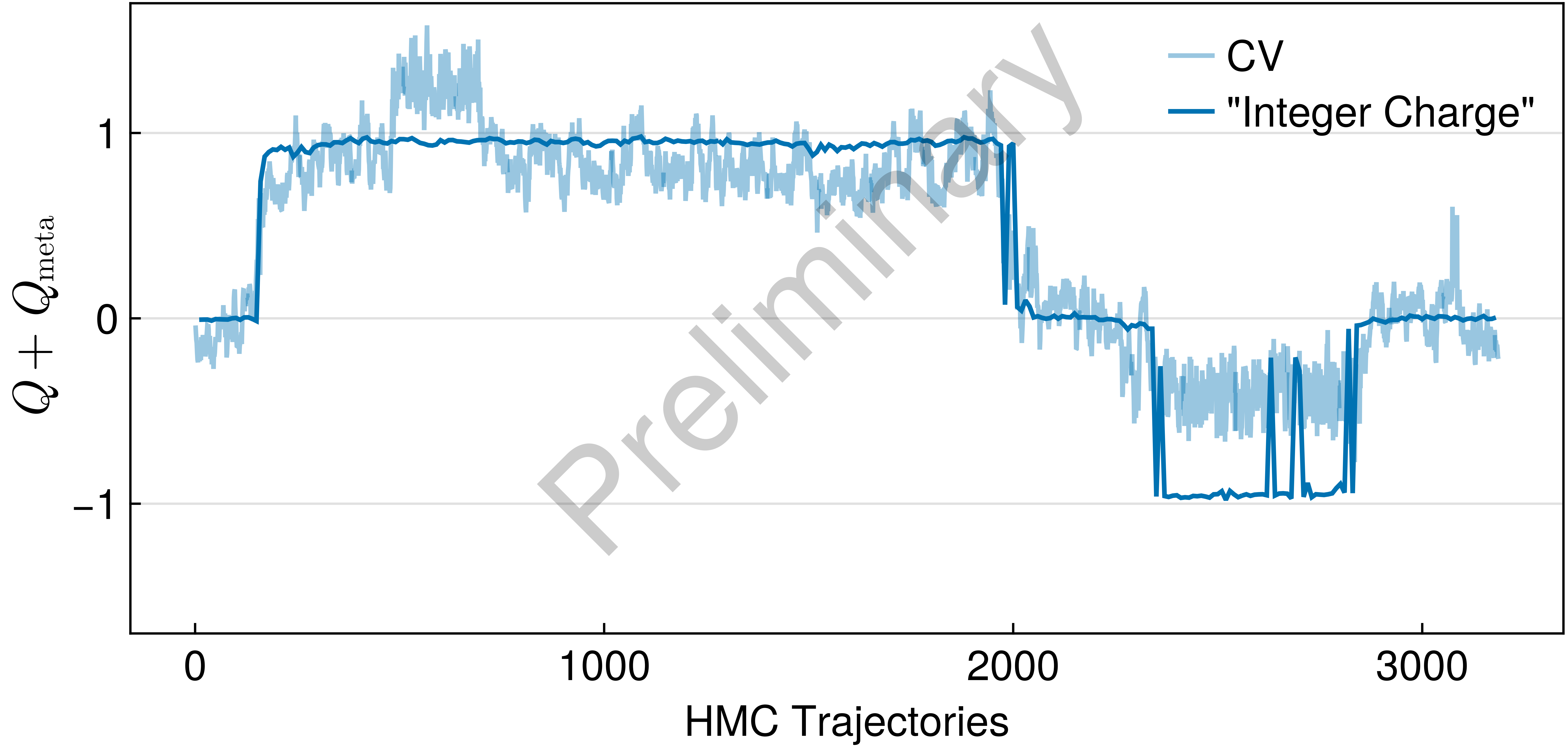
bias potential from 5000 trajectories



Full QCD: first look



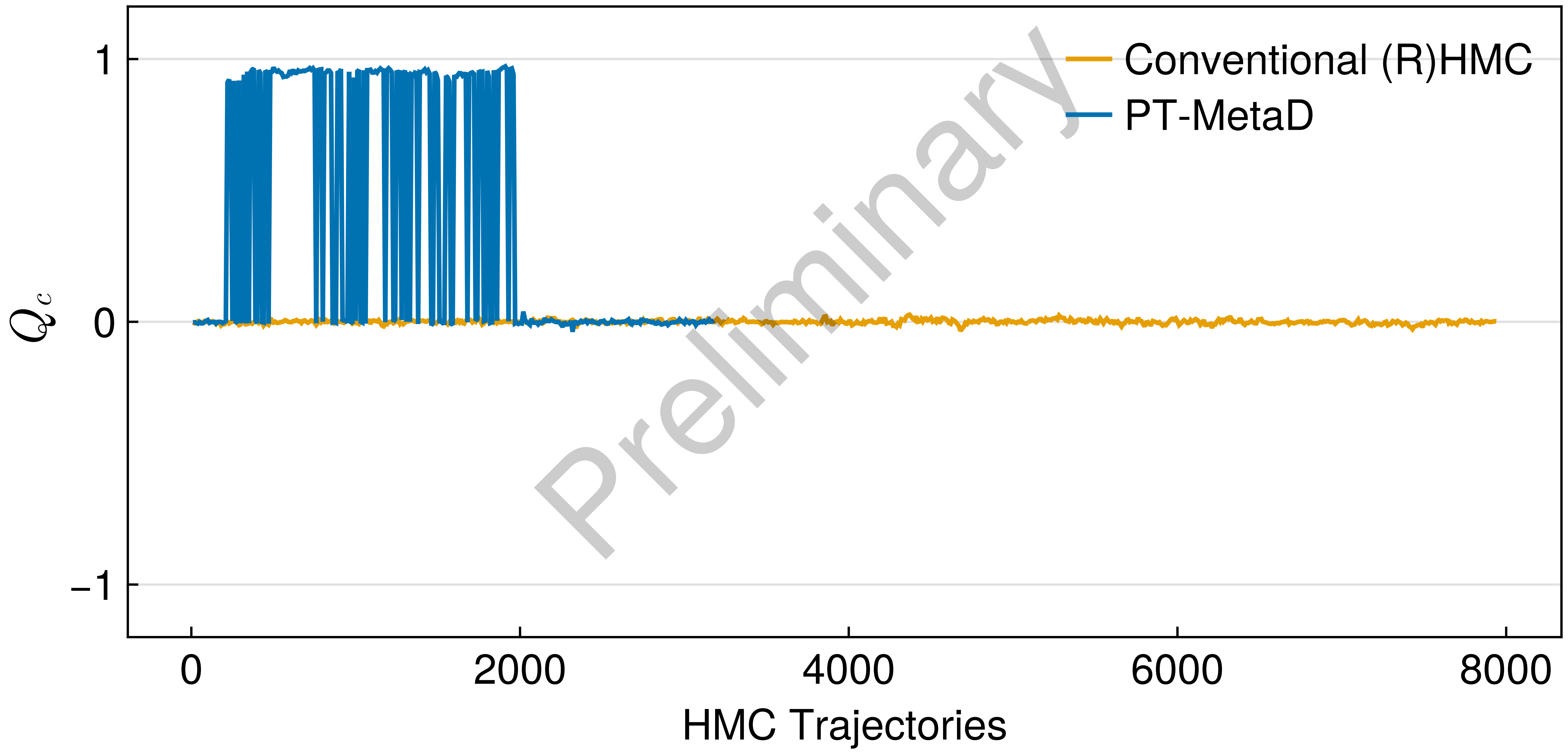
Full QCD: first look



Conclusions and outlook

- PT-MetaD works in QCD!
- Tunnels without ESS reduction
- Easy to implement, parallelize
- Open source codes on GitHub:

  MetaQCD.jl
  Lettuce



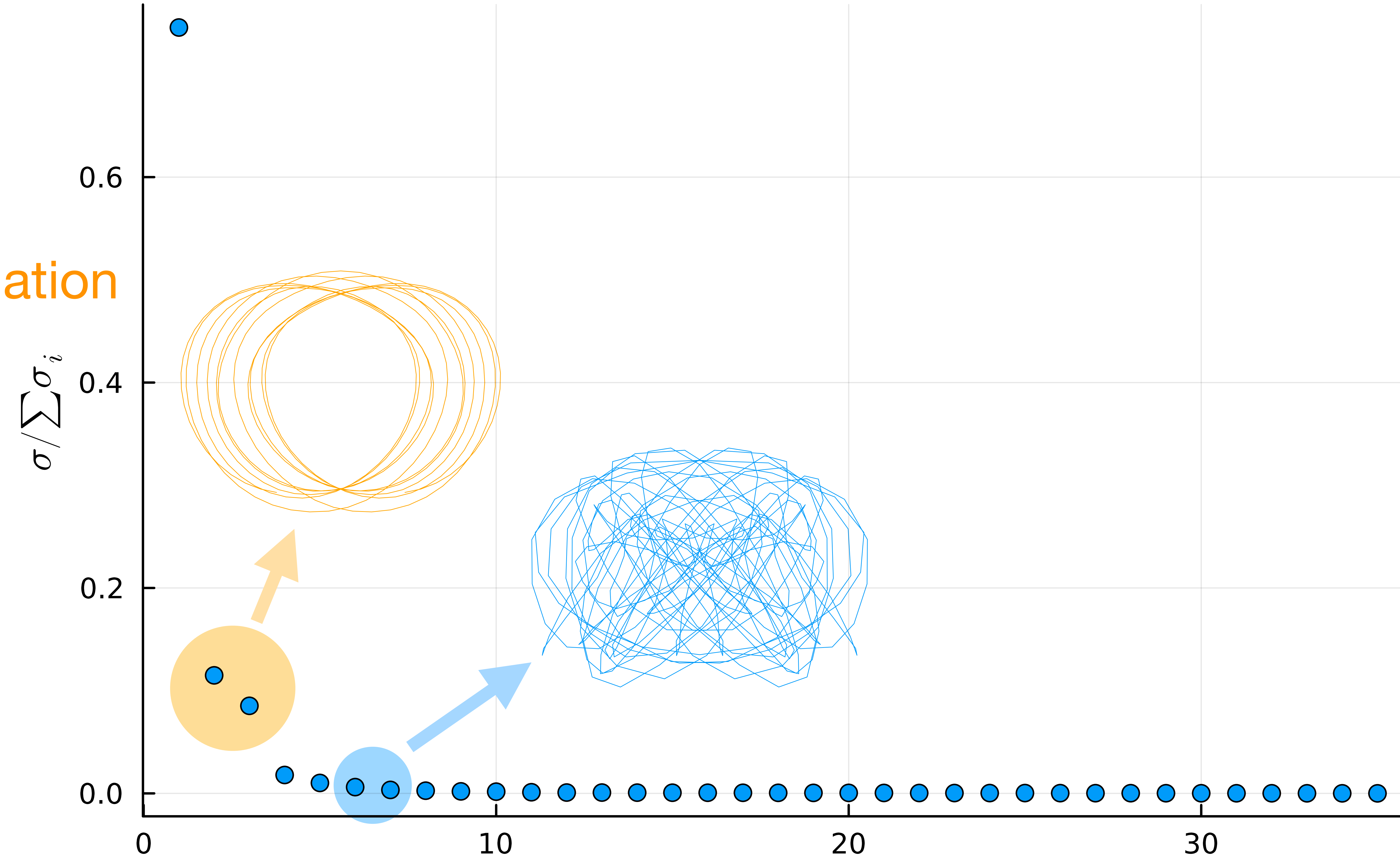
- Potential further improvements: Multiple streams, different potentials, optimize buildup (well-tempered MetaD, OPES, …), parametric potentials, flow-based swaps, …

Thank you!

Backup

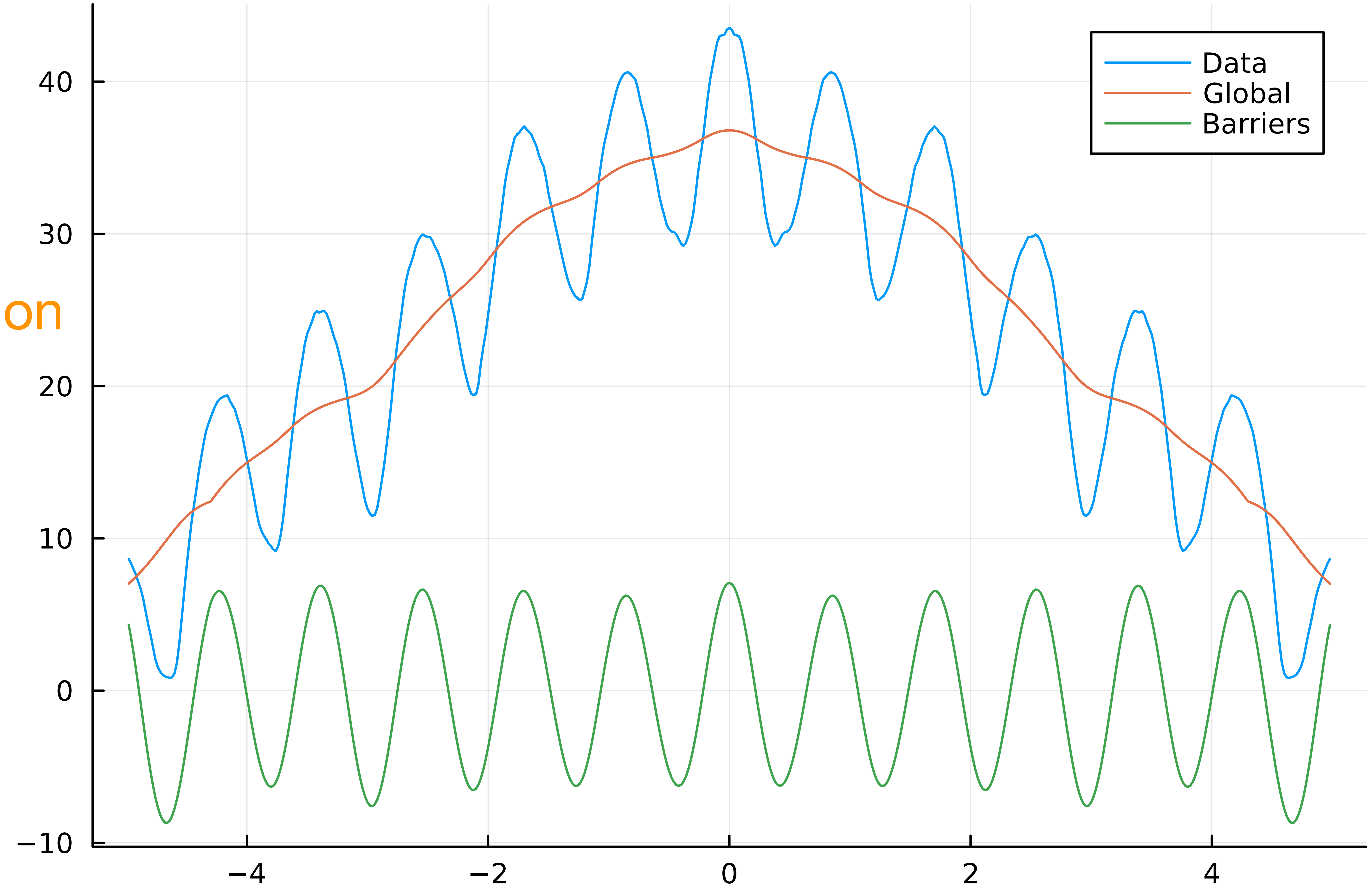
Singular spectrum analysis

- Compute lag-covariance matrix C
- Diagonalize C → orthogonal modes
- Paired modes, $\pi/2$ phase shift → oscillation
- Rest: trend, noise
- Cutoff high modes: reduce noise

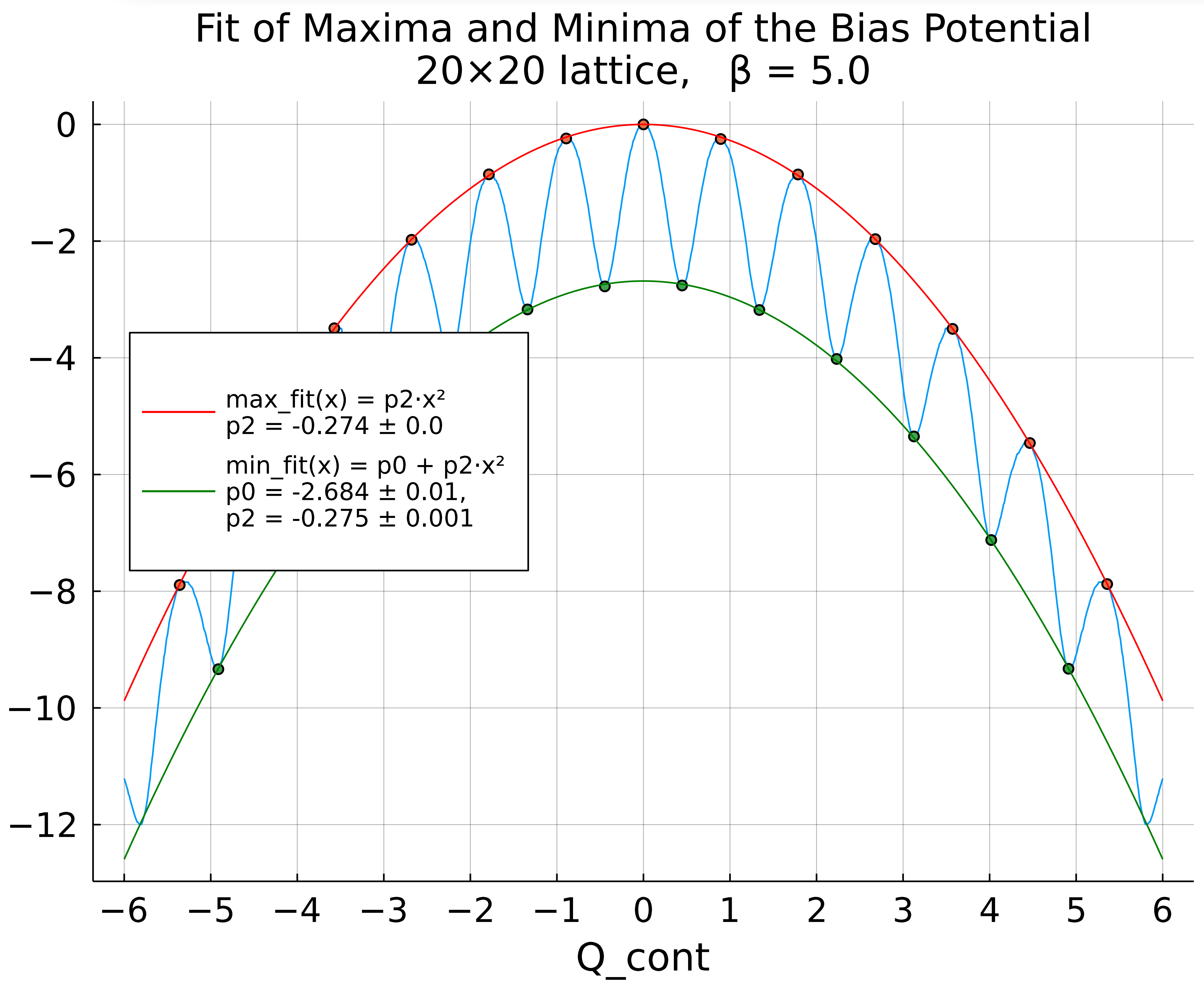


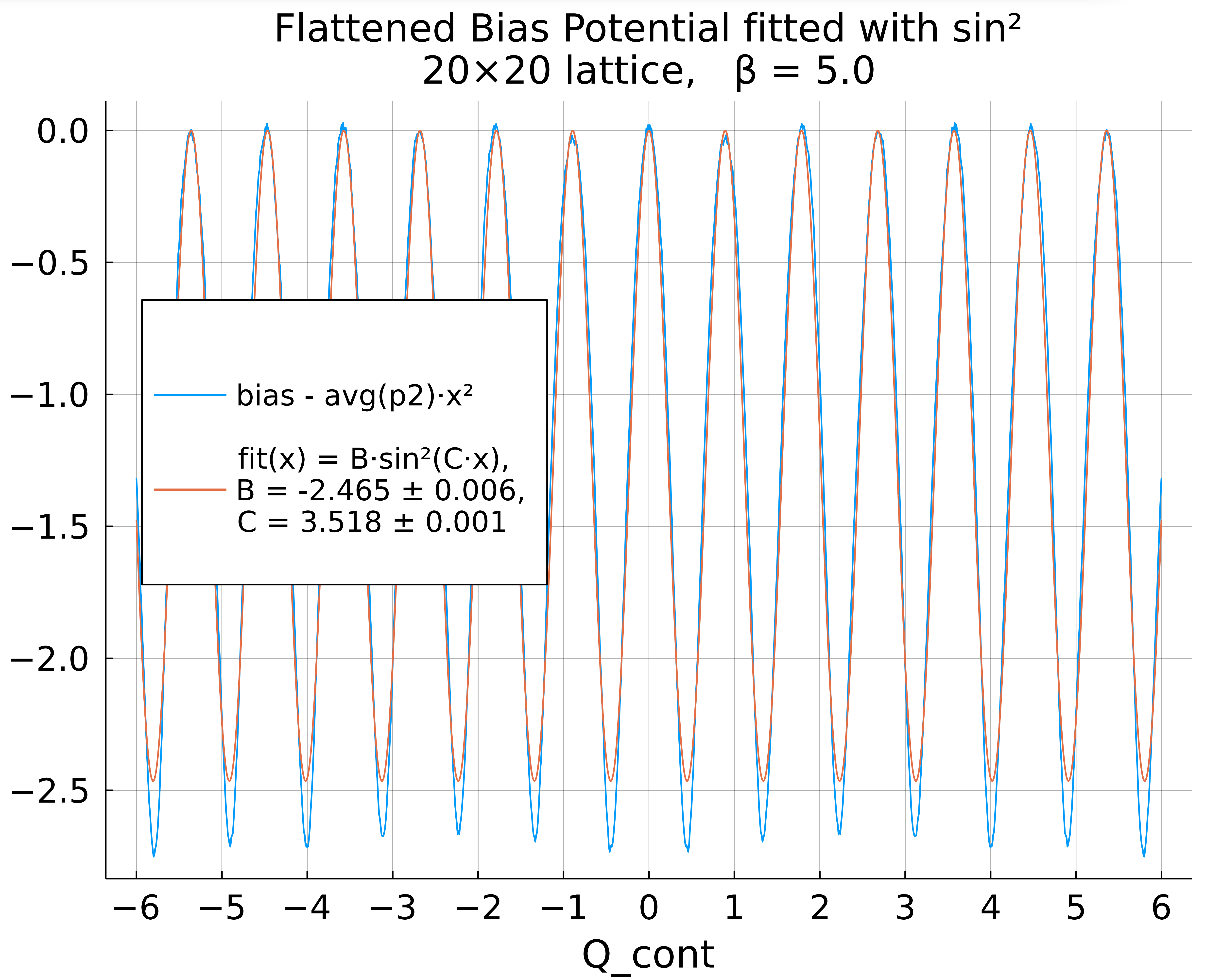
Singular spectrum analysis

- Compute lag-covariance matrix C
- Diagonalize $C \rightarrow$ orthogonal modes
- Paired modes, $\pi/2$ phase shift → oscillation
- Rest: trend, noise
- Cutoff high modes: reduce noise



Quadratic subtraction in 2D U(1)





Fourier analysis in 2D U(1)