Generative Models for the Discrete Hubbard Model

The 43nd International Symposium on Lattice Field Theory
28.07.2026

[arXiv:26MM.XXXXX]

Janik Kreit, Dominic Schuh, and Lena Funcke

Discrete Hubbard Model

Continuous Hubbard Stratonovich Transformation

$$S[\phi] = \frac{1}{2\tilde{U}} \sum_{x,t} \phi_{x,t}^2 - \log \det M[\phi, \mu] - \log \det M[-\phi, \mu]$$

$\phi \in \mathbb{R}^{|\Lambda|}$

Discret

$S[\phi] = -$

$\phi \in \{-1, 1$

Discrete Hubbard Model

ransformation

$$\text{og}\det M[-\phi,\mu]$$

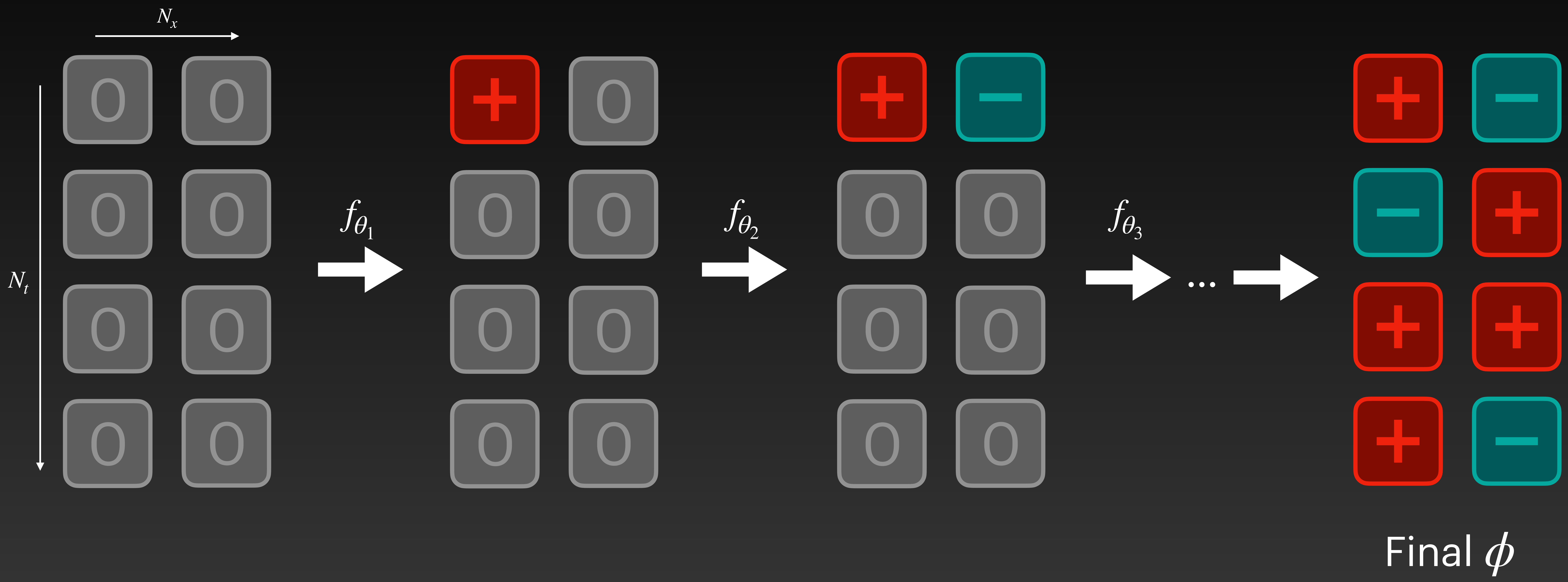
Discrete Hubbard Stratonovich Transformation

$$S[\phi] = -\log\det M[a\phi,\mu] - \log\det M[-a\phi,\mu]$$

$$\phi \in \{-1,1\}^{|\Lambda|}$$

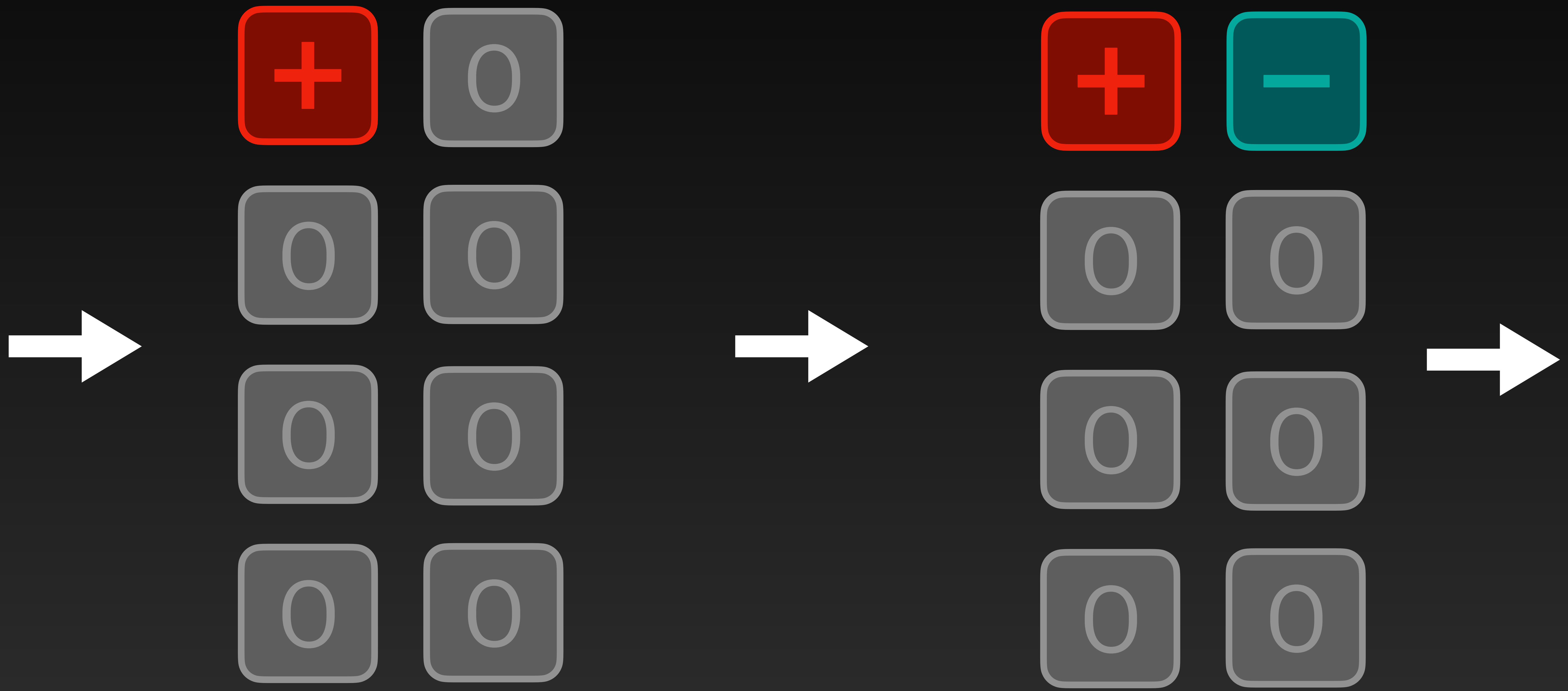
$a = \text{atanh}\left(\sqrt{\tanh(\tilde{U}/4)}\right)$

Generation of Discrete Fields

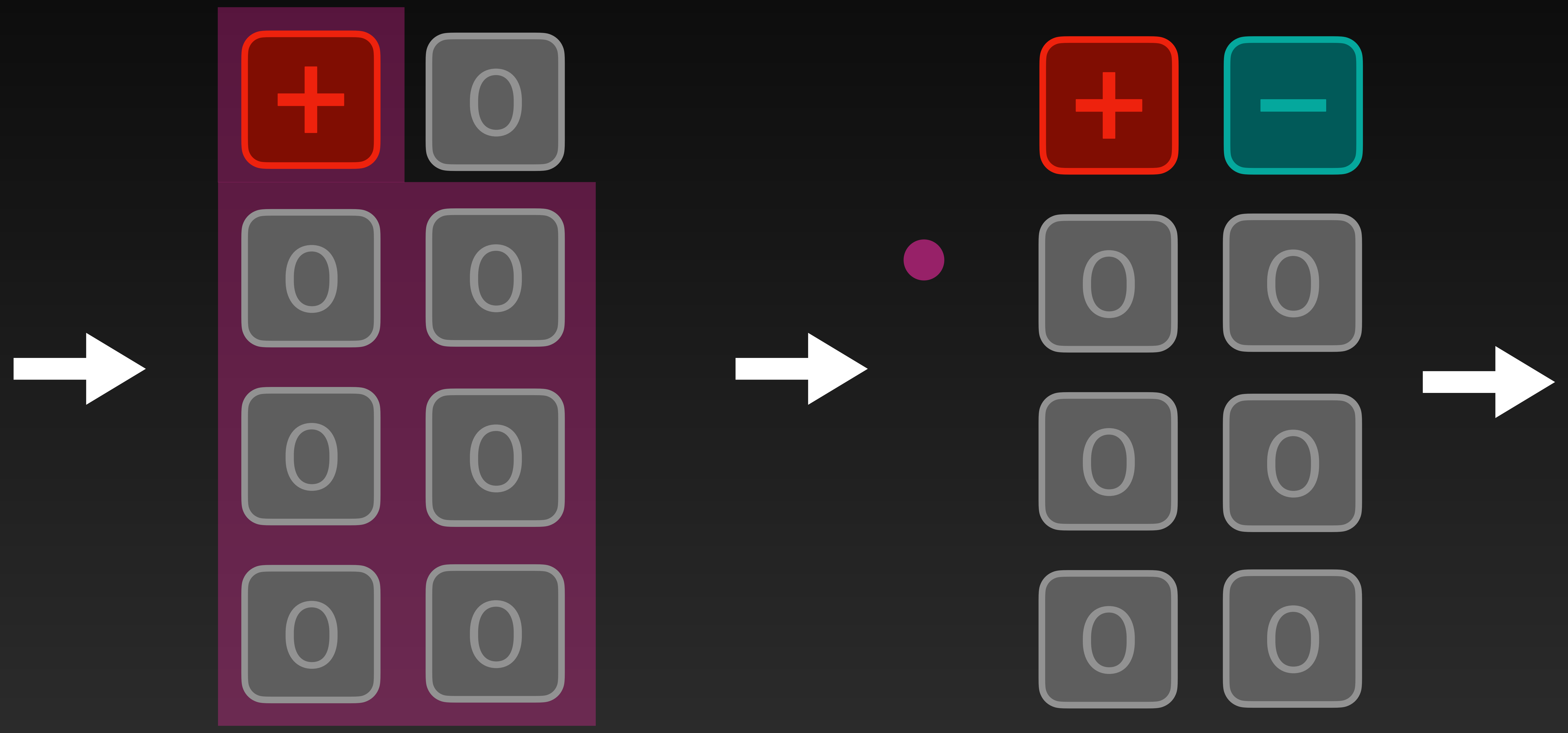


G. Papamakarios et al., NIPS 30 (2017)

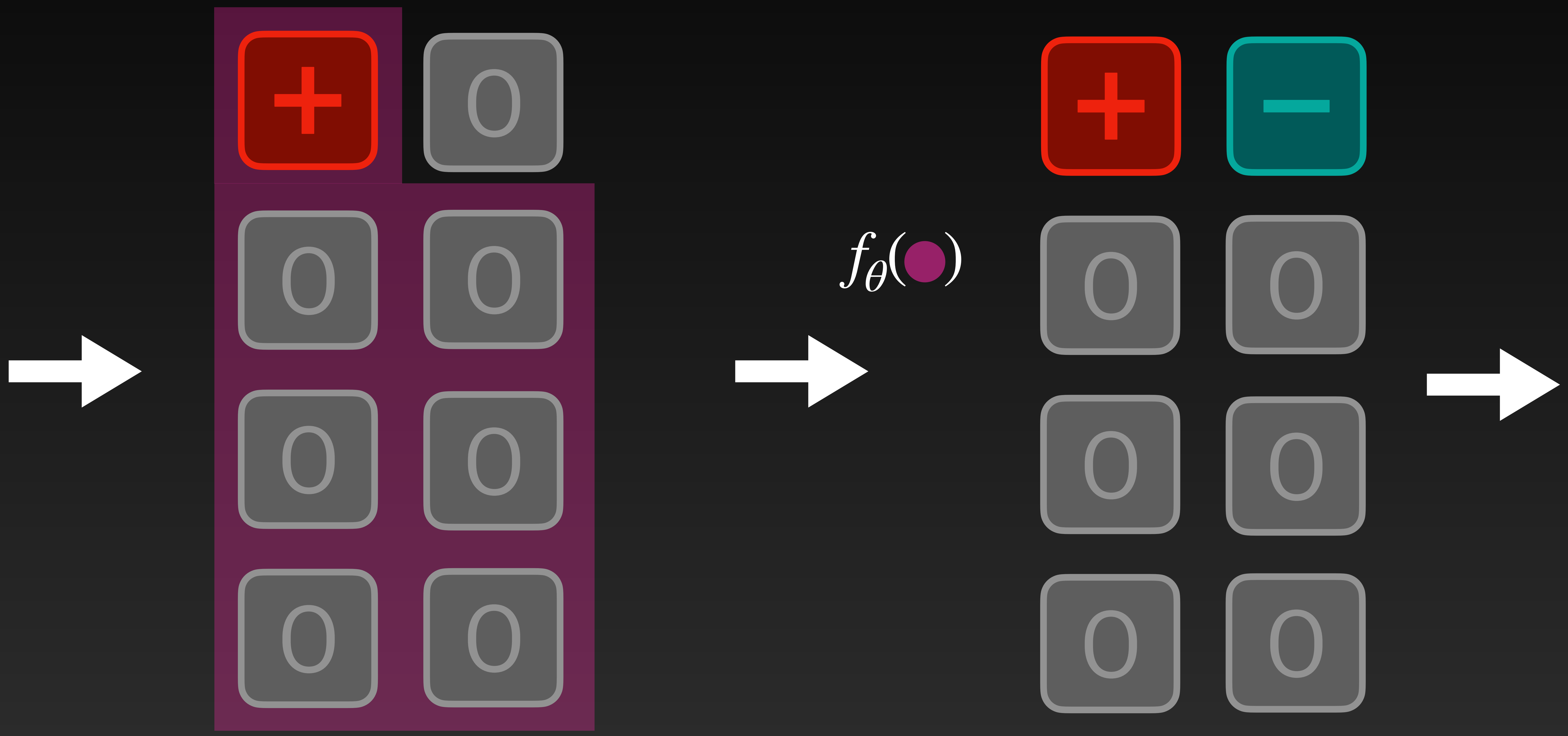
Generation of Discrete Fields



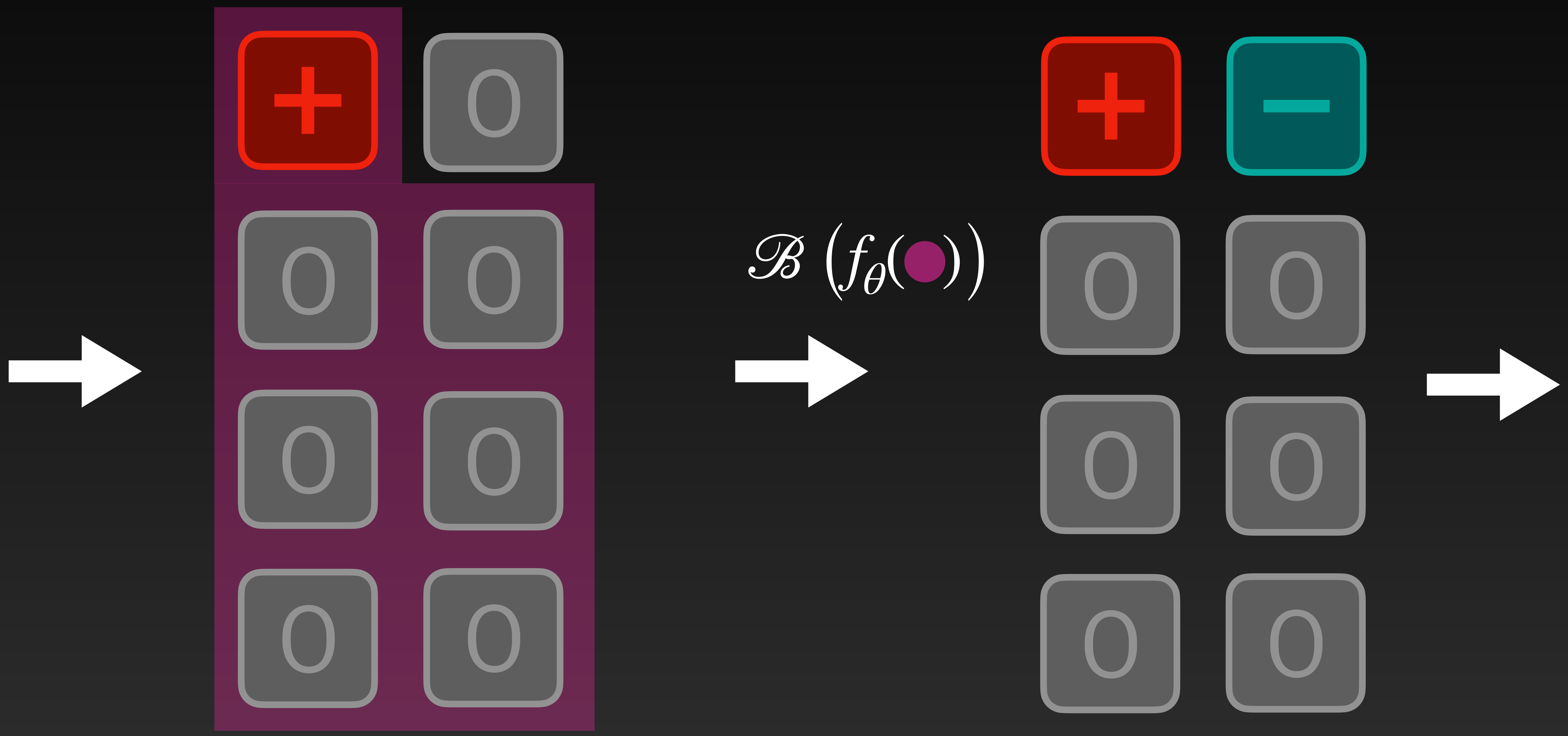
Generation of Discrete Fields



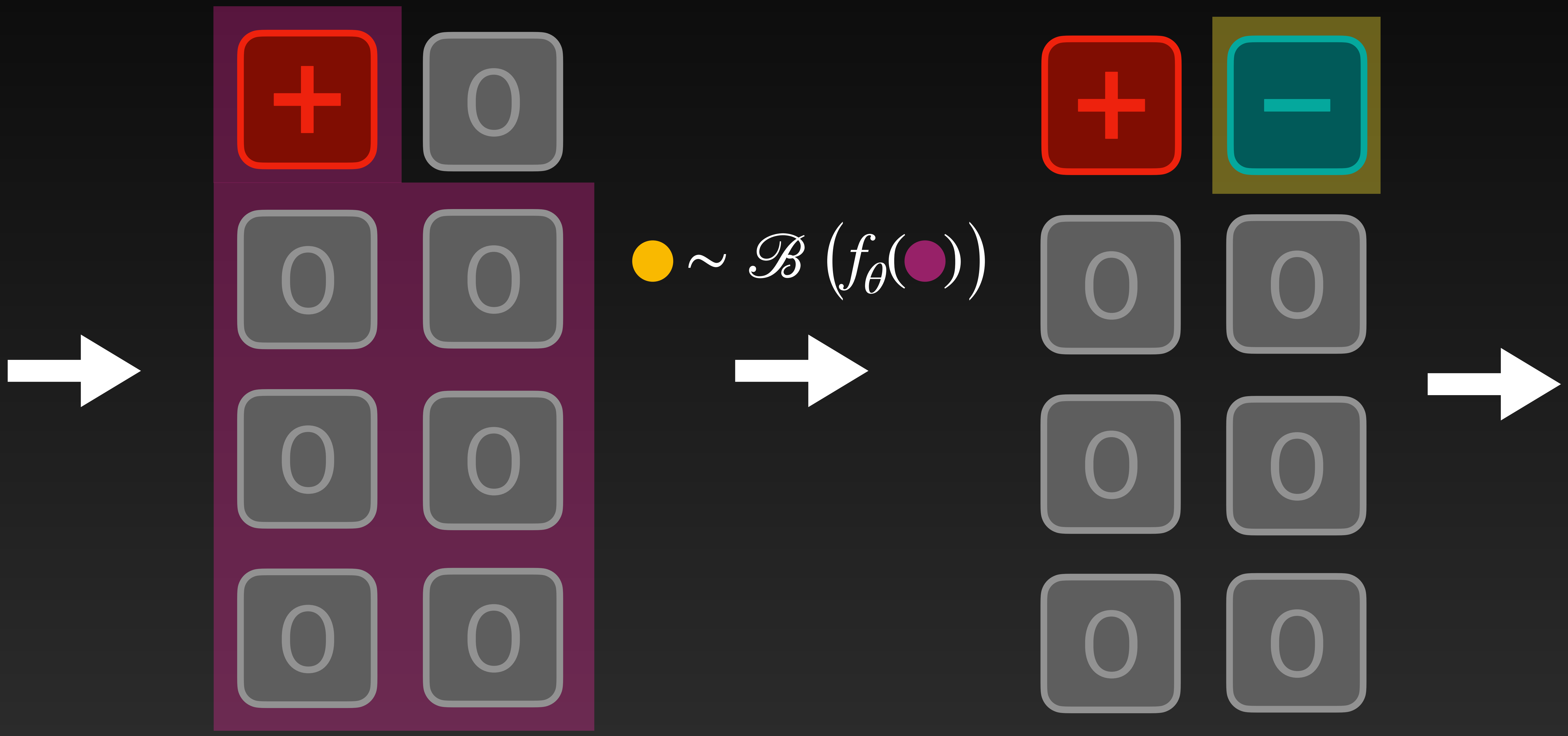
Generation of Discrete Fields



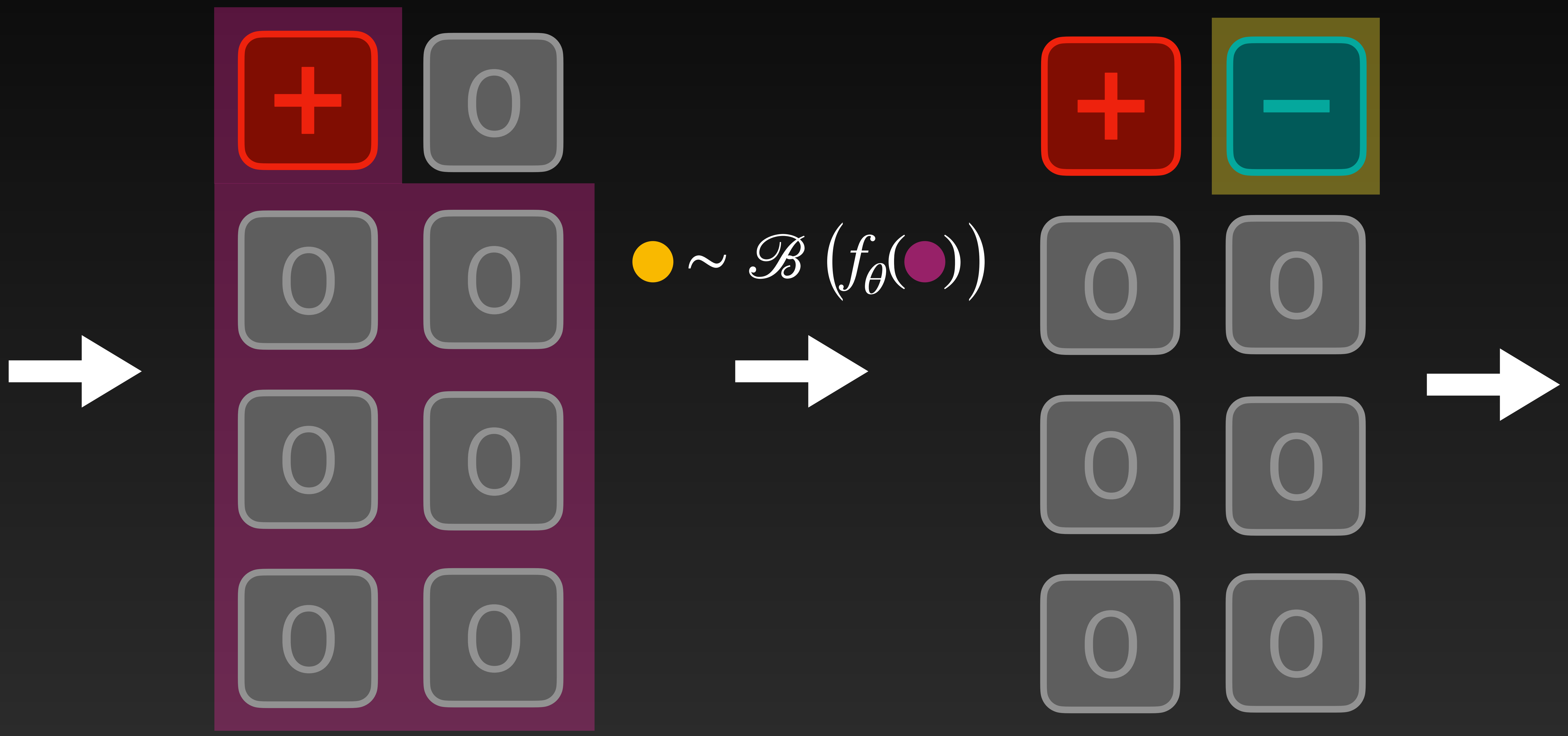
Generation of Discrete Fields



Generation of Discrete Fields

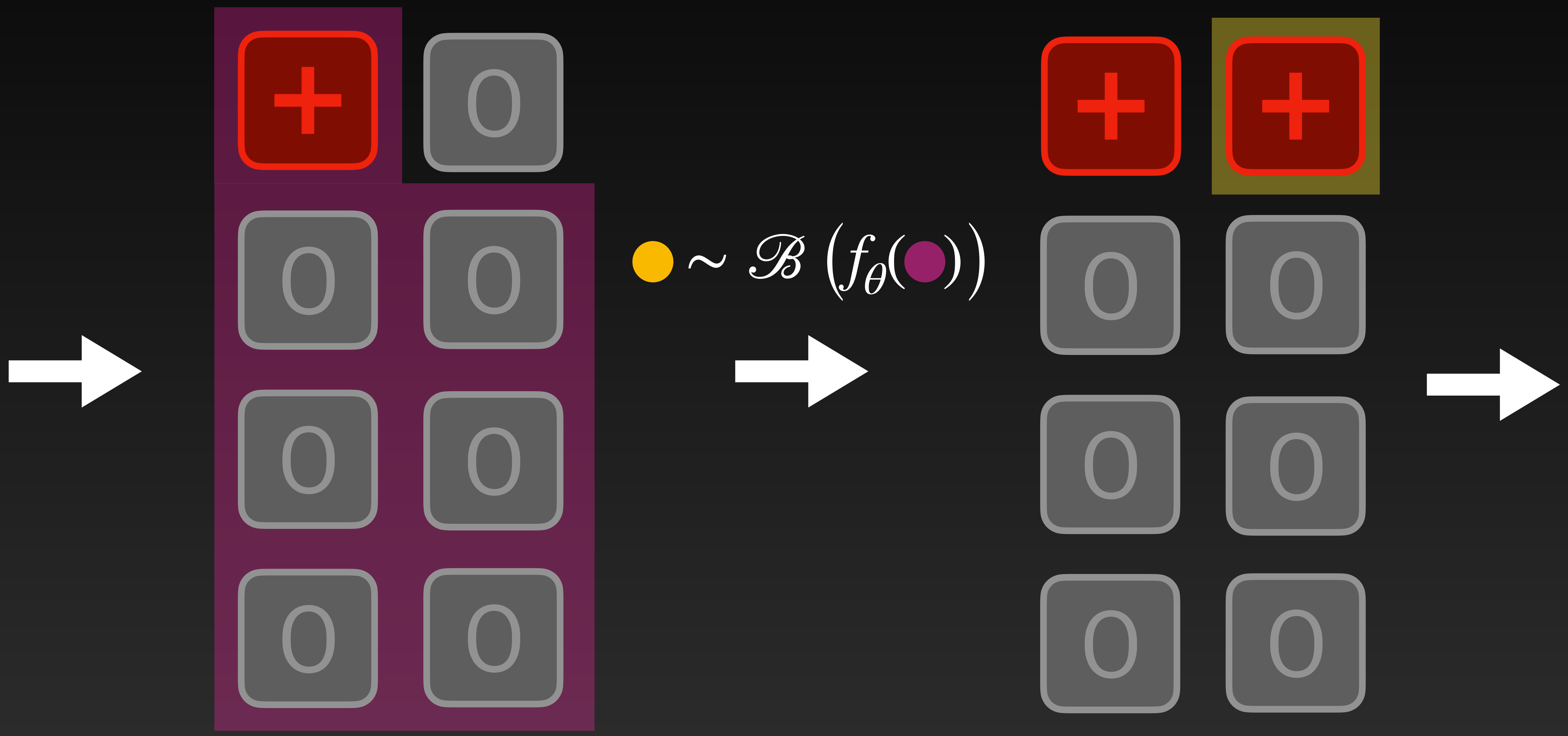


Generation of Discrete Fields



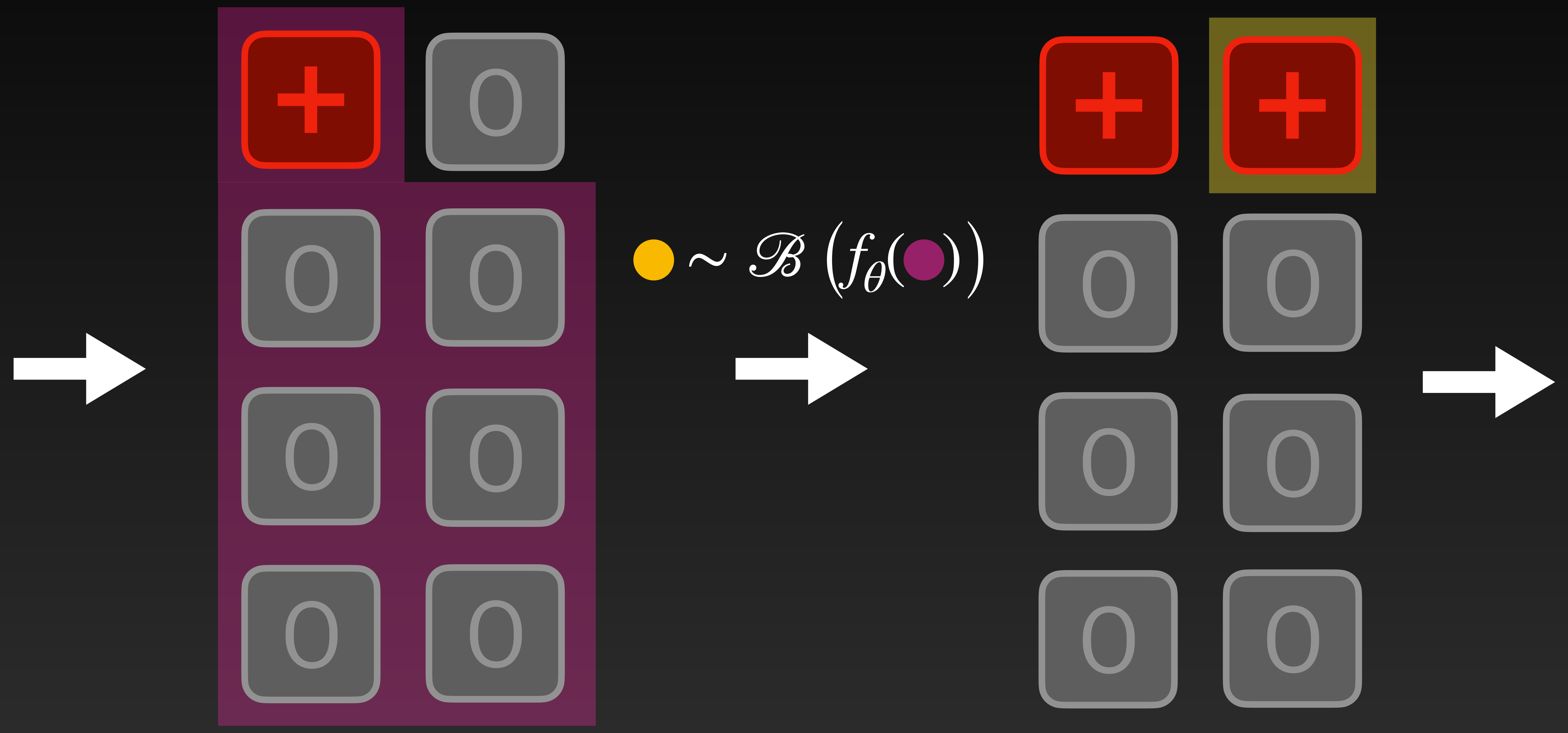
$$p(\bullet) = 1 - f_\theta(\bullet)$$

Generation of Discrete Fields



$$p(\bullet) = f_\theta(\bullet)$$

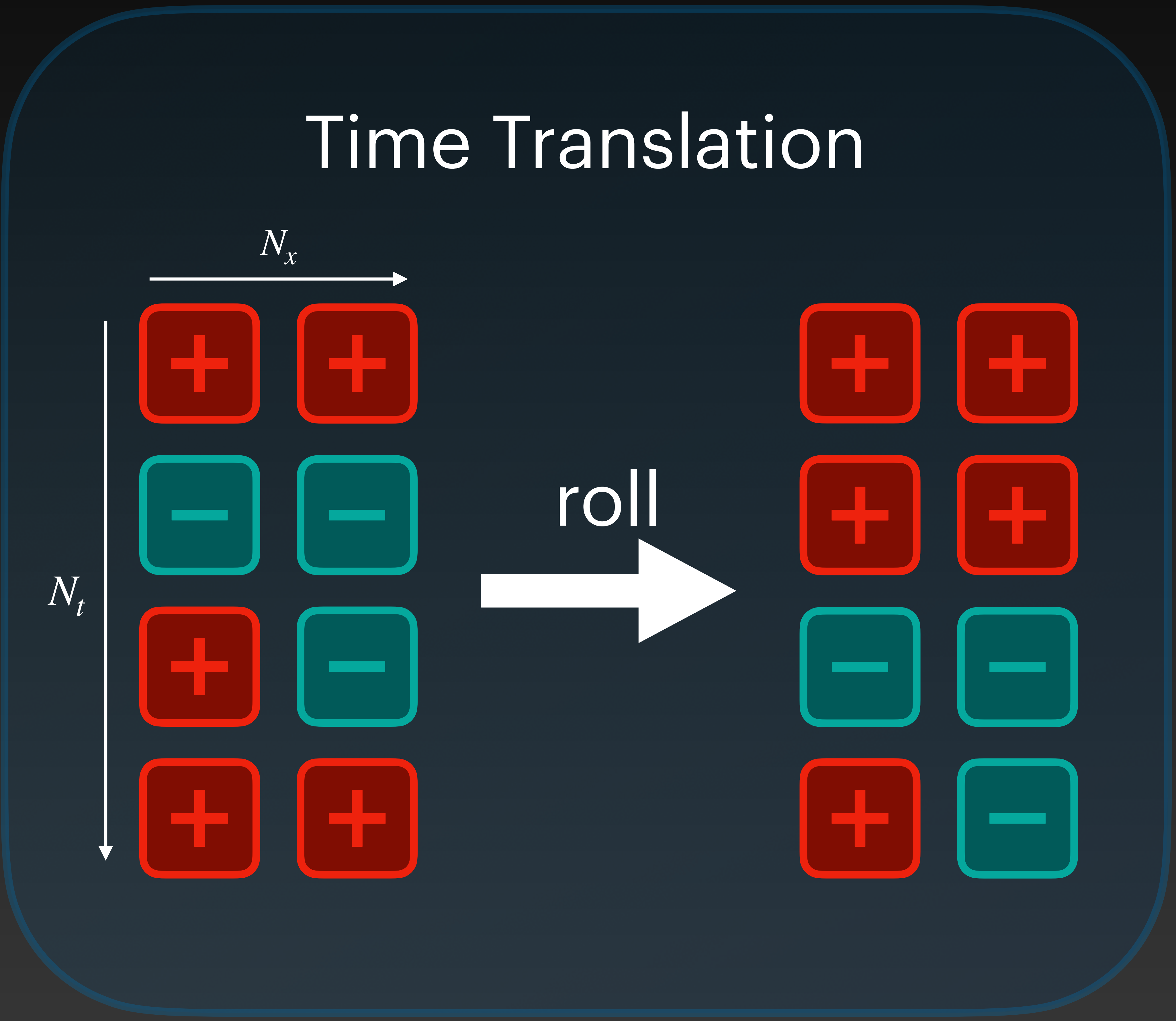
Generation of Discrete Fields



$$p(\bullet) = f_\theta(\bullet)$$

Accumulate p to do **reweighting**

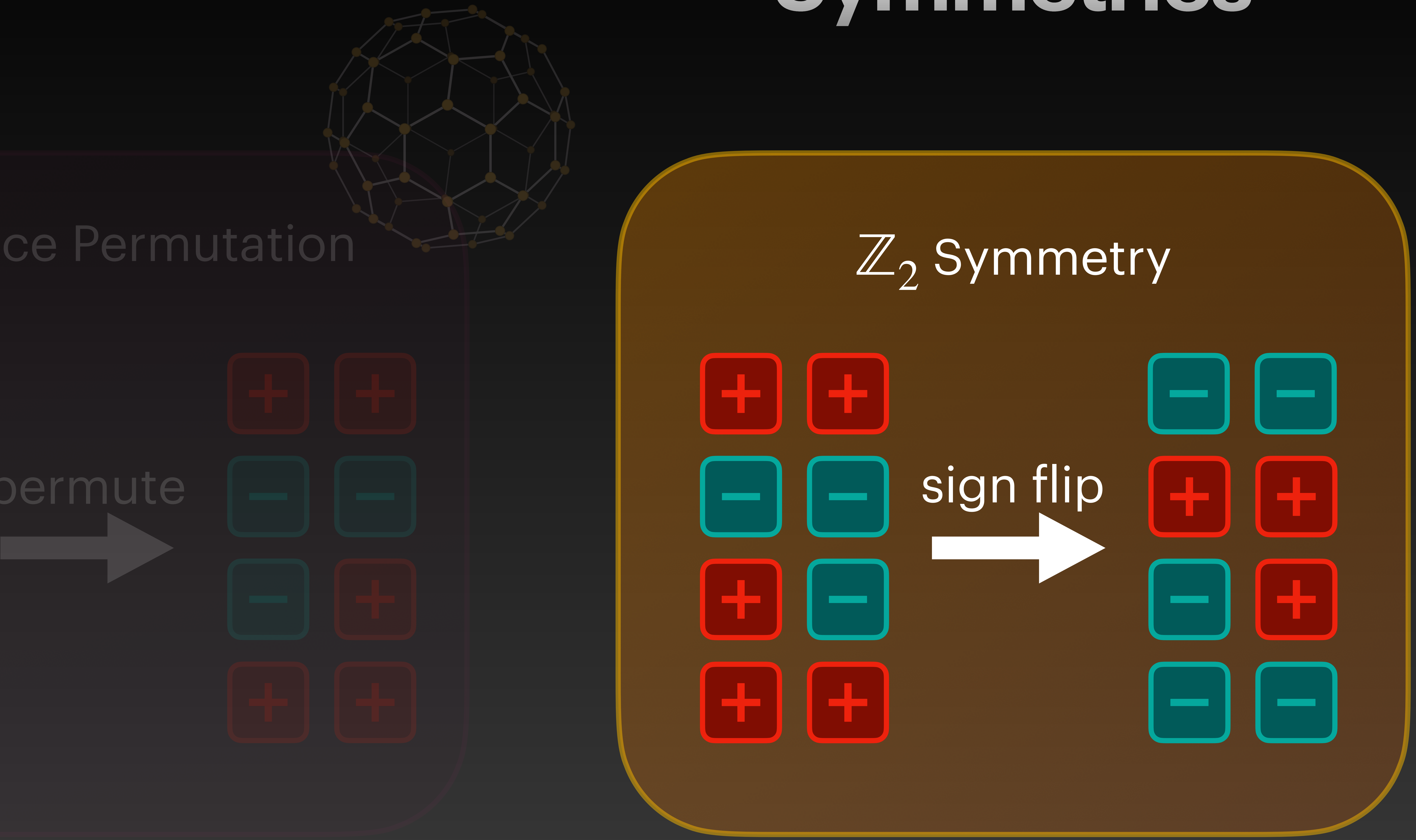
Symmetries



Symmetries

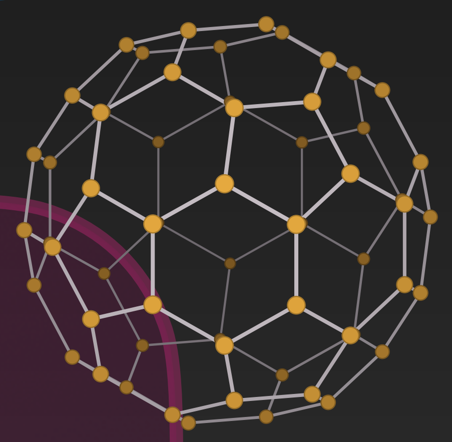
Translation
roll
Lattice Permutation
N_x
N_t
permute
$\mathbb{Z}_2$ Symm
sign f

Symmetries



Symmetries

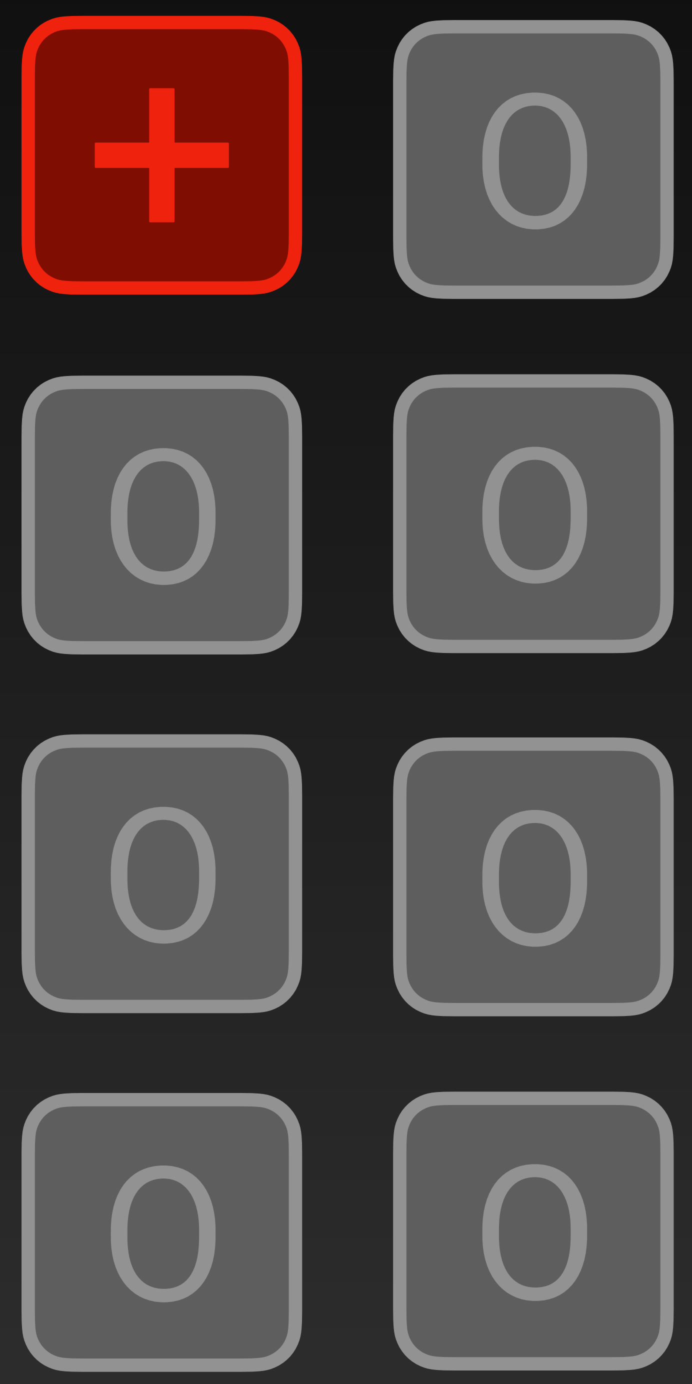
Time Translation



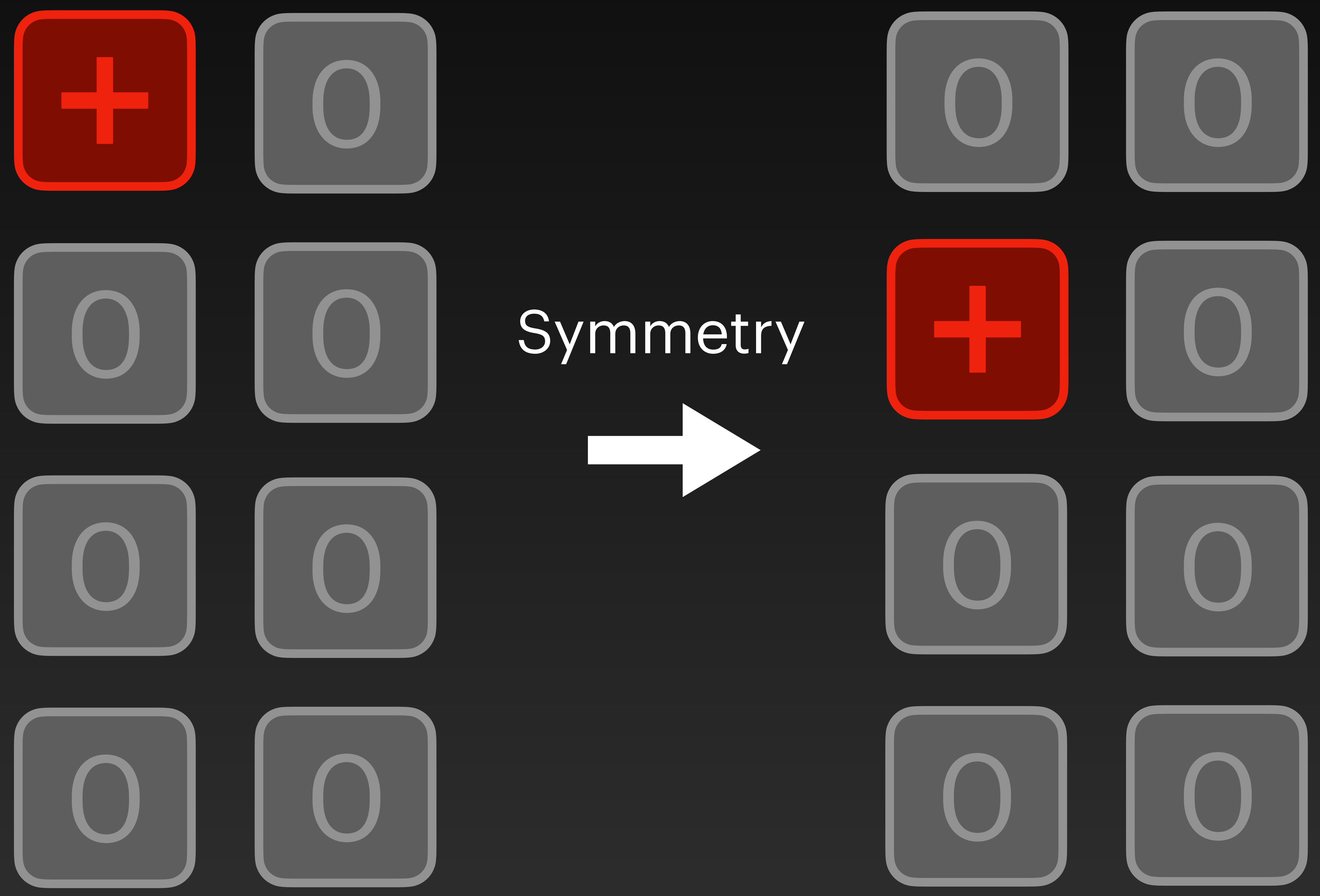
Lattice Permutation

$\mathbb{Z}_2$ Symmetry

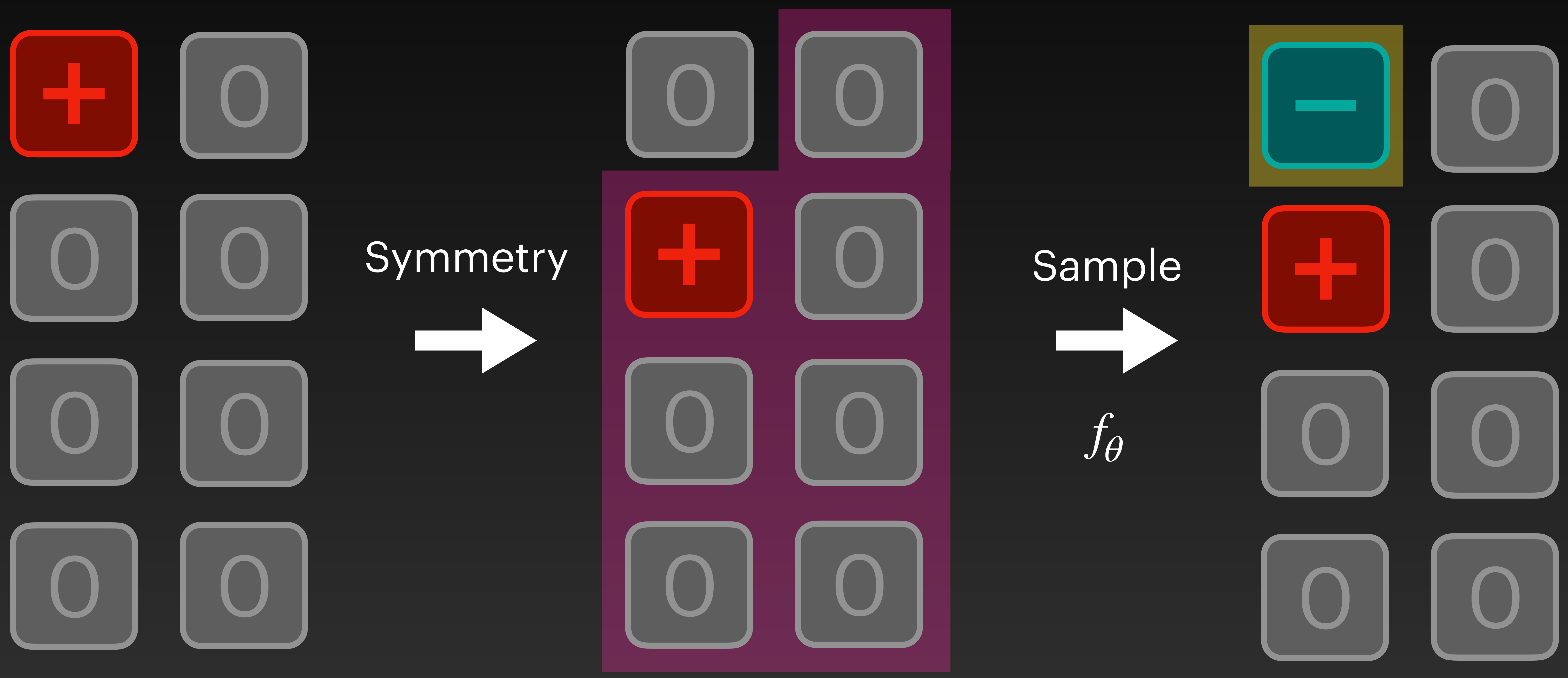
Symmetries



Symmetries

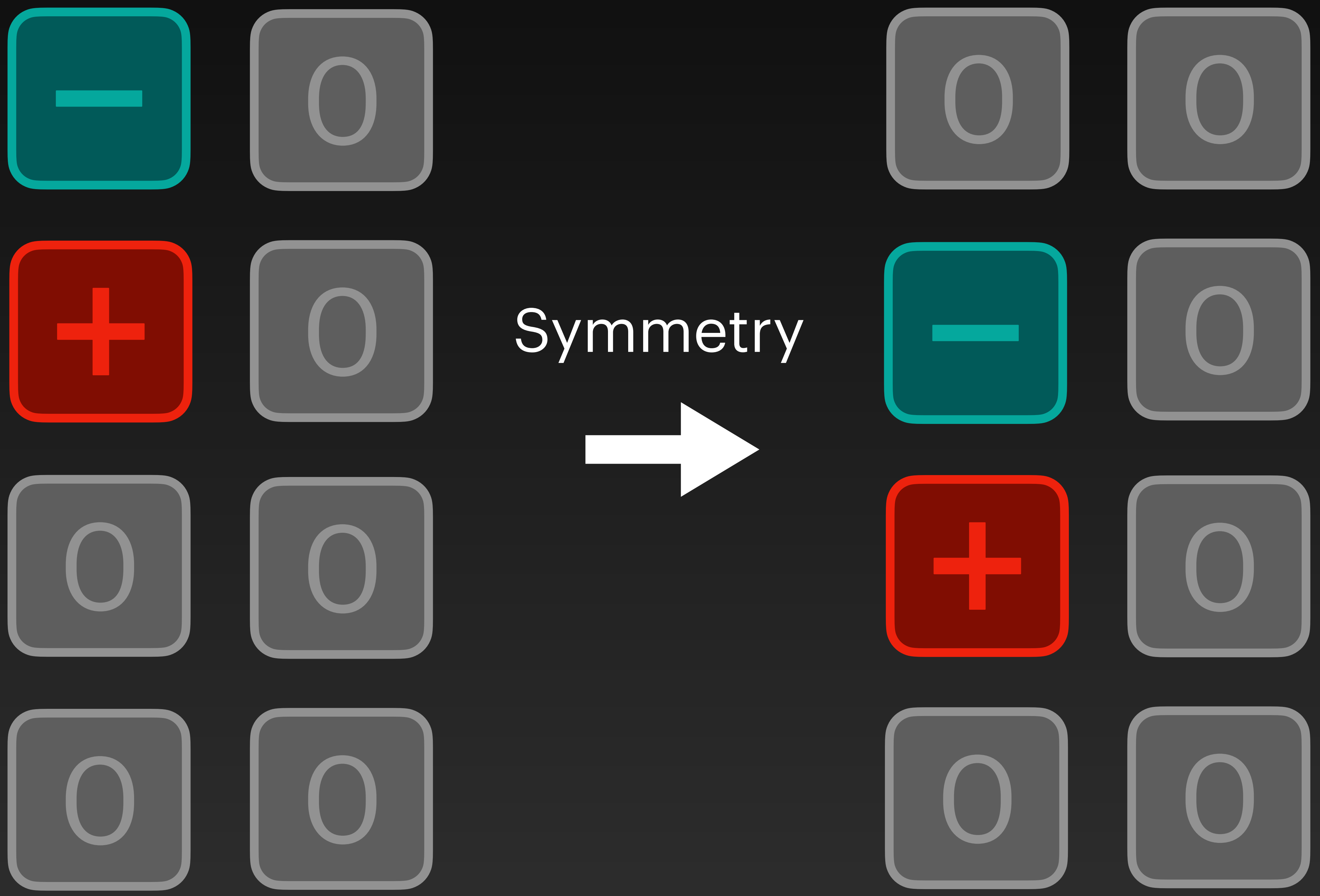


Symmetries

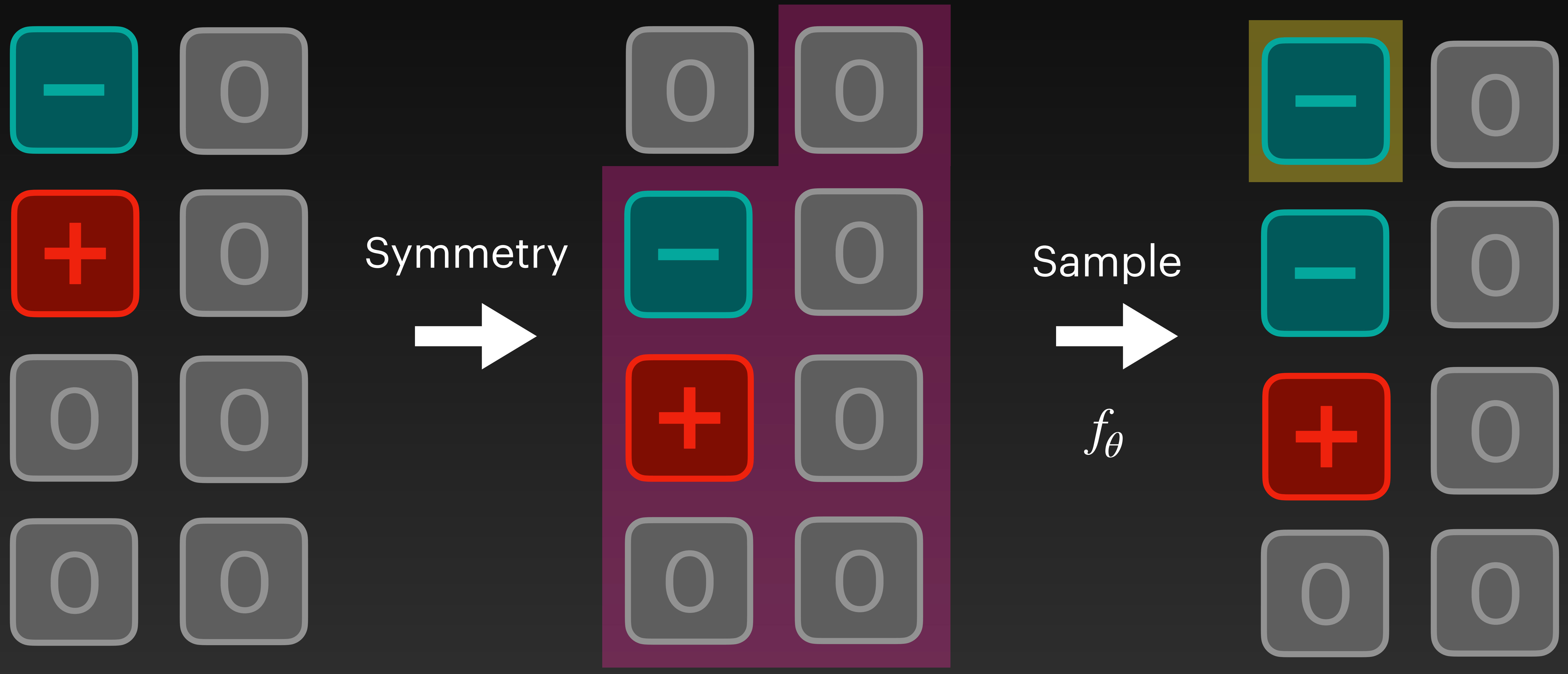


Symmetries

Symmetries

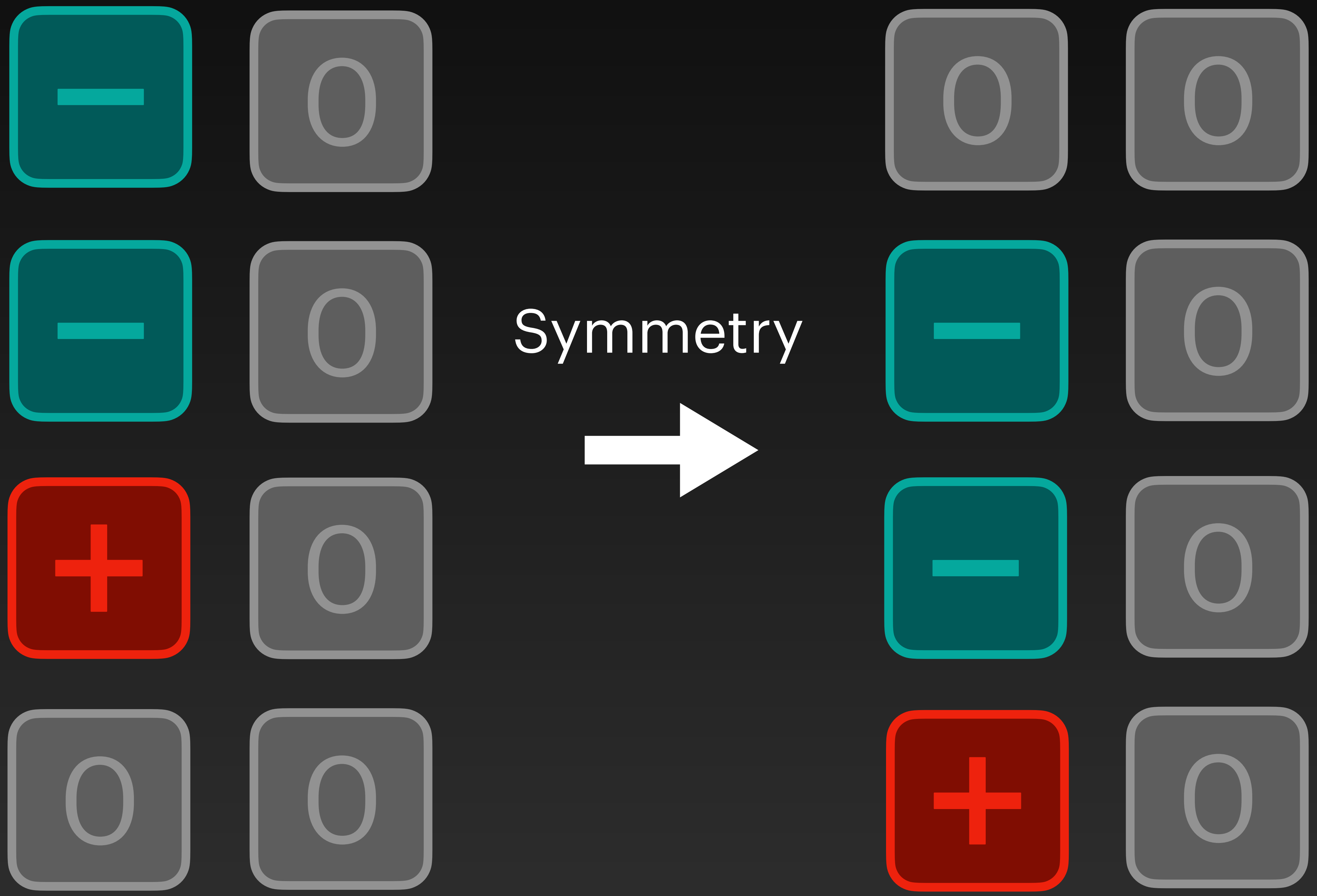


Symmetries

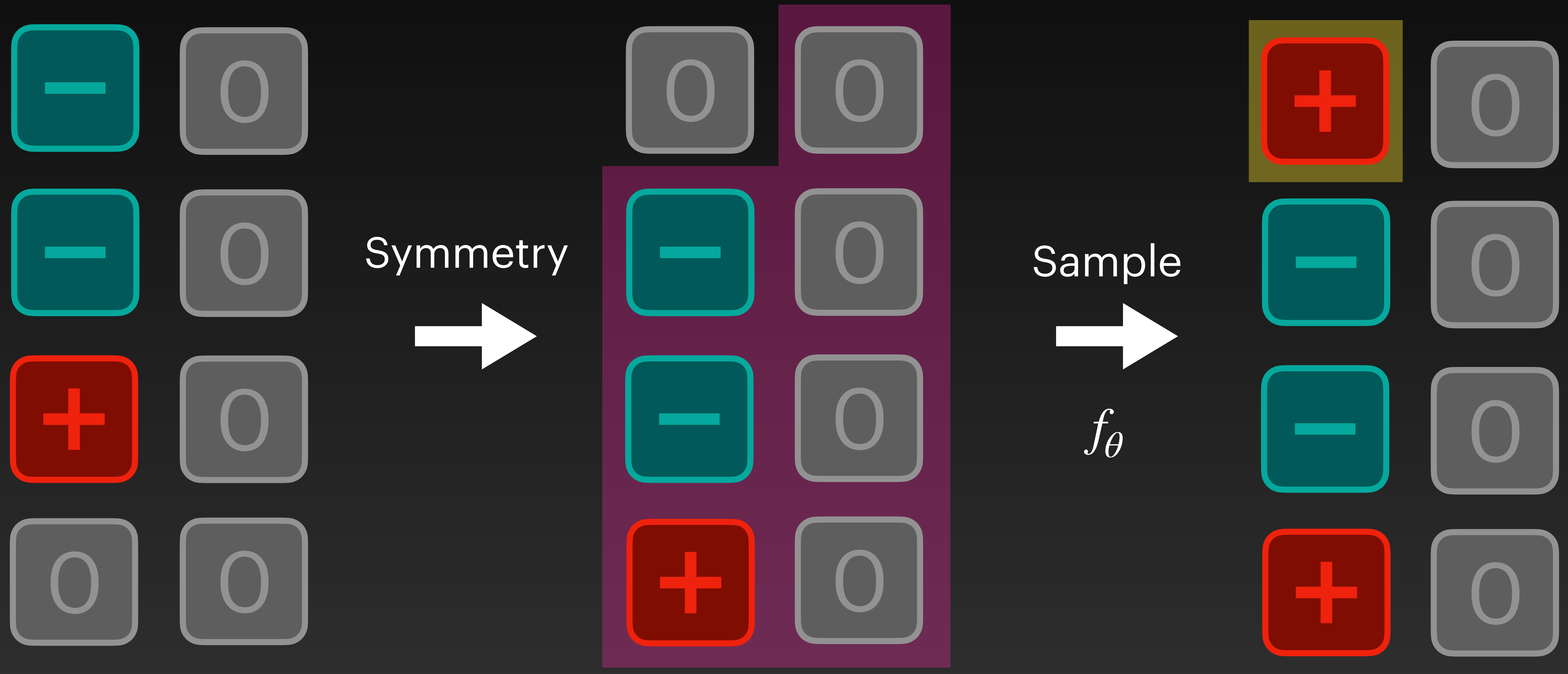


Symmetries

Symmetries



Symmetries



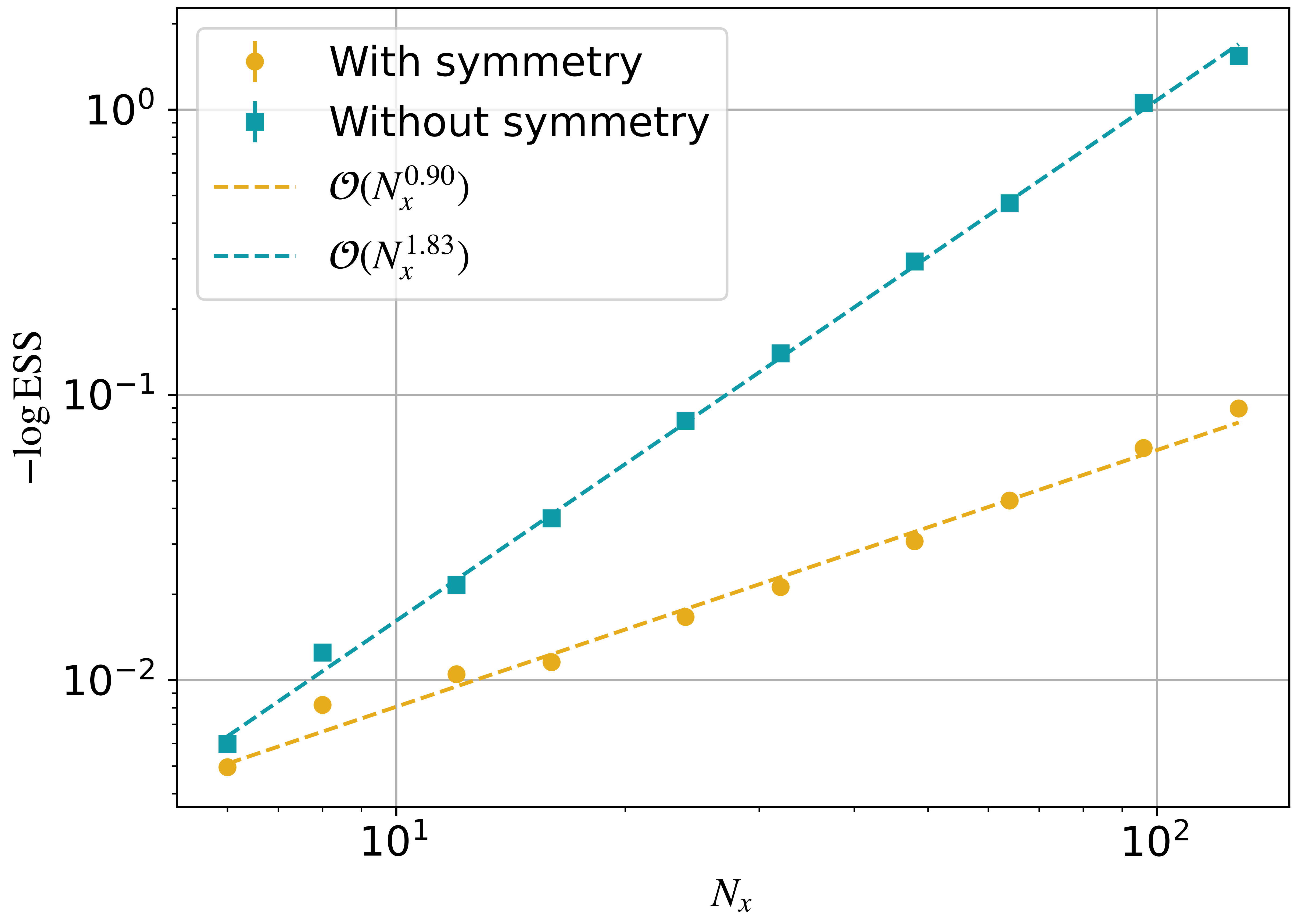
Results

Comparison

Symmetries

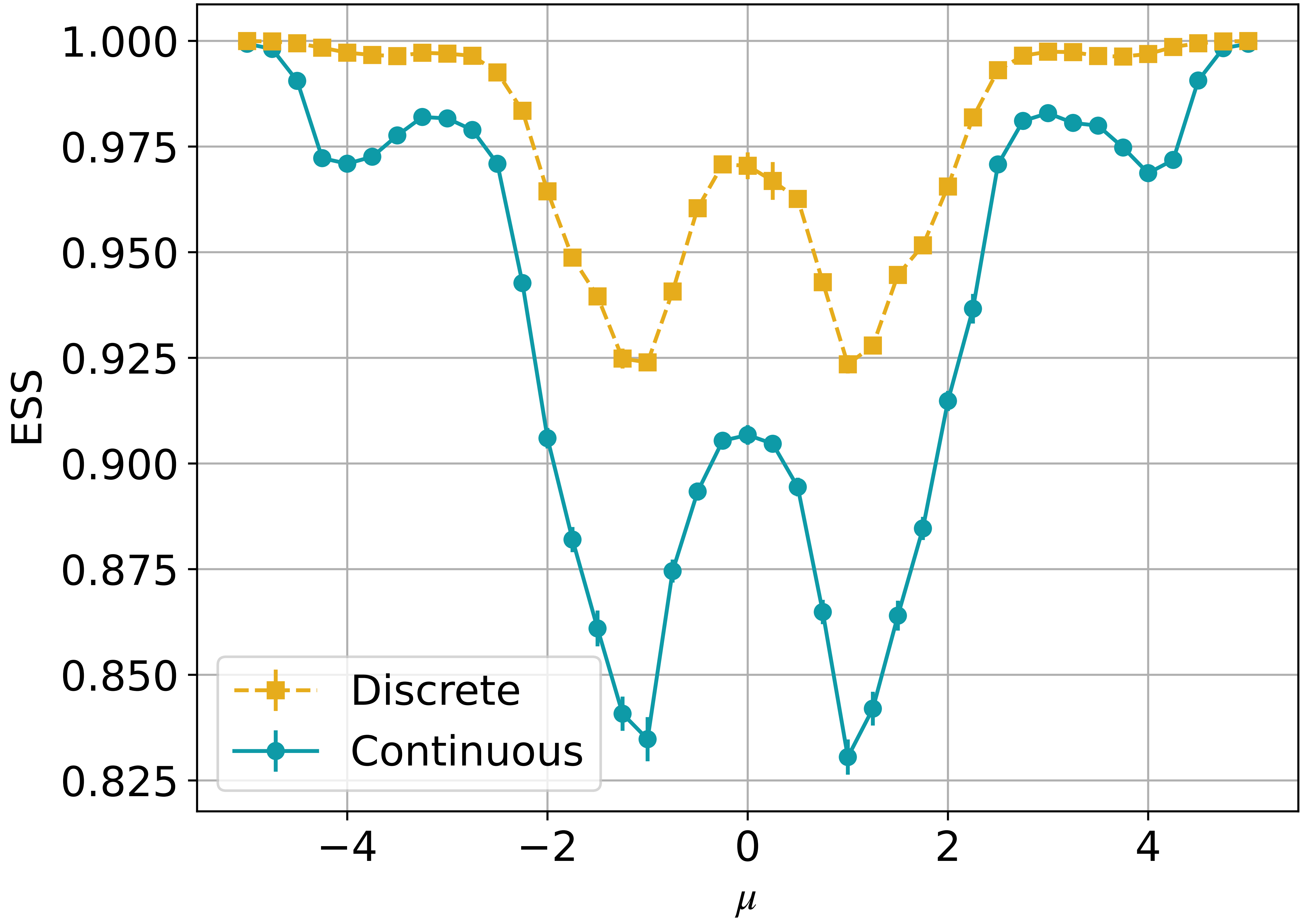
$\mathrm{ESS} \approx \exp\left(-kN_x^2\right)$

$\mathrm{ESS} \approx \exp\left(-kN_x\right)$



$U = 2,\ \beta = 2.5,\ N_t = 10$

Comparison
Continuous / Discrete



Discrete sampler improves training

$N_x = 8$, $U = 2$, $\beta = 5$, $N_t = 20$, all symmetries applied

Scaling

Computational scaling

?

Quality scaling

?

Scaling

Computational scaling

Sampling scaling
$\mathcal{O}(N_x^2 N_t^2)$

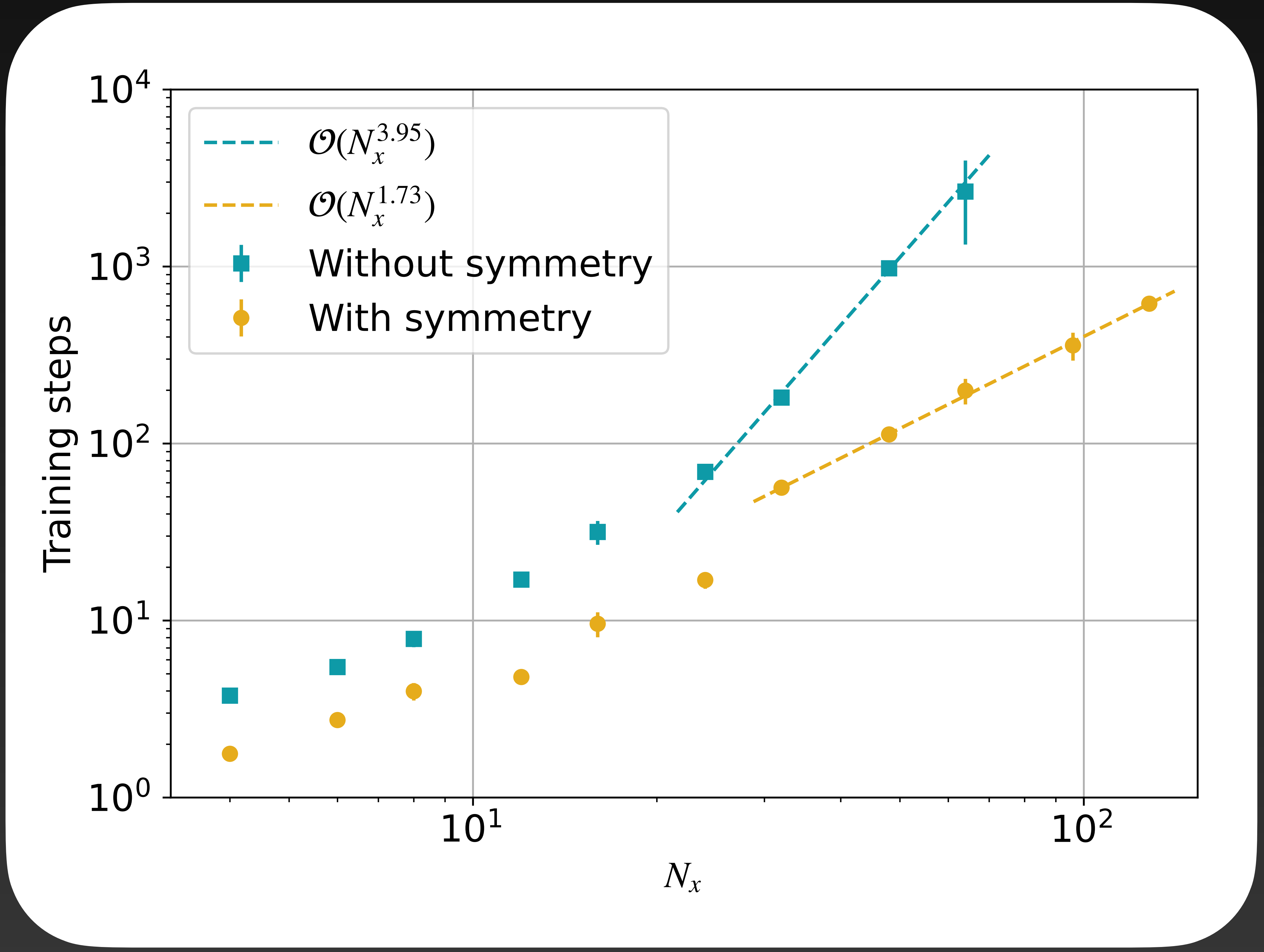
Action scaling
$\mathcal{O}(N_x^3 N_t)$

Quality s

?

Scaling

Computational scaling



Sampling scaling
$\mathcal{O}(N_x^2 N_t^2)$

Action scaling
$\mathcal{O}(N_x^3 N_t)$

How long do I need to train to reach 50% ESS ?

Combined scaling: $\mathcal{O}(N_x^{4.7})$

Quality s

?

$U = 2, \beta = 2.5, N_t = 10$

Scaling

ing

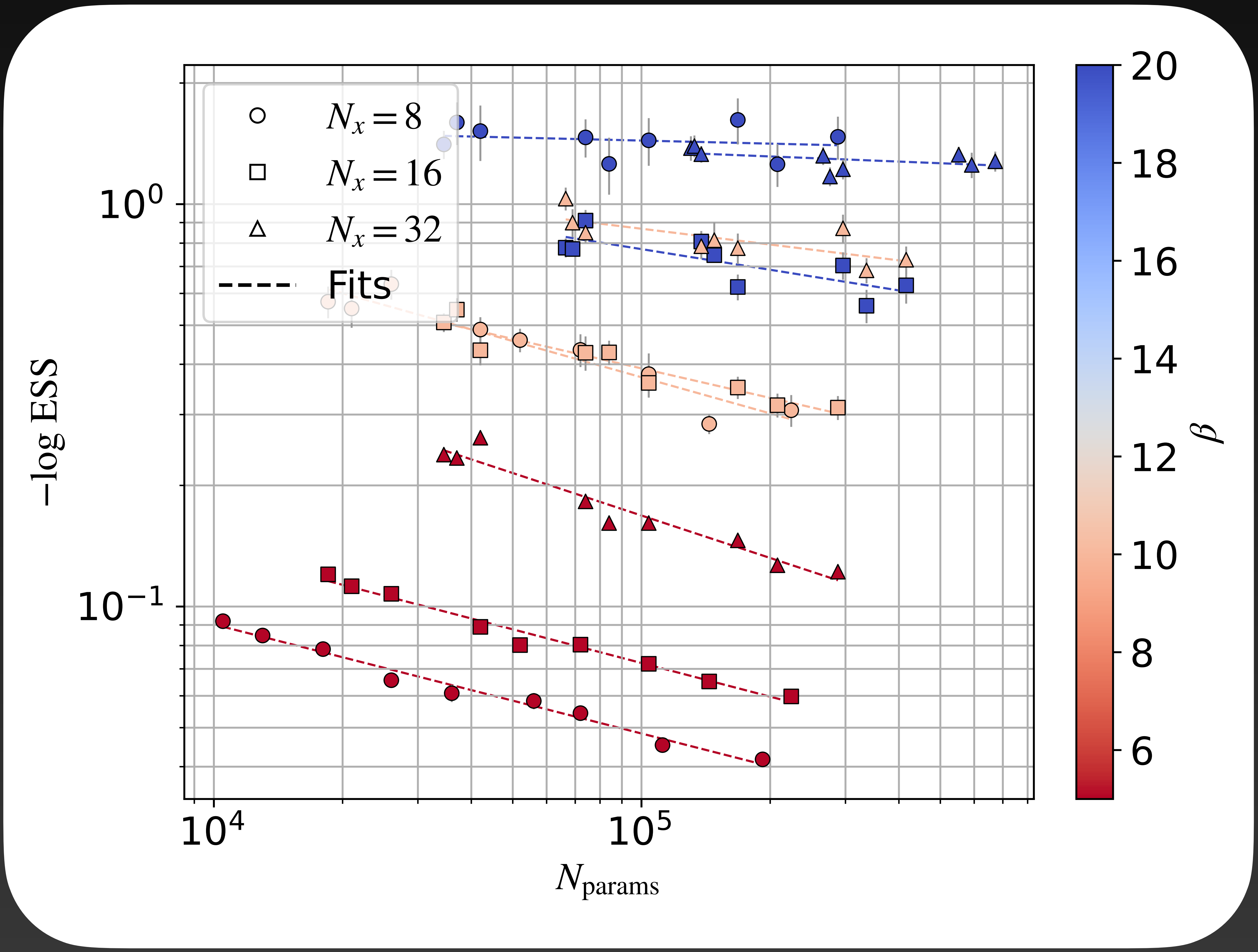
caling

ling

ed to train
ESS ?

g: $\mathcal{O}(N_x^{4.7})$

Quality scaling



$U = 2, N_t = 4\beta$

Scaling

ing

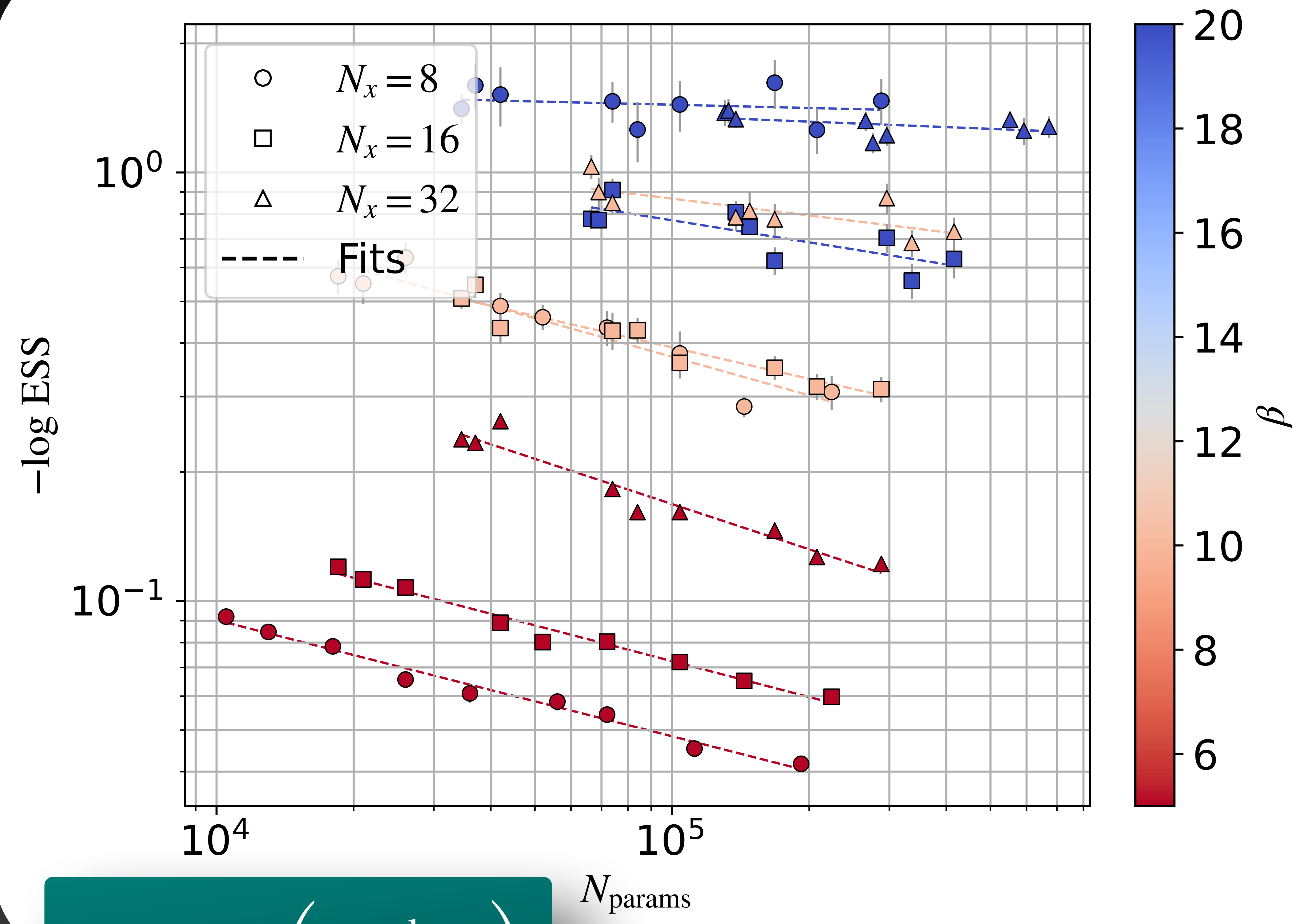
caling

ling

ed to train
ESS ?

g: $\mathcal{O}(N_x^{4.7})$

Quality scaling



$$\mathrm{ESS} \approx \exp\left(-\frac{\mathrm{k}}{N_{\mathrm{params}}^{0.35}}\right)$$

$U = 2,\ N_t = 4\beta$

Scaling

ing

ealing

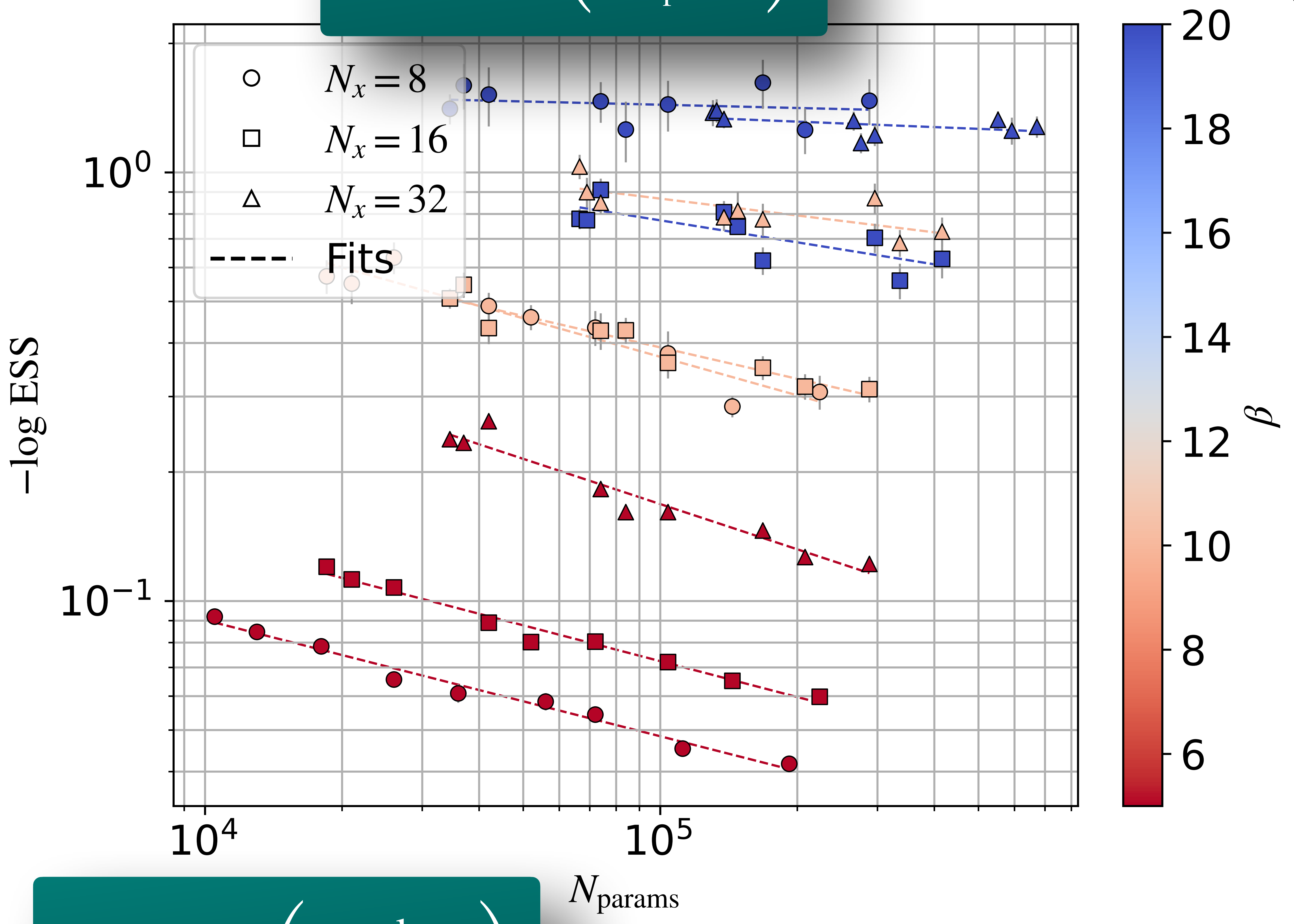
ling

ed to train
 ESS ?

g: $\mathcal{O}(N_x^{4.7})$

Quality scaling

$$\mathrm{ESS} \approx \exp\left(-\frac{\mathrm{k}}{N_{\mathrm{params}}^{0.02}}\right)$$



$$\mathrm{ESS} \approx \exp\left(-\frac{\mathrm{k}}{N_{\mathrm{params}}^{0.35}}\right)$$

$U = 2, N_t = 4\beta$

Scaling

ing

ealing

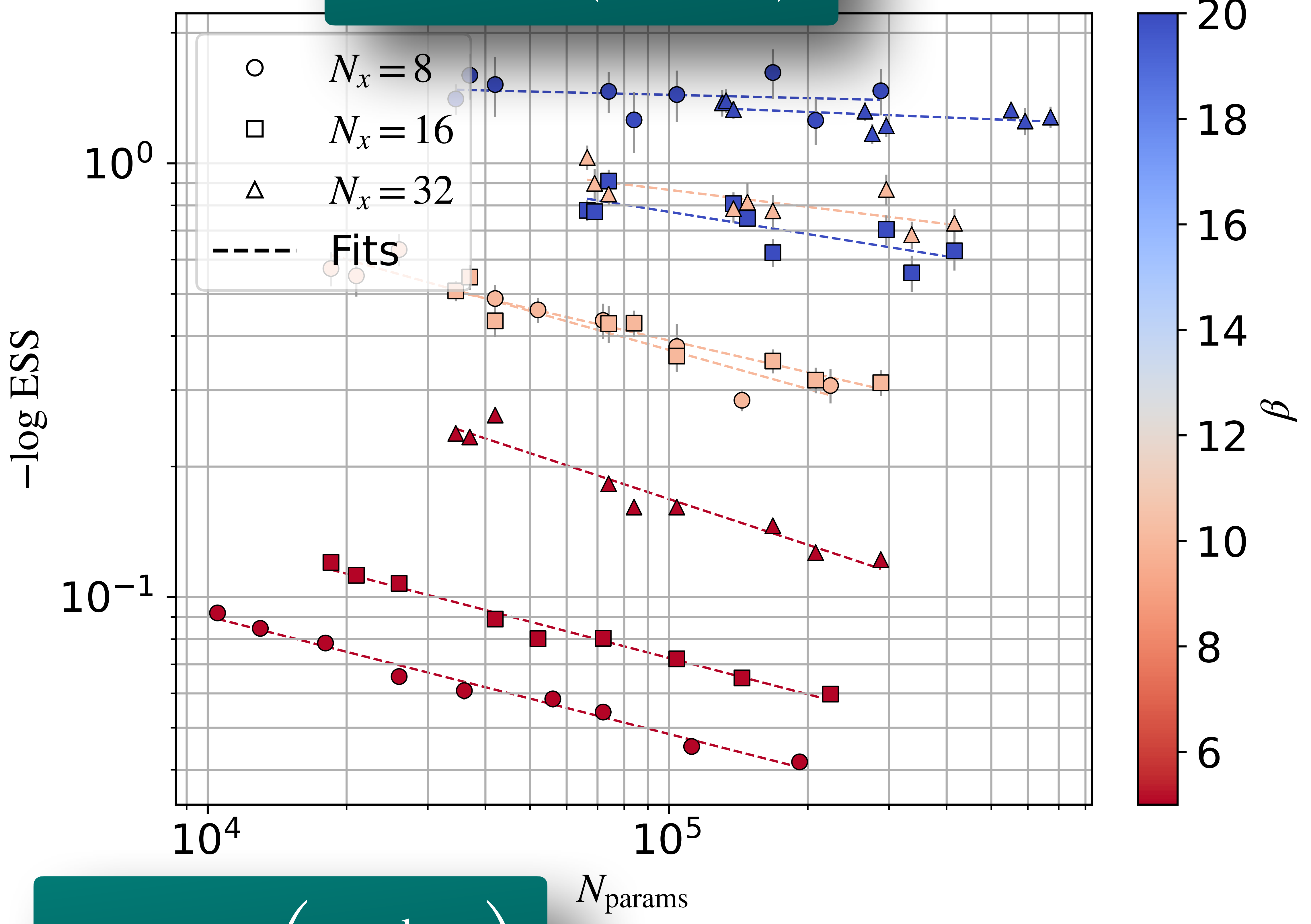
ling

ed to train
ESS ?

g: $\mathcal{O}(N_x^{4.7})$

Quality scaling

$$\mathrm{ESS} \approx \exp\left(-\frac{\mathrm{k}}{N_{\mathrm{params}}^{0.02}}\right)$$



$$\mathrm{ESS} \approx \exp\left(-\frac{\mathrm{k}}{N_{\mathrm{params}}^{0.35}}\right)$$

Very mild scaling with N_{params}

Scaling vanishes at high β

$U = 2,\ N_t = 4\beta$

Scaling

Computational scaling

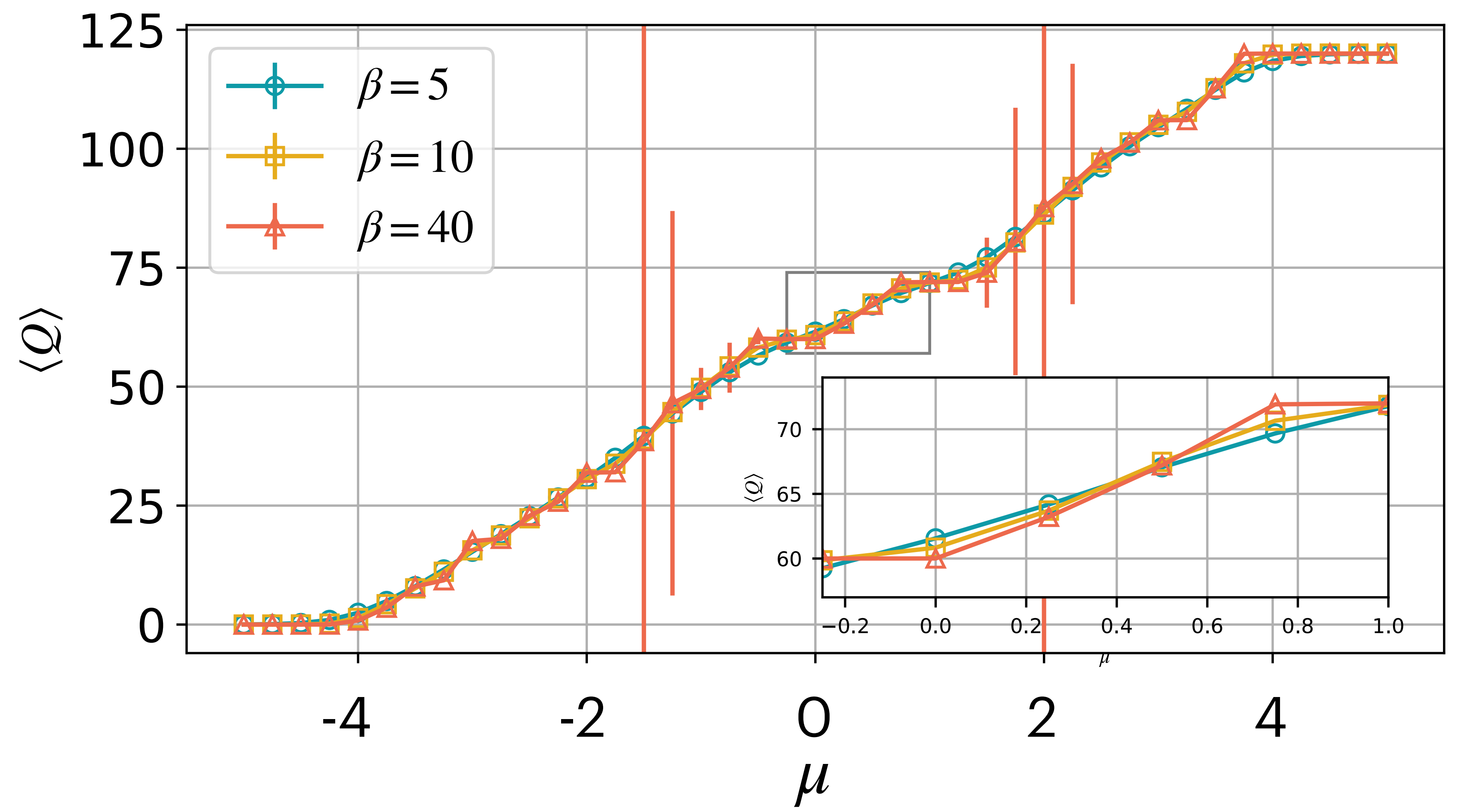
Training time

Combined scaling: $\mathcal{O}(N_x^{4.7})$

Quality scaling

Training quality

Not growing with N_{params}

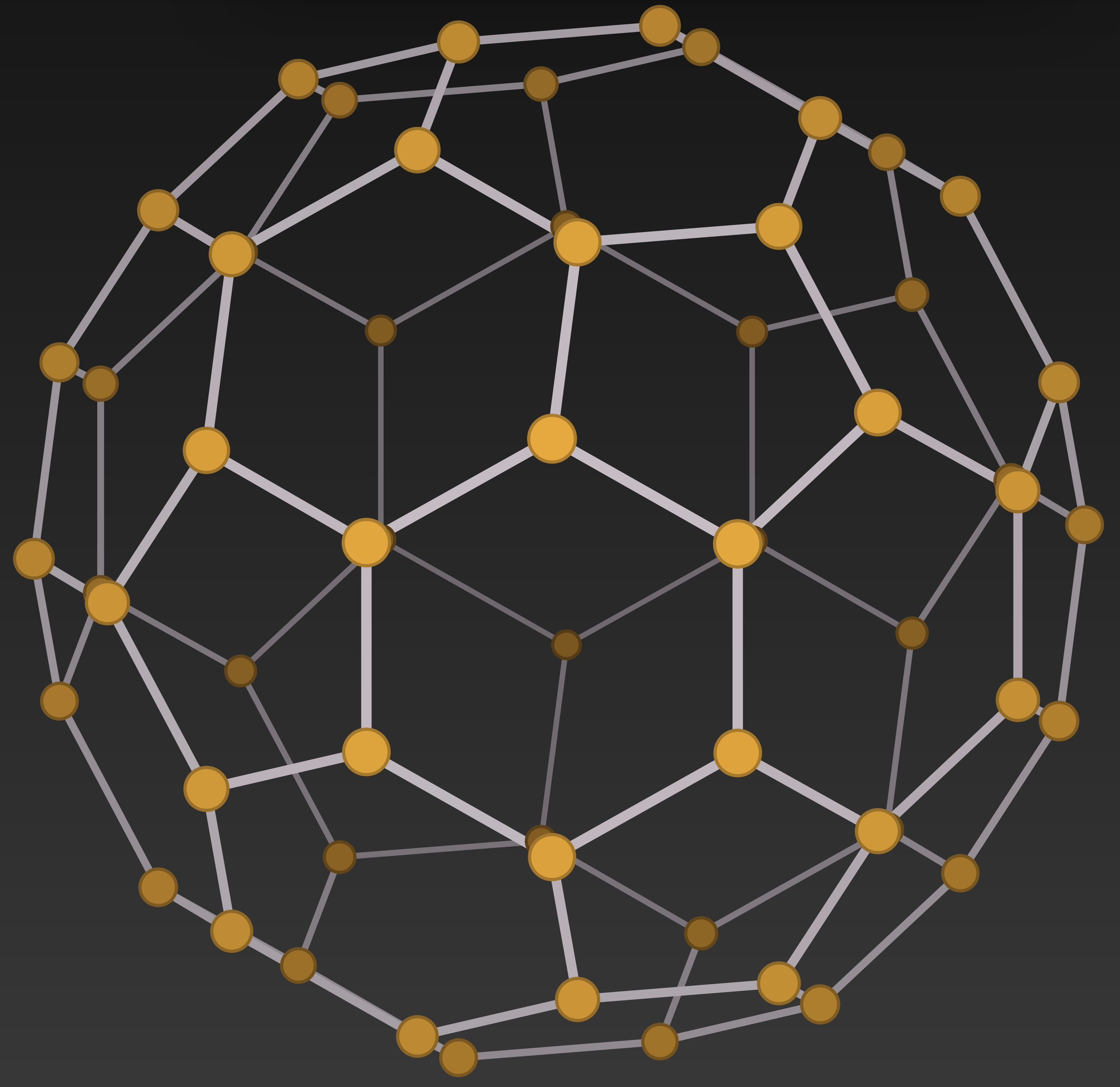


Charge Observable

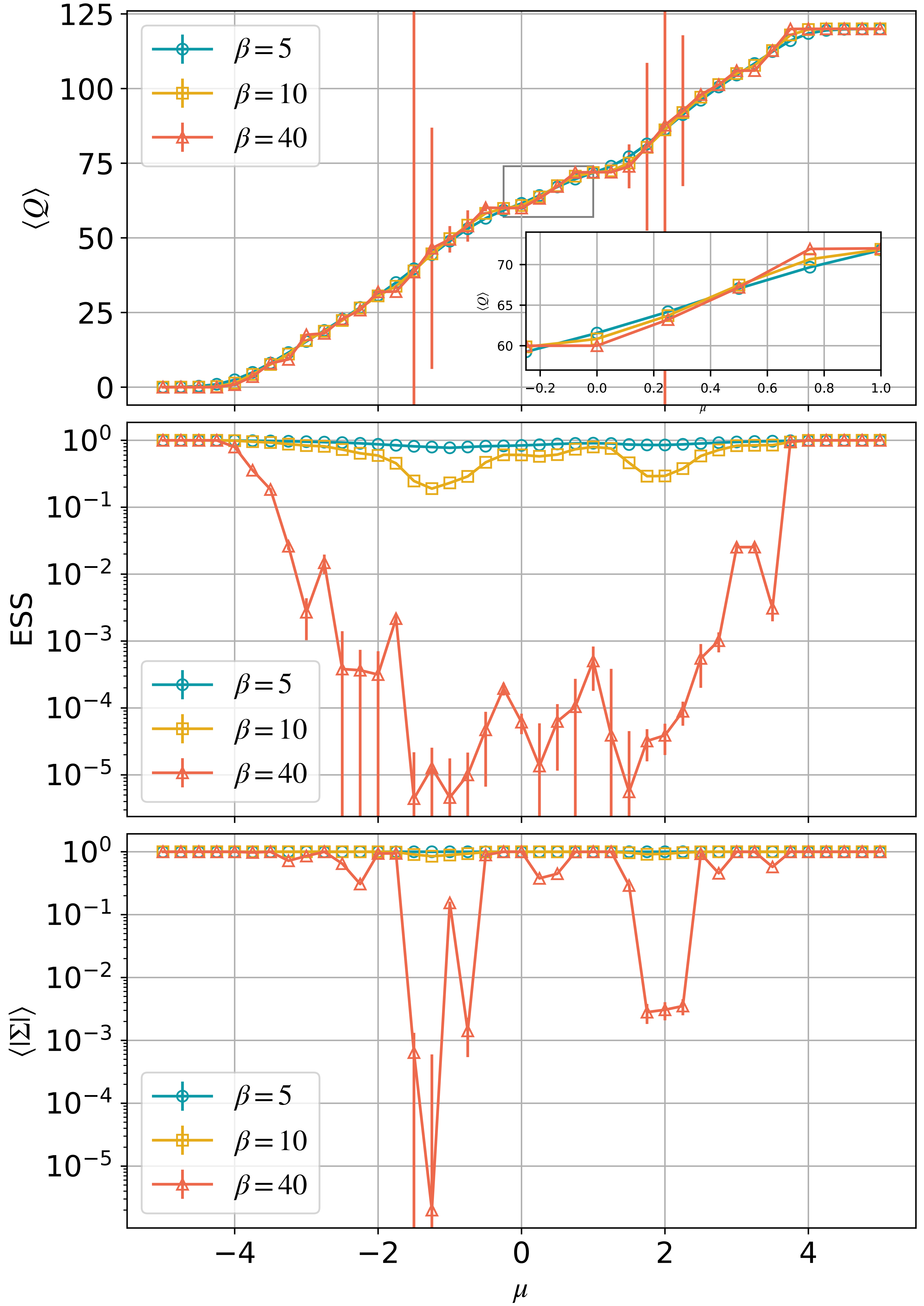
C60 fullerene

60 physical sites

$U = 2 \qquad N_t = 4\beta$



1M samples

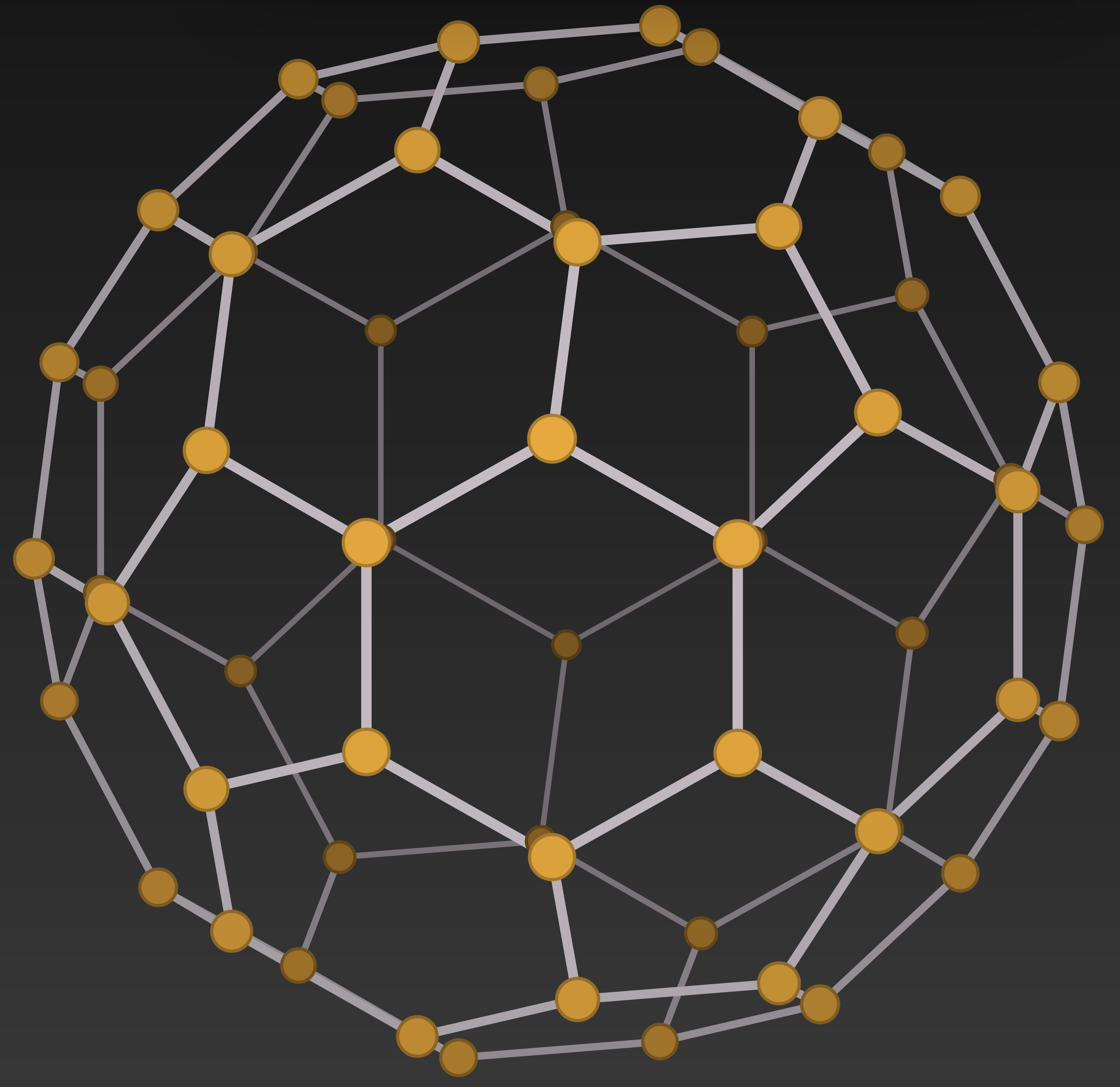


Charge Observable

C60 fullerene

60 physical sites

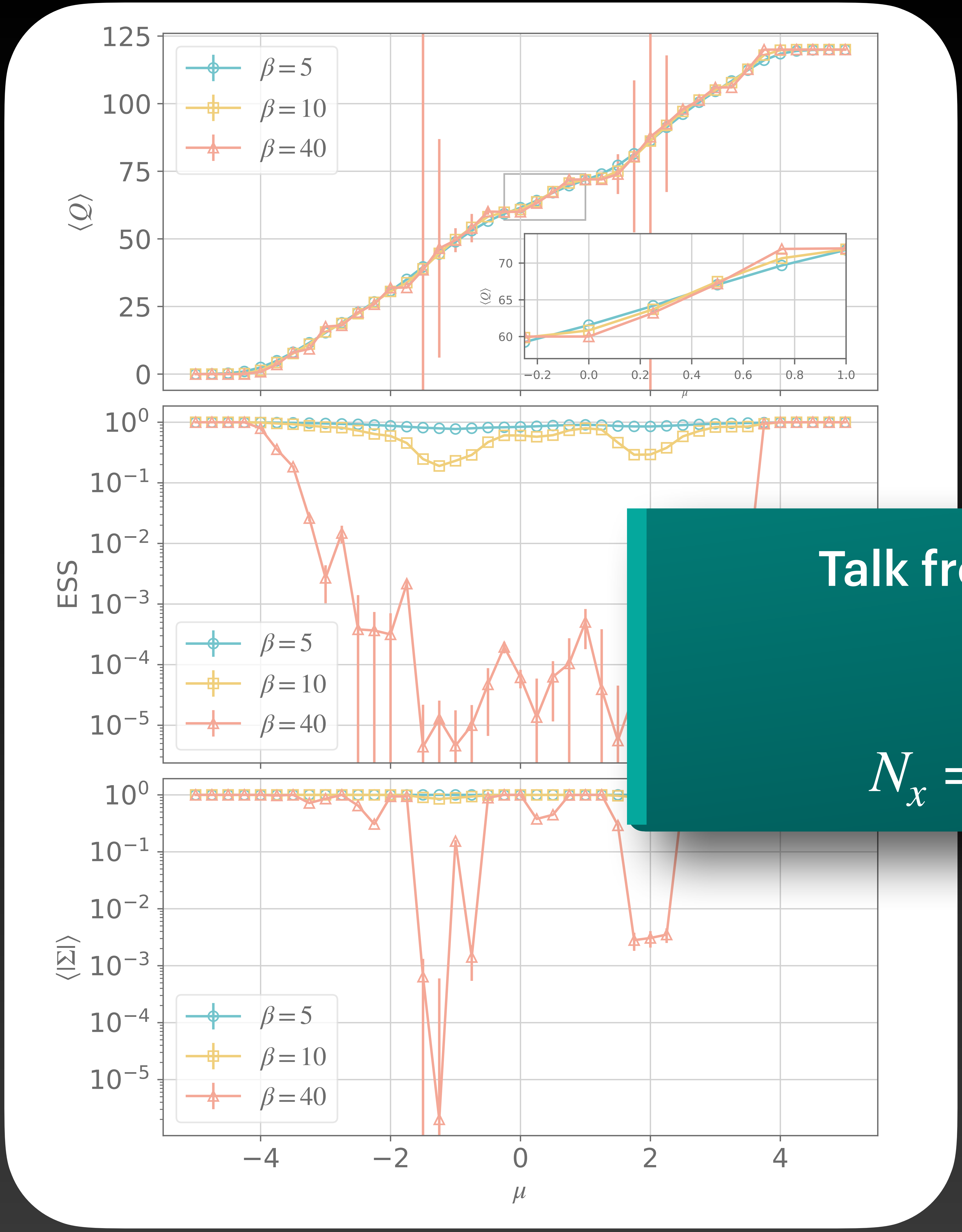
$U = 2 \qquad N_t = 4\beta$



1M samples

Charge Observable

C60 fullerene



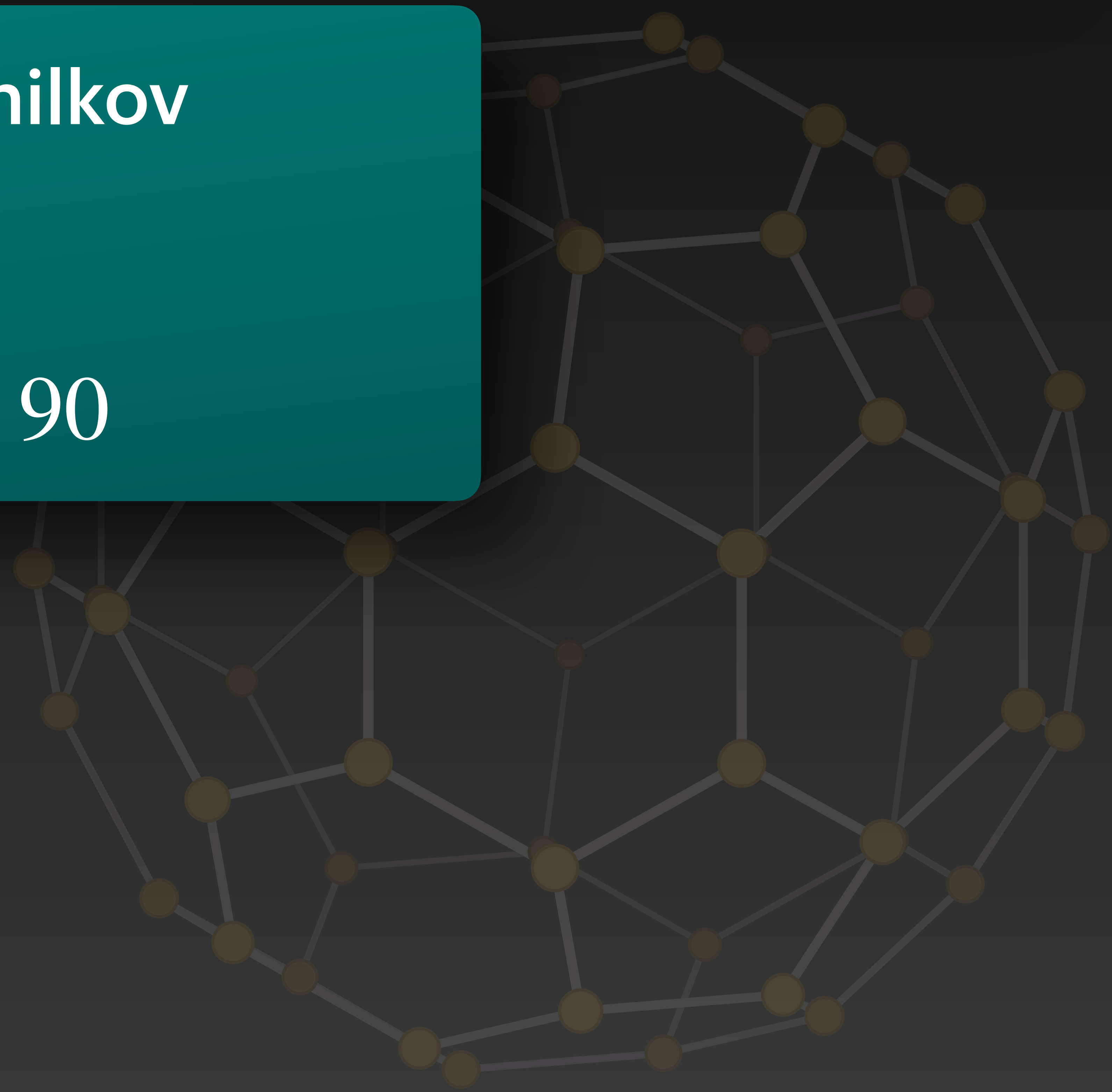
60 physical sites

$U = 2 \quad N_t = 4\beta$

Talk from Petar Sinilkov

HMC at

$$N_x = 20 \qquad \beta = 90$$



1M samples

Summary

Sampling the discrete Hubbard model with machine learning

- Lattice size of up to $N_x \times N_t = 60 \times 160 = 9{,}600$
- Reached inverse temperature $\beta = 40$

- High scaling in training: $\mathcal{O}(N_x^{4.7})$
- Very mild scaling with N_{params}
- Low ESS with high β

Summary

Sampling the discrete Hubbard model with machine learning

- Lattice size of up to $N_x \times N_t = 60 \times 160 = 9{,}600$
- Reached inverse temperature $\beta = 40$

- High scaling in training: $\mathcal{O}(N_x^{4.7})$
- Very mild scaling with N_{params}
- Low ESS with high β

Future direction

Globally update discrete fields $\mathcal{O}(N_x^2 N_t^2) \to \mathcal{O}(N_x N_t)$

Use non-equilibrium dynamics to introduce scaling to higher ESS

Try to simulate Cuprates and Hubbard ladders at low temperature

Thank you!

Lena Funcke

Dominic Schuh

LATTICE 2026

⟨NuMeriQS⟩

HISKP

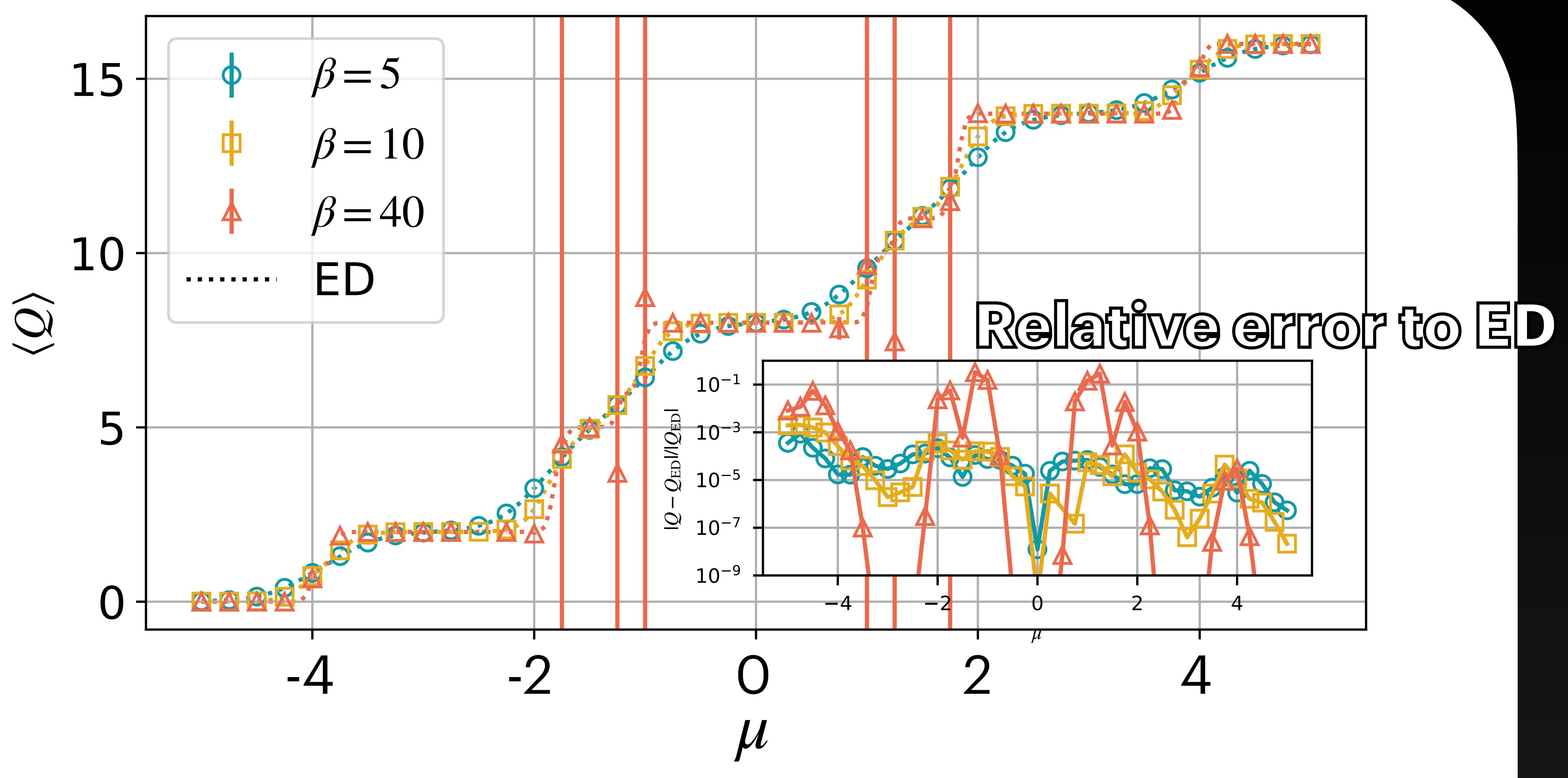
UNIVERSITÄT BONN

RHEINISCHE FRIEDRICH-WILHELMS-UNIVERSITÄT BONN

TRA MATTER

TRANSDISCIPLINARY RESEARCH AREA

Backup

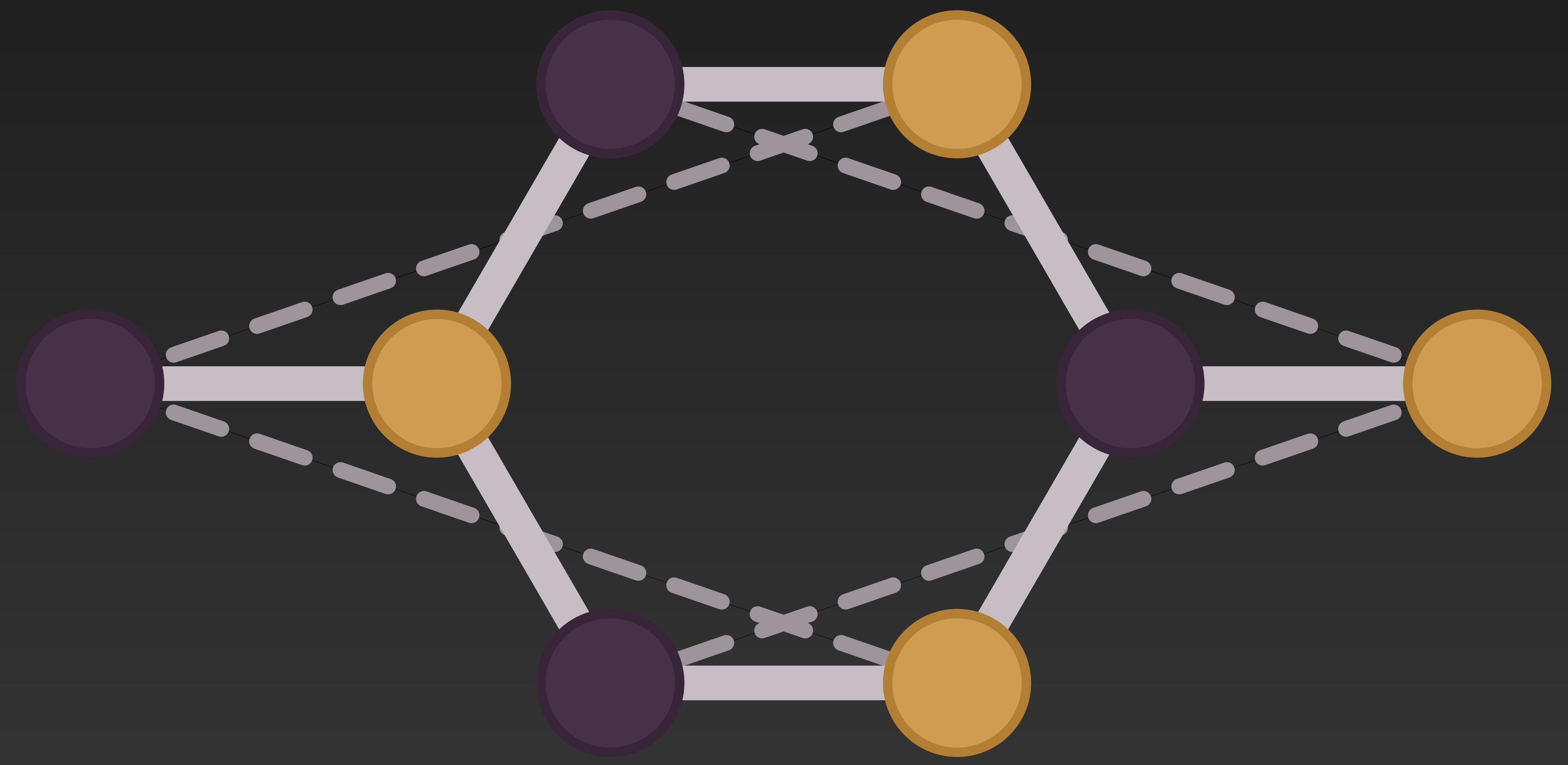


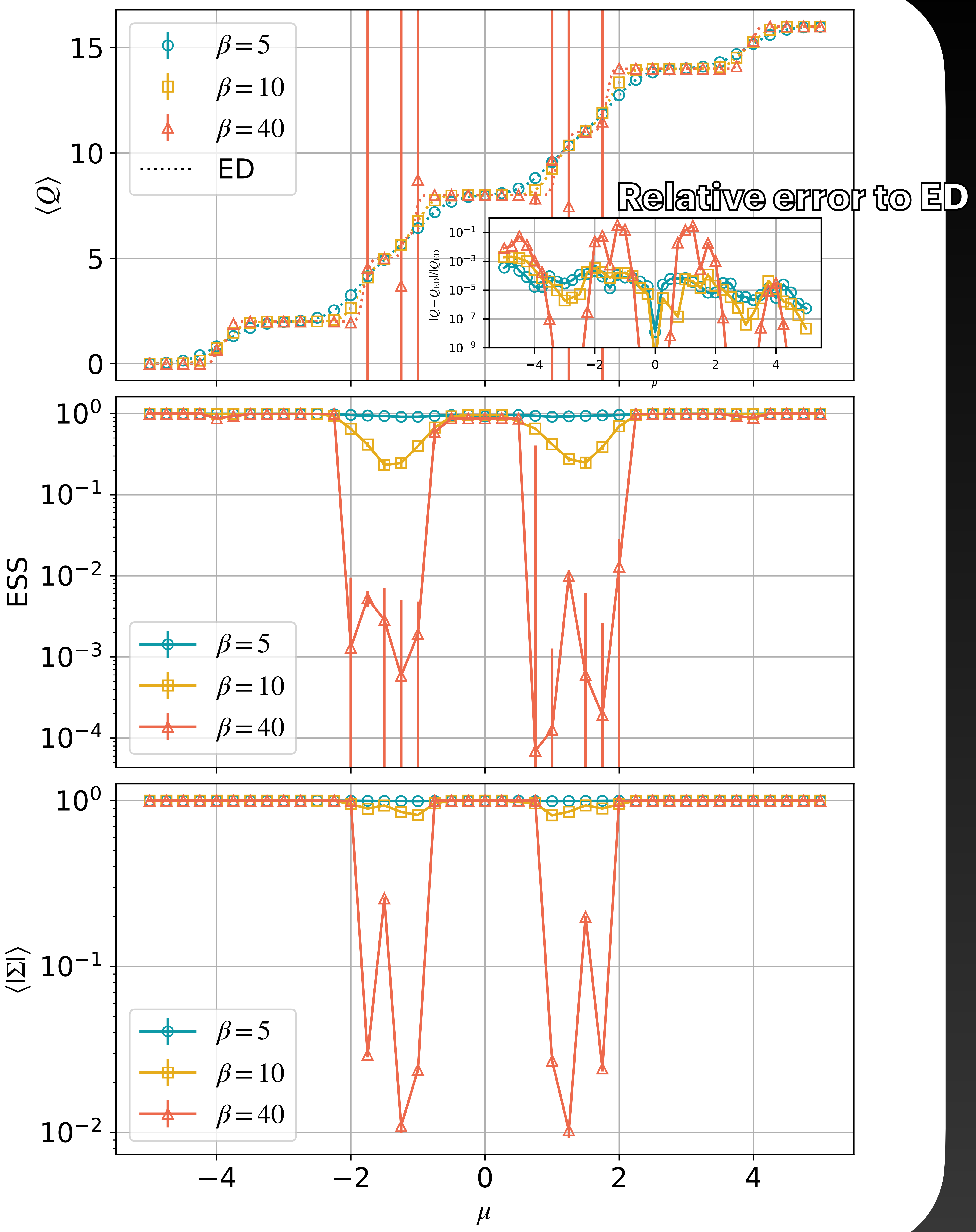
Charge Observable

8 sites hexagonal lattice

8 physical sites

$U = 2 \qquad N_t = 4\beta$



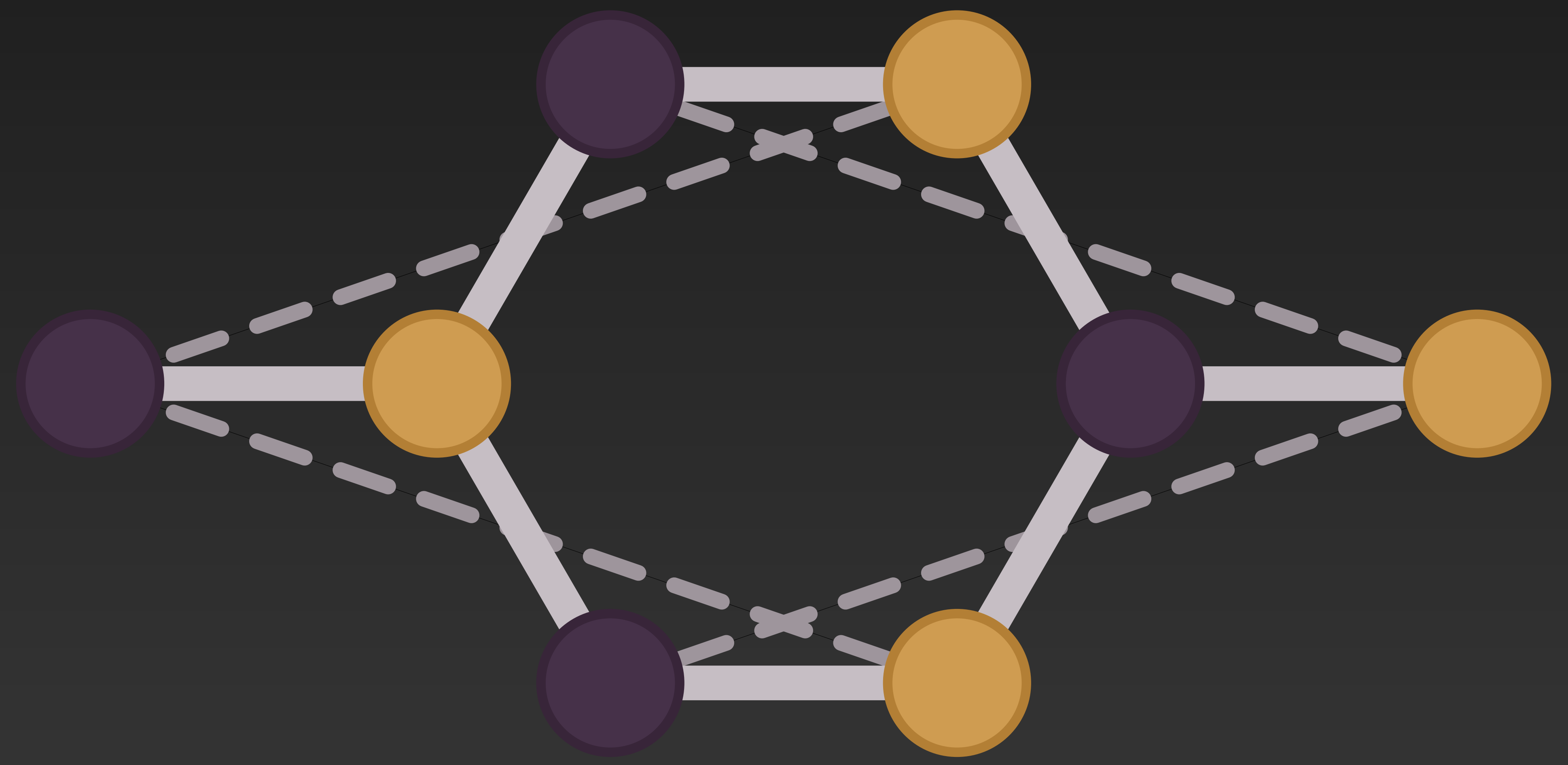


Charge Observable

8 sites hexagonal lattice

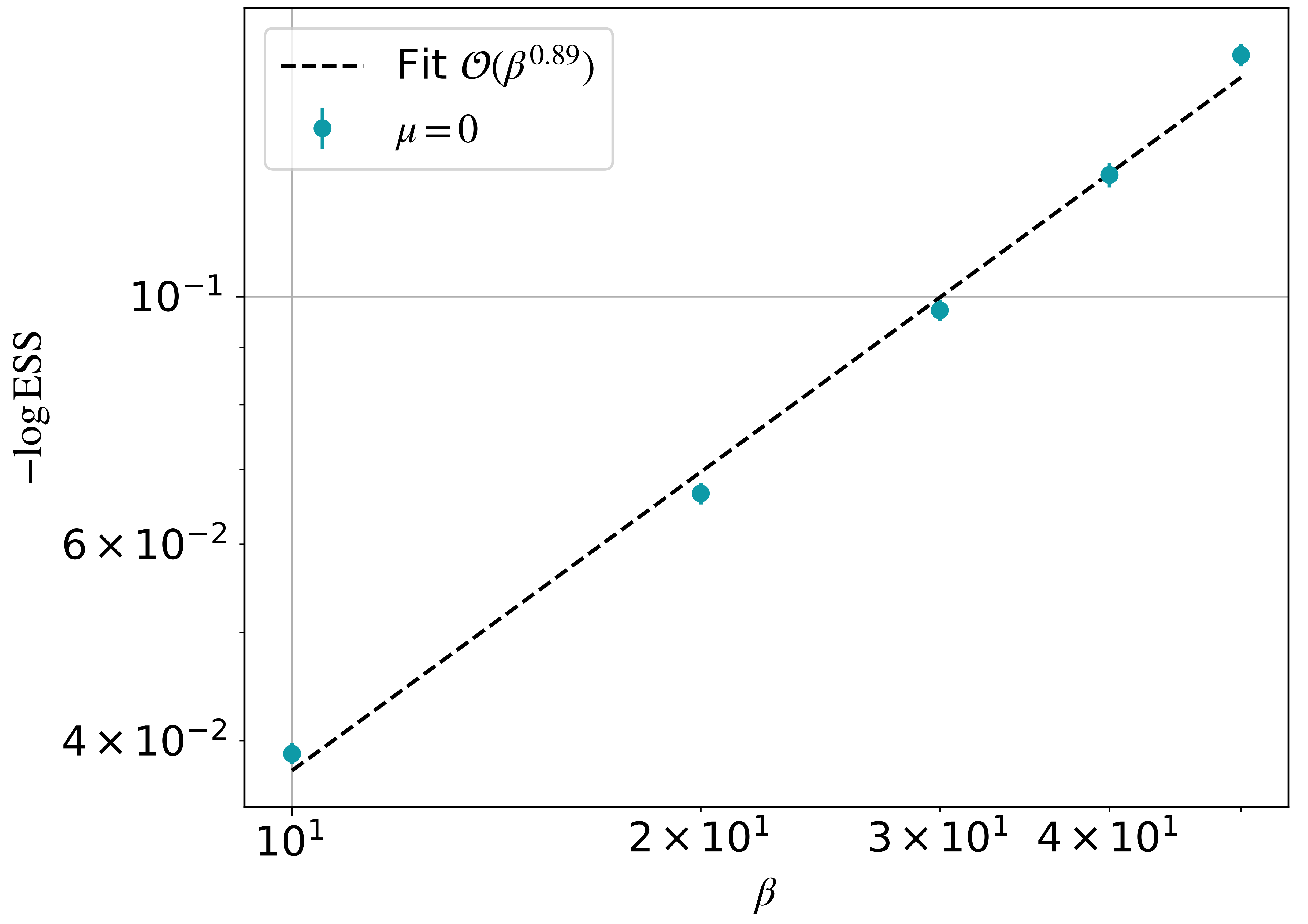
8 physical sites

$U = 2 \qquad N_t = 4\beta$



Scaling

Inverse Temperature

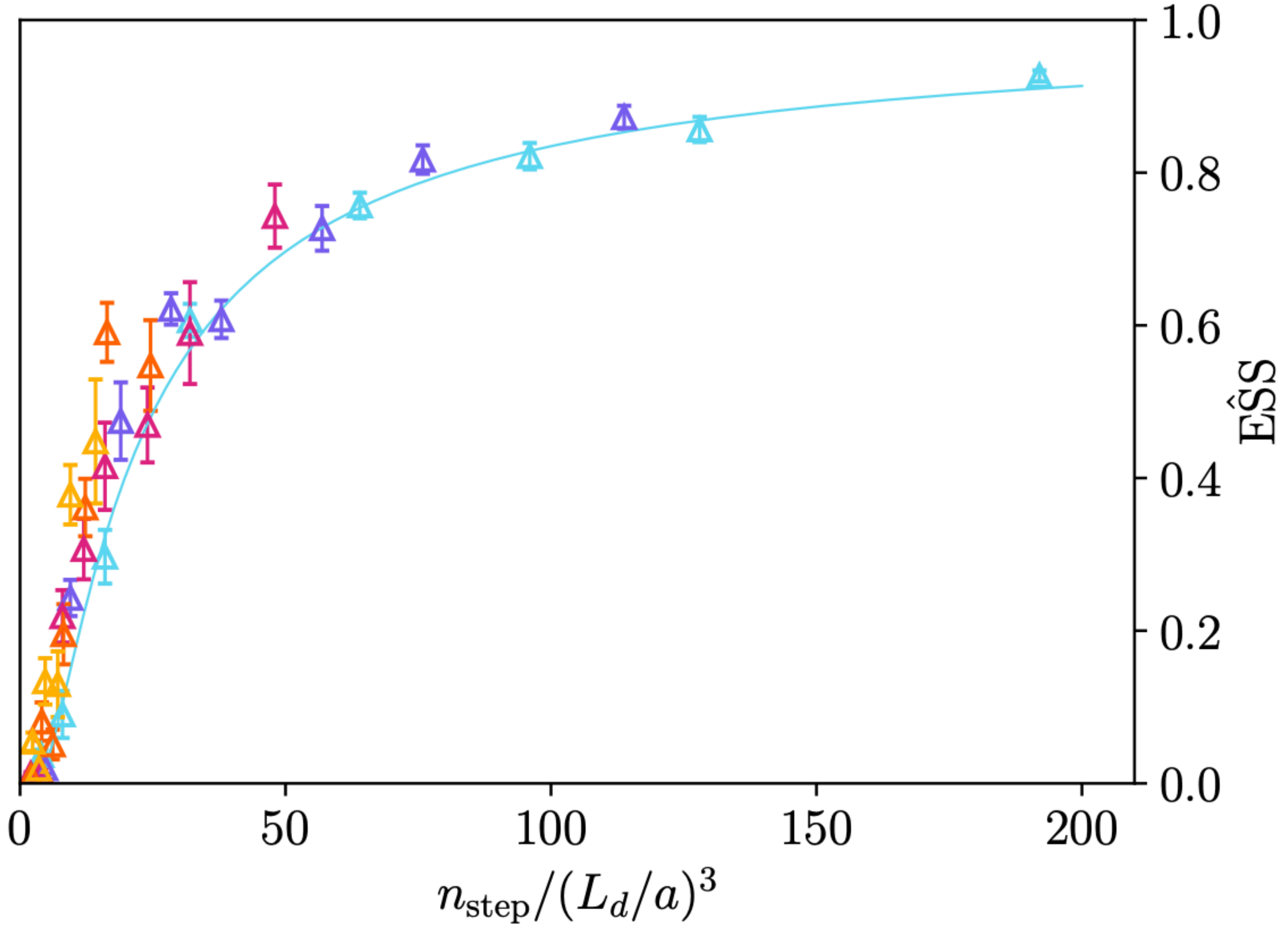


$$\mathrm{ESS} \approx \exp\left(-k\beta\right)$$

$U = 2,\ \beta = 2.5,\ N_t = 10$

Non-equilibrium MC

Quality scaling



ESS improves with n_{step}

C. Bonanno, A. Bulgarelli, E. Cellini, et al., arXiv:2510.25704

Total Computational Scaling

$$\text{Training time} = \left(\underbrace{\mathcal{O}(N_x^3 N_t)}_{\text{Action}} + \underbrace{\mathcal{O}(N_x^2 N_t^2)}_{\text{Sampling}} \right) \cdot \underbrace{\mathcal{O}(N_x^{1.7} N_t)}_{\text{Training}}$$

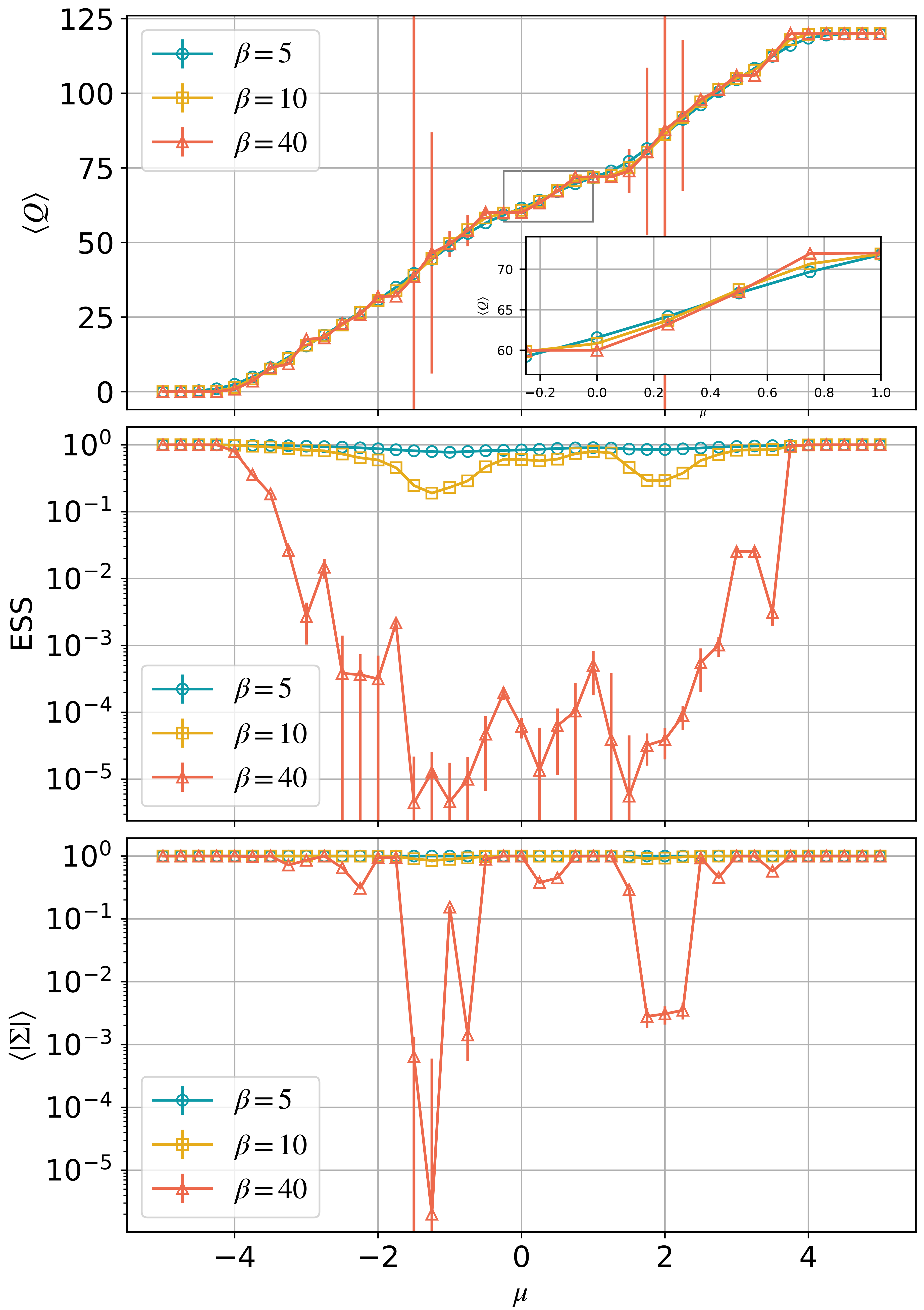
$$\text{Memory} = \underbrace{\mathcal{O}(N_x^2 N_t)}_{\text{Action}} + \underbrace{\mathcal{O}(N_x N_t)}_{\text{Sampling}} + \underbrace{\mathcal{O}(N_x N_t)}_{\text{Training}}$$

Loss Function

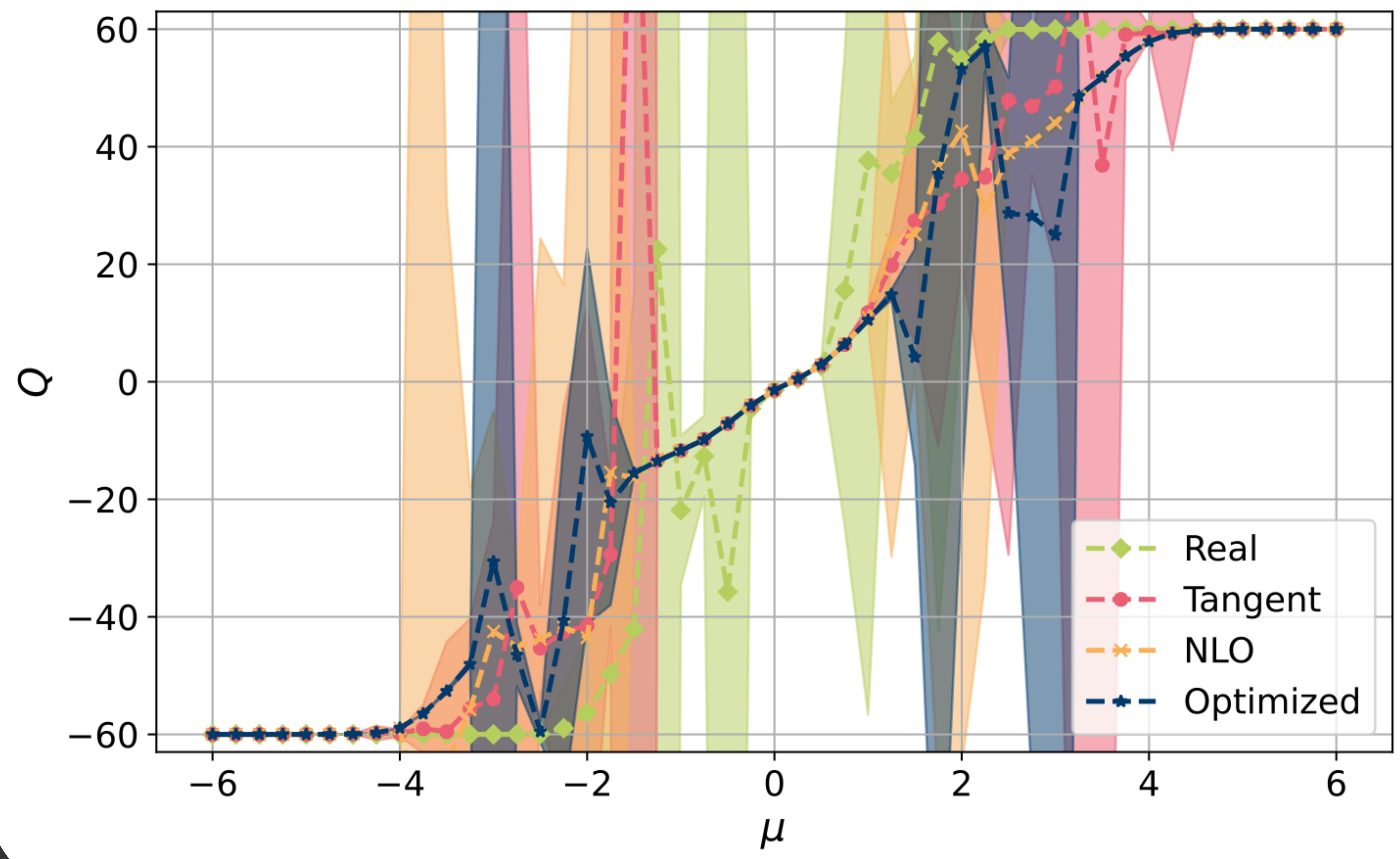
$$\frac{\partial}{\partial \theta}\mathrm{KL}(q_\theta || p) = \mathbb{E}_{\phi \sim q_\theta}\left[\left(\log q_\theta(\phi) + S[\phi]\right) \cdot \frac{\partial}{\partial \theta} \log q_\theta(\phi)\right]$$

Charge Observable

C60 fullerene



Charge basis HMC: $\beta = 6,\ U = 2,\ N_t = 16$



C. Gäntgen et al., PRB 109 (2024)