Swiss National Science Foundation

UNIVERSITÄT BERN

Nesting sampling: SU(3) confinement transition

Gurtej Kanwar
University of Bern

In collaboration with:
Simone Romiti & Urs Wenger

Lattice 2024, Liverpool, UK | Aug 2, 2024

Motivation

Nested Sampling is a **particle Monte Carlo method**
to estimate the **action vs phase space** curve of a theory.

- Estimates of observables at arbitrary couplings
- Gives access to density-of-states / partition function

good for phase transitions

- Cheaper/easier Monte Carlo steps

constrained uniform instead of weighted sampling

- May alleviate topological freezing
- Easily parallelized

benefits of being a particle sampler

Skilling (2004) “Nested sampling”
Skilling (2006) “Nested sampling for general Bayesian computation”

Nested sampling

Skilling (2004) “Nested sampling”
Skilling (2006) “Nested sampling for general Bayesian computation”

Nested sampling

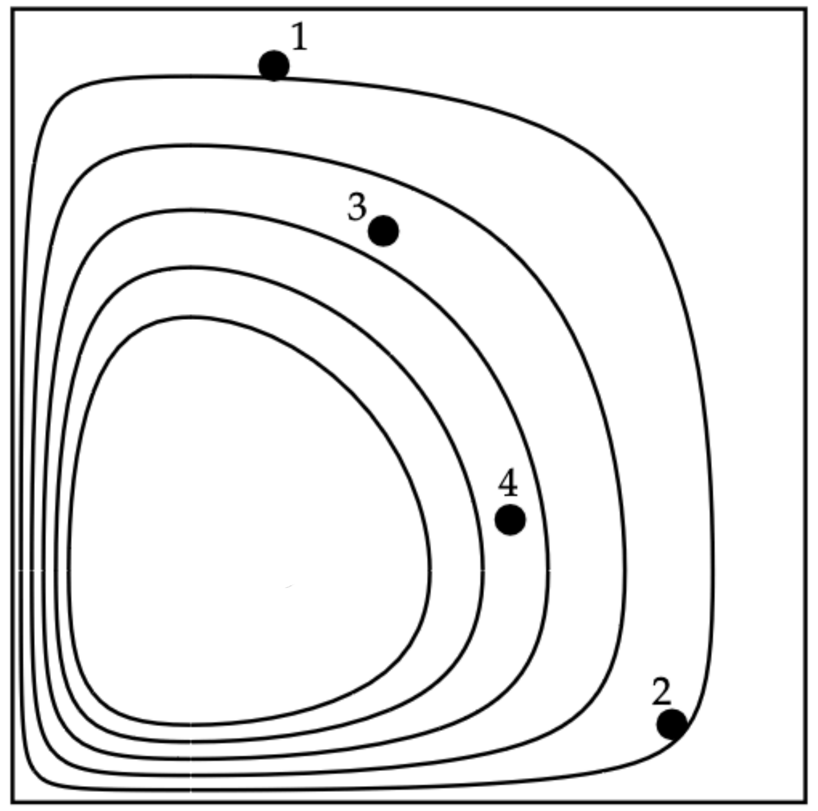
Estimate phase space X within contours of constant likelihood $L = e^{-S}$

Skilling (2004) “Nested sampling”
Skilling (2006) “Nested sampling for general Bayesian computation”

Nested sampling

Estimate phase space X within contours of constant likelihood $L = e^{-S}$

- Initialization: **sample** N_{live} uniformly random configurations (for gauge fields, total space is finite)



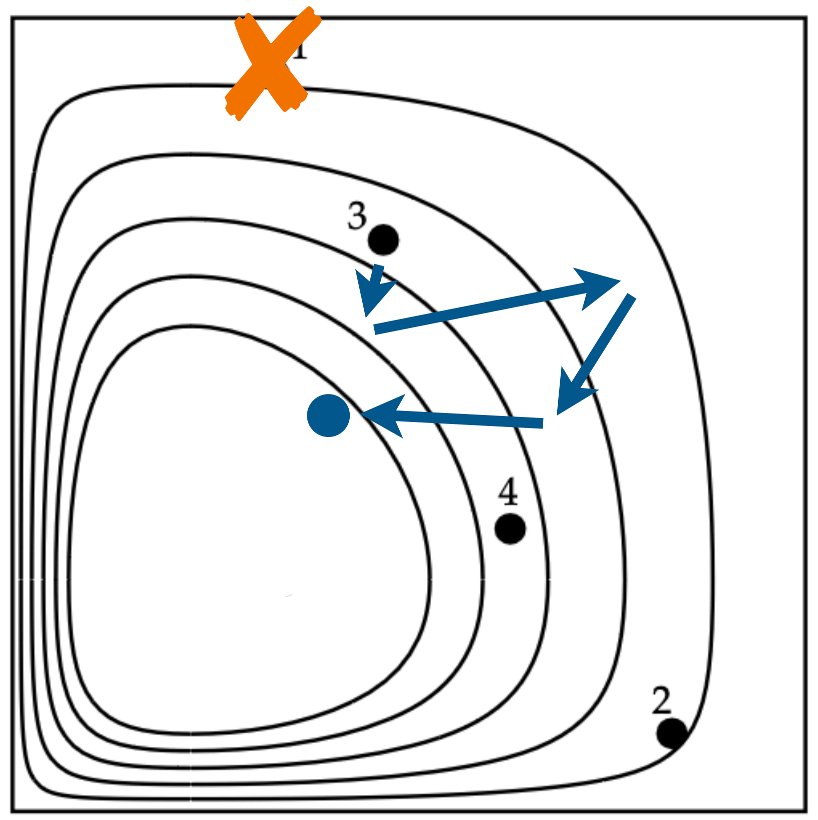
Configuration space

Skilling (2004) “Nested sampling”
Skilling (2006) “Nested sampling for general Bayesian computation”

Nested sampling

Estimate phase space X within contours of constant likelihood $L = e^{-S}$

- Initialization: **sample** N_{live} uniformly random configurations (for gauge fields, total space is finite)
- Step i: **record** largest action S_i, **drop** this sample, **resample** uniformly within $S < S_i$



Configuration space

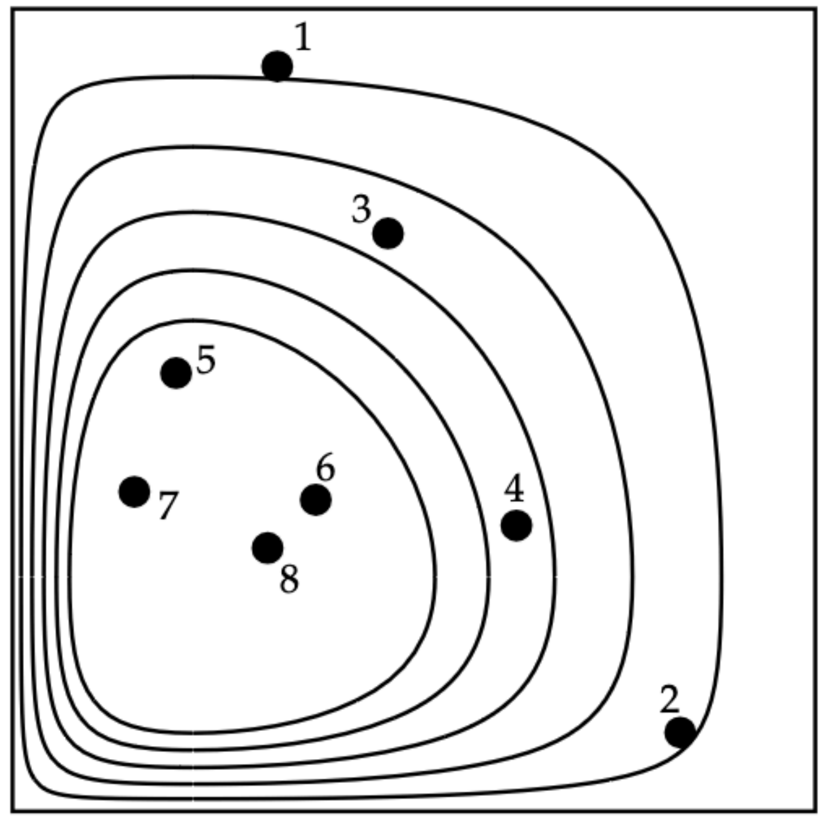
Skilling (2004) "Nested sampling"
Skilling (2006) "Nested sampling for general Bayesian computation"

Nested sampling

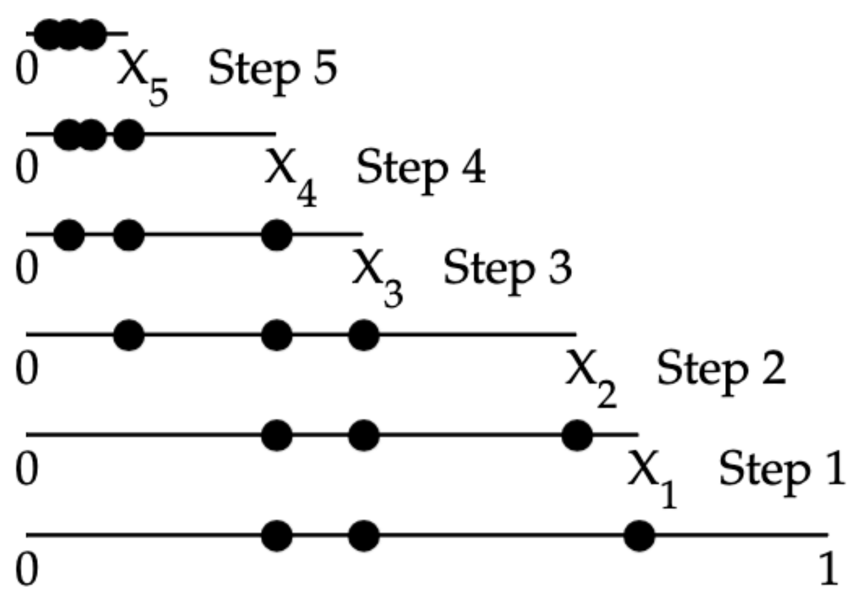
Estimate phase space X within contours of constant likelihood $L = e^{-S}$

- Initialization: **sample** N_{live} uniformly random configurations (for gauge fields, total space is finite)
- Step i: **record** largest action S_i, **drop** this sample, **resample** uniformly within $S < S_i$
- Analysis: **Compression factor** X_{i+1}/X_i follows a Beta distribution
  bootstrap or central value to build $L(X)$ curve

$$t \equiv X_{i+1}/X_i \in [0,1]$$
$$p(t) \propto t^{N_{\text{live}}-1}$$



Configuration space



Enclosed phase space X

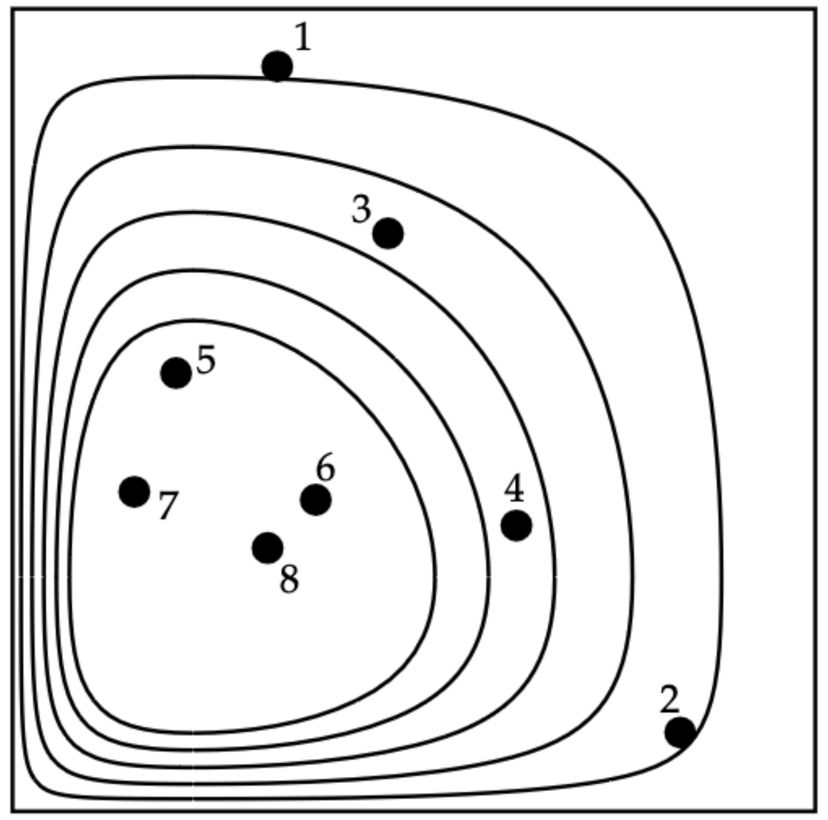
Skilling (2004) “Nested sampling”
Skilling (2006) “Nested sampling for general Bayesian computation”

Nested sampling

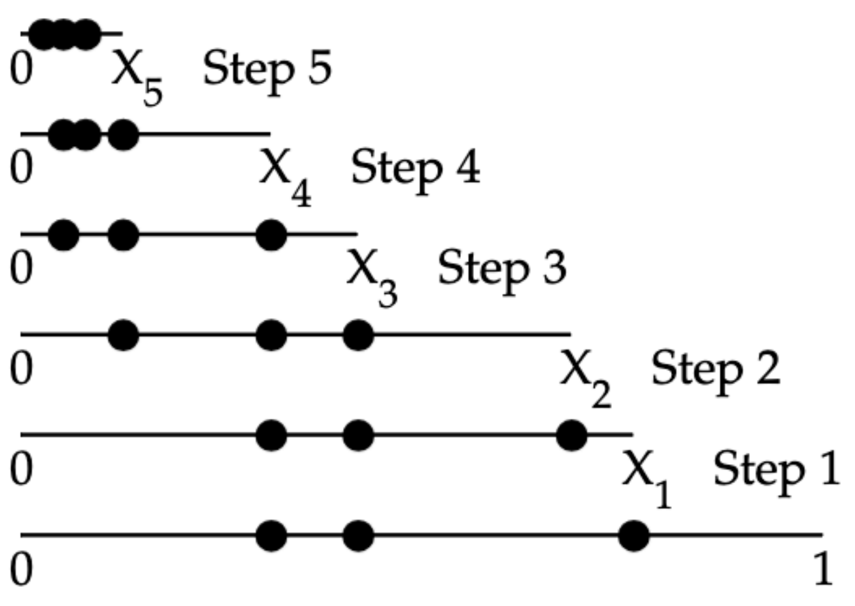
Estimate phase space X within contours of constant likelihood $L = e^{-S}$

- Initialization: **sample** N_{live} uniformly random configurations (for gauge fields, total space is finite)
- Step i: **record** largest action S_i, **drop** this sample, **resample** uniformly within $S < S_i$
- Analysis: **Compression factor** X_{i+1}/X_i follows a Beta distribution
  bootstrap or central value to build $L(X)$ curve

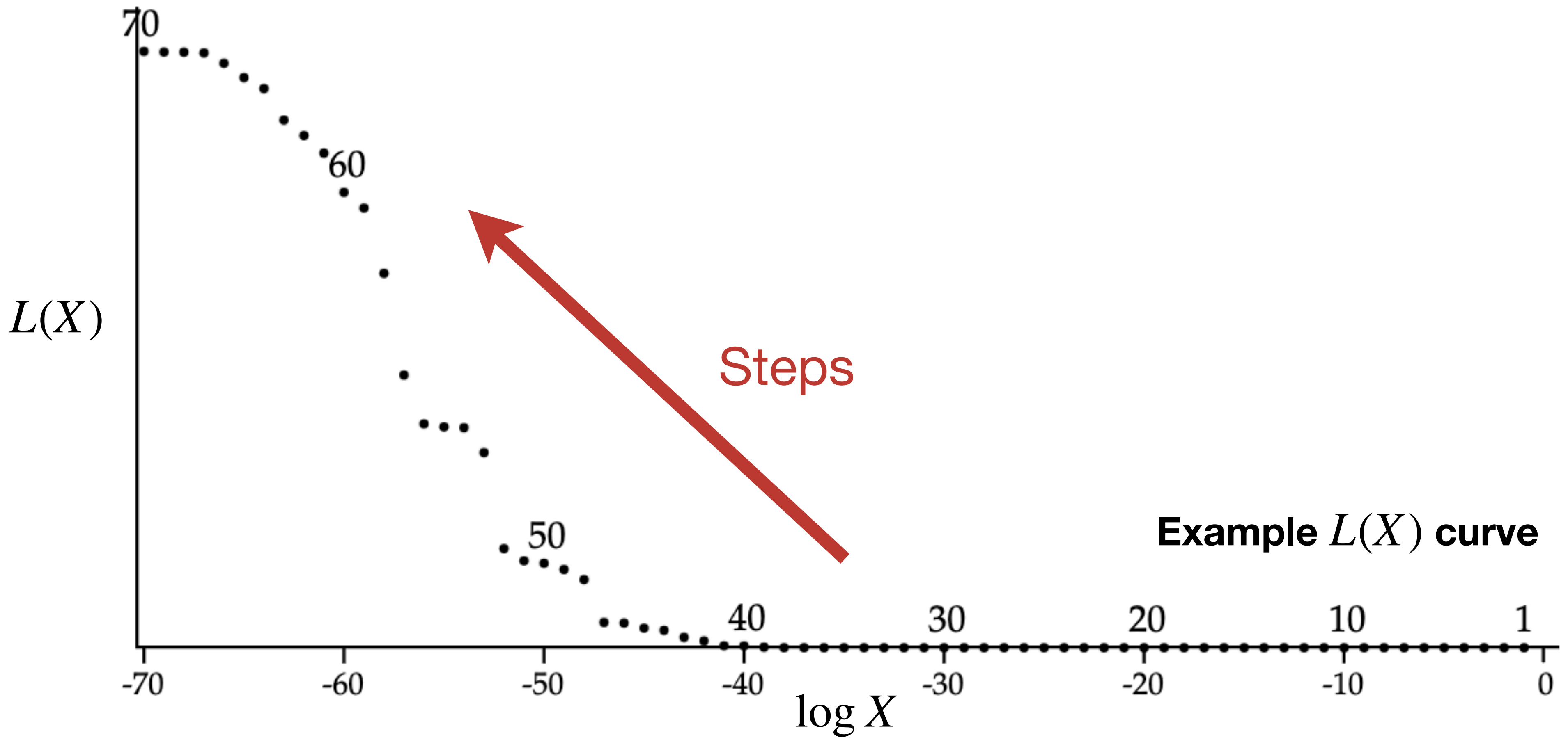
$$t \equiv X_{i+1}/X_i \in [0,1]$$
$$p(t) \propto t^{N_{\text{live}}-1}$$



Configuration space



Enclosed phase space X



Example $L(X)$ curve

Skilling (2004) “Nested sampling”
Skilling (2006) “Nested sampling for general Bayesian computation”

Nested sampling

Estimate phase space X within contours of constant likelihood $L = e^{-S}$

- Partition function and observables at multiple choices of β

$$Z(\beta) = \int_0^1 dX L(X)^\beta \qquad \langle O(\theta) \rangle_\beta = \frac{1}{Z(\beta)} \int_0^1 dX L(X)^\beta \langle O \rangle_{L(X)}$$

  - Relevant region of L (and X) depends on β See previous talks by S. Romiti, U. Wenger

- Density of states $\rho(S) = \dfrac{dX}{dS} = -\dfrac{dX}{d\log L}$

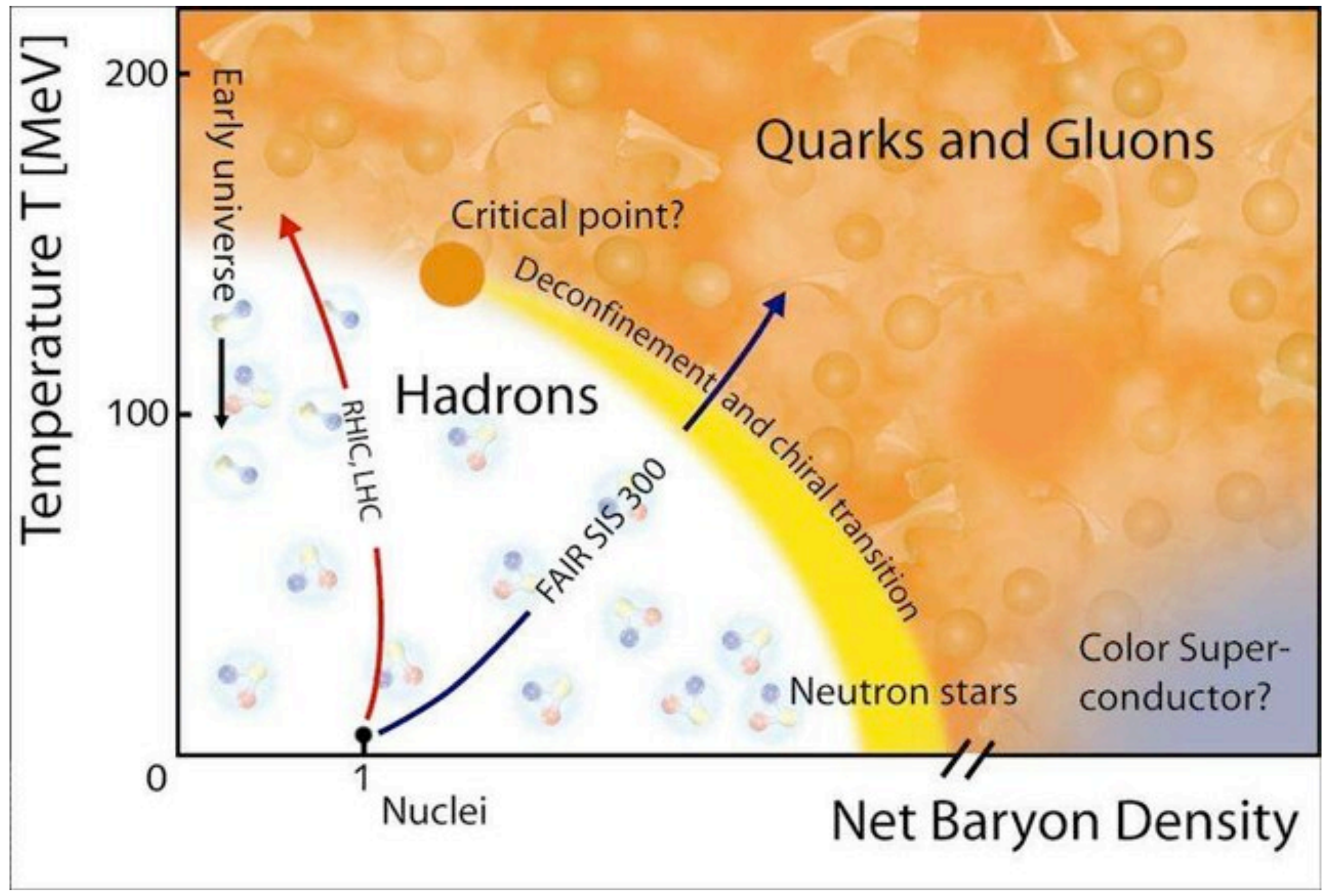
  - Universal function independent of β

- Useful to restrict sampling to important regions to improve statistics

SU(3) confinement transition

Aarts, et al. 1412.0847

First-order transition in full QCD



Center symmetry

Rotate temporal links on one timeslice by $z = e^{2\pi i k/3} \in \mathbb{Z}_3$

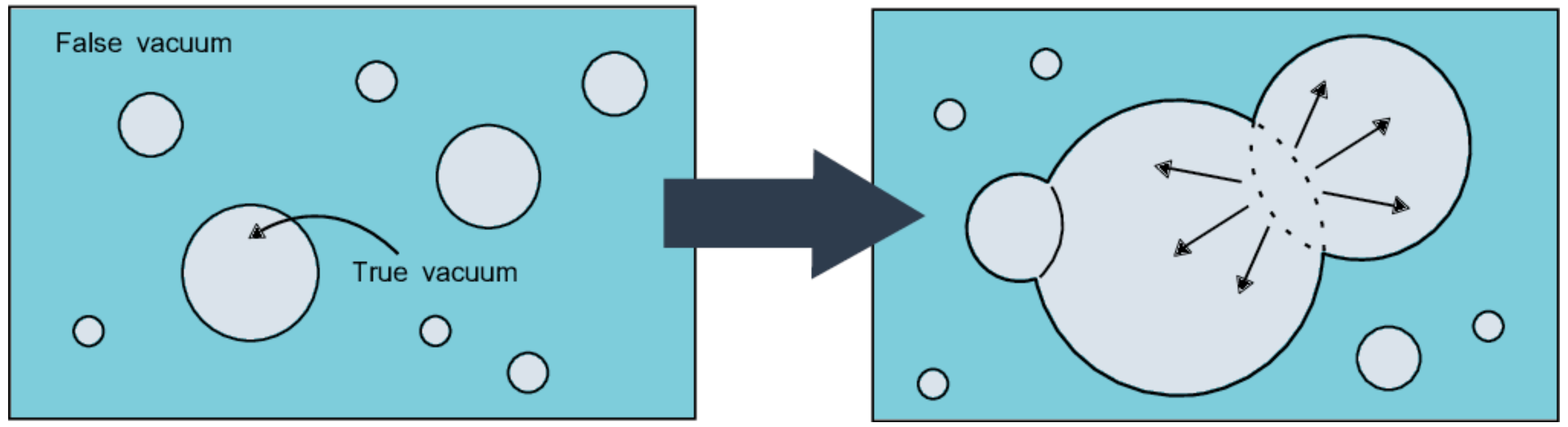
Polyakov loop is a good order param

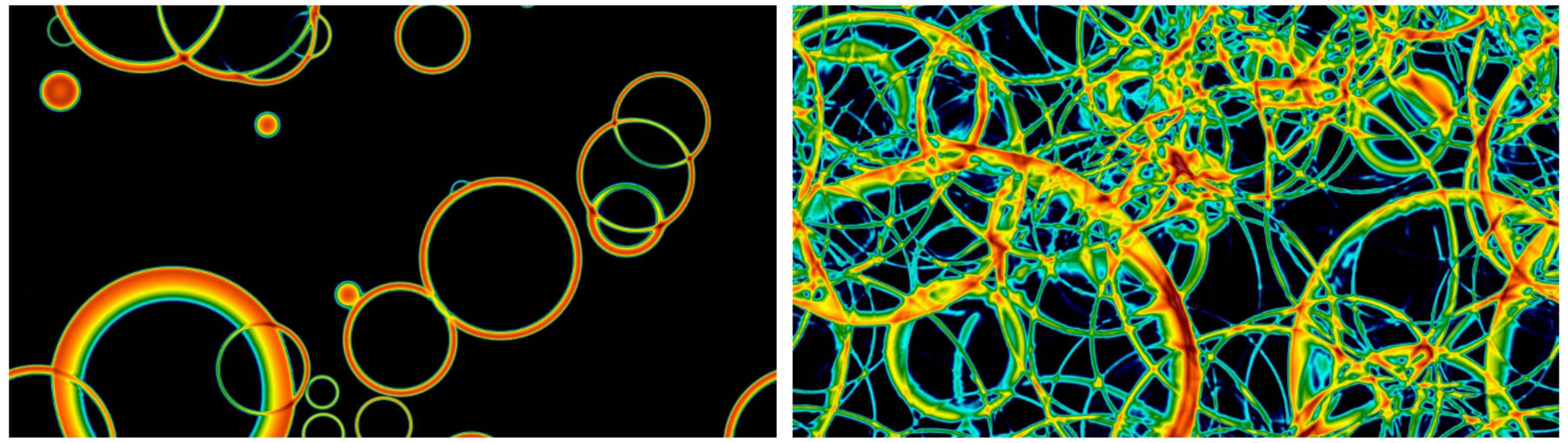
$$\ell(\vec{x}) \equiv \frac{1}{N_c}\mathrm{Tr}[\prod_t U_0(\vec{x},t)] \longrightarrow P \equiv \frac{1}{N_s^3}\sum_{\vec{x}} \ell(\vec{x})$$

Confined phase:
$|P| = 0$

1st order transition

Deconfined phase:
$|P| \neq 0, P \propto \{1, z, z^2\}$





First-order transition generating bubble dynamics in the early universe

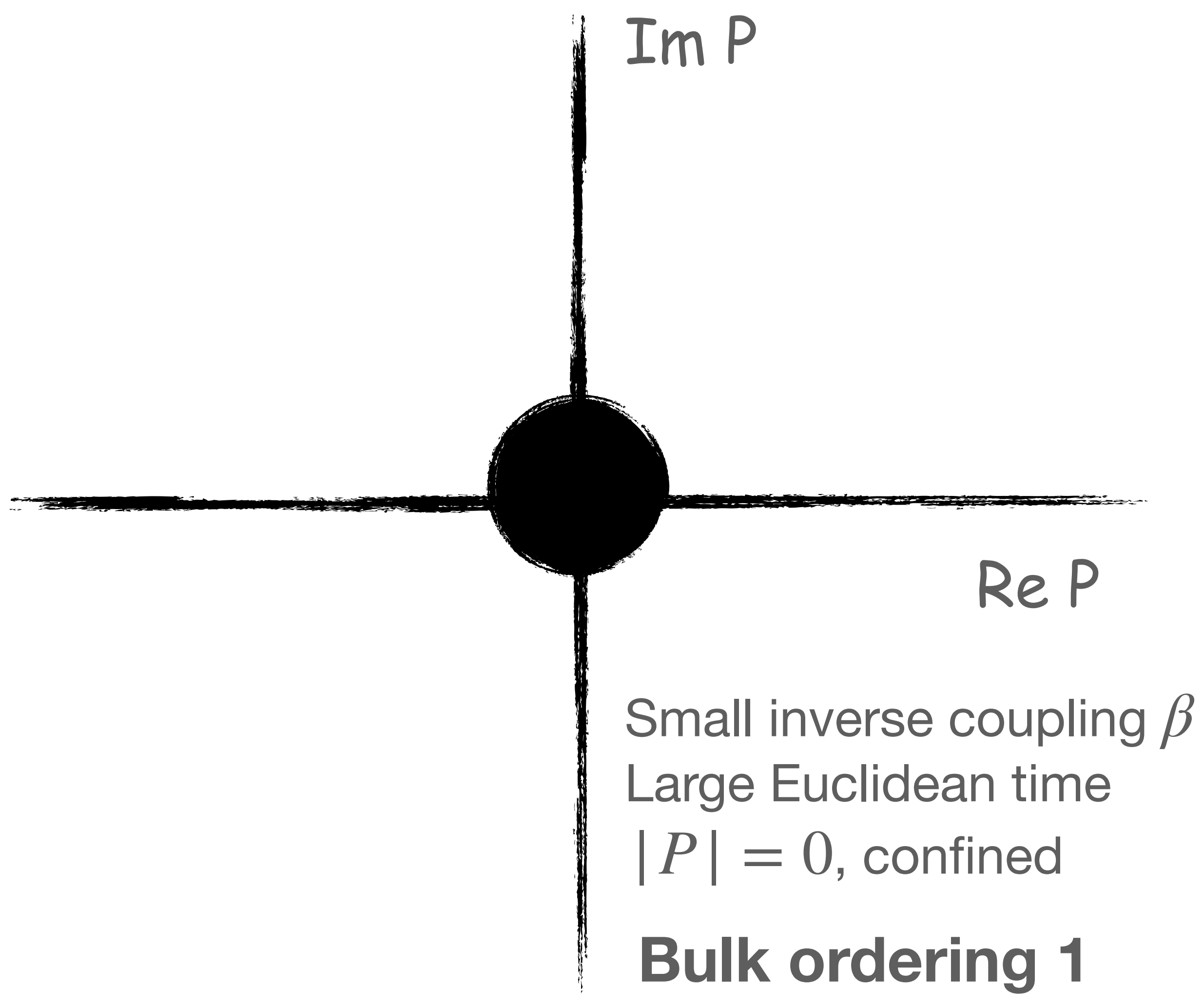
D. Weir (2023) 1705.01783

SU(3) confinement transition

First order phase transition ($N_c > 2$)

- **Bulk ordering 1:** Polyakov loops disordered
  Entropically favored, **energetically** disfavored

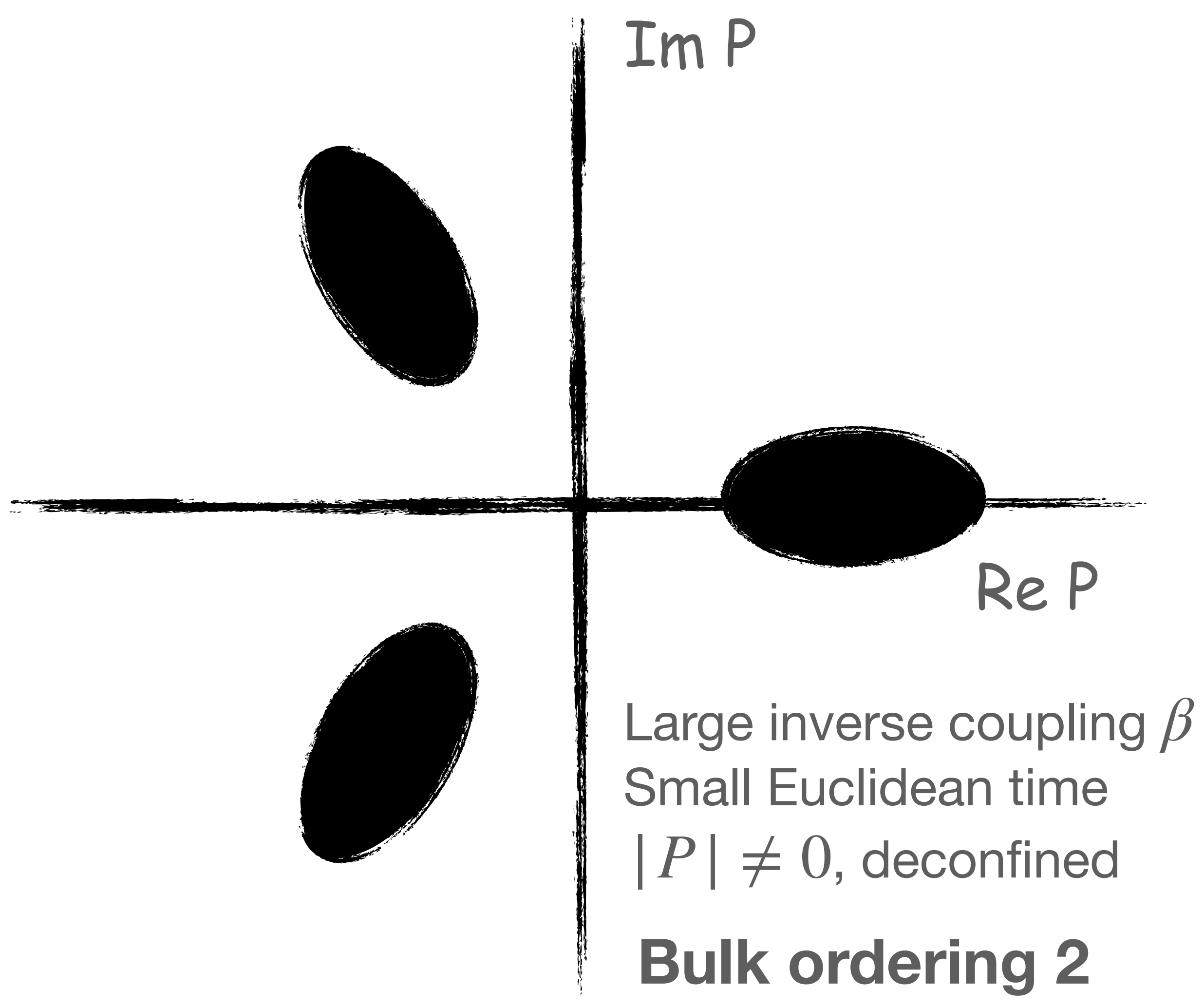
- **Bulk ordering 2:** Polyakov loops ordered
  Energetically favored, **entropically** disfavored



Study using thermodynamic lattices ($N_s^3 \times N_t$, $N_t < N_s$) with varying inverse coupling β

Various existing lattice results

- Standard MC
  Kajantie, et al. (1981)
  Çelik, et al. (1983)
  Gottlieb, et al. (1985)
  …
- Parallel tempering
  Borsanyi, et al. 2202.05234
- LLR method
  Lucini, et al. 2305.07463
  Wed 12:15 D. Mason
- Shifted BCs
  Fri 14:15 L. Virzi



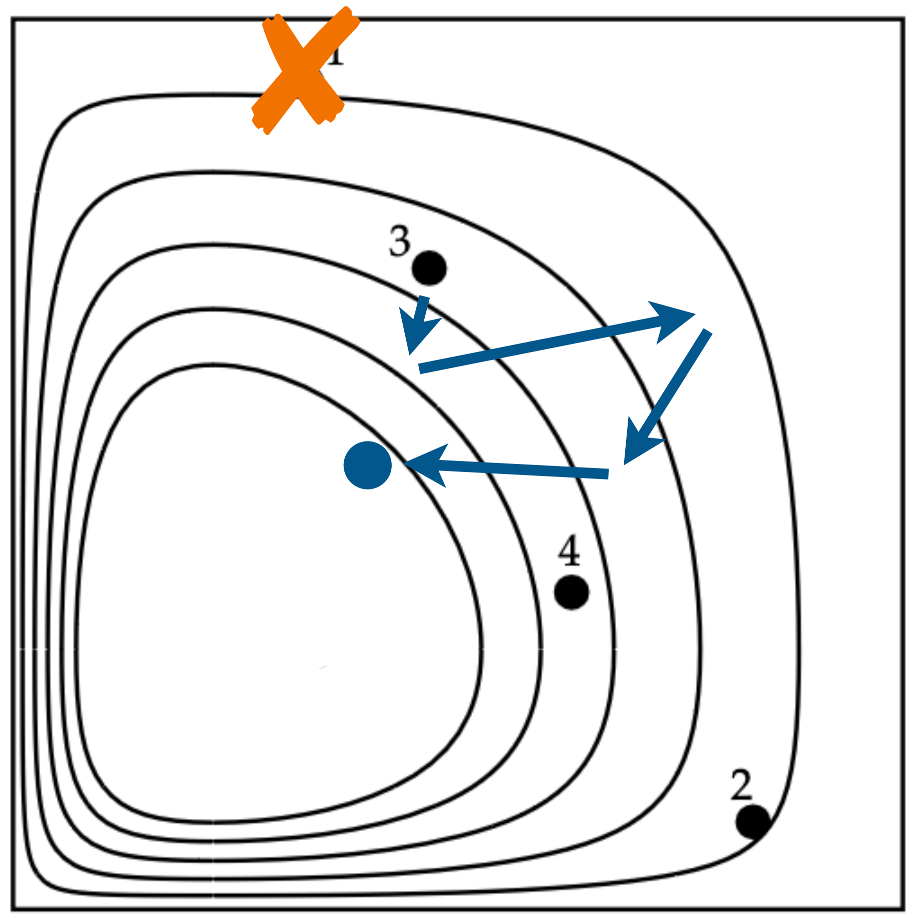
Nested sampling for SU(3)

Executed 16 fully independent “streams”

- $N_{\text{live}} = 8192$ walkers for high-stats $L = 8$, $L_t = 4$ run
- $N_{\text{live}} = 256$ walkers for exploratory $L = 12$, $L_t = 4$ run
- Bootstrap over streams for Monte Carlo errors, compression errors still required

Constrained Monte Carlo $p(U) \propto \Theta(S^* - S(U))$

- Initialize each resampling step copying another walker in stream
- Local constrained Metropolis updates mix sufficiently well
- Constrained HMC also possible Betancourt (2010) 1005.0157 Skilling (2012)



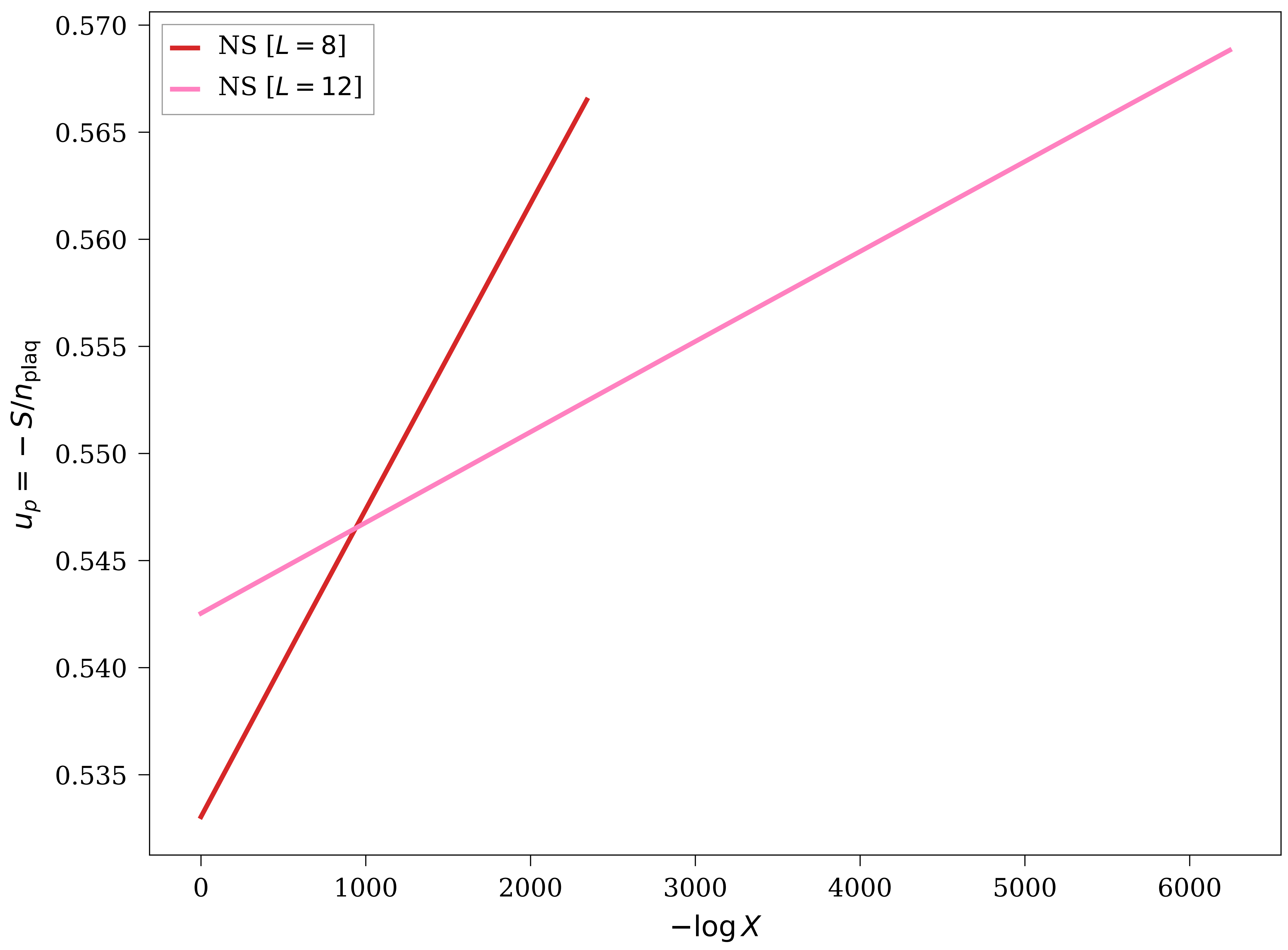
Configuration space

Results: Action vs phase space

- Smooth movement through action values

- Nearly linear vs $-\log X$

- MCMC appears to be performing well



Results: Polyakov evolution vs beta (L=8)

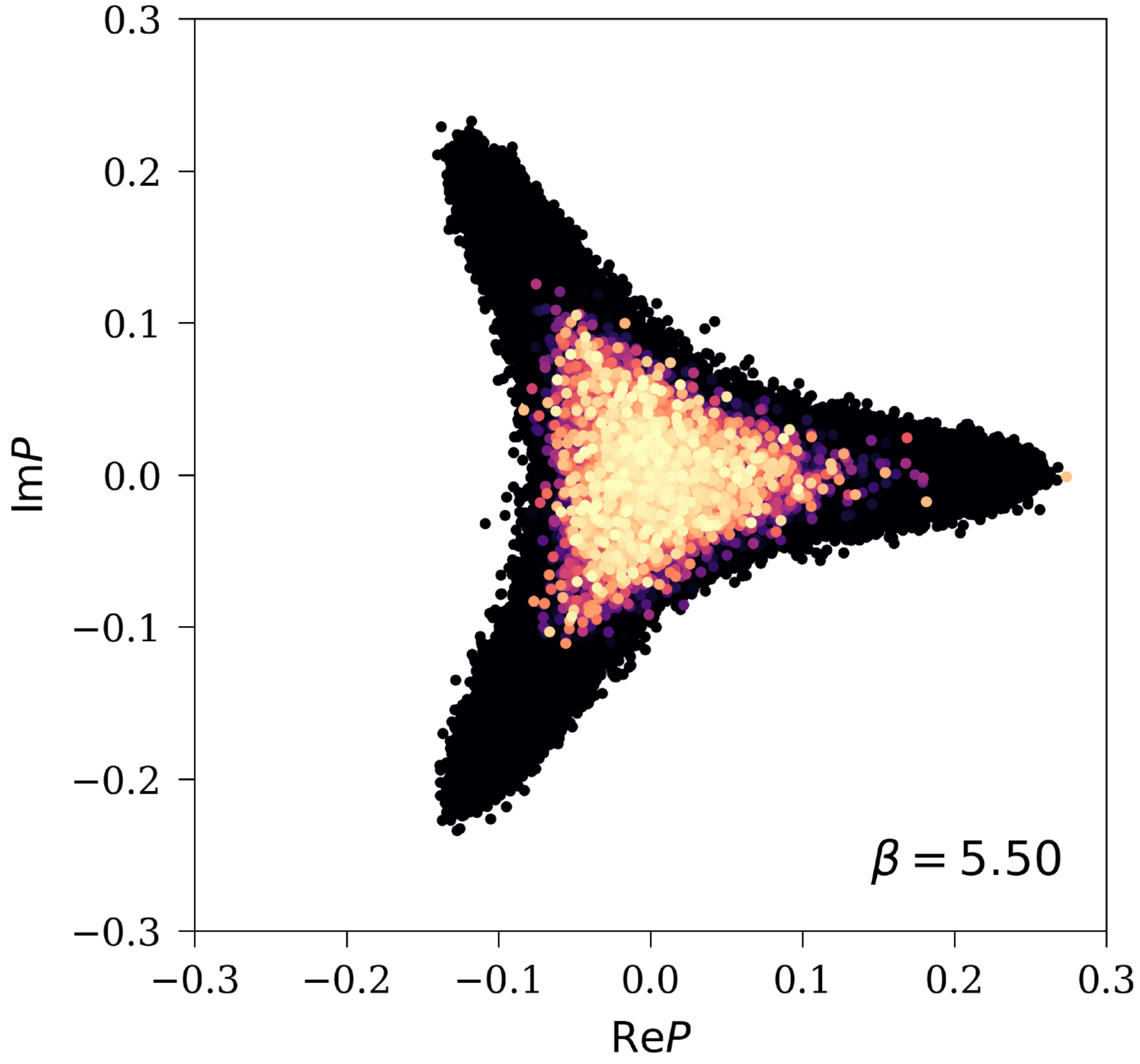
All measurements of P during NS run

Color based on weight in NS integral:

$$\langle O(\theta) \rangle_\beta = \frac{1}{Z(\beta)} \int_0^1 dX L(X)^\beta \langle O \rangle_{L(X)}$$

↓

$$w_i \propto (X_{i+1} - X_i) e^{-\beta S_i}$$



Results: Polyakov evolution vs beta (L=8)

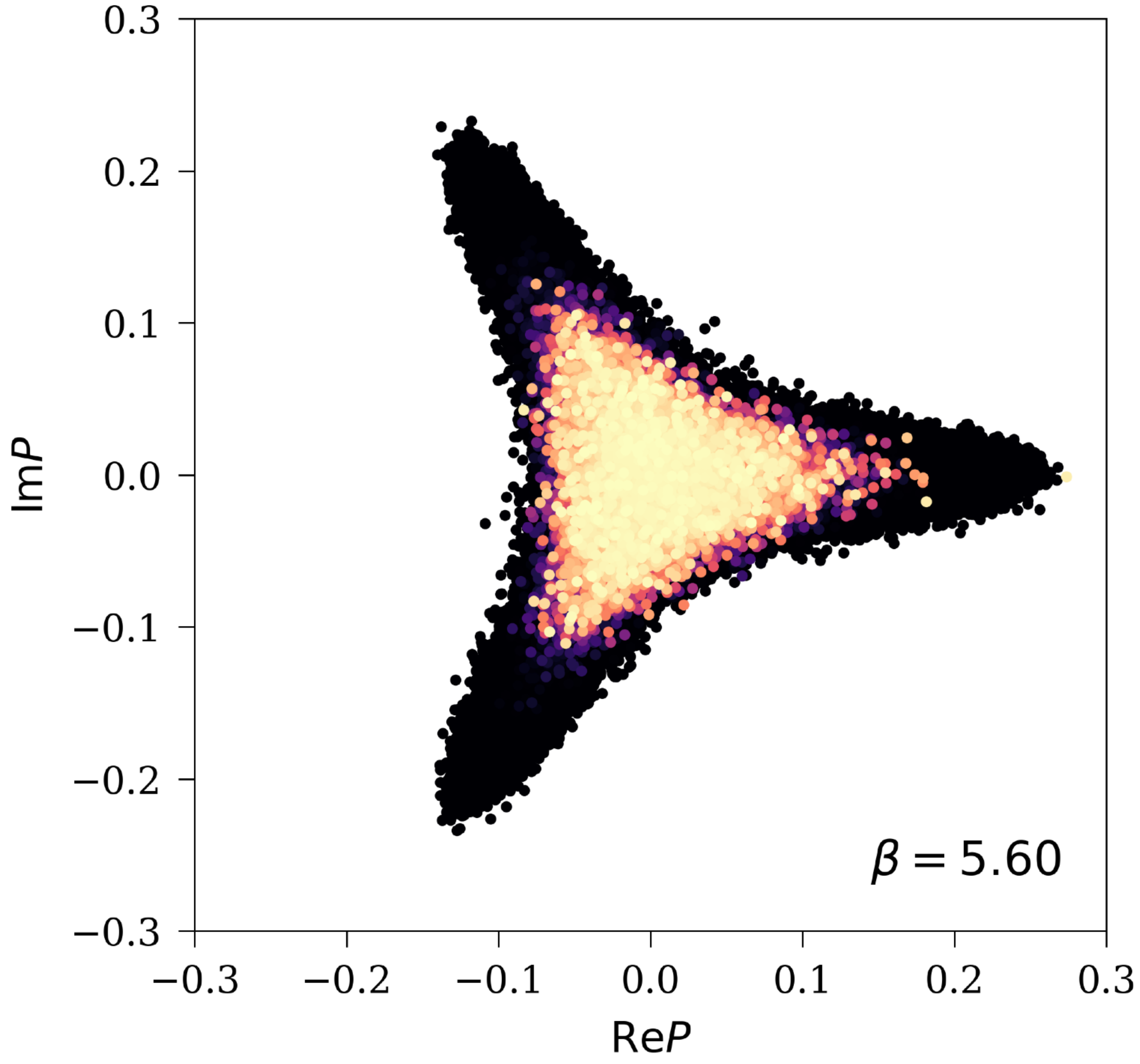
All measurements of P during NS run

Color based on weight in NS integral:

$$\langle O(\theta)\rangle_\beta = \frac{1}{Z(\beta)}\int_0^1 dX L(X)^\beta \langle O\rangle_{L(X)}$$

↓

$$w_i \propto (X_{i+1} - X_i)e^{-\beta S_i}$$



Results: Polyakov evolution vs beta (L=8)

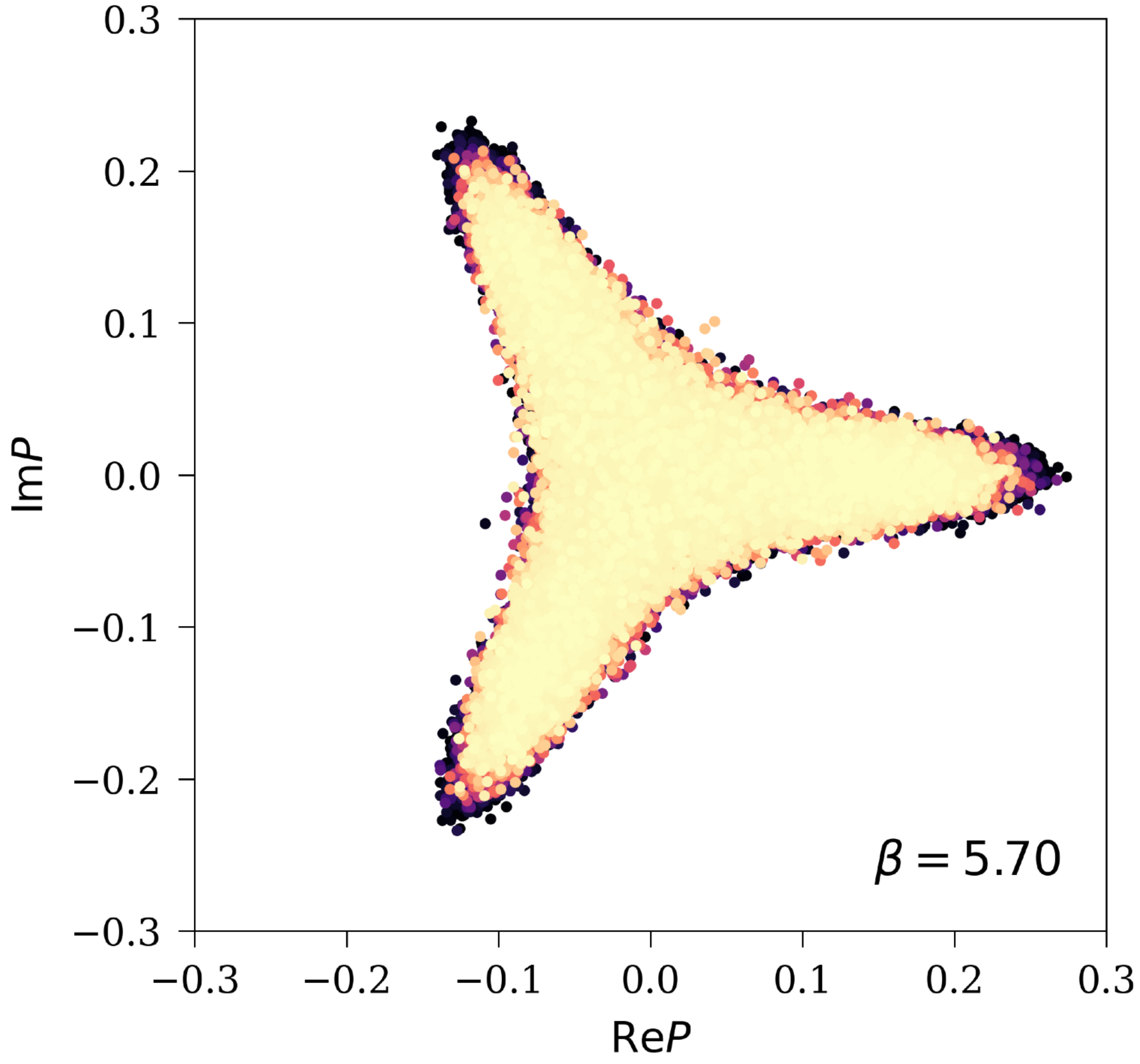
All measurements of P during NS run

Color based on weight in NS integral:

$$\langle O(\theta)\rangle_\beta = \frac{1}{Z(\beta)}\int_0^1 dX L(X)^\beta \langle O\rangle_{L(X)}$$

↓

$$w_i \propto (X_{i+1} - X_i)e^{-\beta S_i}$$



Results: Polyakov evolution vs beta (L=8)

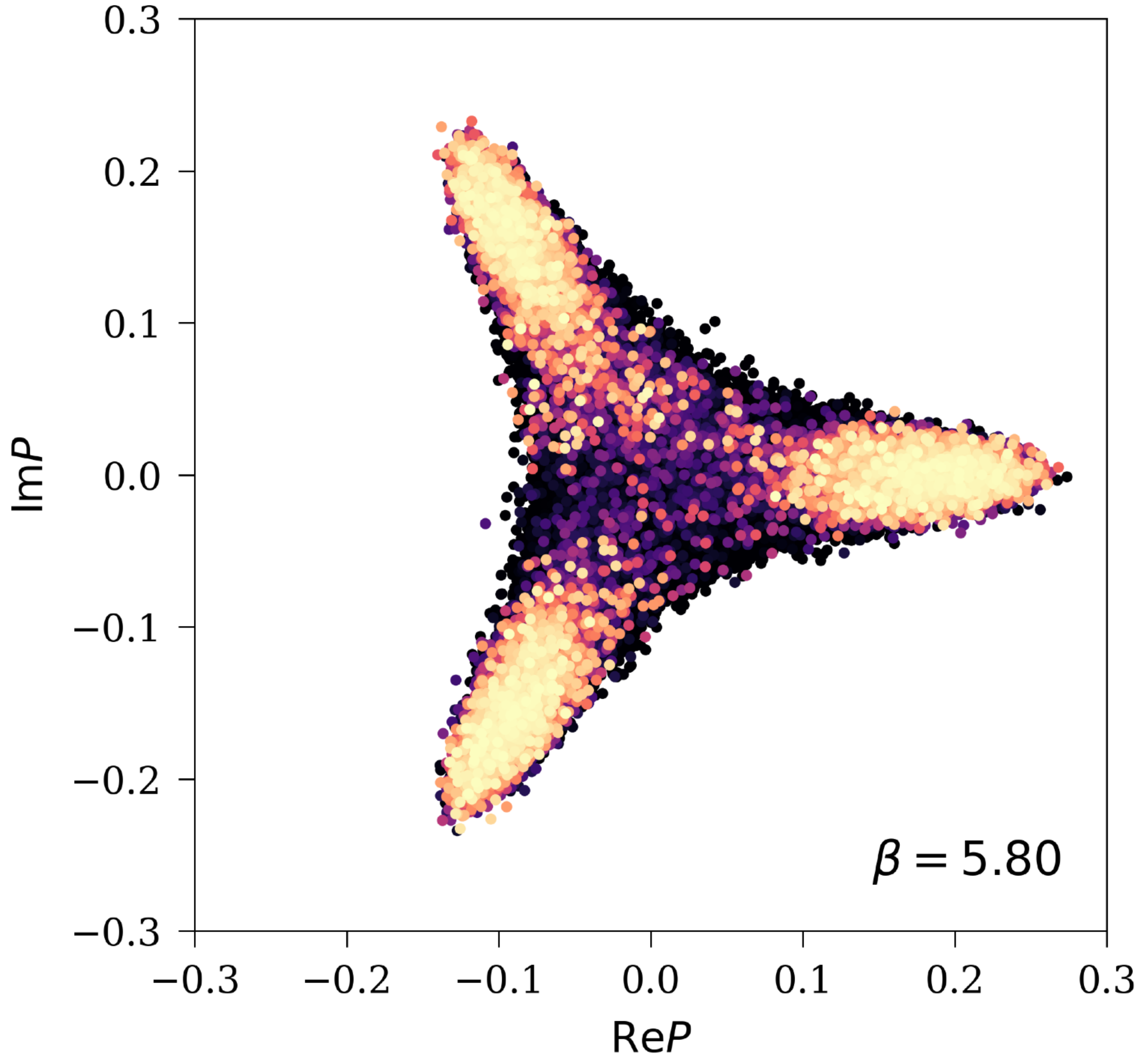
All measurements of P during NS run

Color based on weight in NS integral:

$$\langle O(\theta)\rangle_\beta = \frac{1}{Z(\beta)}\int_0^1 dX L(X)^\beta \langle O\rangle_{L(X)}$$

↓

$$w_i \propto (X_{i+1} - X_i)e^{-\beta S_i}$$



Results: Polyakov evolution vs beta (L=8)

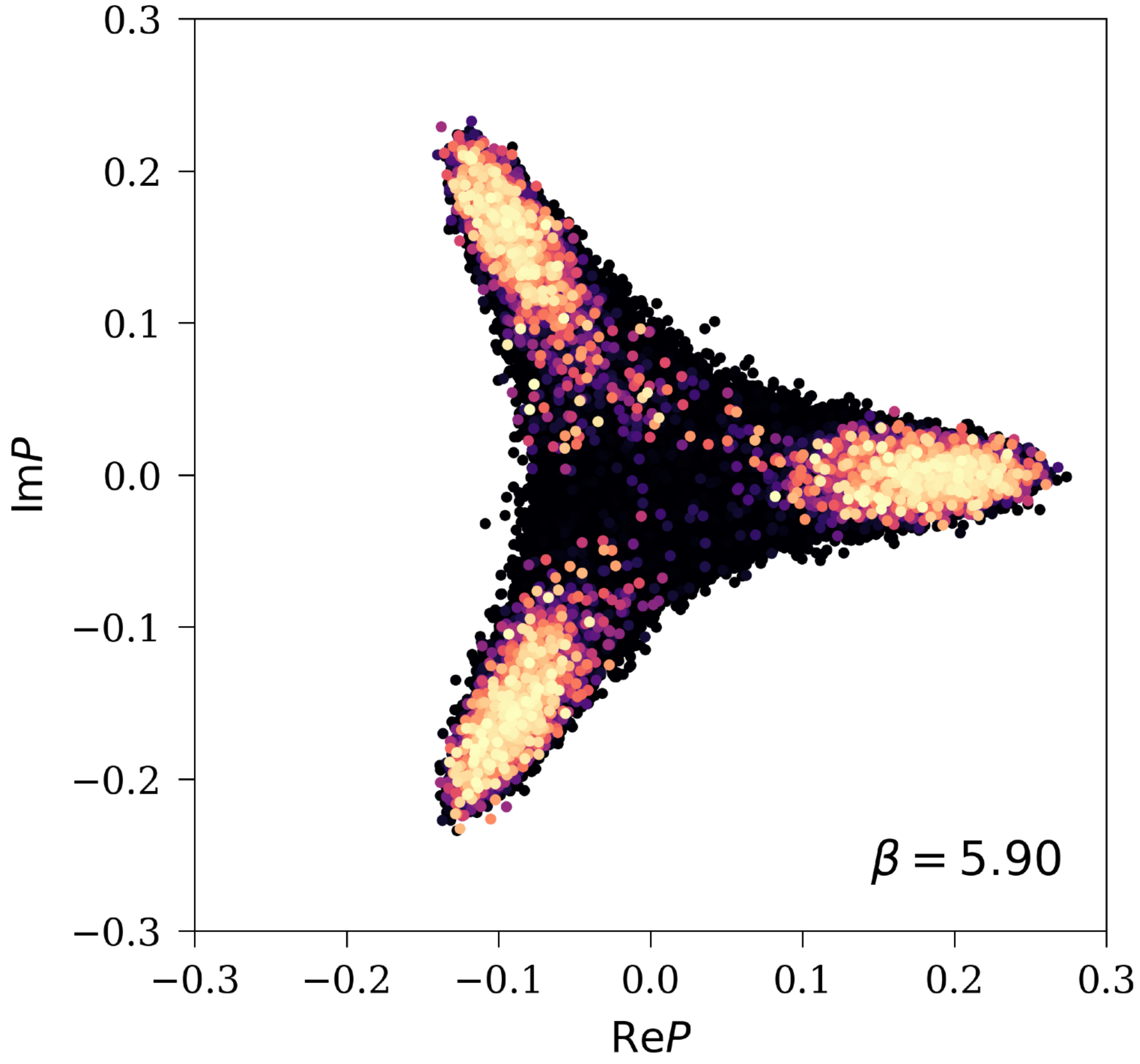
All measurements of P during NS run

Color based on weight in NS integral:

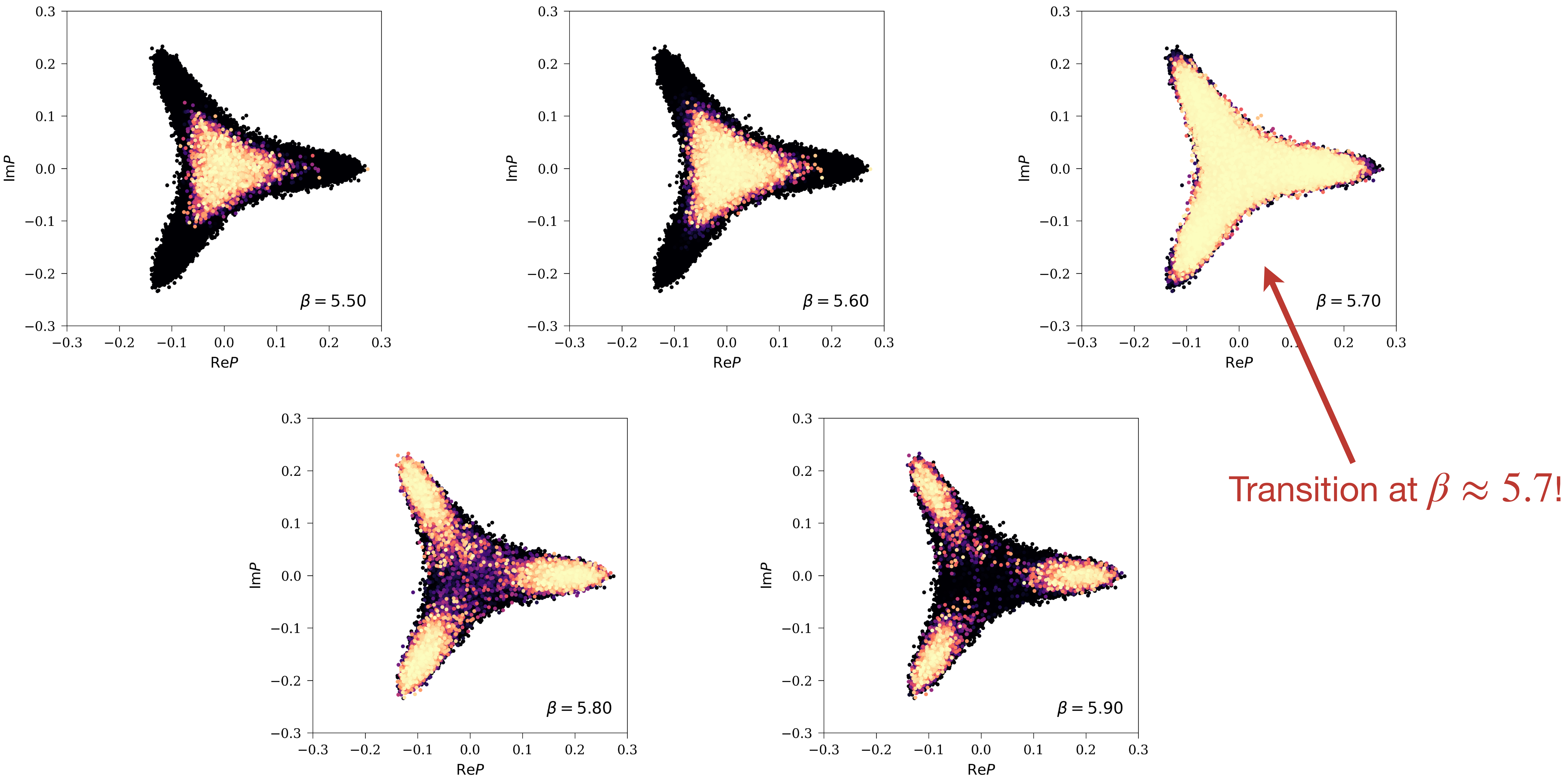
$$\langle O(\theta)\rangle_\beta = \frac{1}{Z(\beta)}\int_0^1 dX L(X)^\beta \langle O\rangle_{L(X)}$$

↓

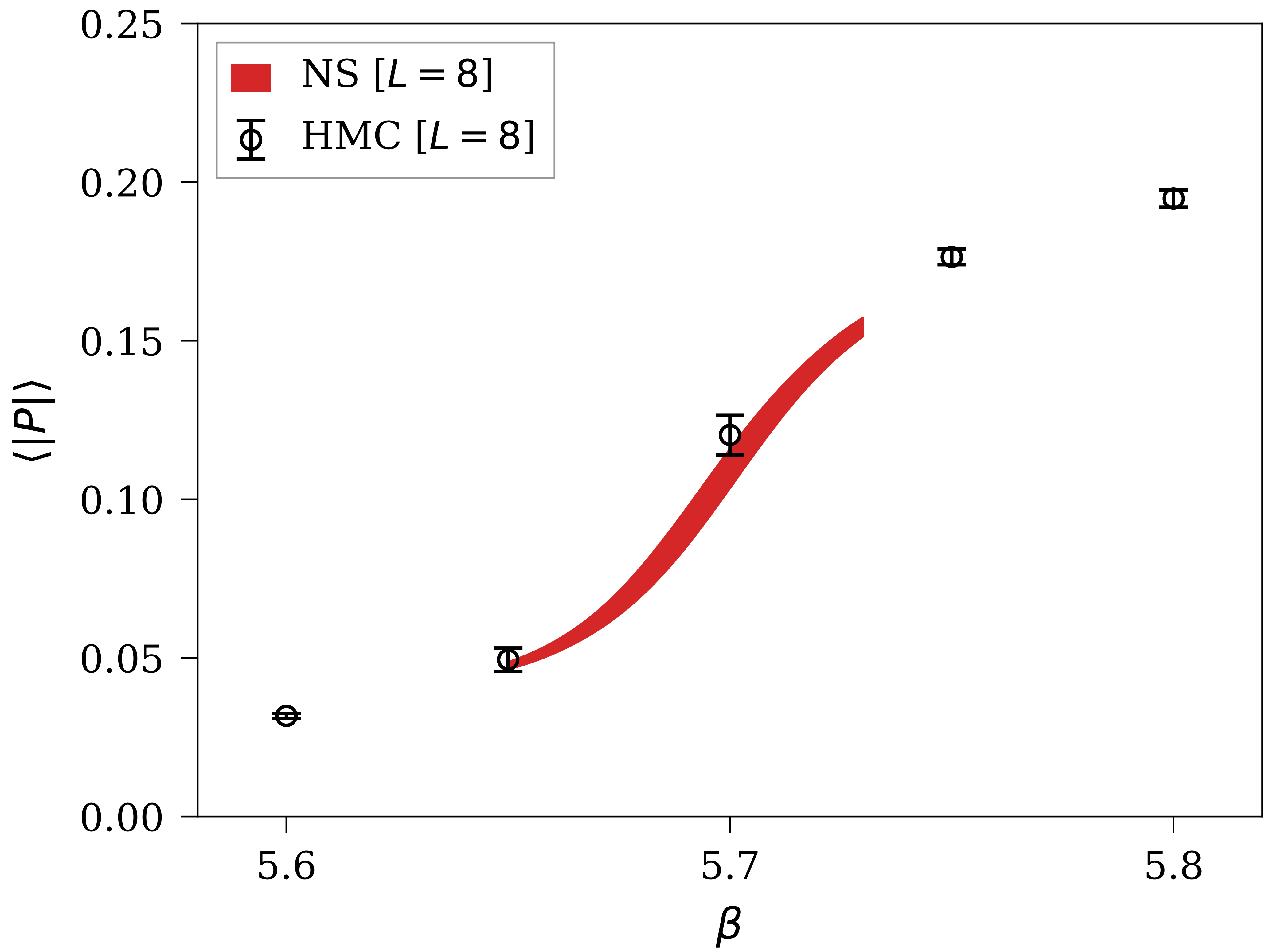
$$w_i \propto (X_{i+1} - X_i)e^{-\beta S_i}$$



Results: Polyakov evolution vs beta (L=8)

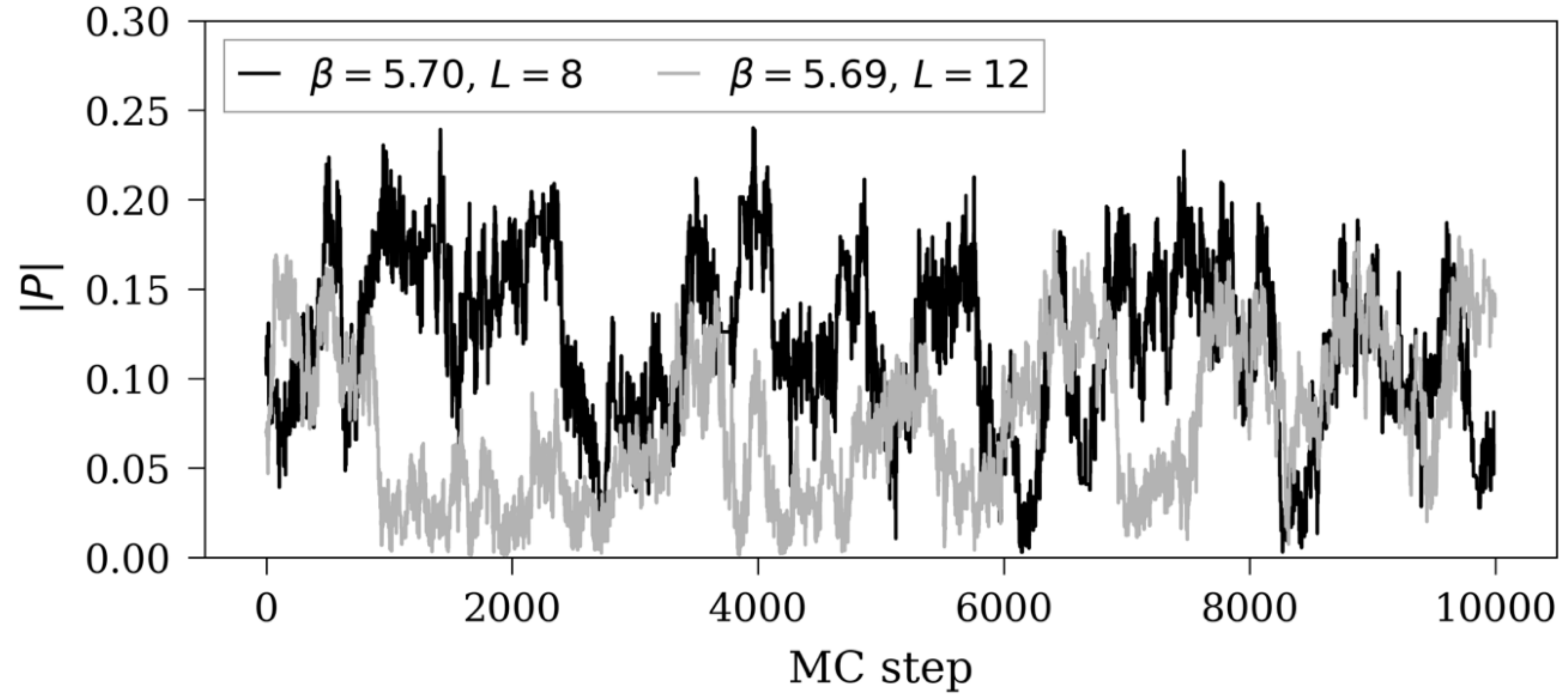


Results: Polyakov order param

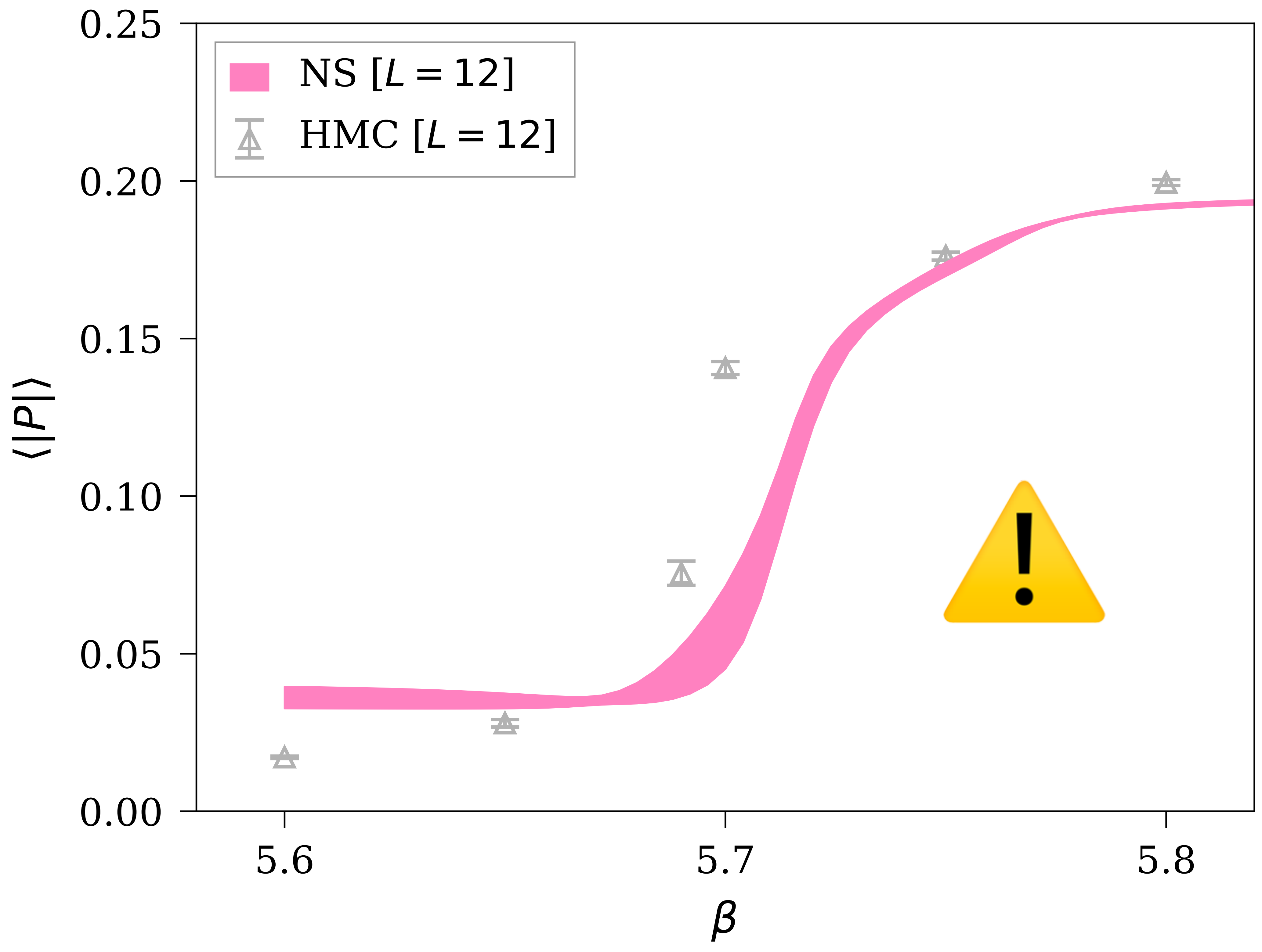


Challenges for HMC:

- Long autocorrelation times
- Hysteresis near transition

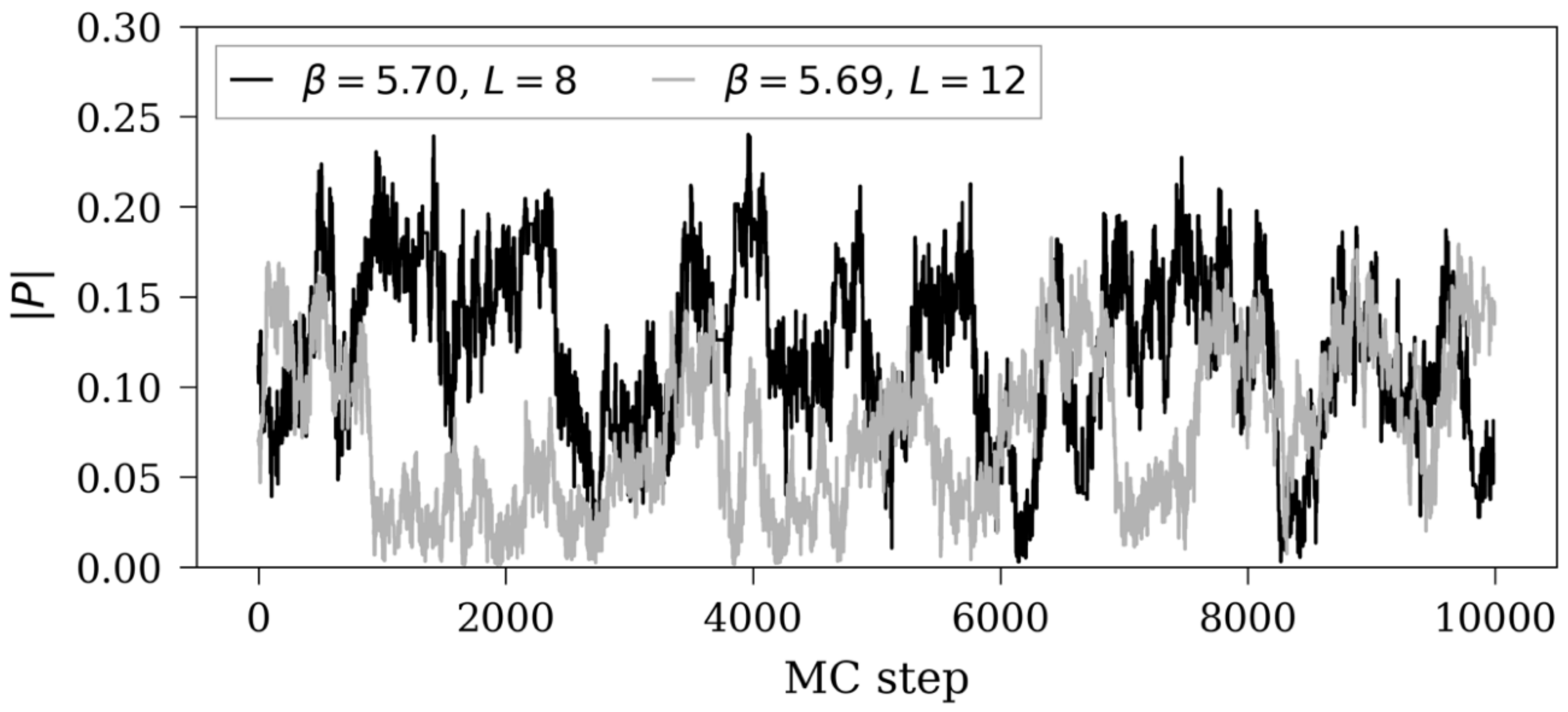


Results: Polyakov order param

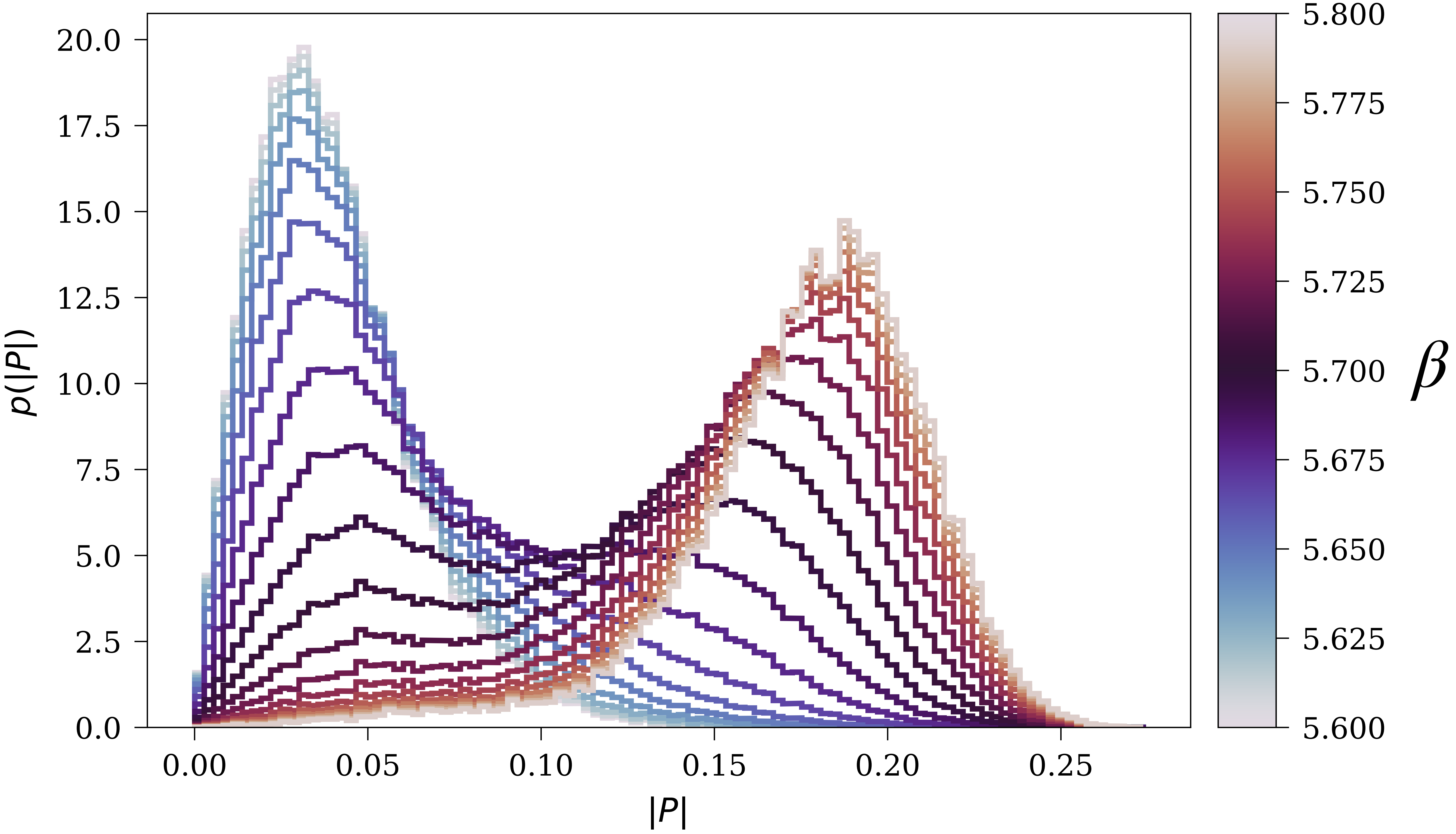


Challenges for HMC:

- Long autocorrelation times
- Hysteresis near transition



Results: Polyakov histograms (L=8)



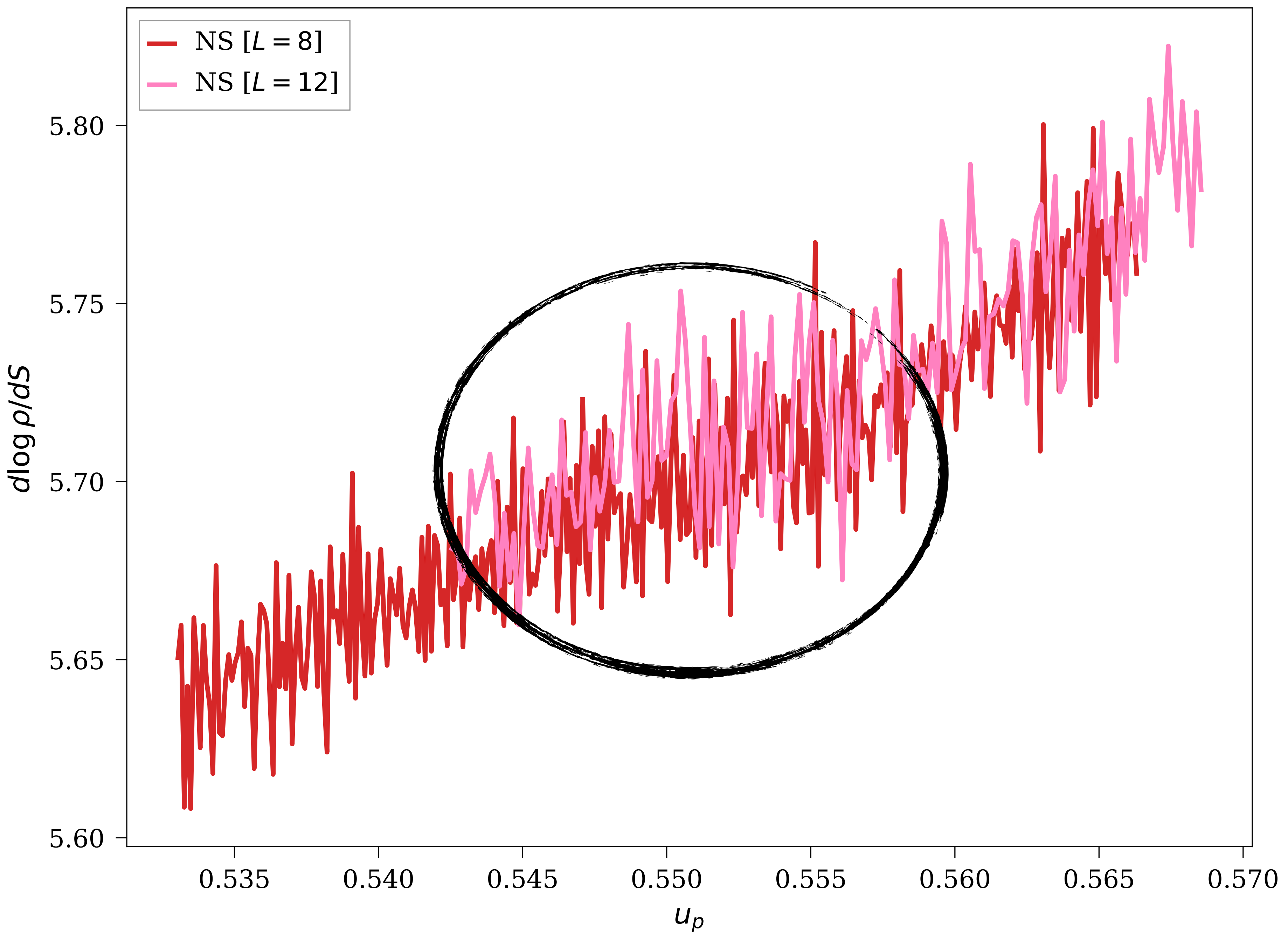
Results: Density of states

Density of states:

$$\rho(S) = \frac{dX}{dS} = -\frac{dX}{d\log L}$$

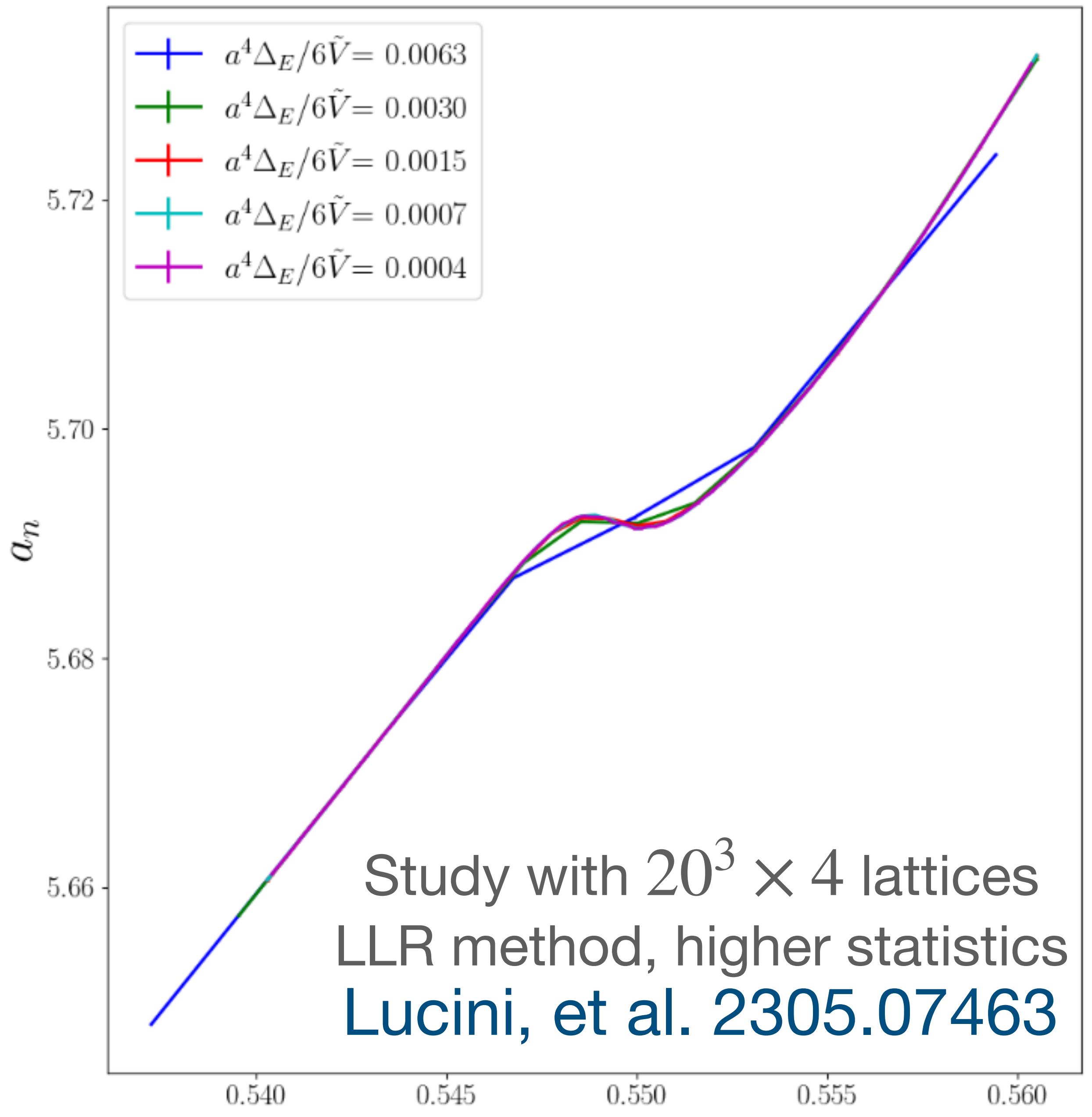
Microcanonical temperature:

$$\frac{1}{t} = \frac{d\log\rho}{dS}$$



$L \to$ large

FUTURE WORK



Comments: Monte Carlo challenges

1. Local Metropolis was much easier to tune than HMC

- Acceptance rate remains relatively constant over whole run
- Hard constraint for HMC requires (a) costly reflection calculation or (b) higher rejection rate unless $d\tau$ tuned well Betancourt (2011)
- Soft constraint for HMC possible Habeck (2015)

2. Even with many live walkers, fluctuations between 3 degenerate modes of broken symmetry phase are large

- Not a problem for relevant observables, such as $|P|$
- Relevant if inequivalent modes (e.g. instanton sectors) See previous talk by U. Wenger

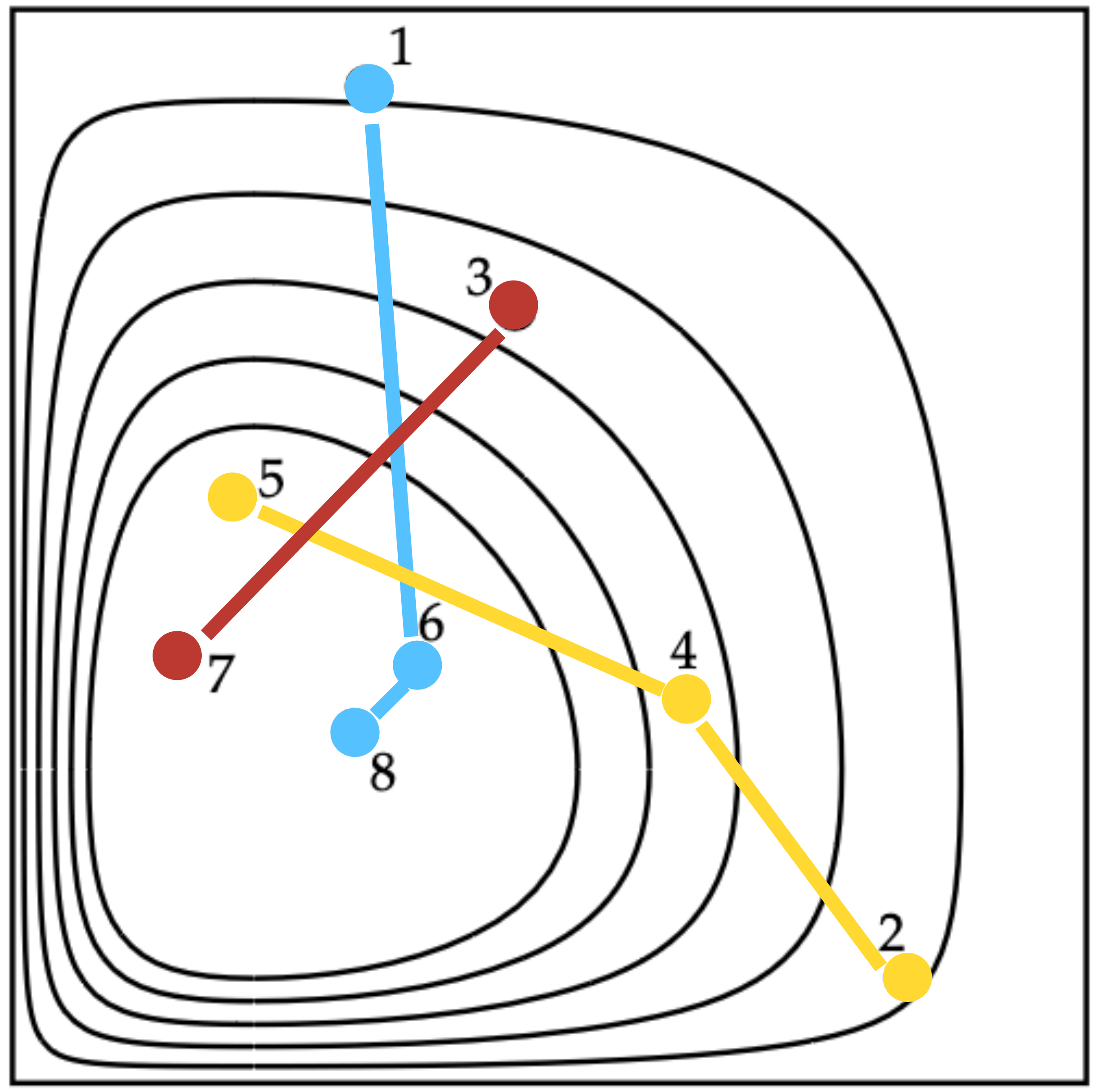
Comments: Parallelization

1. Nested sampling splits into independent "threads"

Higson, et al. (2018)

- Each resampled U' only depends on action S_i of deleted U_i
- Other walkers still useful as a resource to initialize resampling, but $k < N_{\mathrm{live}}$ can be deleted and resampled in parallel

Note: In contrast to existing literature, we resample with respect to distinct $S*$ values per deleted walker, avoiding loss of resolution



Configuration space

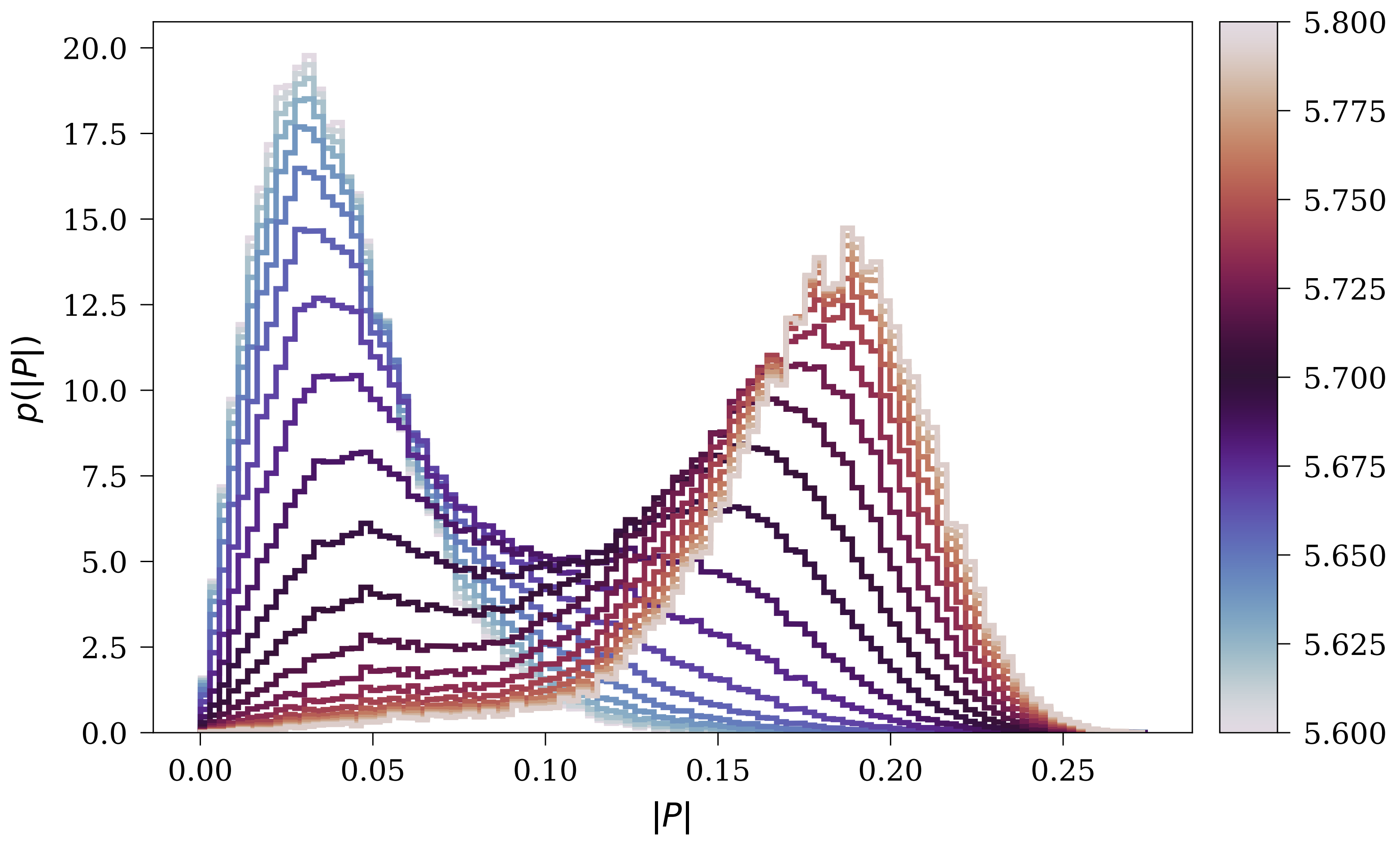
2. Threads can be combined post-hoc

- Binning and bootstrapping over threads for error estimates even with only one stream
- Statistics can be incrementally generated as usual

Summary / outlook

1. Nested sampling is a promising new Monte Carlo method
   - Particularly useful for phase transitions
   - Very different sampling strategy (uniform within contours)

2. Early results for $SU(3)$ confinement transition



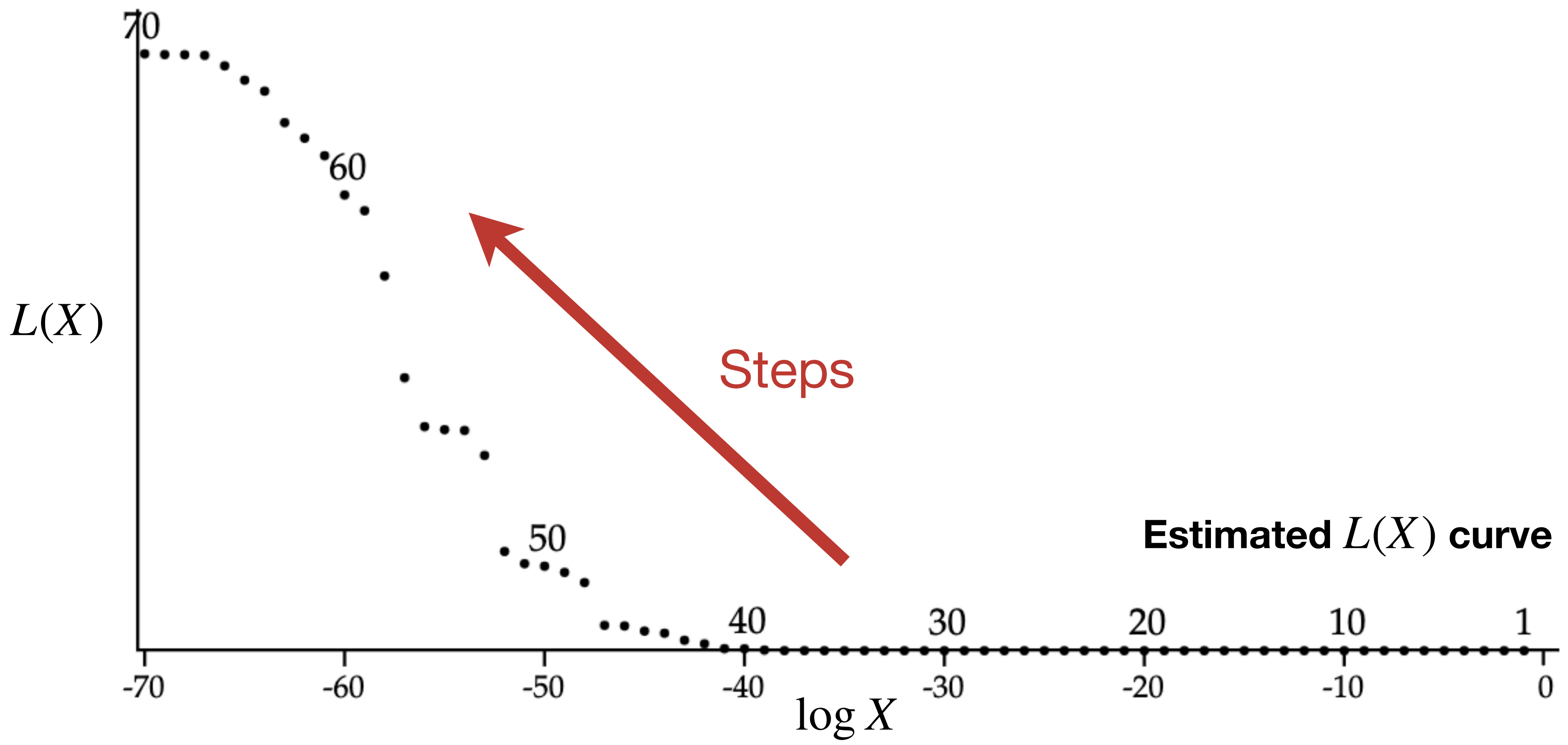
3. More developments to be done
   - Continue to control statistical and systematic error
   - Better understand volume scaling

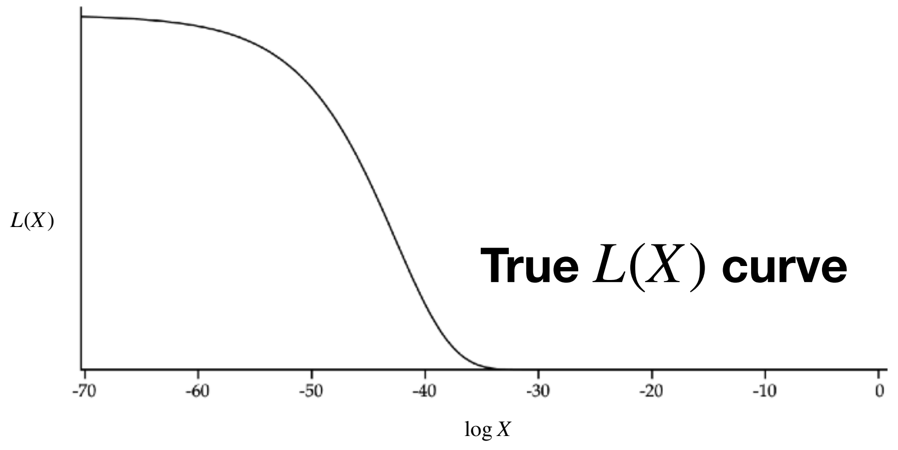
Backup slides

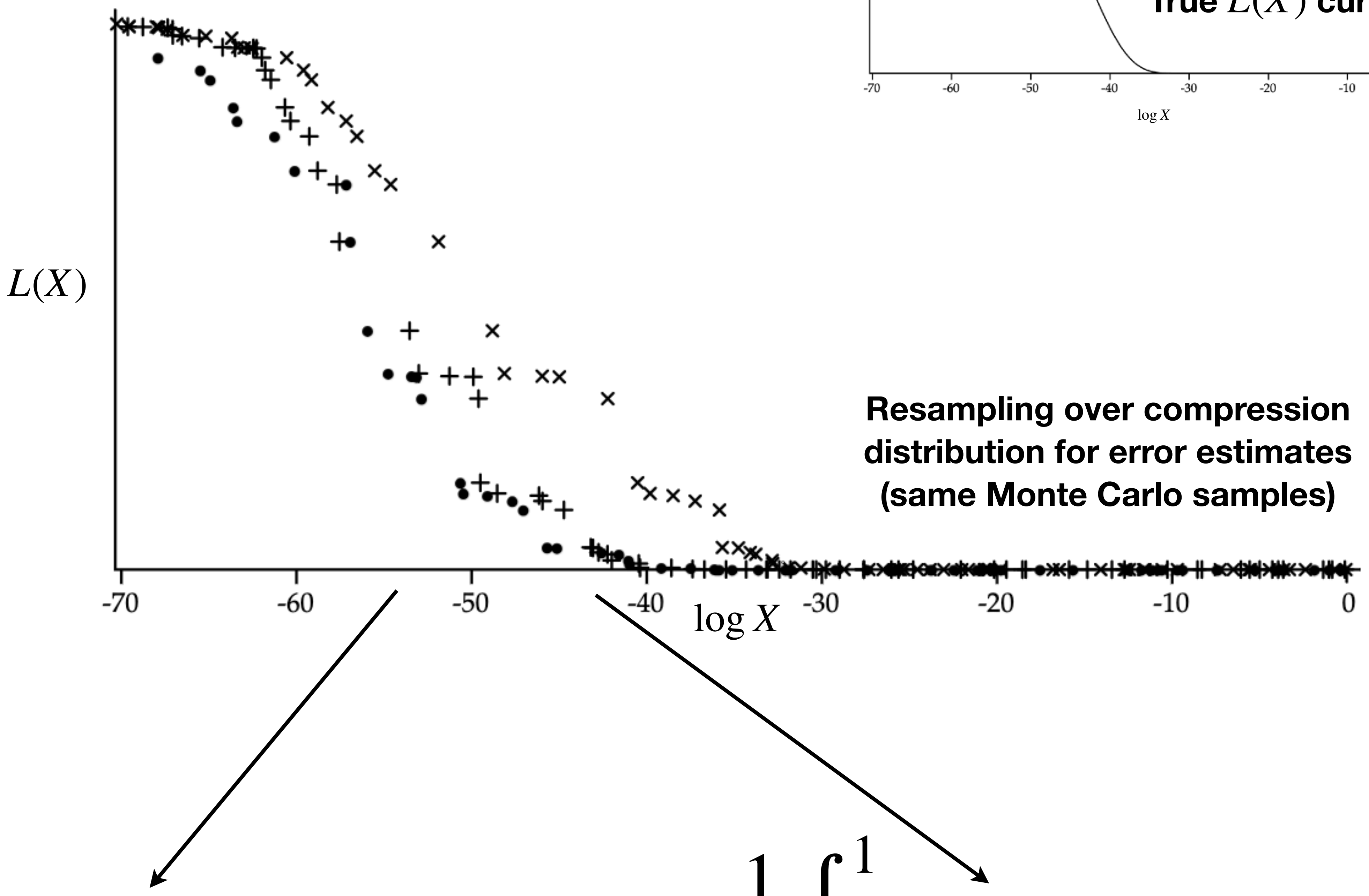
Nested sampling details

Estimate phase space X within contours of constant likelihood $L = e^{-S}$

- Example: N-dim Gaussian $L(\theta) = e^{-\frac{1}{2\sigma^2}\sum_{i=1}^{N}\theta_i^2}$







- Partition function and observables

$$Z = \int_0^1 dX L(X) \qquad \langle O(\theta)\rangle = \frac{1}{Z}\int_0^1 dX L(X) \langle O\rangle_{L(X)}$$

Undersampling is apparent for L=12