Conditional Normalizing Flow for Lattice Field Theory

21st International Workshop on Advanced Computing and Analysis Techniques in Physics Research

Ankur Singha,
PhD Scholar,
Dept. of Physics,
IIT Kanpur, India

INTRODUCTION

Quantum Field Theory describes physical world at the smallest scales.

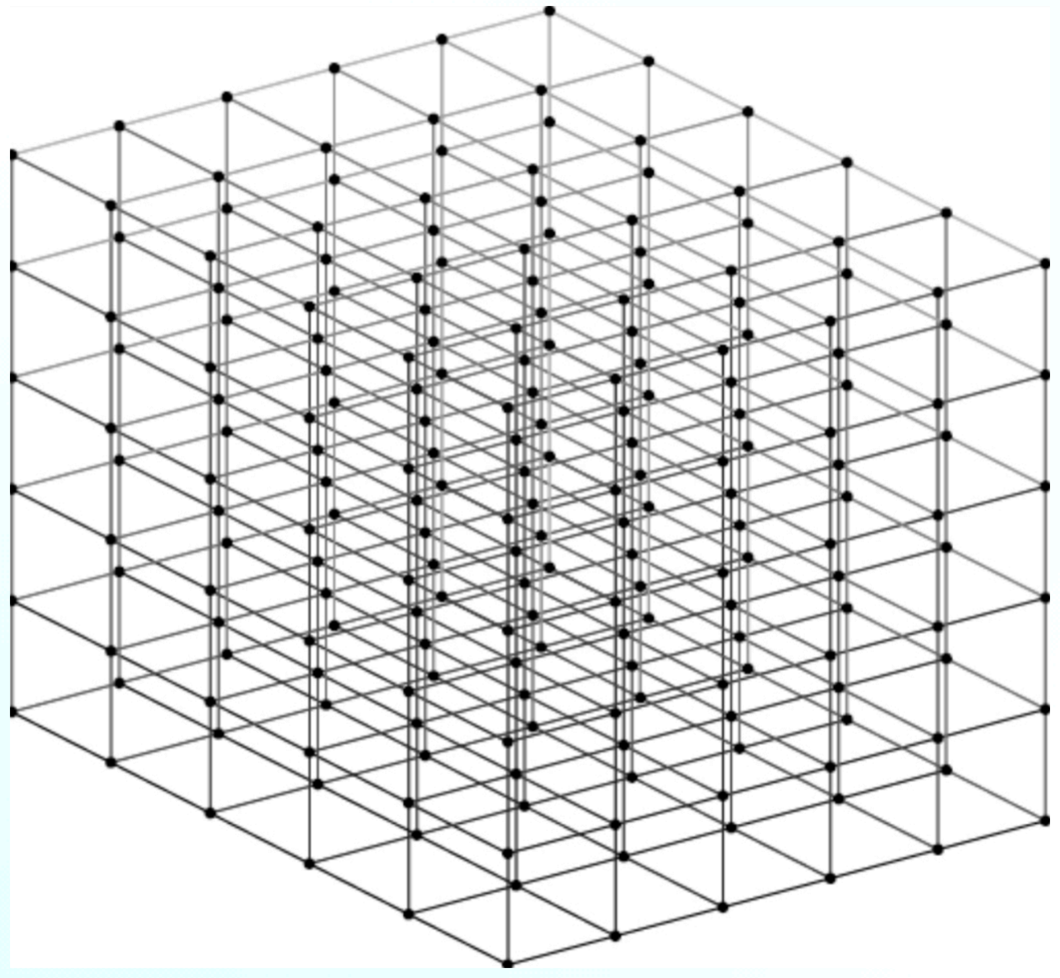
Theory: defined by Action or Lagrangian ;

$$\mathscr{L}(\phi, g_i)$$

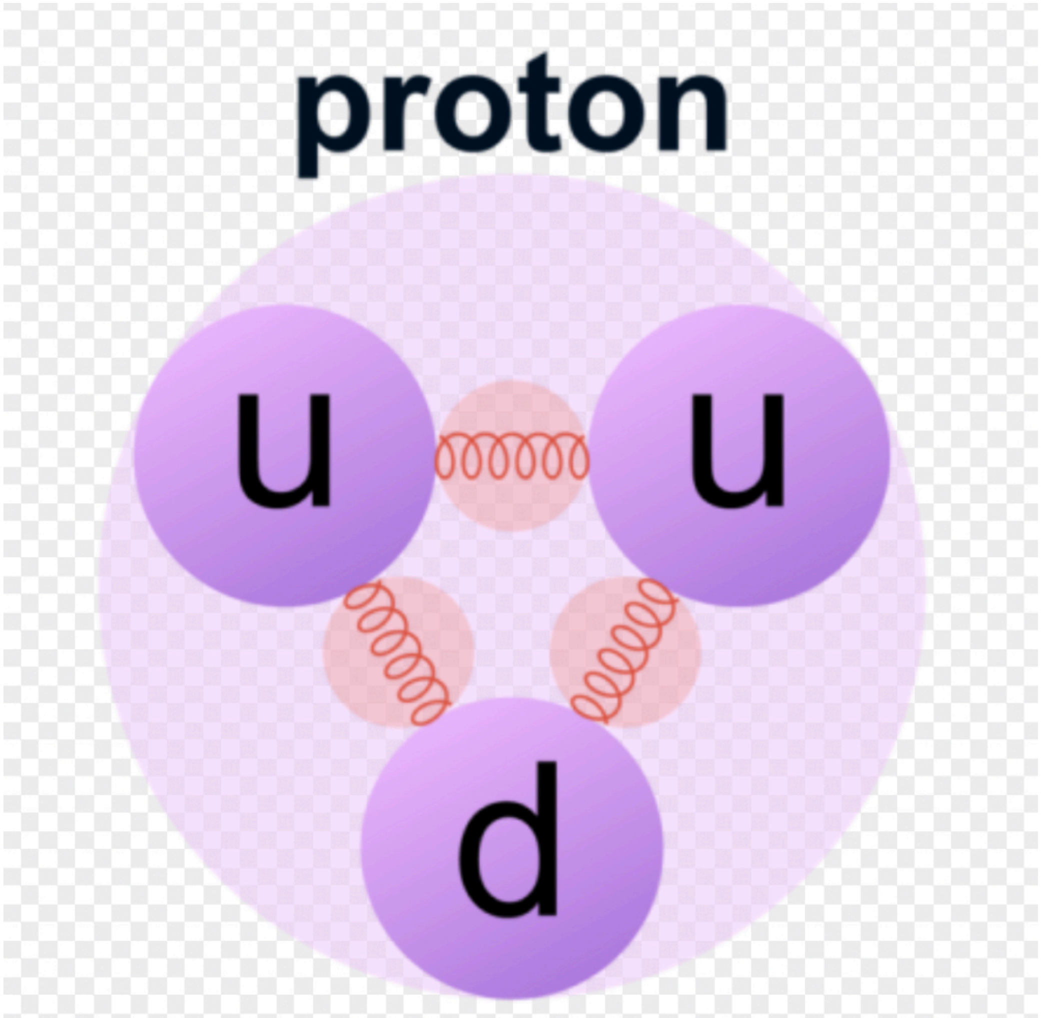
Strong Coupling → **Lattice Field Theory**

Weak Coupling → **Perturbative Methods**

Discrete space time



Computational approach → **QCD: Strong Interaction**



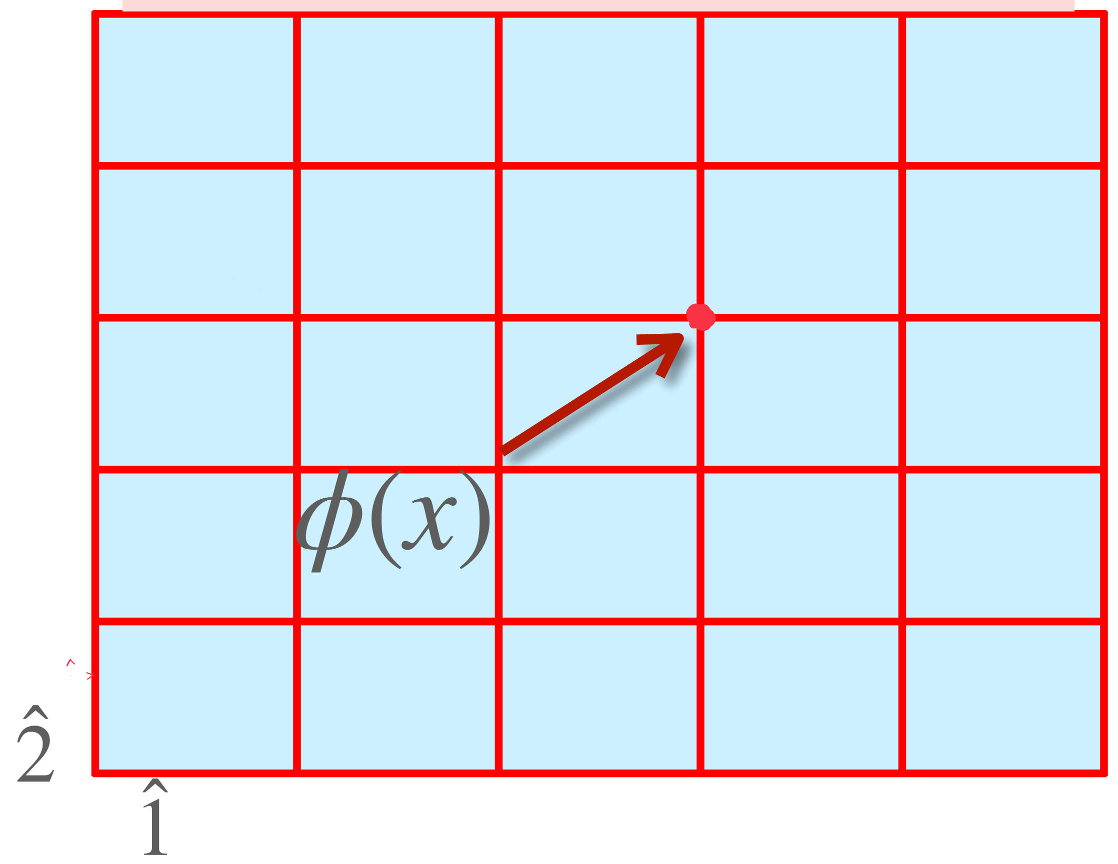
Scalar Lattice Field Theory

Discretized lattice Action:

$$S[\phi, m, \lambda] = \sum_{x} \sum_{\mu=1,2} [(2+m)\phi^2(x) - \phi(x)\phi(x+a\hat{\mu}) - \phi(x)\phi(x-a\hat{\mu}) + \lambda\phi^4(x)]$$

where, m and λ are the parameters of the theory.

2d Lattice



$$x \rightarrow (x_1\hat{\mu}, x_2\hat{\nu})$$

Observable:

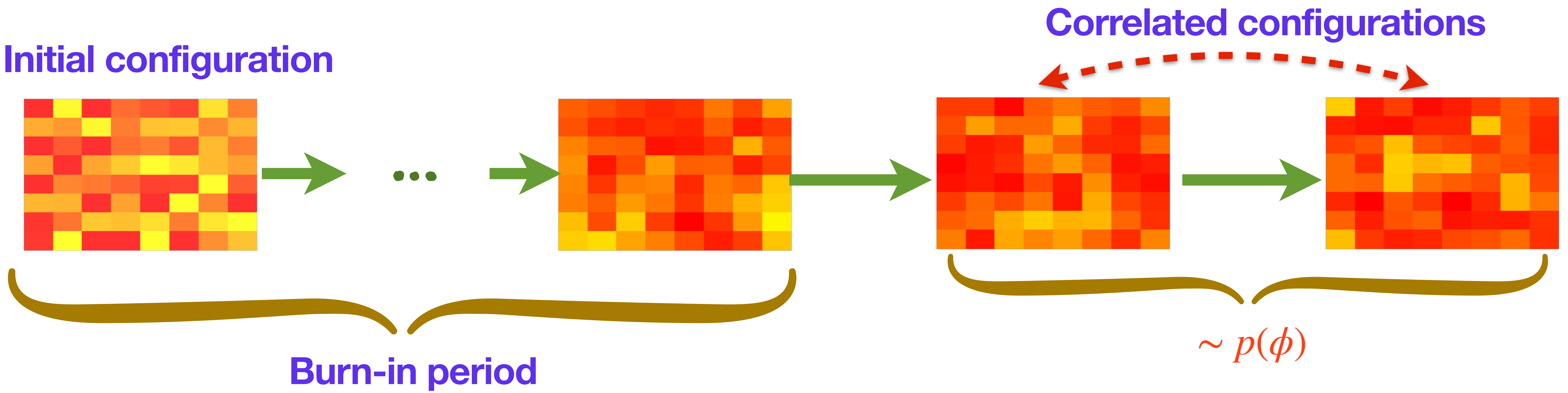
$$\langle O \rangle = \sum_{\phi_i} O(\phi_i) \frac{e^{-S(\phi_i)}}{Z} \rightarrow P(\phi_i)$$

Lattice Field Theory → **Sampling Task**

Distribution

$$P(\phi) = \frac{e^{-S(\phi)}}{Z}$$

MCMC & Critical Slowing Down (CSD)

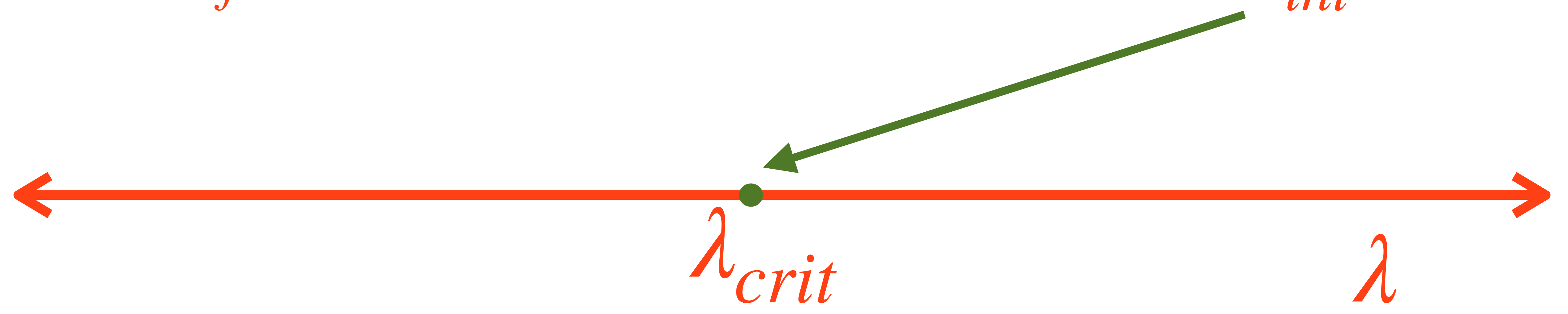


Integrated autocorrelation time:

$$\tau_{int} = \frac{1}{2} + \lim_{\tau \to \infty} \sum_{\tau=1}^{\tau_{max}} \frac{\rho(\tau)}{\rho(0)}$$

For a lattice action : $S(\phi, m_{fixed}, \lambda)$

τ_{int} **Divergent, CSD dominates**



ML based methods for Sampling

- **Generative Adversarial Network (GAN):**
  - Learns a distribution from training samples and generates new unseen data.
  - Can not estimates the density explicitly.

- **Normalizing Flow (NF):**
  - Learns a distribution without any samples as well as with training samples and generates new unseen data.
  - We can estimates the density explicitly.

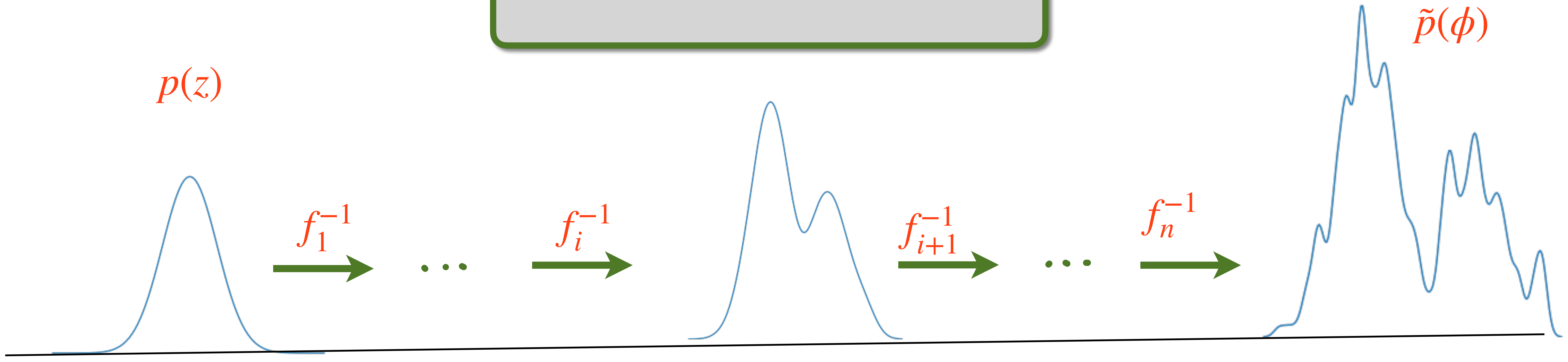
Normalising Flow

$$\phi = f^{-1}(z)$$

Prior distribution: $p(z)$

$$\tilde{p}(\phi) = \left| det \frac{f^{-1}(z)}{\partial z} \right| p(z)$$

Target distribution: $p(\phi)$



$f_i^{-1} \longrightarrow$ **Neural Network**

$\tilde{p}(\phi) \longrightarrow \tilde{p}(\phi;\theta)$

Optimise model so that:
$$\tilde{p}(\phi;\theta) \approx p(\phi)$$

Normalising Flow

Training the NF model can done by minimizing the KL divergence between the $\tilde{p}(\phi;\theta)$ and $p(\phi)$.

- **Forward KL:**

$$\mathscr{L}_R = D_{KL}\big(p(\phi)\,||\,\tilde{p}(\phi;\theta)\big)\,.$$

Require training samples from $p(\phi)$

- **Reverse KL:**

$$\mathscr{L}_F = D_{KL}\big(\tilde{p}(\phi;\theta)\,||\,p(\phi)\big)\,.$$

Require samples from $\tilde{p}(\phi;\theta)$

Application to lattice phi4 theory

Phi4 theory: m **and** λ **fixed** .

$$S[\phi, m, \lambda] = \sum_{x} \sum_{\mu=1,2} [(2+m)\phi^2(x) - \phi(x)\phi(x+a\hat{\mu}) - \phi(x)\phi(x-a\hat{\mu}) + \lambda\phi^4(x)]$$

Generates training data from HMC simulation → **10000 lattice configurations**

✓ **Training the model**

$$z = f(\phi; \theta_{opt})$$

✓ **Generation of lattices**

$$\phi = f^{-1}(z)$$

z **can be easily sampled**

Metropolis-Hastings

Samples from the NF model $\tilde{p}(\phi;\theta)$ cannot be considered for observable calculation.

Cause: → Biases in observable.

We use the samples from NF model as proposal to construct a Markov chain.

Metropolis-Hastings step:

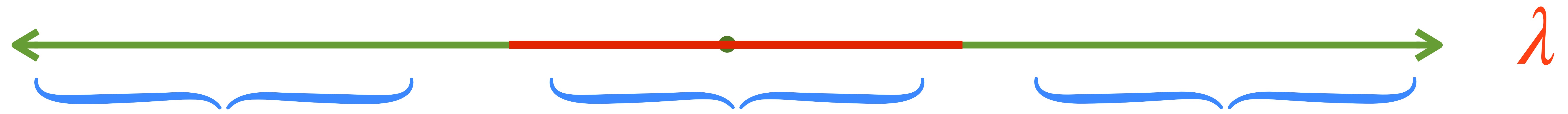
$$A(\phi^{i-1}, \phi') = min\Big(1, \frac{\tilde{p}(\phi^{(i-1)})p(\phi')}{p(\phi^{(i-1)})\tilde{p}(\phi')}\Big)$$

Provides the exactness of the distribution.

Simulation at multiple λ values

For a lattice action : $S(\phi, m_{fixed}, \lambda)$ → $p(\phi\,|\,\lambda_i)$

action parameter



Non-Critical Region

Critical Region

Non-Critical Region

τ_{int} Divergent, simulation cost is high

Utilise information from non-critical regions → Generate lattice configurations in critical regions

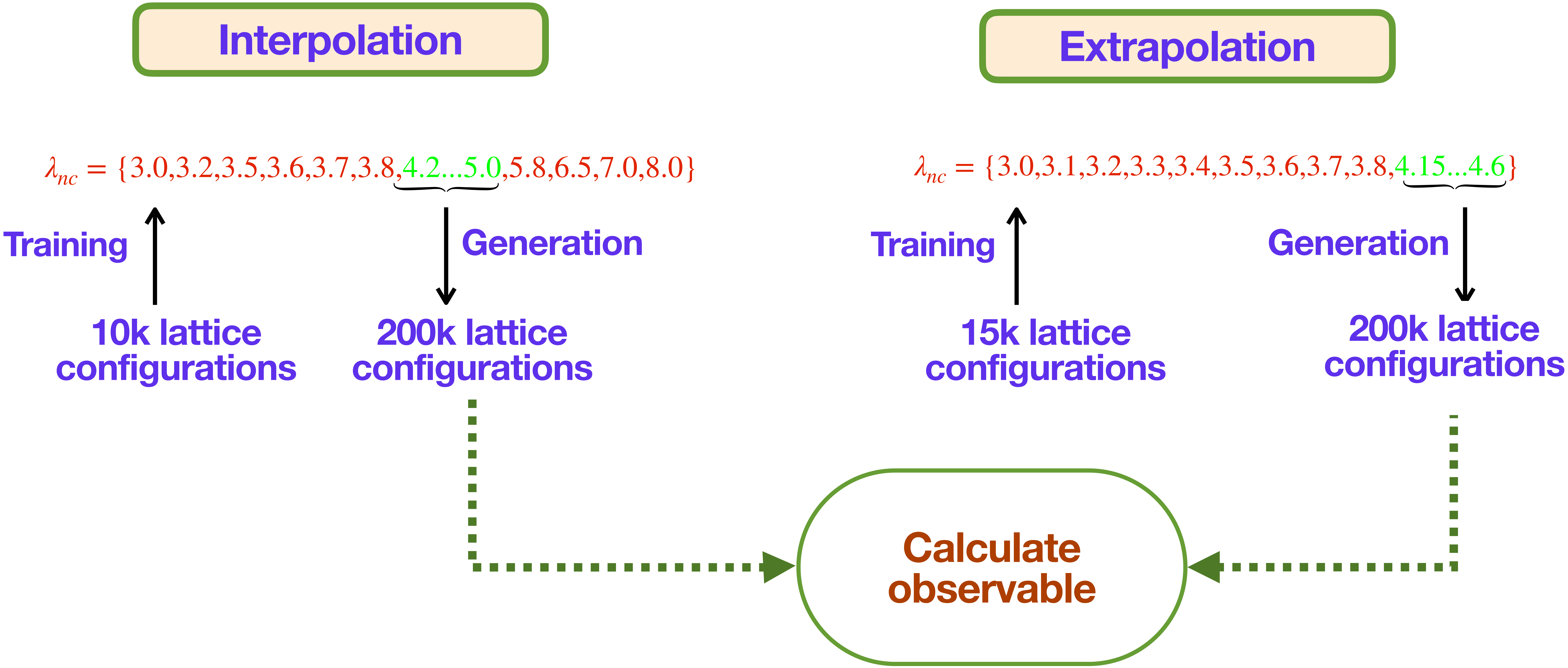
$\phi_{prop} \sim \tilde{p}(\phi; \lambda_c, \theta_{opt})$

Conditional NF

We studied two cases:

- **Interpolation**: generates samples on **both sides** of critical region.
  - Training samples $\phi \sim p(\phi \mid \lambda_{nc})$ where λ_{nc} belong to both non-critical regions.
- **Extrapolation**: generates samples on a **single side** of critical region.
  - Training samples $\phi \sim p(\phi \mid \lambda_{nc})$ where λ_{nc} belong to a single non-critical regions.

C-NF: Training & Generation



Observable

Observable calculated on lattice from HMC and C-NF Model are:

1. $\langle \tilde{\phi}^2 \rangle$: $\tilde{\phi} = \frac{1}{V} \sum_x \phi(x)$
2. Zero momentum Correlation Function:

   $C(t) = \sum_{x_1} G_c(x_1, t)$

   where, $x = (x_1, t)$ and

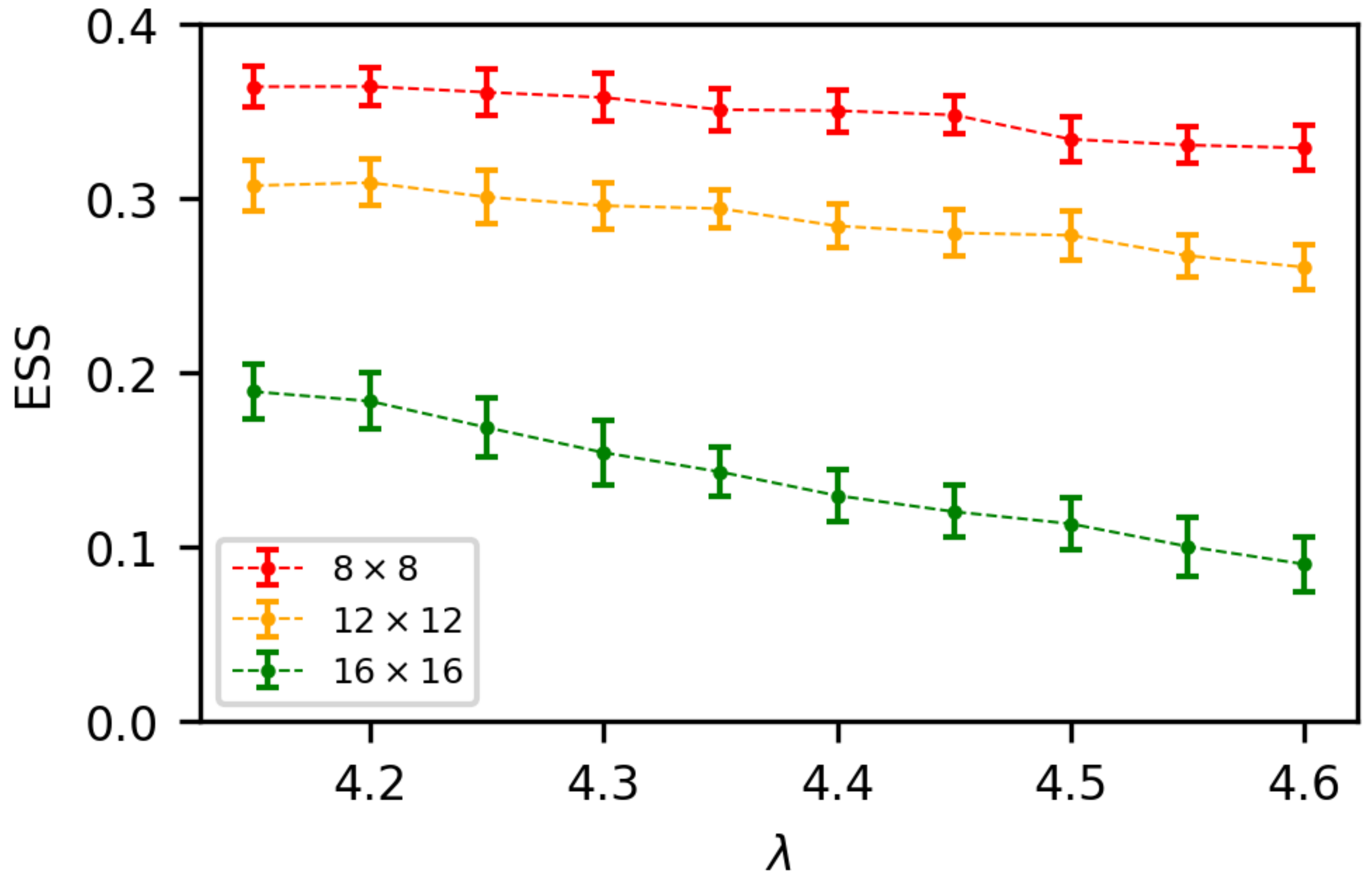
   $G_c(x) = \frac{1}{V} \sum_y [\langle \phi(y)\phi(x+y) \rangle - \langle \phi(y) \rangle \langle \phi(x+y) \rangle]$

3. Two Point Susceptibility: $\chi = \sum_x G(x)$

Results: Effective Sample Size (ESS)

C-NF model quality: ESS

$$ESS = \frac{1}{N}\frac{\left(\sum_{i=1}^{N} p(\phi_i;\lambda)/\tilde{p}(\phi_i;\lambda,\theta)\right)^2}{\sum_{i=1}^{N}\left(p(\phi_i;\lambda)/\tilde{p}(\phi_i;\lambda,\theta)\right)^2}$$

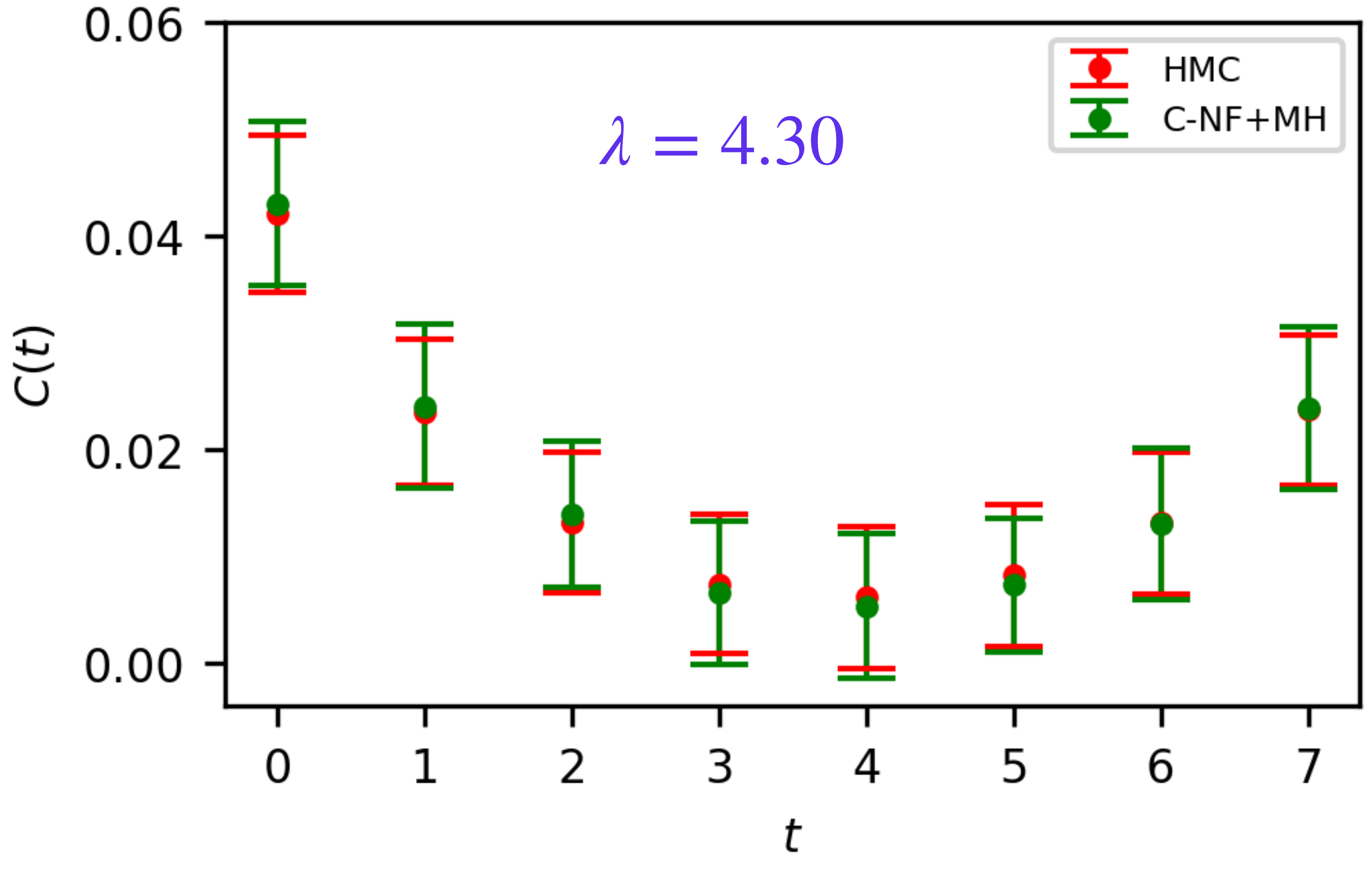


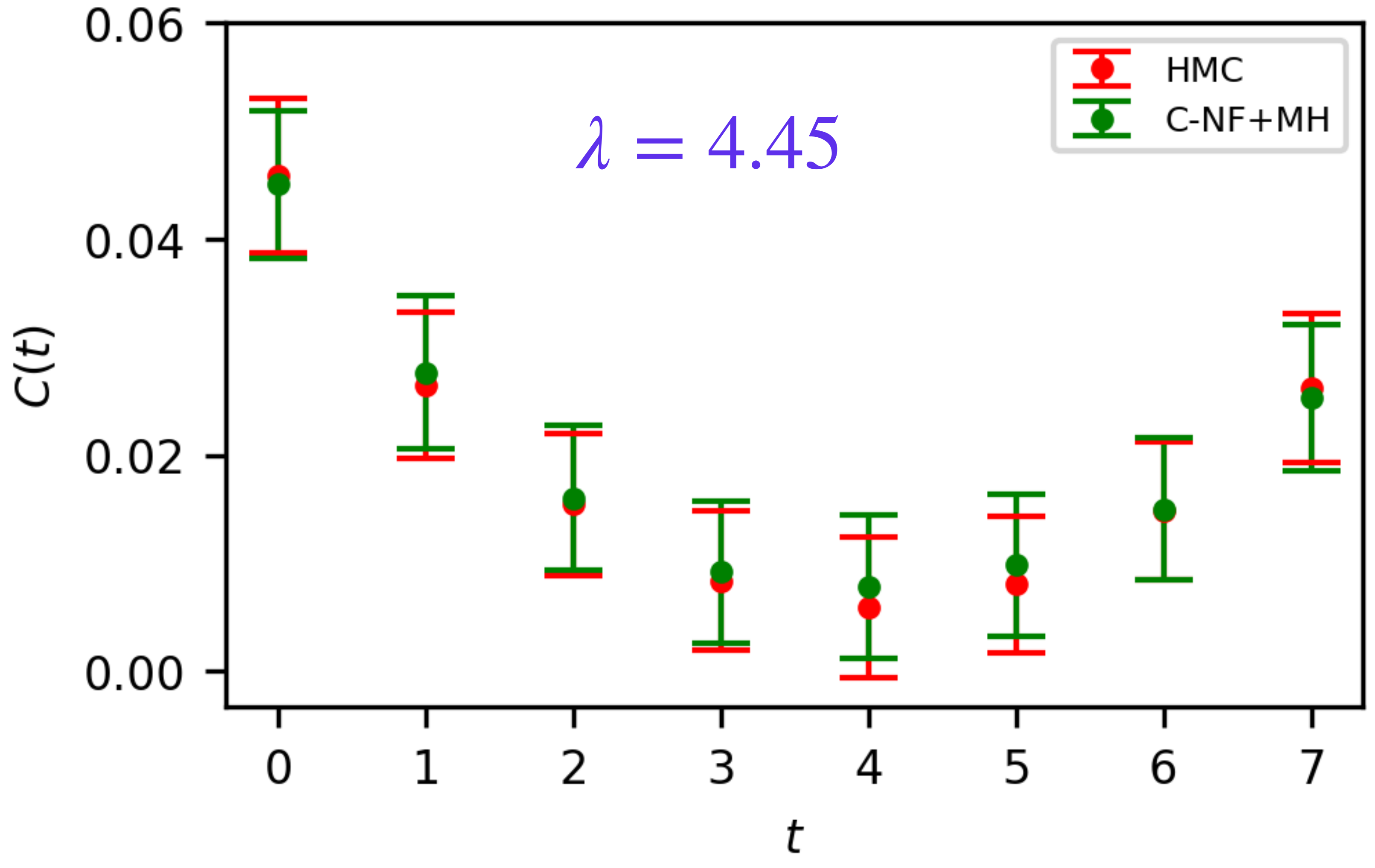
Higher ESS

Better Model → **Higher Acceptance** → **Low Autocorrelation**

Results: Correlation Functions

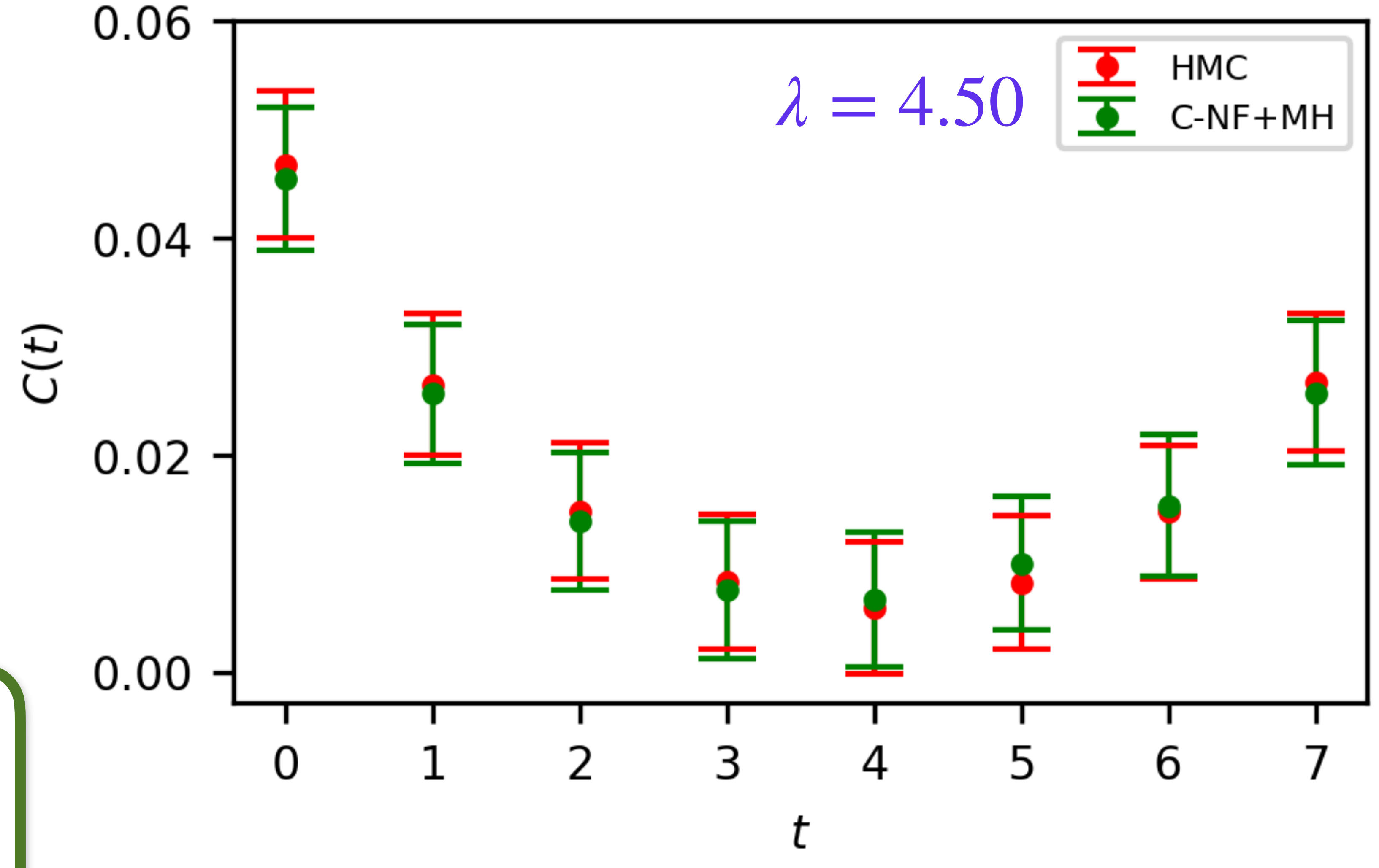
Extrapolation

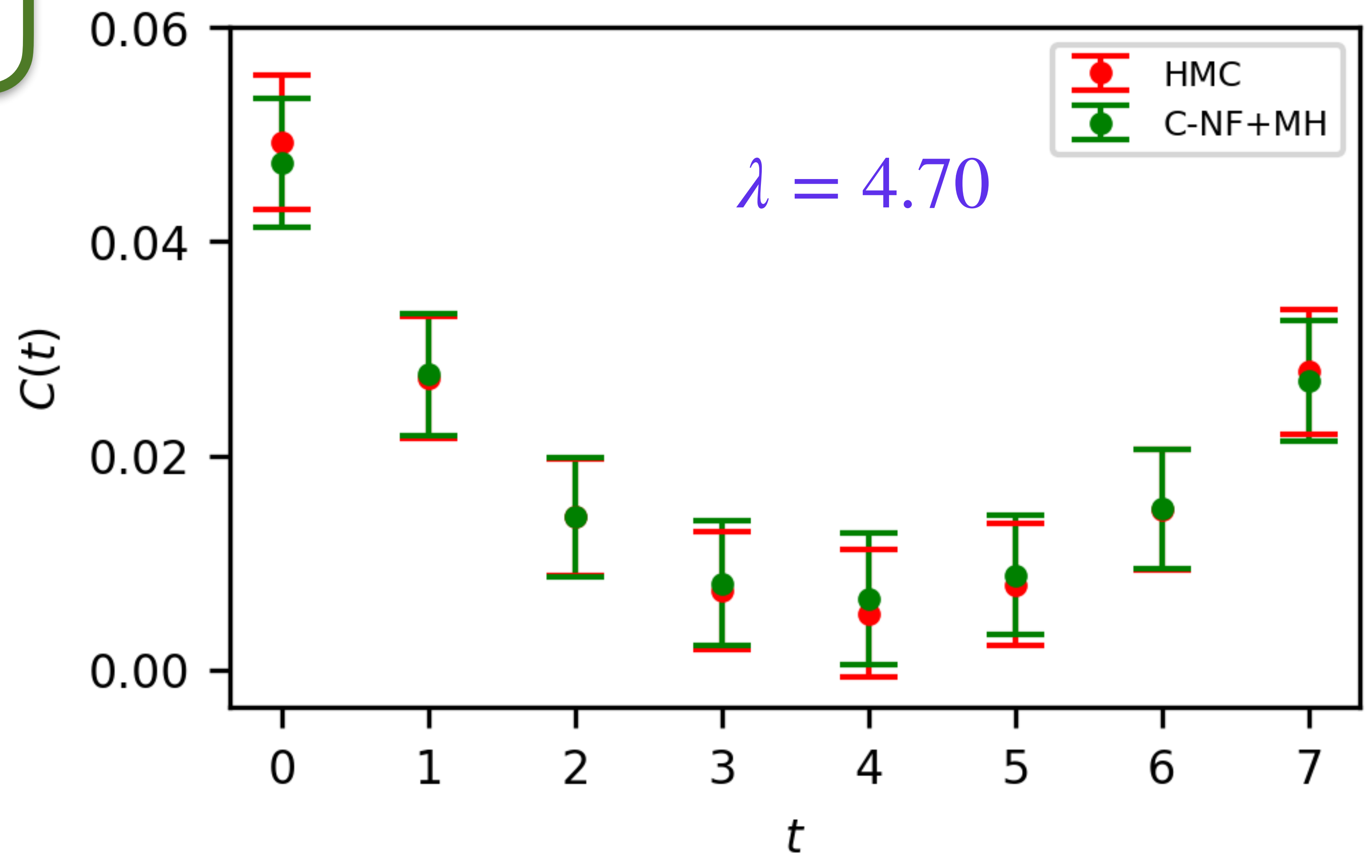




Results only from the critical region

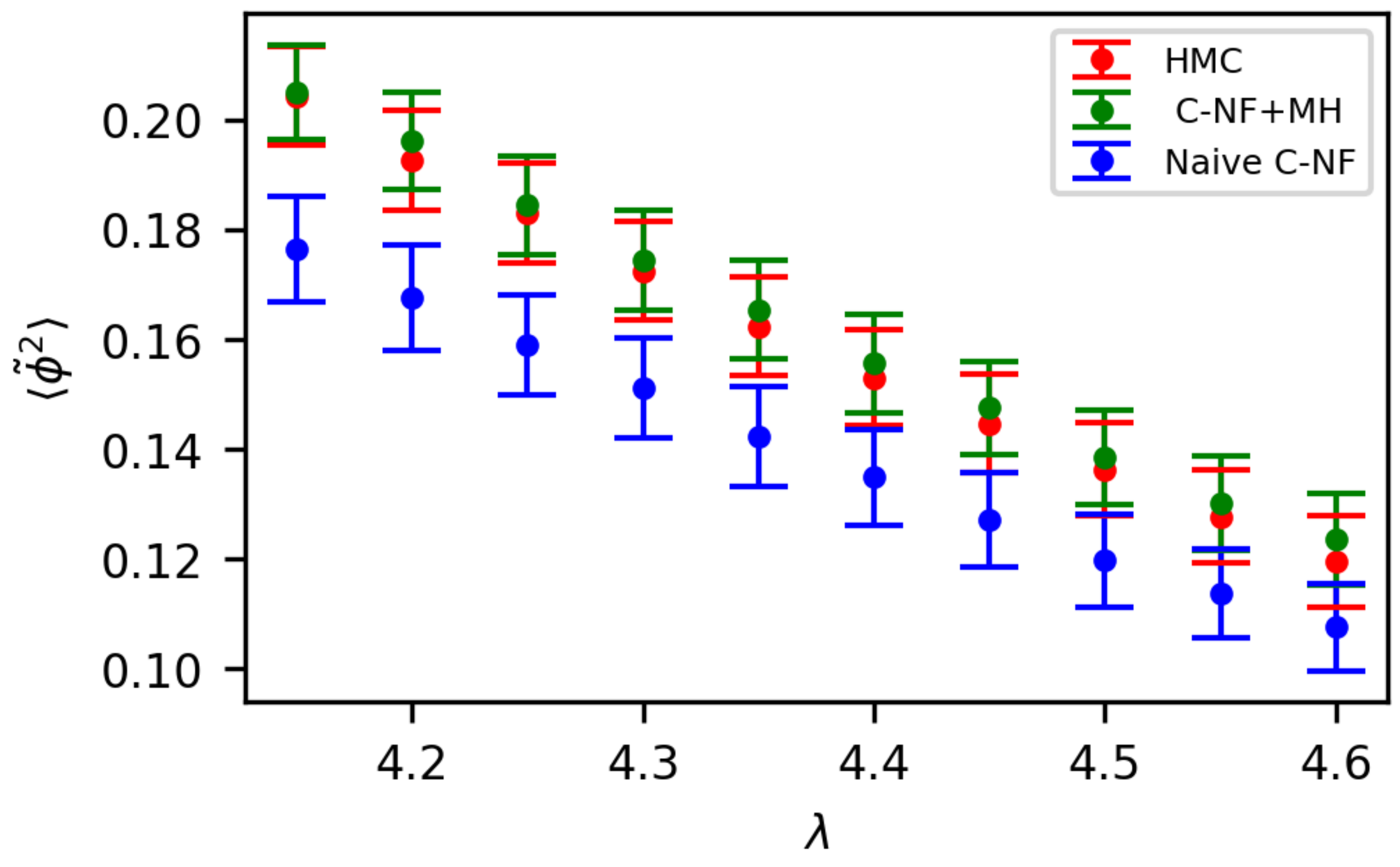
Interpolation

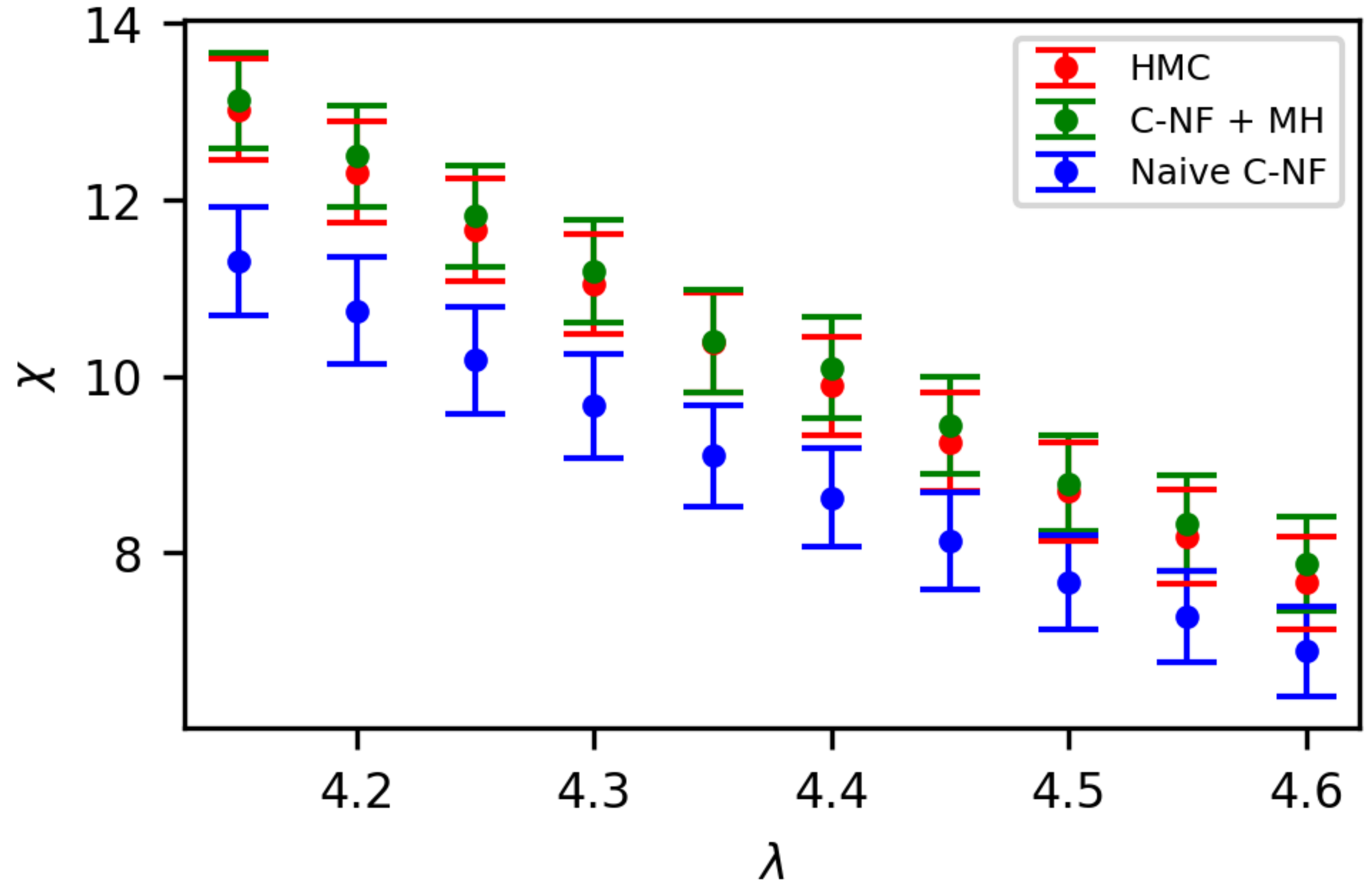




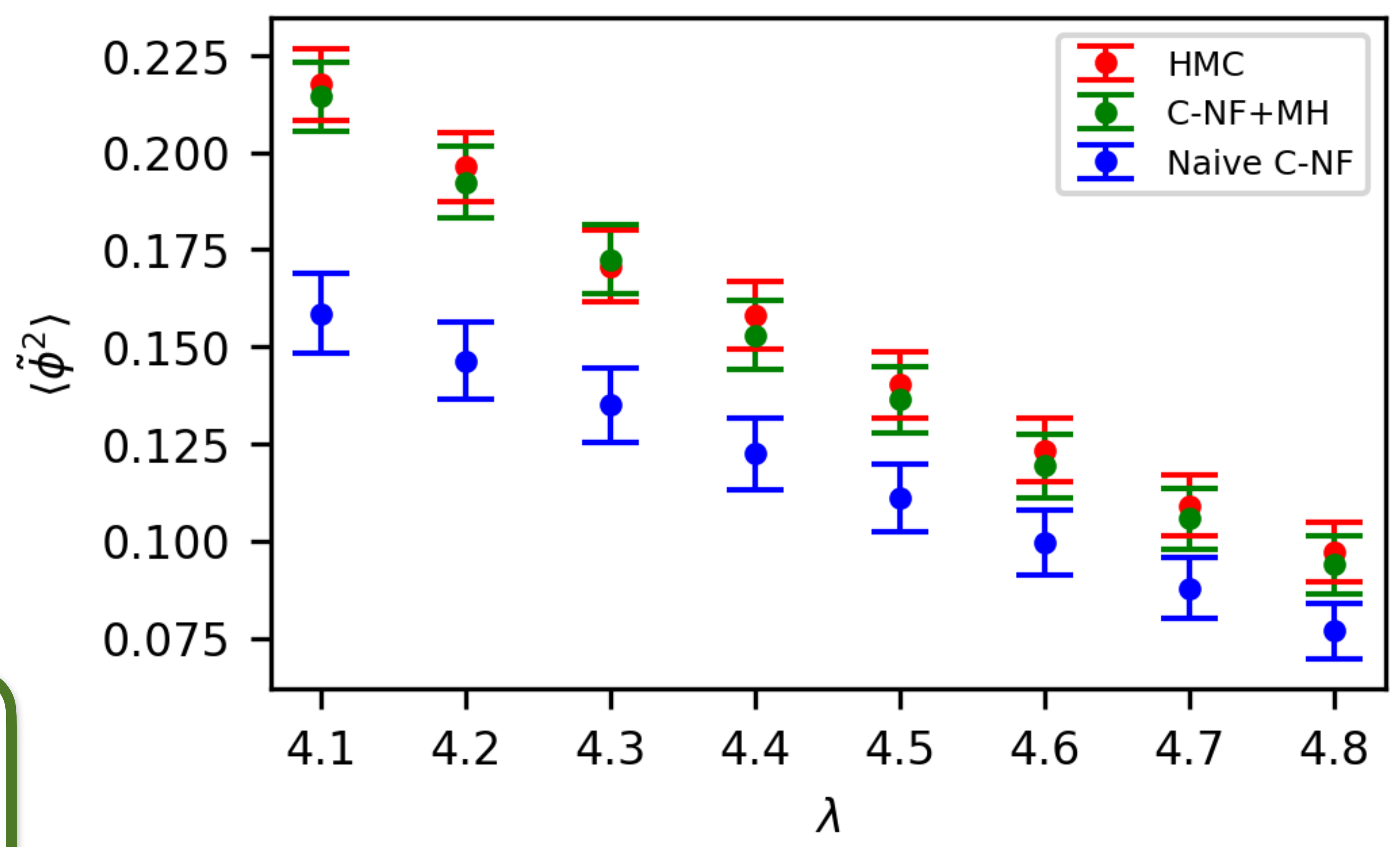
Results : χ *and* $\langle \tilde{\phi}^2 \rangle$

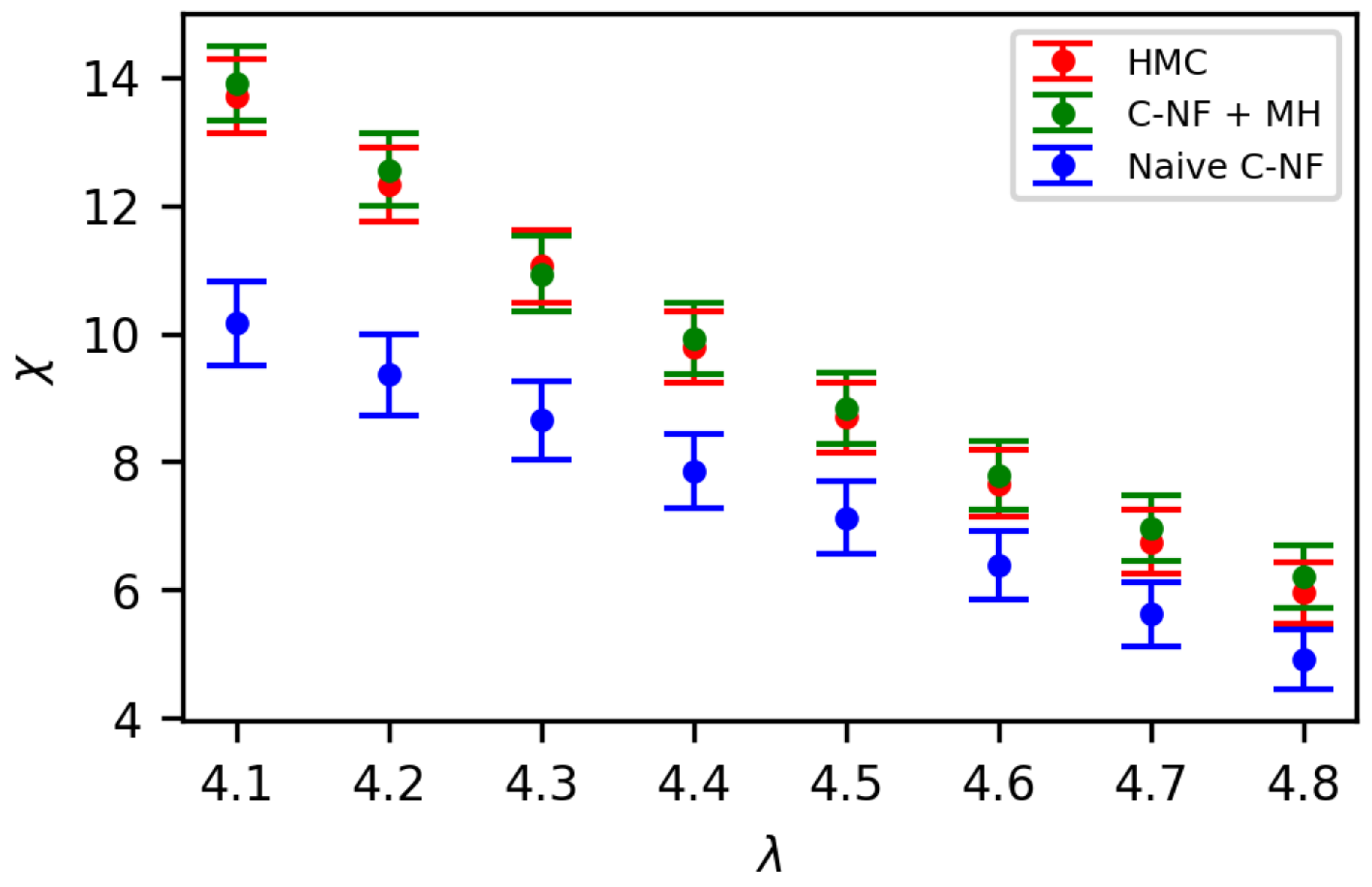
Extrapolation





Interpolation

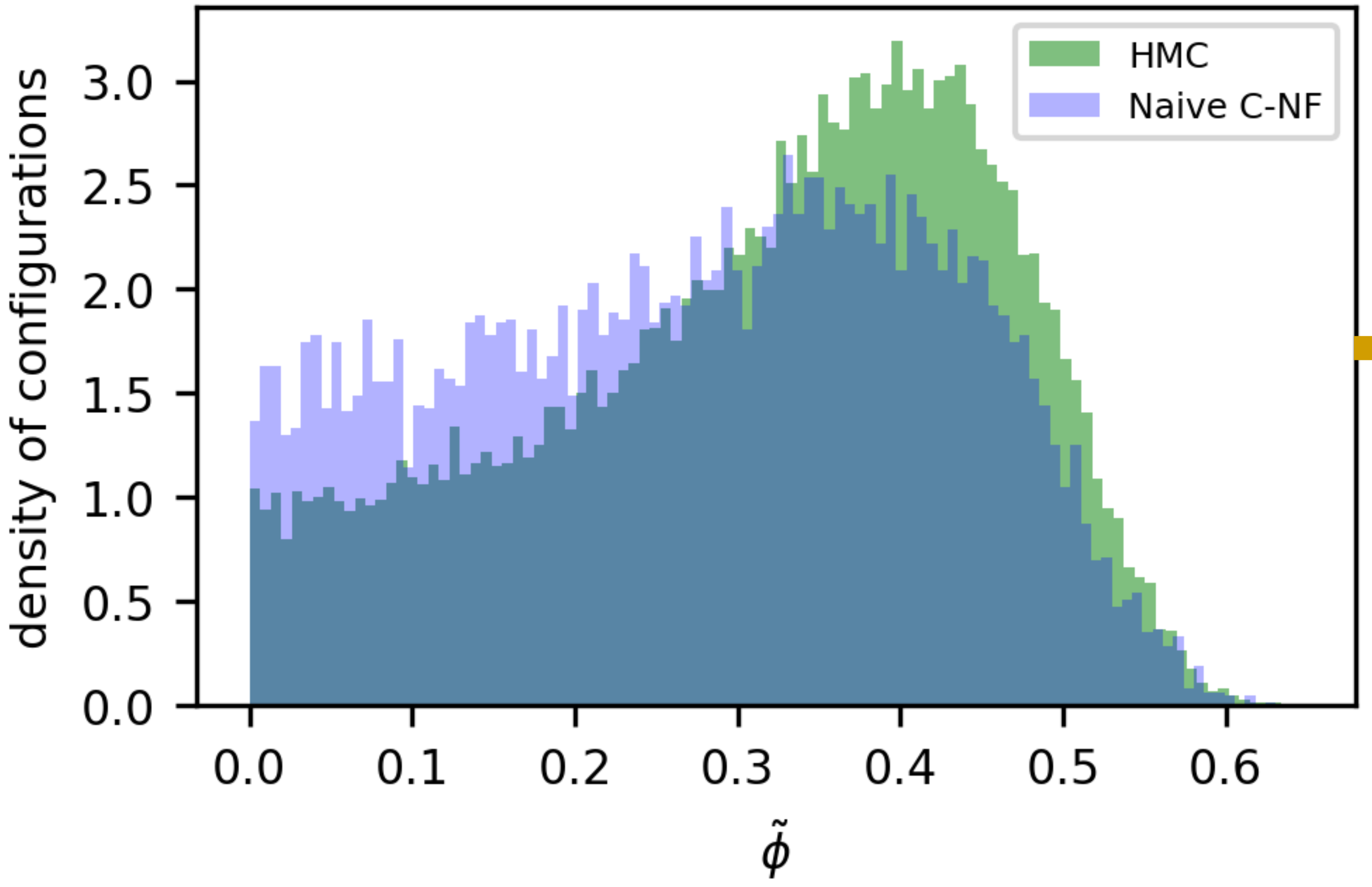




Results matches within statistical uncertainty

Results: artefacts removed by MH

Naive model



MH algorithm

Naive model + MH

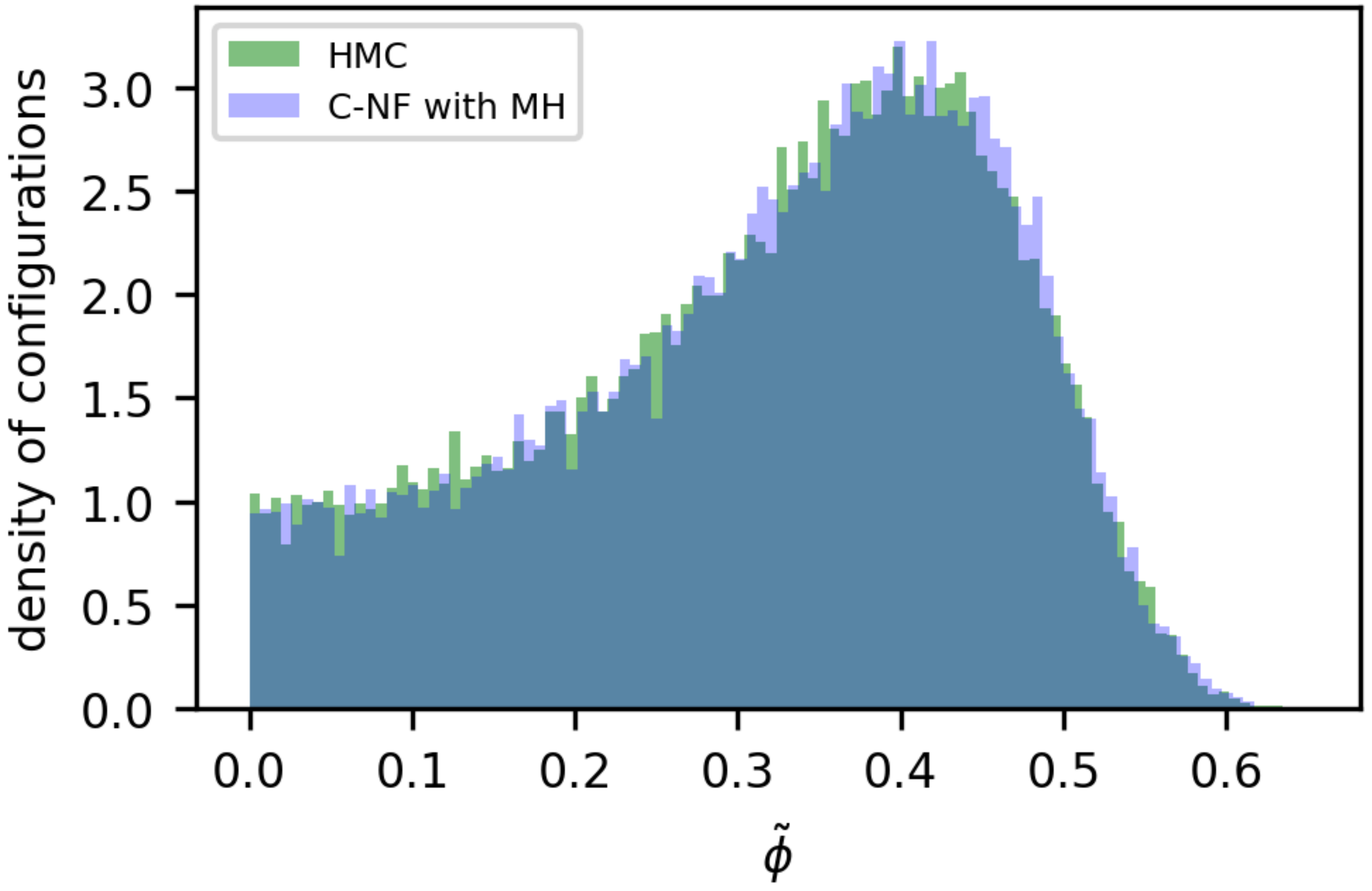


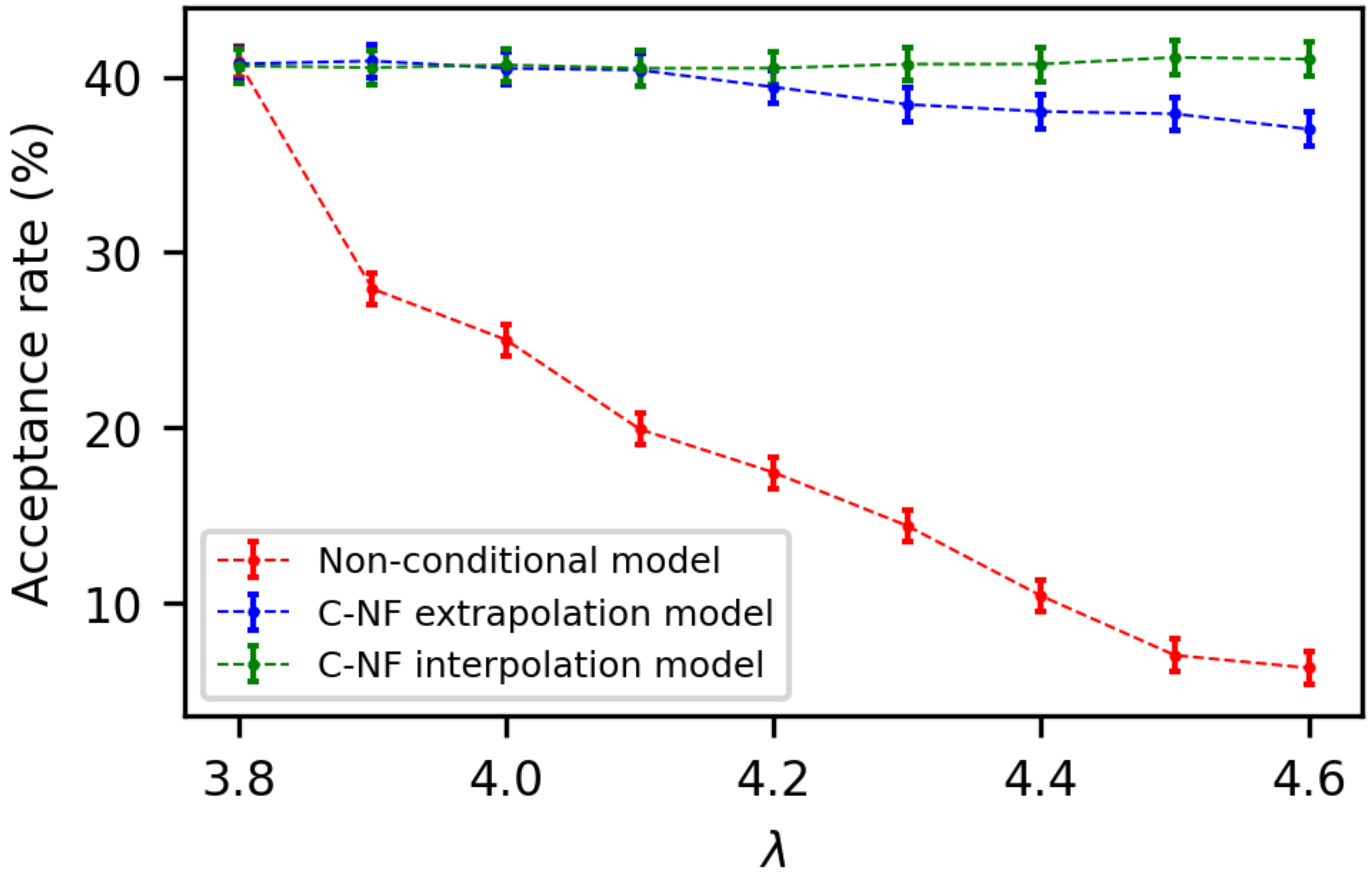
Figure: Histogram of density of lattice configurations for average field value ϕ on a lattice.

Results: compared to a non-conditional model

Acceptance rate in MH: a non-conditional model trained at $\lambda = 3.8$ vs conditional models.

✓ **For interpolation and extrapolation the acceptance rate is almost constant over action parameter.**

✓ **For a non-conditional model the acceptance rate drops faster over the action parameter.**



Summary

Problem: **MCMC simulations are severely affected by CSD.**

Proposed: A conditional NF model.

https://arxiv.org/abs/1904.12072

NF generate proposals for MH

Low Simulation Cost

Utilise samples from non-critical region

Conditional NF generate MH proposals for wide range of action parameter.

THANK YOU !