Normalising flows for the density of states: from gauge theories to the doped Hubbard model

Simran Singh

University of Bonn

Bridging Machine Learning and Statistical Field Theory through generative and diffusion processes @ ECT* Trento

September 10, 2026

Brief outline

I: Motivation for DoS

II: Normalising flows based DoS

III: Doped Hubbard model (two sites for now)

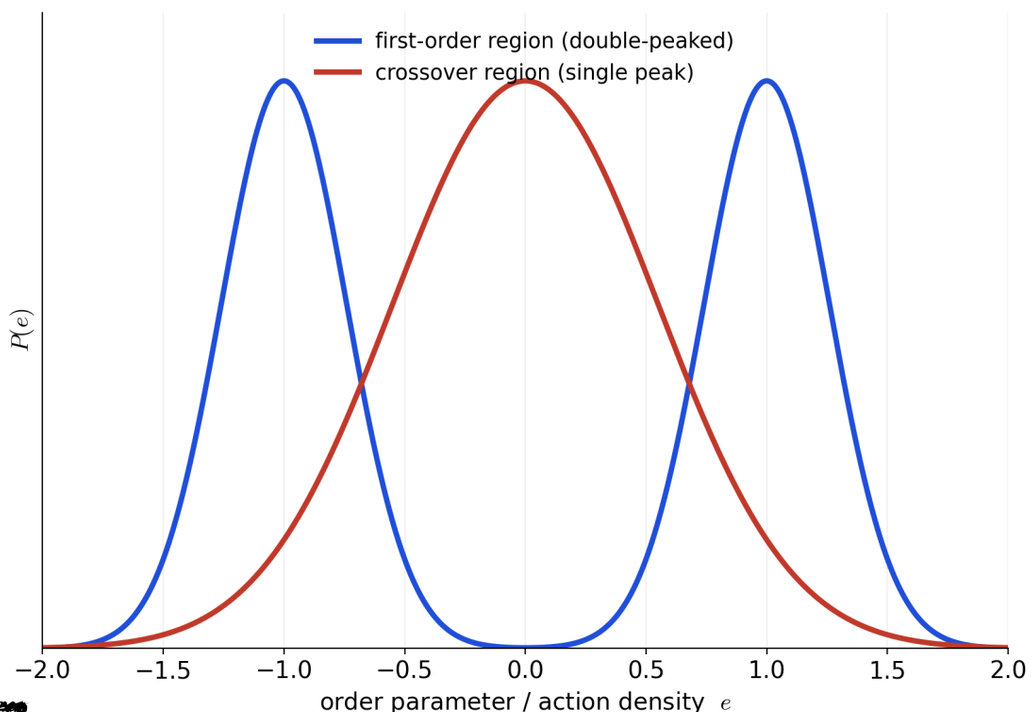
IV: $U(1)$ gauge theory in $2D$

Going beyond Boltzmann distributions

Challenges of sampling from a Boltzmann distribution ?

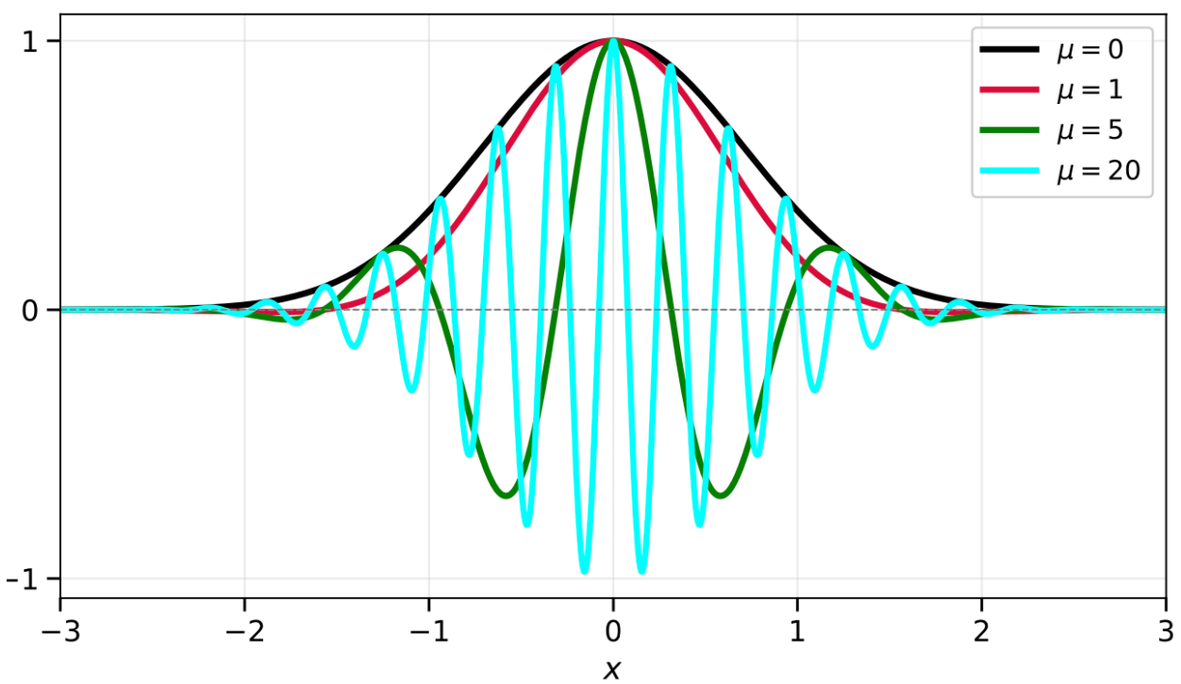
Multi-modality in real actions

- Across a phase transition
- Fermion determinant zeroes
- ...



Sign problem in complex actions

- QCD at finite density (μ)
- Gauge theories with θ-term
- ...



The original approach: DoS for real actions

this talk U(1) 2D

$$Z(\beta) = \int DU \, e^{-\beta S_e[U]} \longrightarrow U \sim p_\beta(U)$$

Instead of sampling from full **Boltzmann** distribution → split computation

Step 1: Compute $\beta-$independent DoS

$$\rho(E) = \int DU \, \delta(S_e[U] - E) \qquad U \sim p_E(U)$$

Step 2: Perform 1D integral to recover partition function

$$Z(\beta) = \int dE \; \rho(E) \, e^{-\beta E}$$

[F. Wang and D. P. Landau (2001)], [K. Langfeld, B. Lucini and A. Rago (2012)]

Generalisation of DOS to complex actions

this talk doped Hubbard

Sign-problem-afflicted action: $\beta S_e[U] = S_R[U] + i\,h\,S_I[U]$



$$U \sim e^{-\beta S_e[U]}$$



real distribution!

Generalised DoS: $\rho(c) = \int DU e^{-S_R[U]}\,\delta(c - S_I[U])$

[K. Langfeld, B. Lucini (2014)]

Recover partition function via a Fourier Transform!

$$Z(\mu) = \int dc\,\rho(c)\,e^{-i\mu c}$$

How does one compute the DoS in practice?

How does one sample: $U \sim \delta(c - S_I[U])$?



Regularise the delta: $\lim_{P\to\infty} e^{-\frac{P}{2}(x-a)^2 - \log N} = \delta(x-a)$

P-dependent DoS*: $\rho_P(c) = \int \mathscr{D}U\, e^{-S_{c,P}[U]}$

with $S_{c,P} = S_r[U] + \frac{P}{2}\left(c - S_I[U]\right)^2 + \log N$

*some discussion on the magnitude of P can be found here - w.r.t the “integerness” of the topological charge finite P

S. Borsanyi and D. Sexty (2021)

Method: Using normalising flows to estimate DoS

First application of NF-DoS [J. M. Pawlowski, J. M. Urban (2023)]

$$\rho_P(c) = \int \mathscr{D}U\, e^{-S_{c,P}[U]}$$

$q_{c,P}[U]$

Normalised Prior → NN model → Target

Trick: insert identity

K. A. Nicoli et. al., (2021)

$$\rho_P(c) = \int \mathscr{D}U\, q_{c,P}(U) \frac{e^{-S_{c,P}[U]}}{q_{c,P}(U)}$$

$$\rho_P(c) \approx \langle e^{-S_{c,P}[U] - \ln q_{c,P}(U)} \rangle_{U \sim q_{c,P}(U)}$$

Two applications

After the first successful application to ϕ^4 in $0, 1, 2$ D

[J. M. Pawlowski, J. M. Urban (2023)]

Conjecture: If you have an NF setup for sampling the Boltzmann distribution of the original theory … then DoS can be applied

gauge theory: $2D$ $U(1)$

- Publicly available gauge equivariant NF code

[G. Kanwar et. al (2021)], [M. S. Albergo et.al. (2021)]

doped Hubbard model

- Access to efficient NFs thanks to J. Kreit & D. Schuh

[D. Schuh, J. Kreit et.al., PoS Lattice 2025]

Application to the doped Hubbard model

work in progress with:

U. Bonn

U. Bielefeld



Felicitas Freche :
main actor for
Hubbard DoS project



Lena Funcke



Janik Kreit



Dominic Schuh



Christian
Schmidt-Sonntag

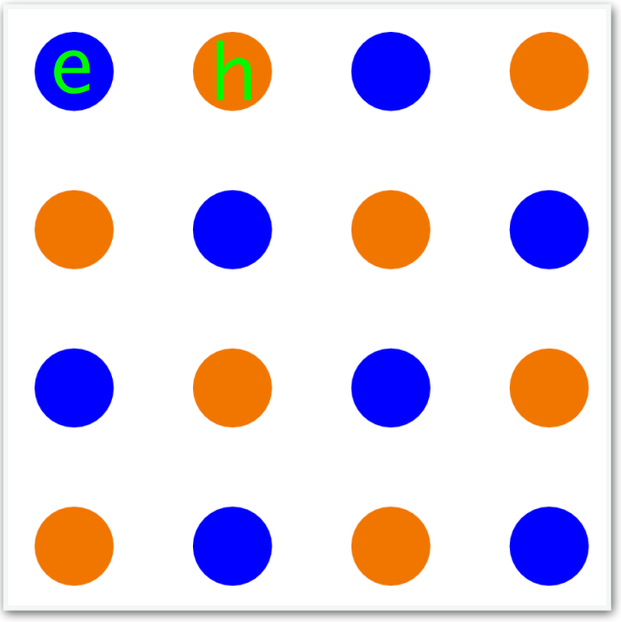
original implementation of NFs for Hubbard

The doped Hubbard model (charge basis)

System of strongly correlated electrons, on fixed spatial lattices

$$H = -\kappa \sum_{\delta_{\langle x,y\rangle},x} (p_x^\dagger p_y - h_x^\dagger h_y) + \frac{U}{2} \sum_{\delta_{x,y},x} q_x q_y + \mu \sum_x q_x$$

where $q_x = p_x^\dagger p_x - h_x^\dagger h_x$



$\mu = 0$

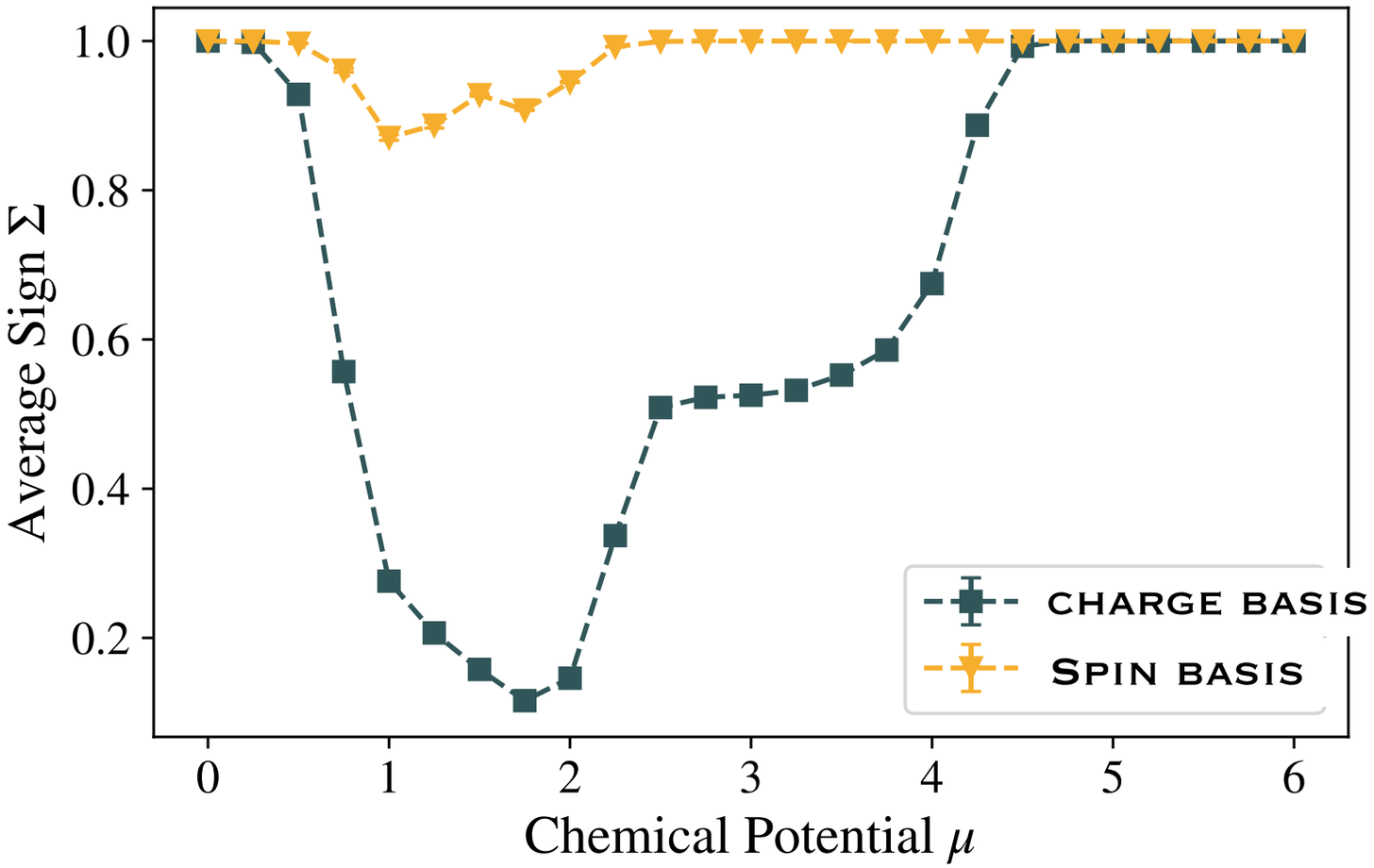
study the problem in the action-formalism

$$S = -\underbrace{\ln \det M[\phi, \kappa, \mu] M[-\phi, -\kappa, -\mu]} + \frac{1}{U} \sum_{x,t} \phi_{x,t}^2$$

complex ! 

The doped Hubbard model (charge basis)

Lena's talk: choose spin basis for mild sign problem



[D. Schuh, L. Funcke, J. Kreit, T. Luu, **S.S.**, arXiv:2603.18205]

Why apply ML-DoS to Hubbard in charge basis?

- simpler distributions

- $S[\phi, \mu] = S_r[\phi] + i\pi_{\phi,\mu}$ *

- benchmark exits (next slide)

*treat it as a schematic equation

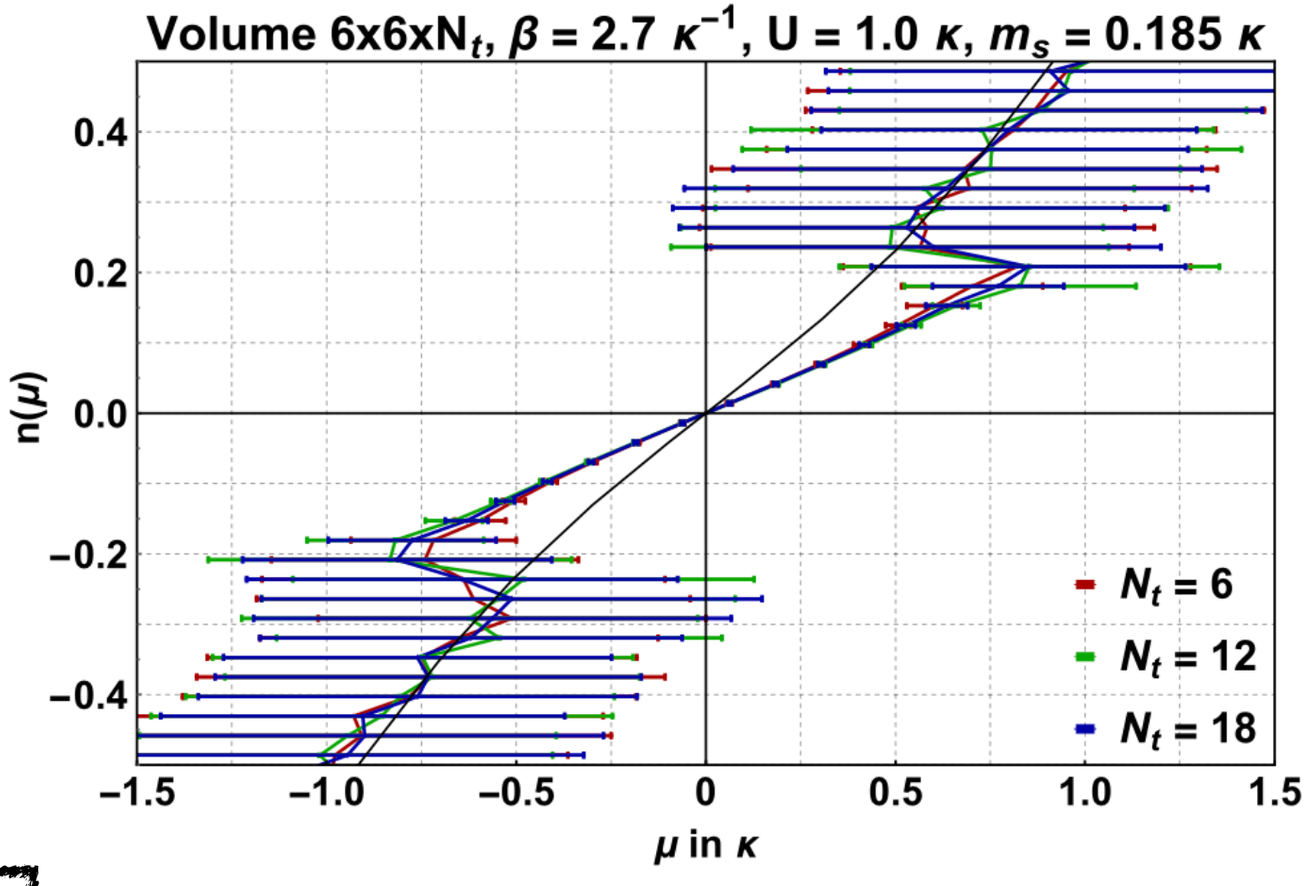
DoS ansatz for the Doped Hubbard model

Can the generalised DoS method be used here?

PHYSICAL REVIEW D **102,** 054502 (2020)

Density of states approach to the hexagonal Hubbard model at finite density

Michael Körner,[1] Kurt Langfeld,[2] Dominik Smith,[1] and Lorenz von Smekal[1]



Our original aim:

Can the NF-based DoS method reduce these error bars from LLR?

DoS ansatz for the Doped Hubbard model

Can the generalised DoS method be used here?

[M. Körner, K. Langfeld, D. Smith, L. v Smekal, PRD (2020)]

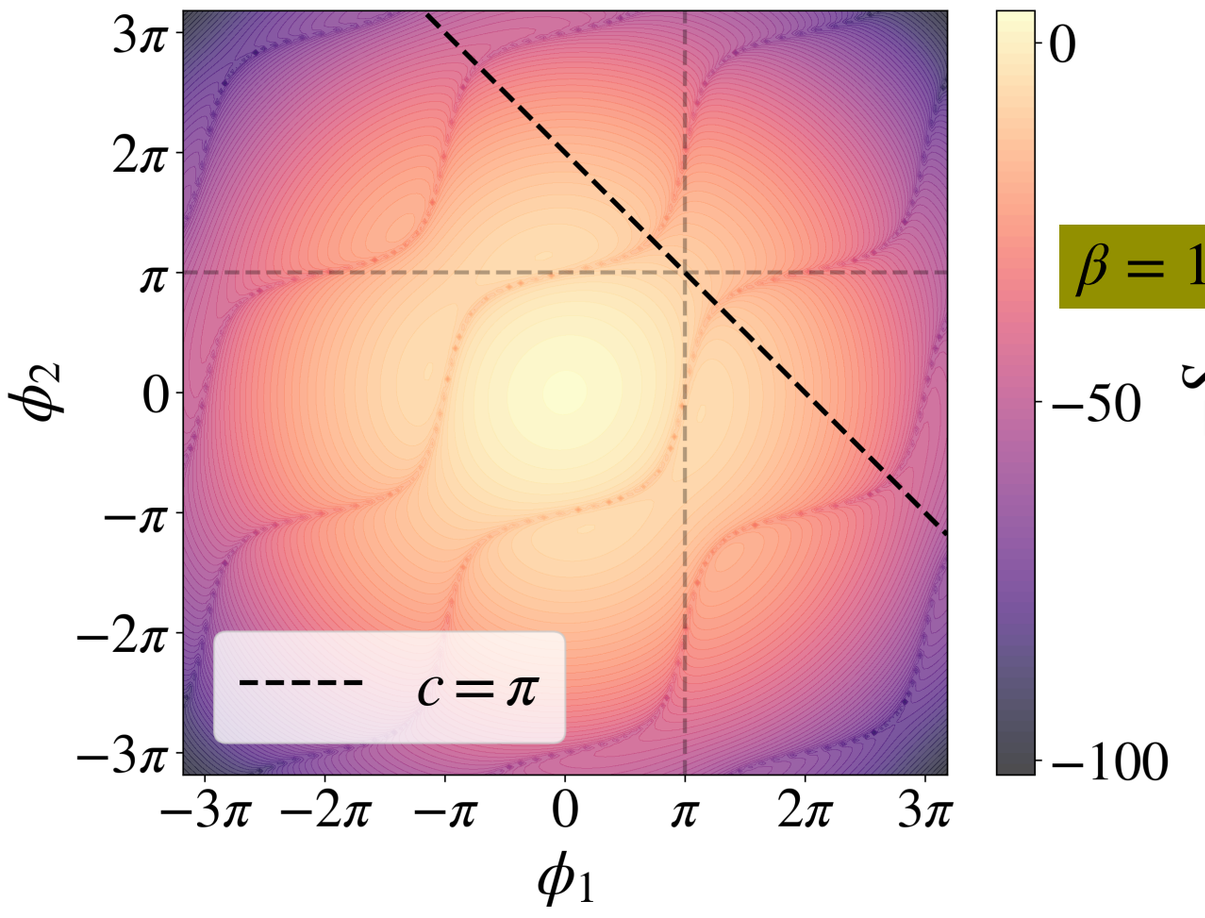
$$\phi \to \phi + i\mu$$

$$S = S_{\text{Hub}}[\phi, \mu = 0] + i\frac{\mu N_x N_t}{U}\Phi$$

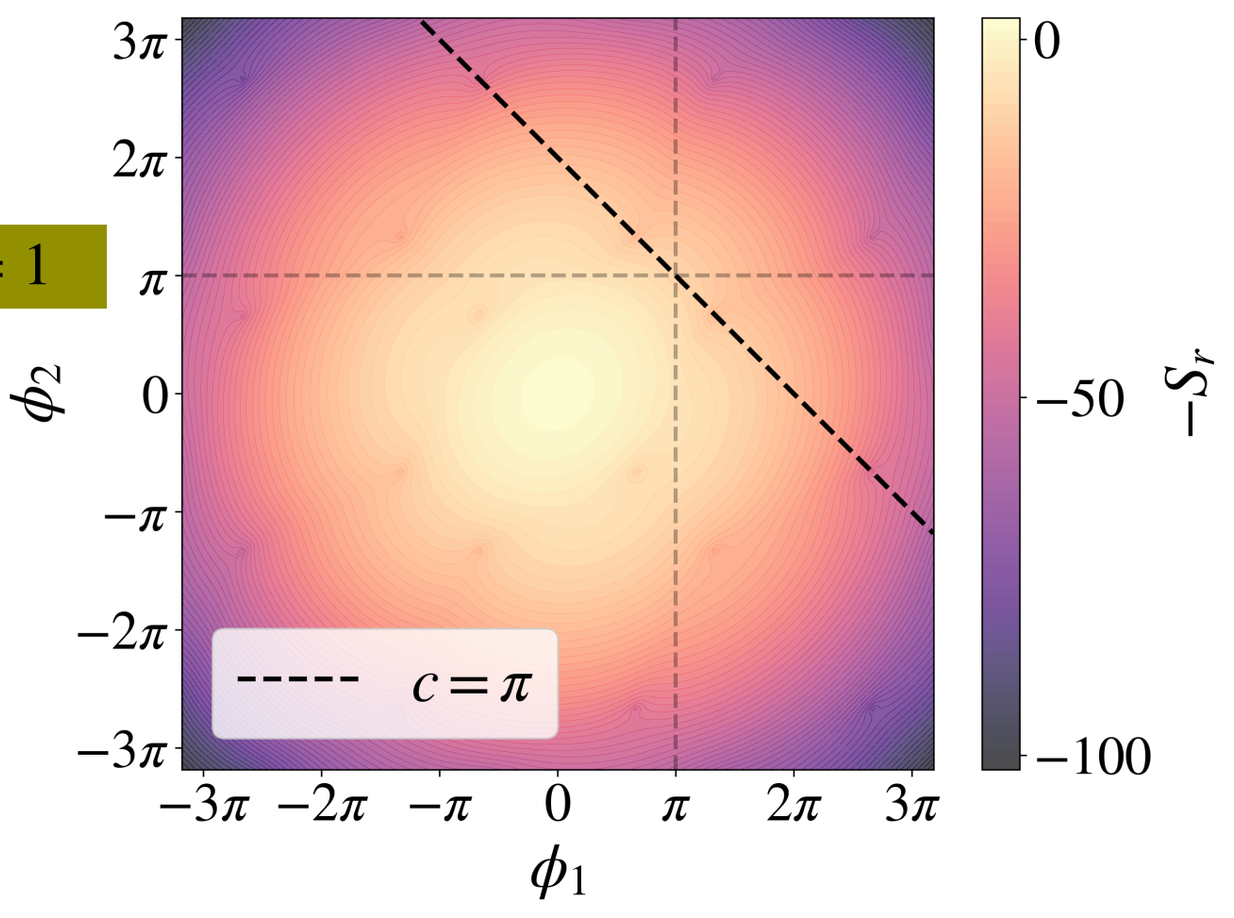


all plots in Hubbard section Felicitas Freche

$\Phi = \frac{1}{V}\sum_{x,t}\phi_{x,t}$ = c becomes the constraint!



$\beta = 1, N_t = 1$



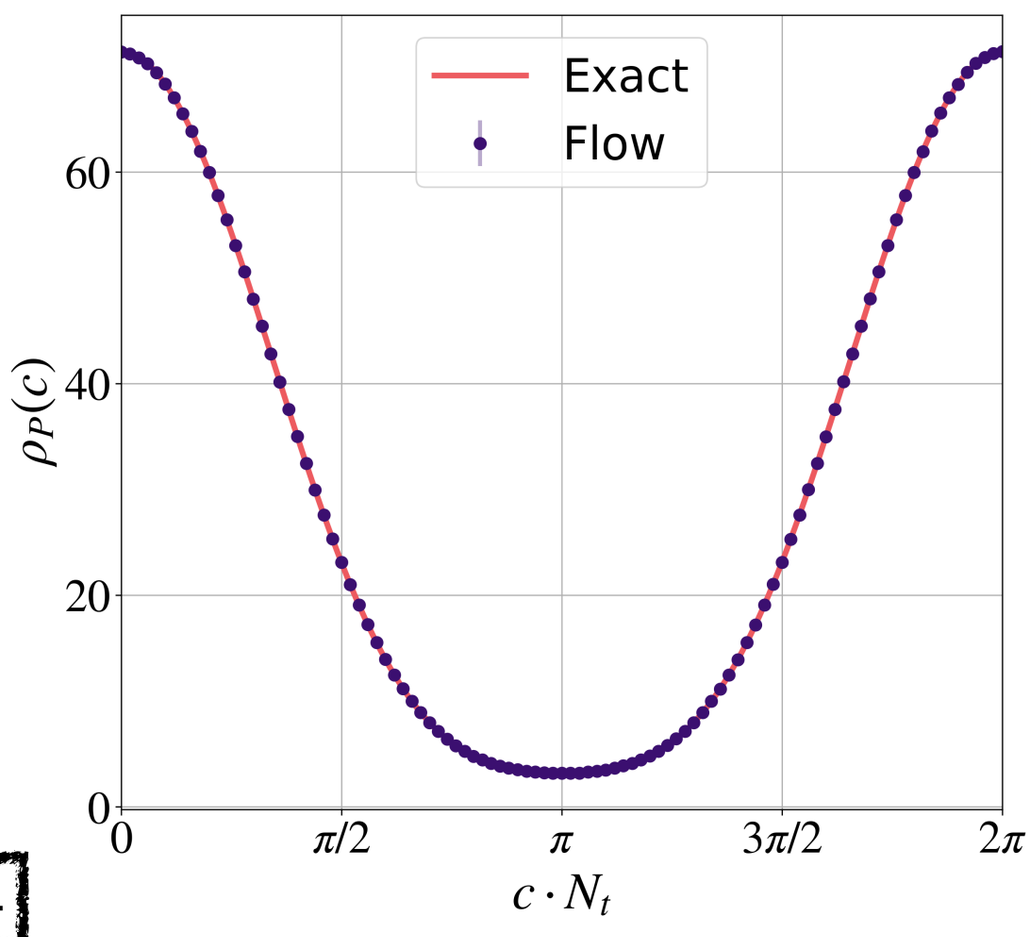
NF-based DoS for the Hubbard Model

DoS "effective" action: $S_{c,P}(\phi) = S_{\text{Hub}}[\phi, \mu = 0] + \frac{P}{2}(c - \Phi)^2 + \log \mathcal{N}$

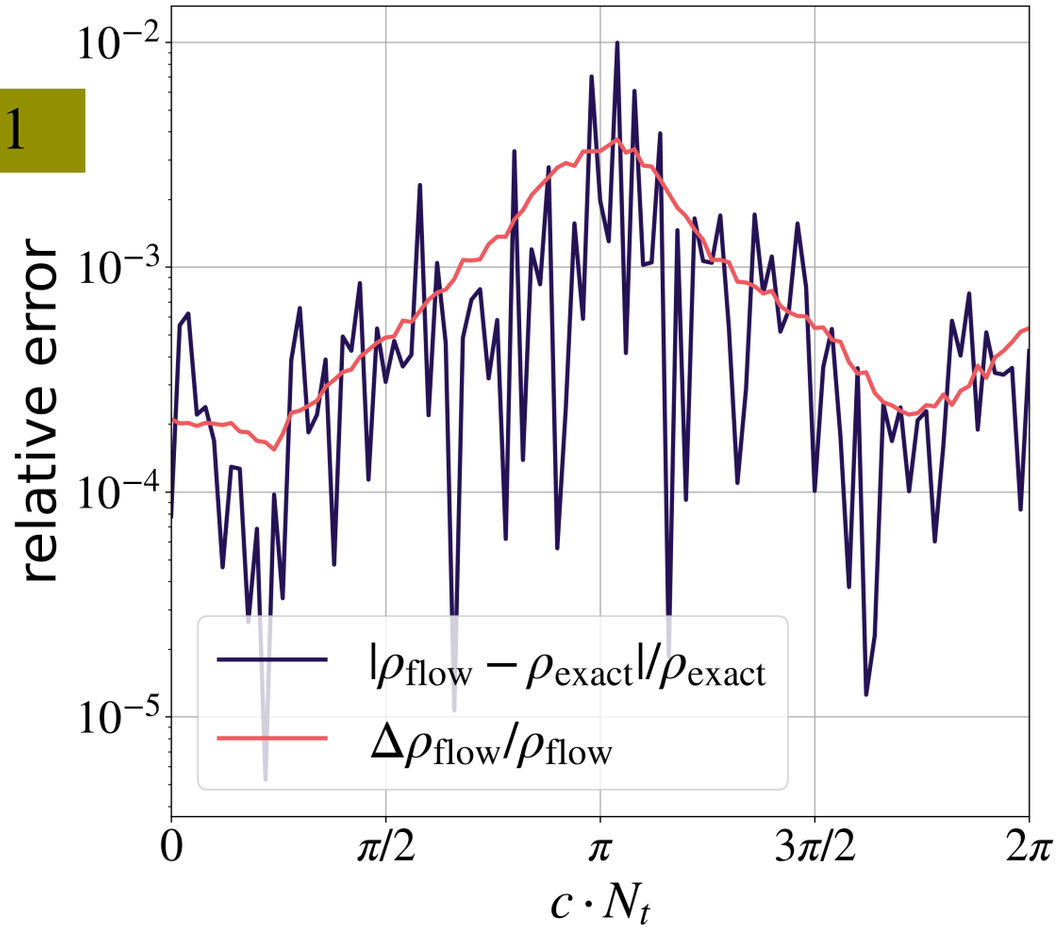
$$\rho_P^{\text{exact}}(c) = \int \mathscr{D}\phi e^{-S_{c,P}(\phi)}$$

$$\rho_P^{\text{flow}}(c) = \langle e^{-S_{c,P}(\phi) - \log q_P(\phi|c)} \rangle_{\phi \sim q_P(\phi|c)}$$

- Prior: multivariate Gaussian
- architecture: real NVP, ESS ~[0.8,0.99]
- c's drawn uniformly for training



$\beta = 1, N_t = 1$



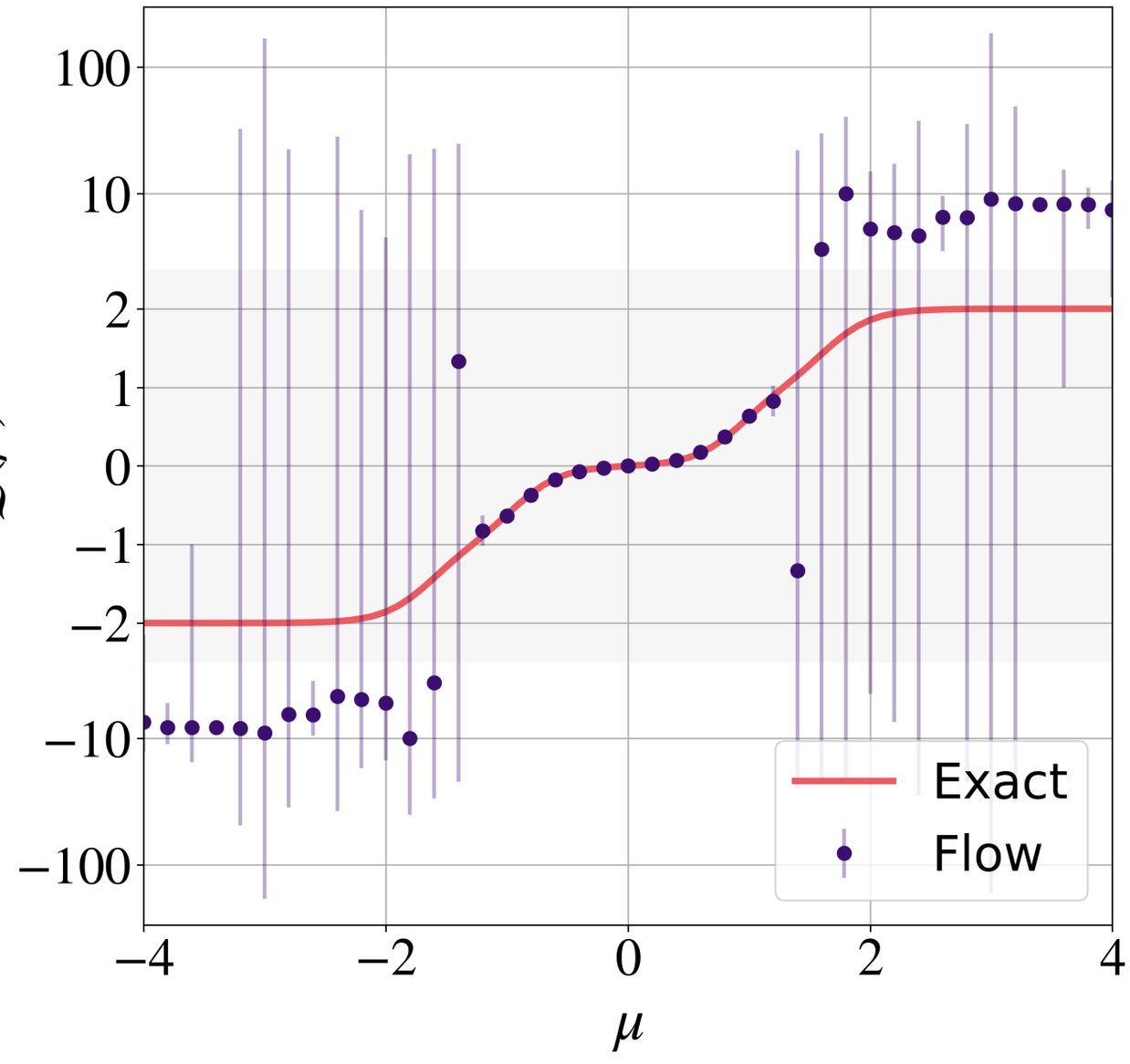
NF-based DoS for the Hubbard Model

Brute-forcing the $1D$ oscillatory integral - not such a good idea

$$N(\mu) = T \frac{d \ln Z(\mu)}{d\mu}$$

$$Z(\mu) = \int dc\, \rho_P(c)\, e^{-if(c,\mu,V,U)}$$

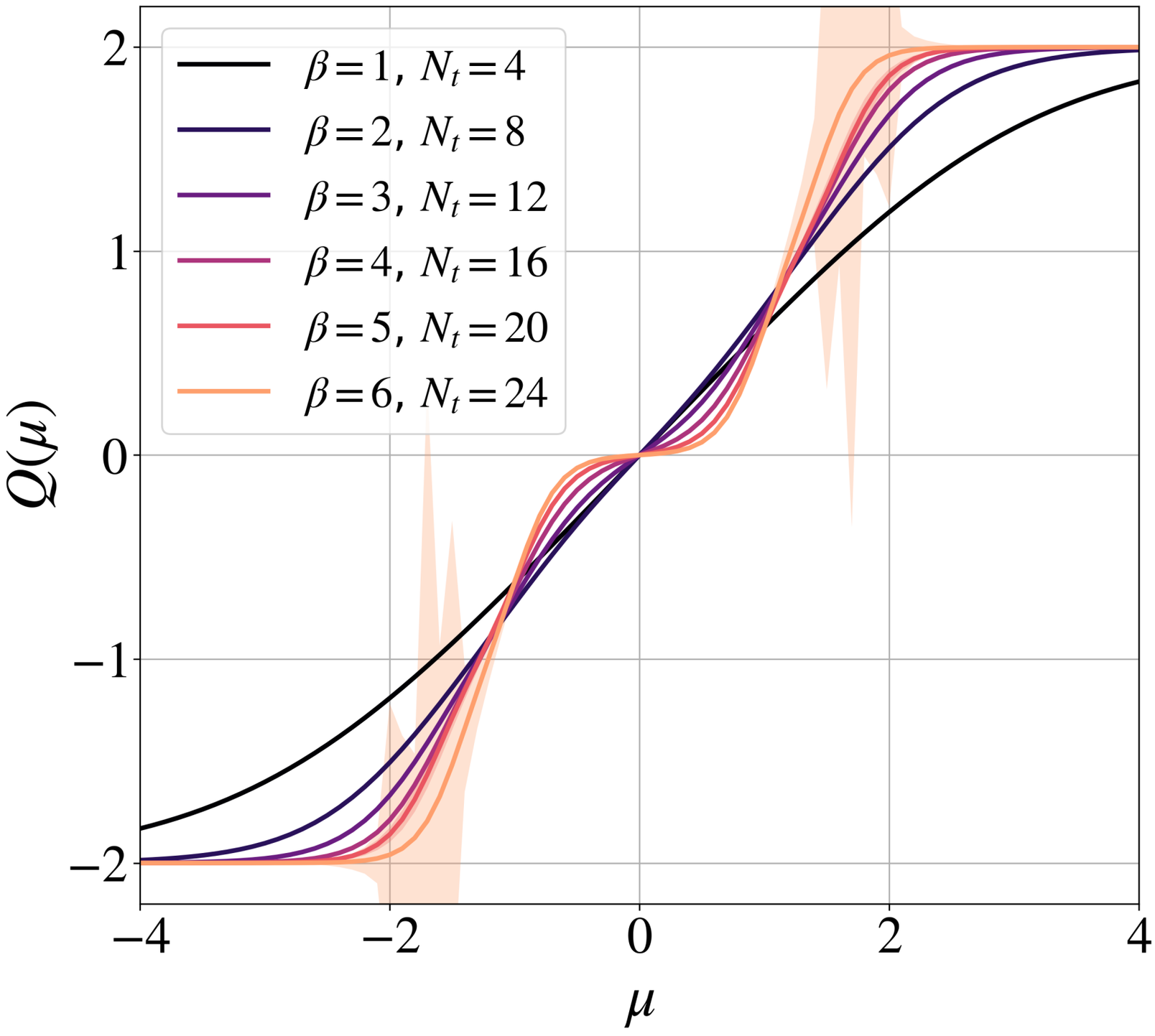
$\beta = 5, N_t = 20$



Hint: Leverage the **periodicity of the DoS**, compute leading Fourier coefficients from Flow -> **integrate analytically***

*semi-analytic integration [K. Langfeld, B. Lucini (2014)]

NF-based DoS for the Hubbard Model

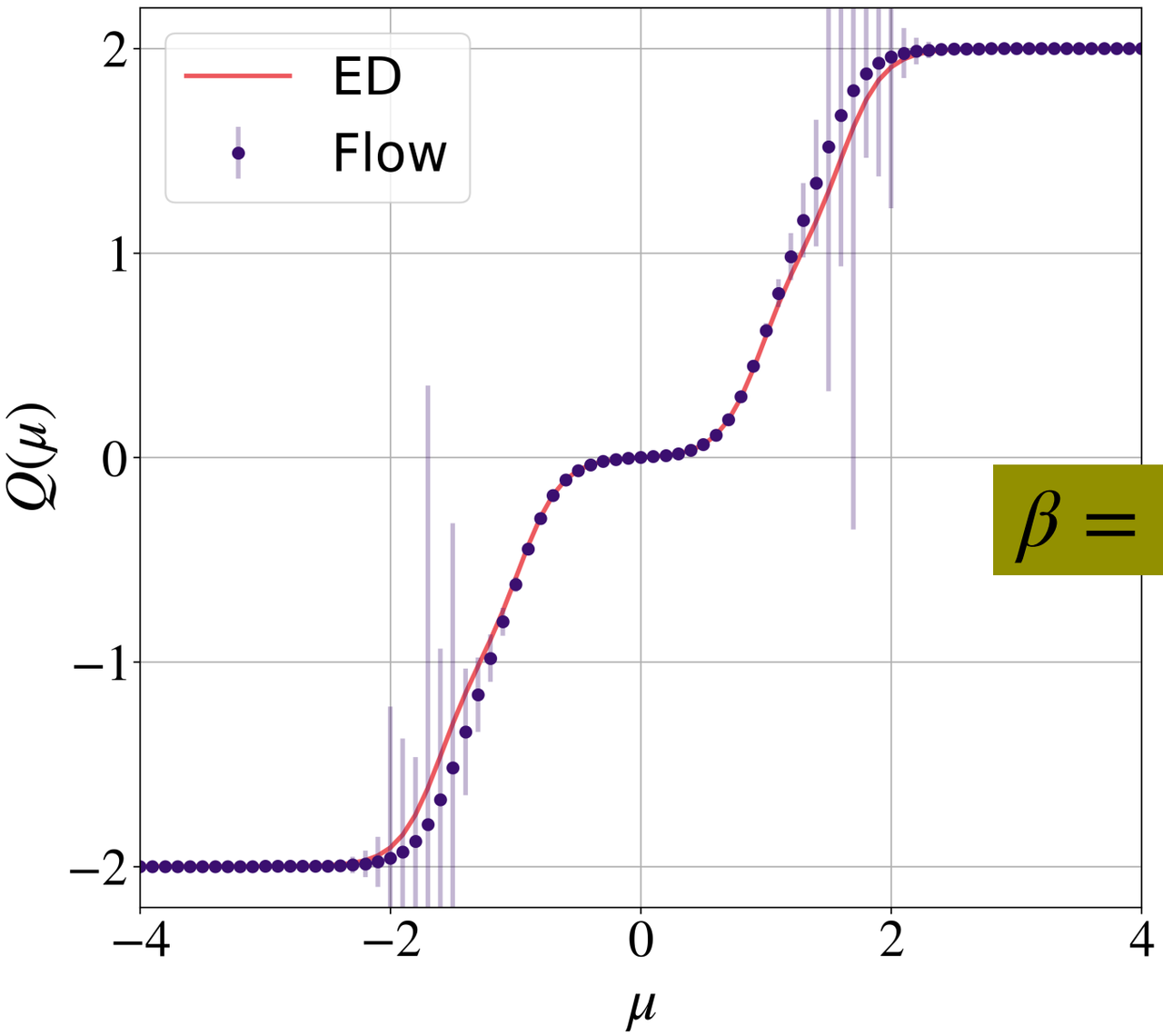


fixed discretisation:

$$\delta = \beta / N_t = 0.25$$

Seems promising until $\beta \leq 5$, cannot yet resolve (crucial) information $\beta = 6$ …

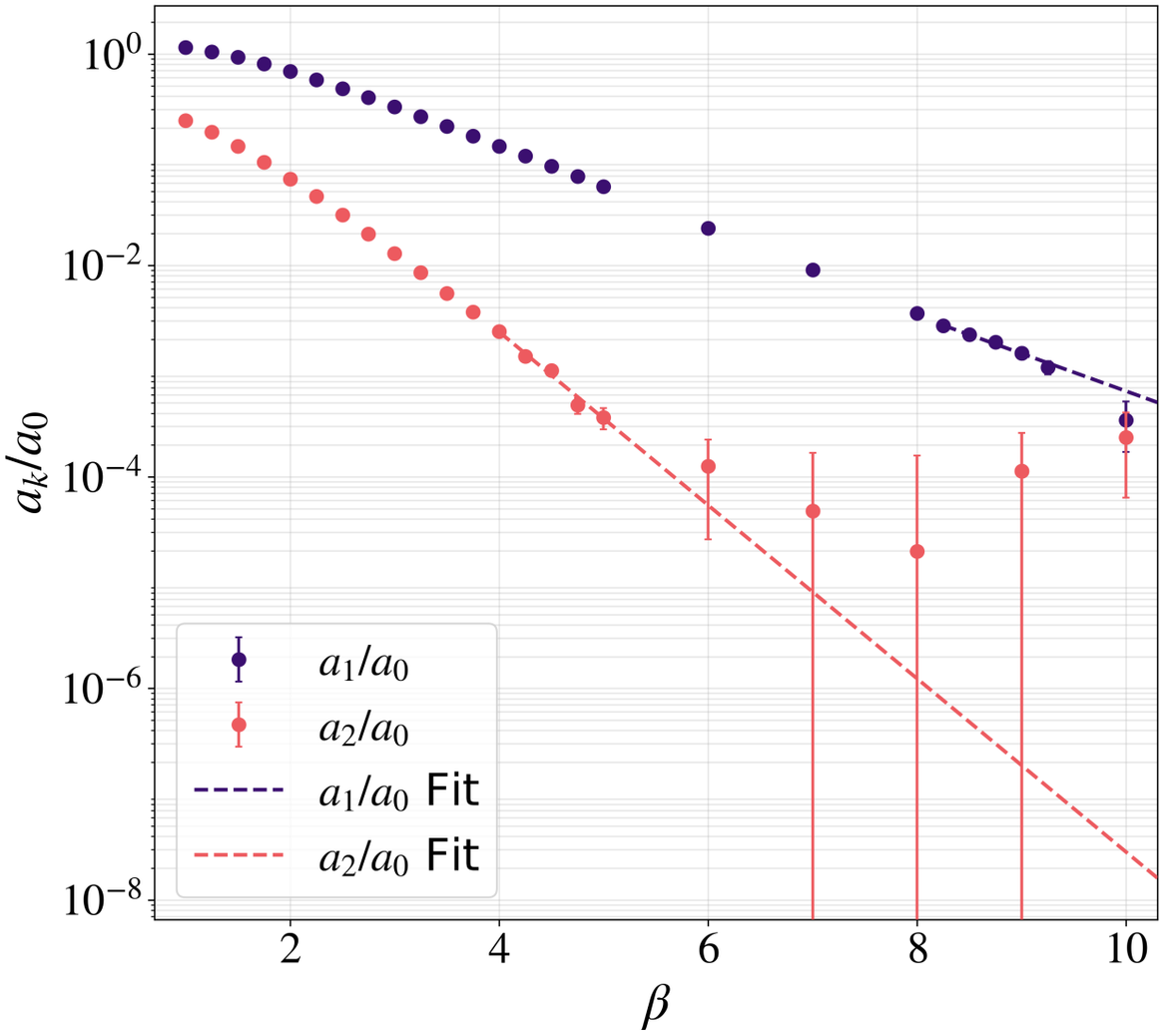
NF-based DoS for the Hubbard Model



Recall: We fit the flow DoS to

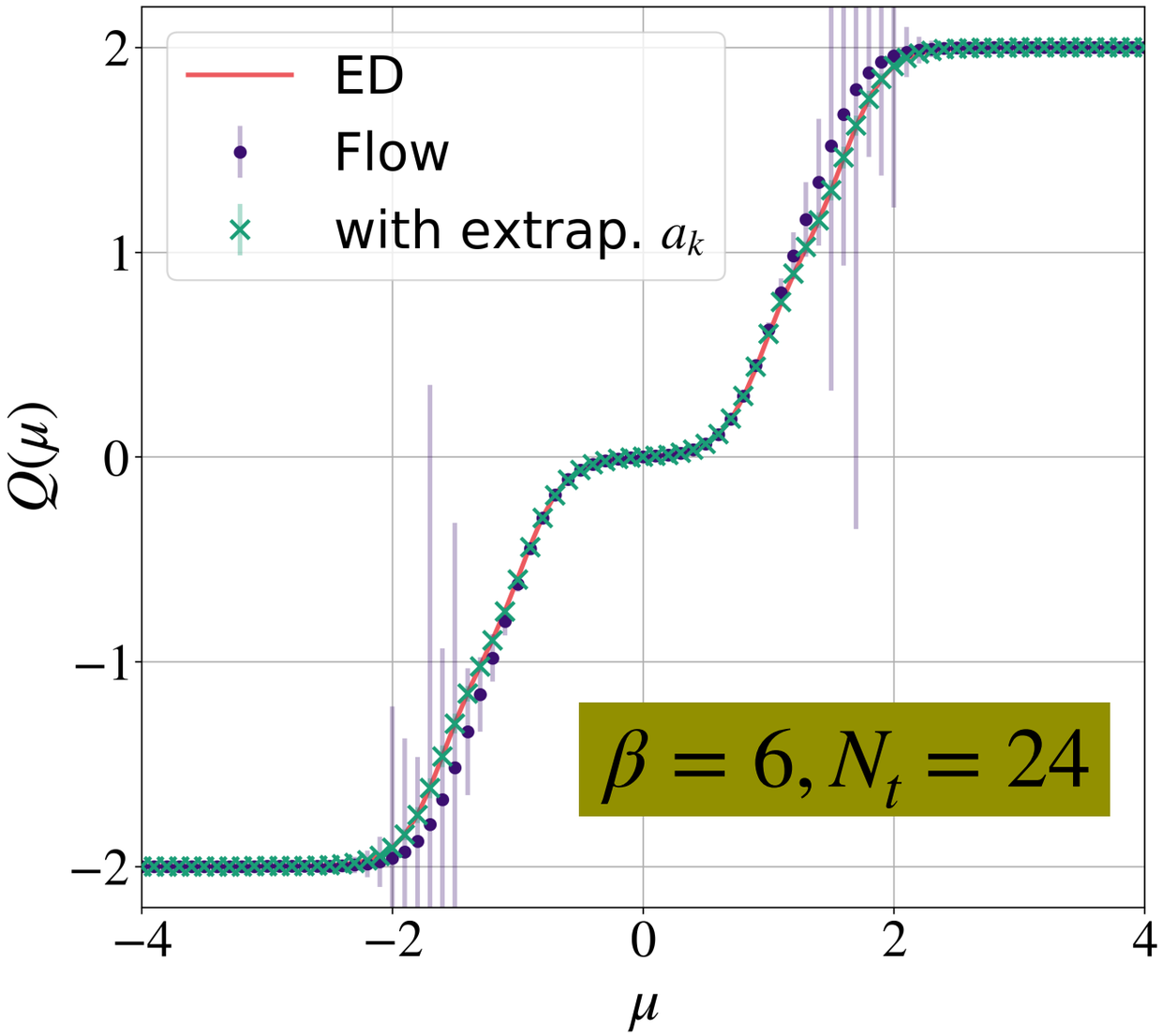
$$\rho(c) = \sum_{N_x} a_k \cos(kN_t c)$$

$\beta = 6, N_t = 24$

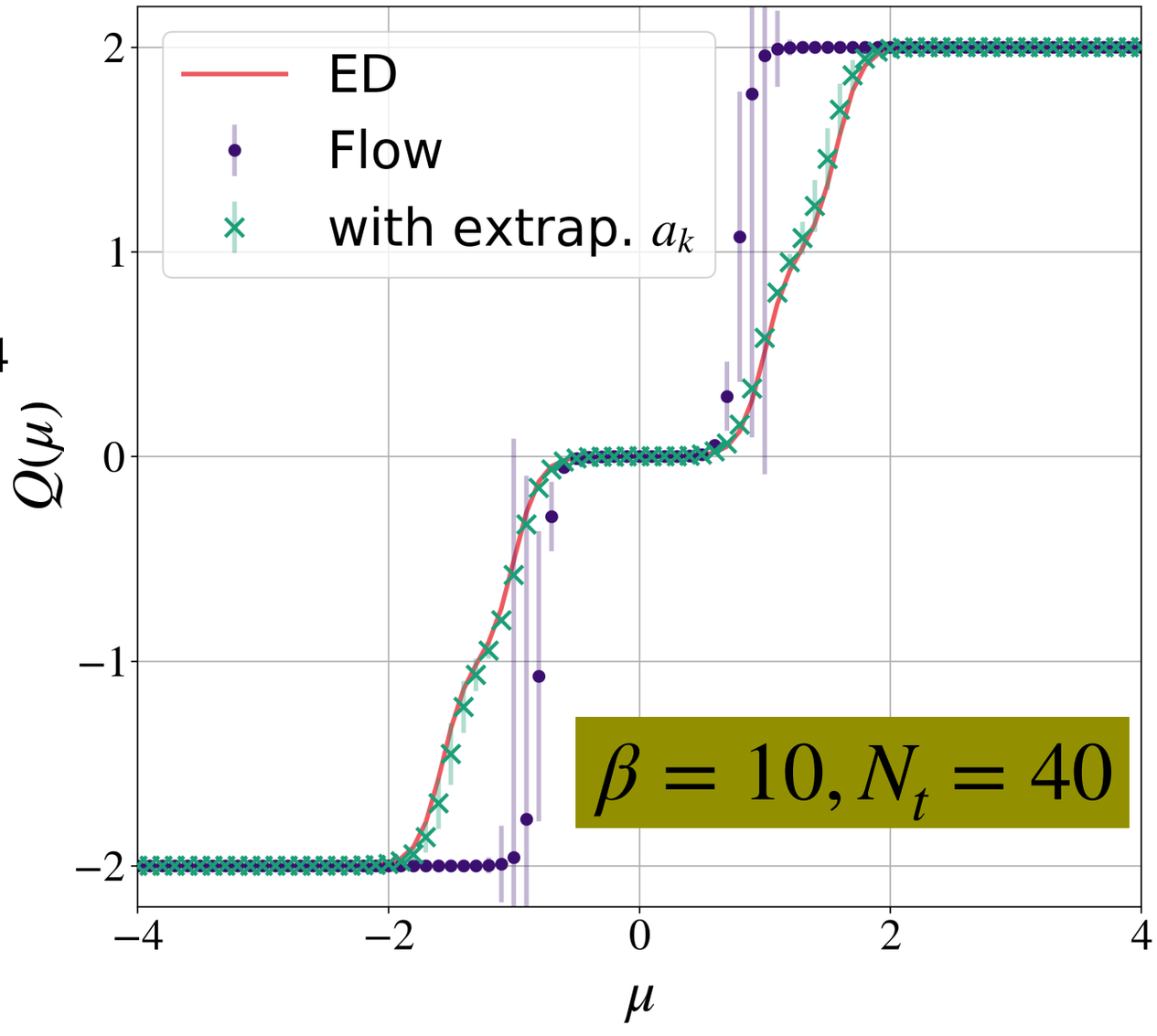


Next step: (Theoretically motivated) extrapolation of a_k

NF-based DoS for the Hubbard Model



semi-analytic + a_k extrapolated

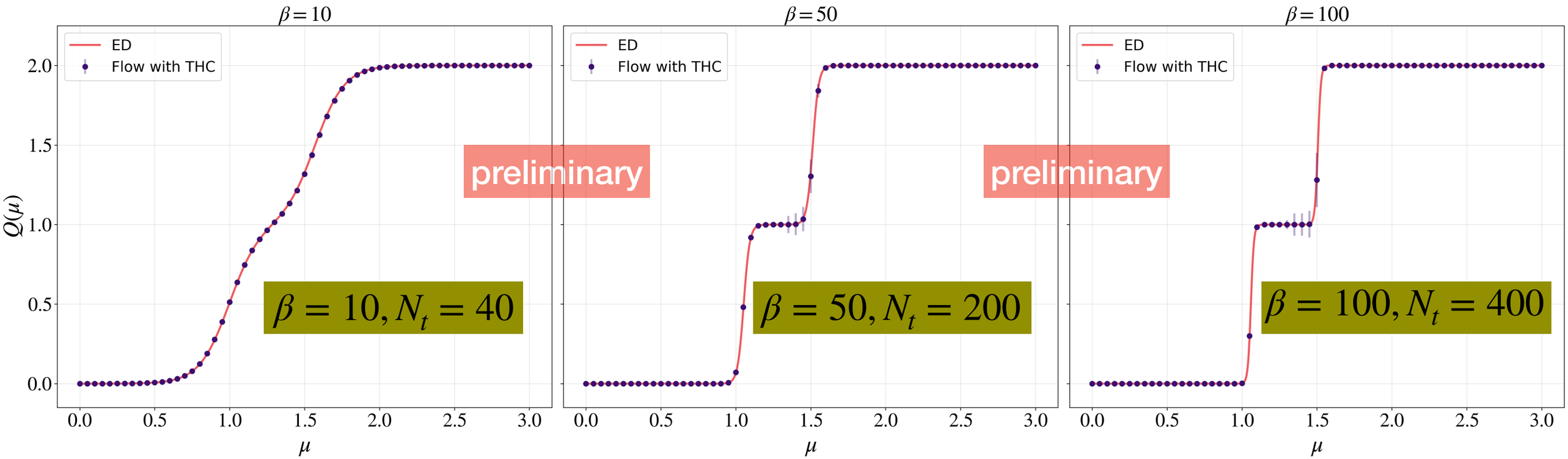


NF-based DoS for the Hubbard Model

Pushing to lower temperatures with $\delta = 0.25$

The Truncated Hankel Correlator Method

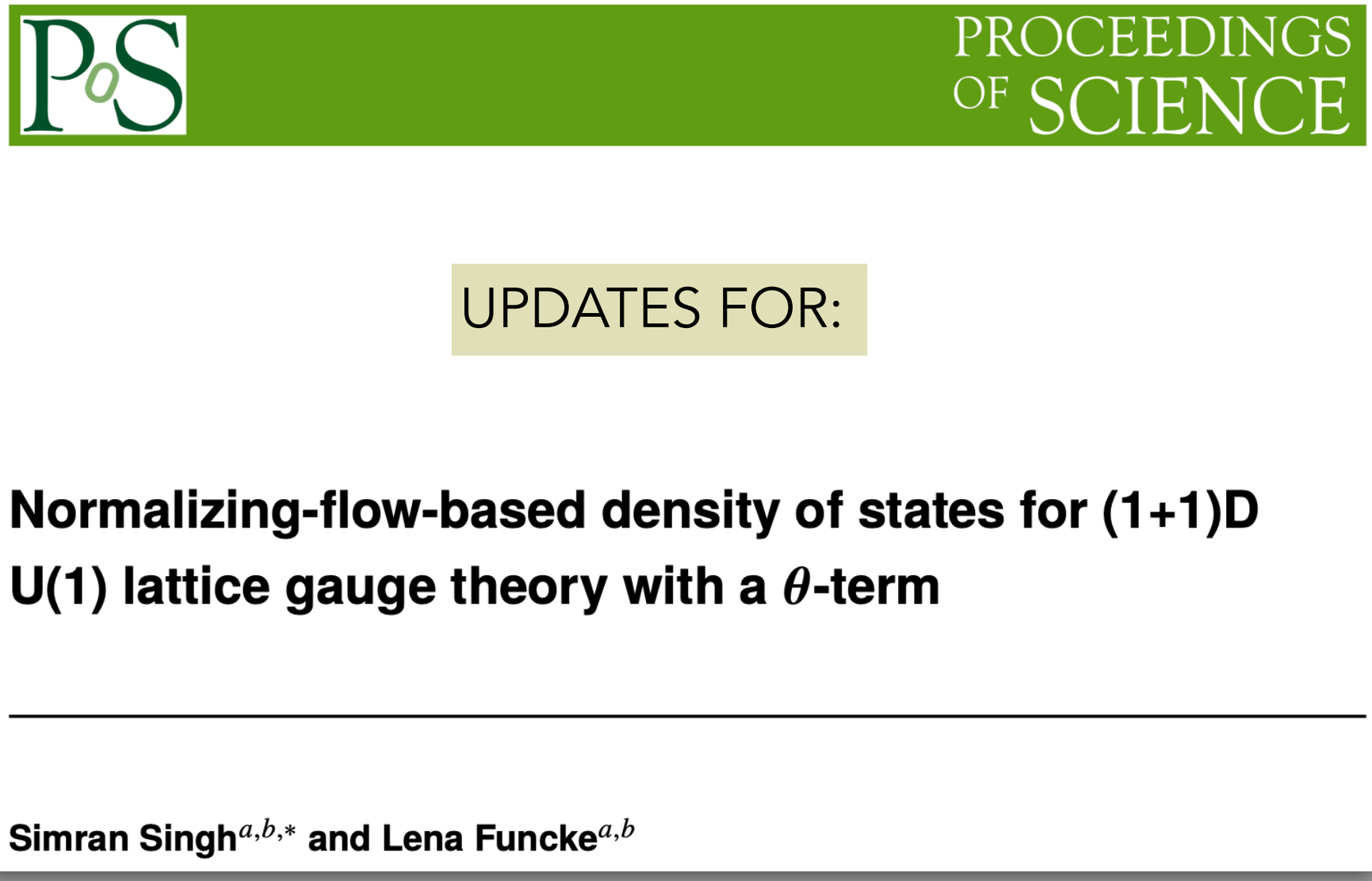
J. Ostmeyer[1] and C. Urbach[1]



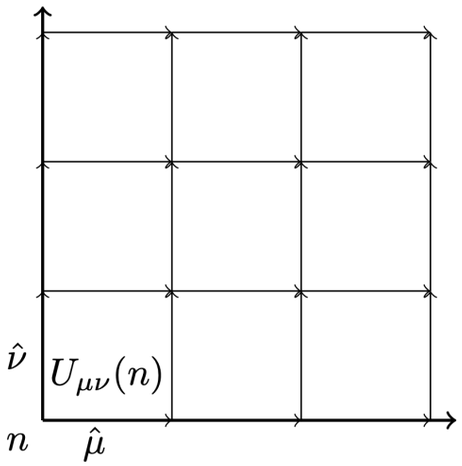
Main message: The Flow DoS data had enough information for (at least) a two-site Hubbard with $N_x = 2, N_t = 400$ lattice, whether one retrieves it depends on choice of 1D integral

… stay tuned for more results (also $N_x = 4$) in Master thesis of Felicitas Freche!

Application to $2D$ $U(1)$: with and without a θ-term



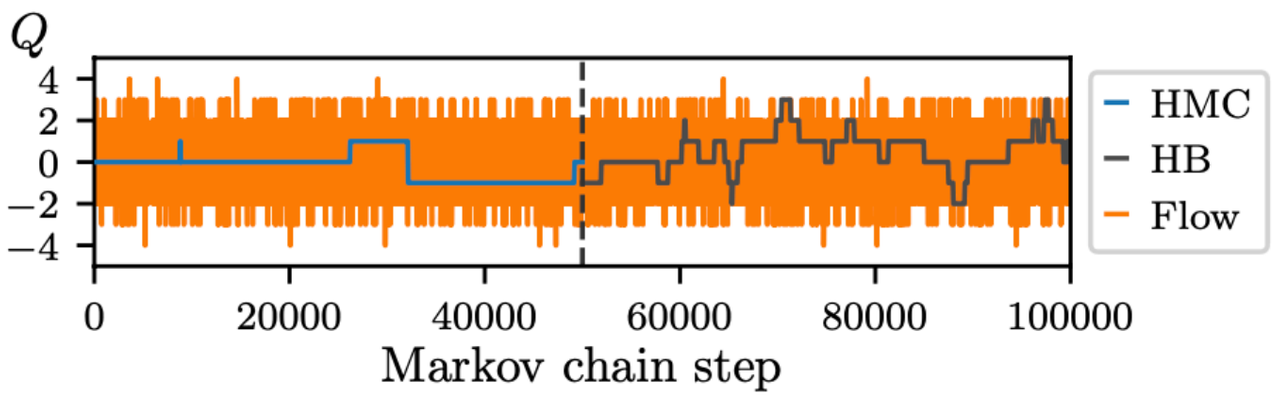
$2D\ U(1)$ Lattice Gauge Theory



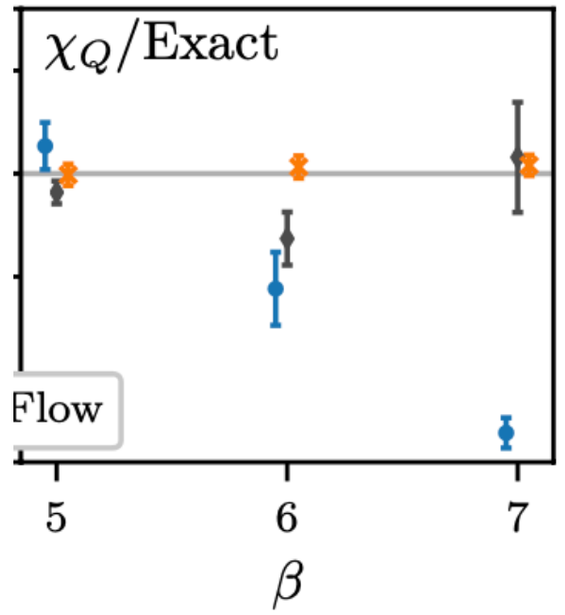
Why 2D U(1) ?

- exactly solvable, with a well defined topological charge, has sign problem with θ-term

- Any advantage beyond NF sampling at high β by decoupling beta?



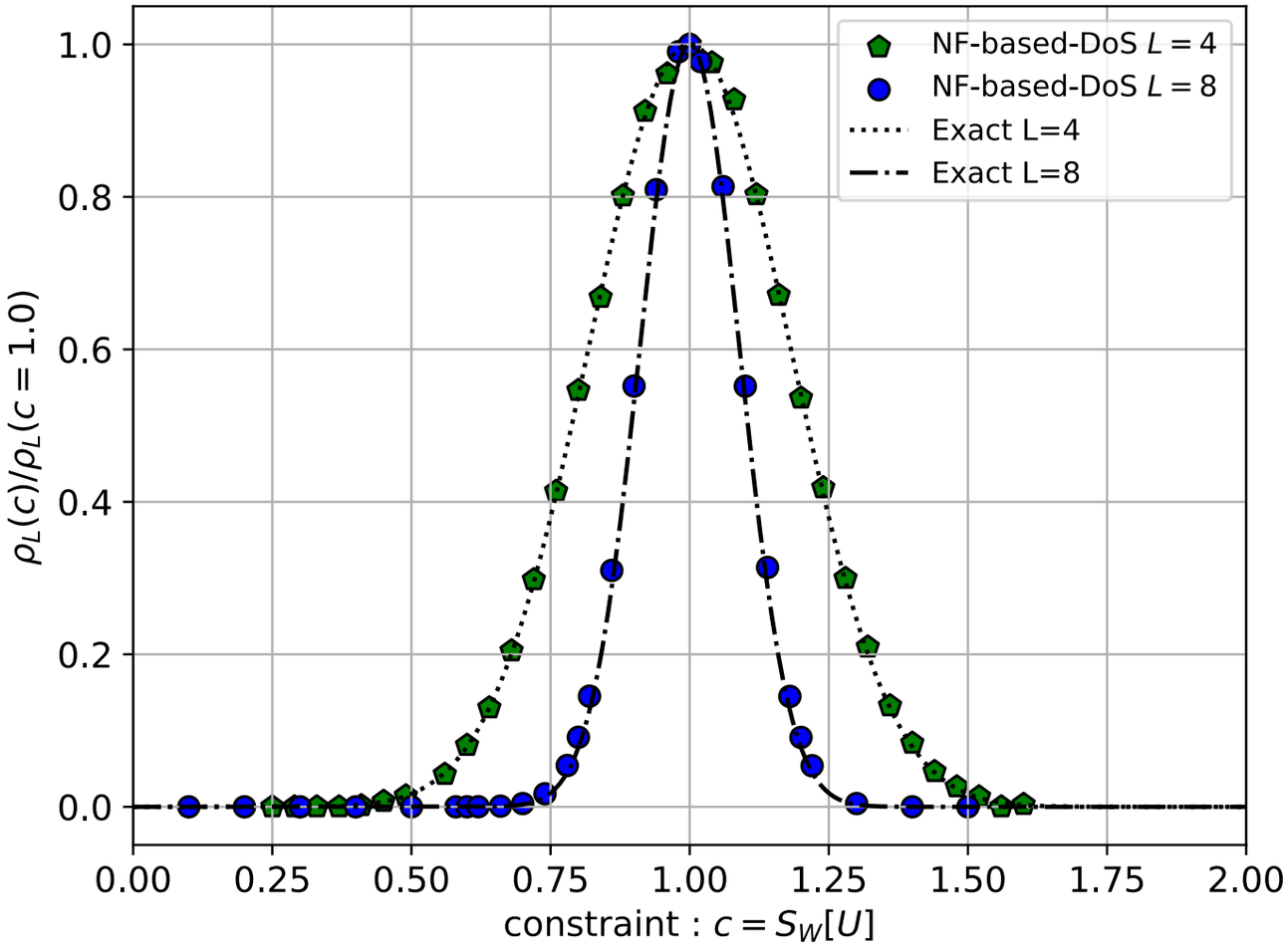
$L = 16, \beta = 7$



[G. Kanwar et. al (2021)], [M. S. Albergo et.al. (2021)

Computing the DoS: 2D U(1) without θ-term

$$S_G[U] = \sum_p \Big(1 - \cos(\phi_p)\Big)/N_p \in [0,2] \qquad S_{c,P} = \frac{P}{2}(c - S_G[U])^2 + \log \mathcal{N}, \quad \rho_P(c) = \int DU\, e^{-S_{c,P}[U]}$$



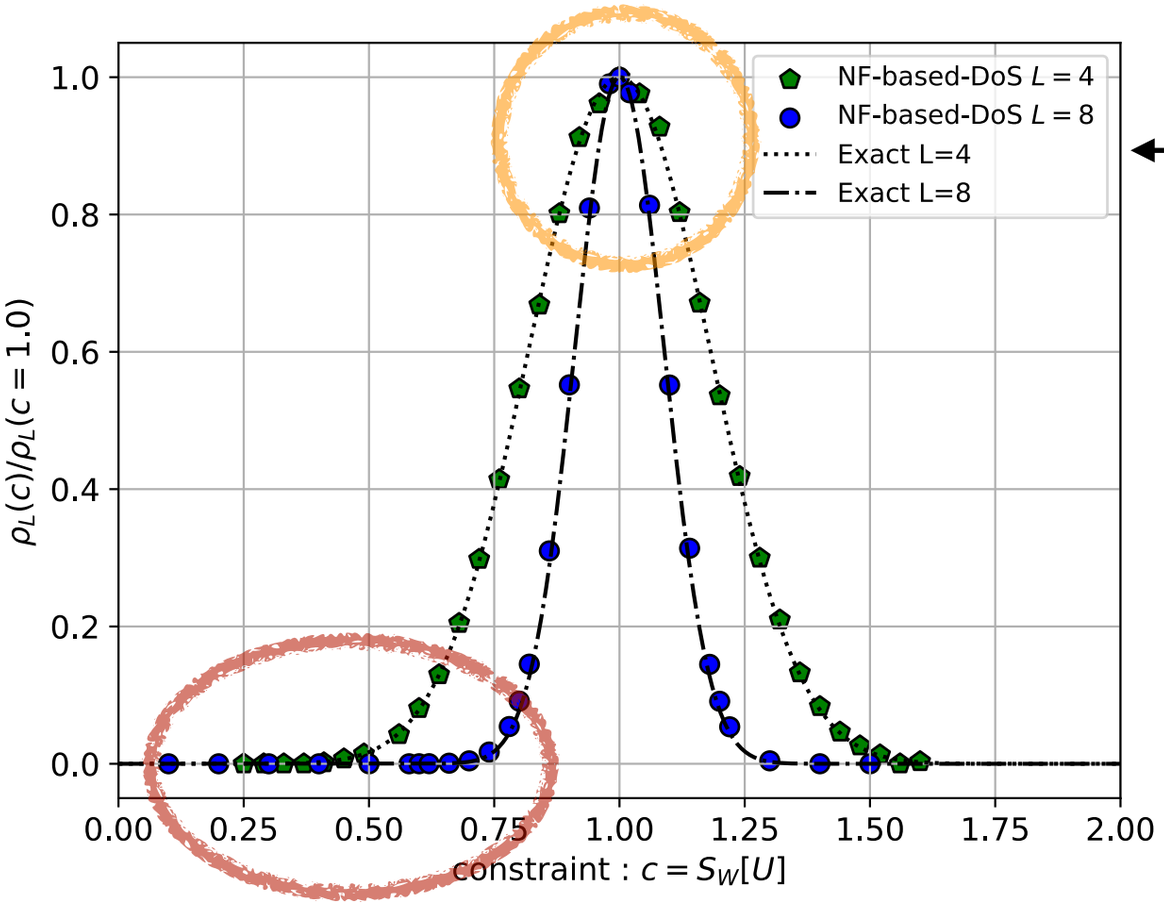
status from last year!

- One trained model per c, per L

- (Qualitatively) reproduced DoS for $L = 4,8$ and $P = 2000$

Code adapted from [M. S. Albergo et.al. (2021)] L=4, Layers: 20, hidden: [8,8], kernel: 5, L=8, double

Computing the DoS: 2D U(1) without θ-term



Effective sample size (ESS):

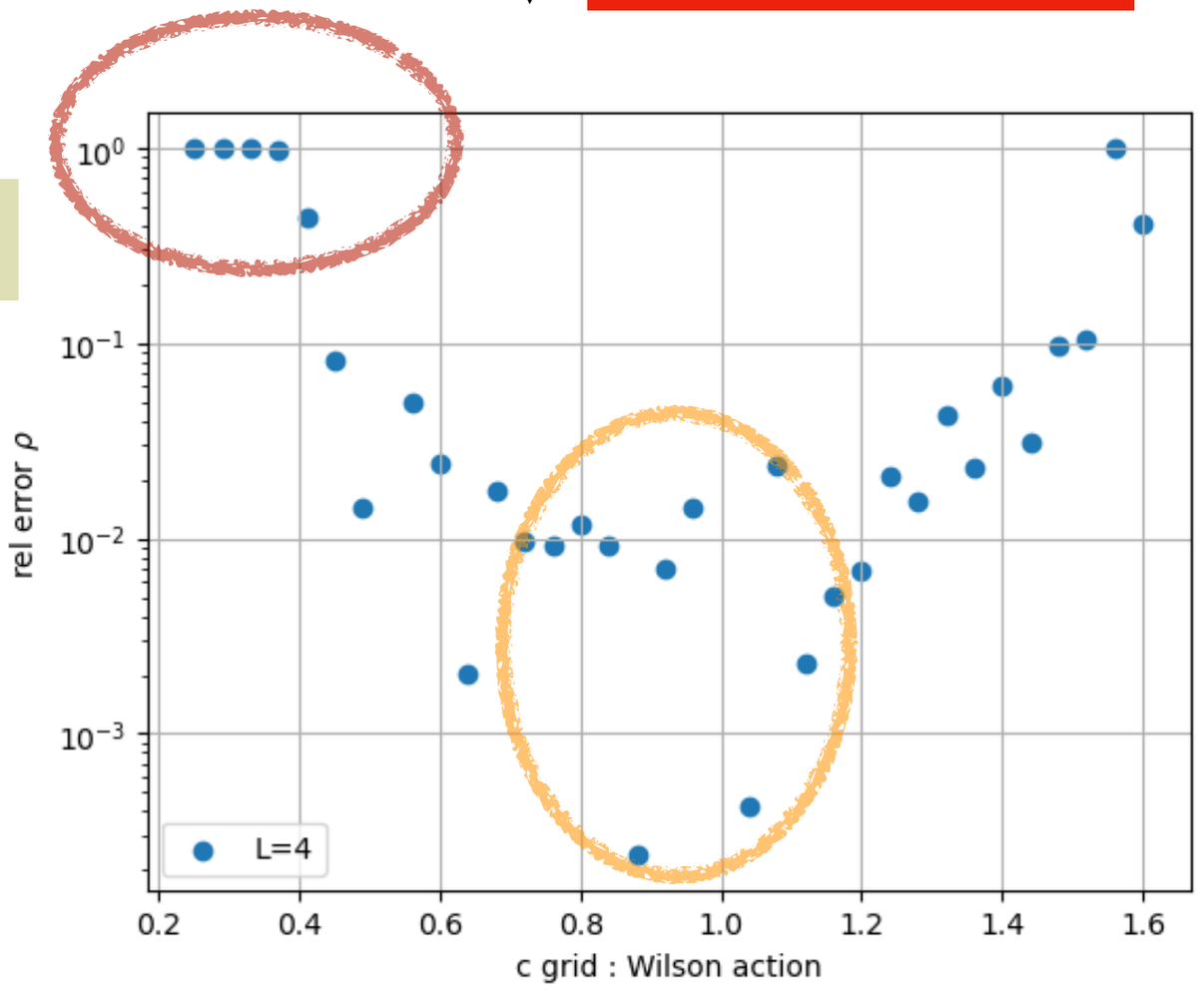
~30% at c = 1.0,

~0.003% at c ≈ 0.2

status from last year!

20 models trained at different $c \in [0.1, 1.5]$

… to access high β we need precise tails



Computing the DoS: Annealing in P + stitching

Main difficulties:

-learning the constrained target $\frac{P}{2}(c-S)^2$, more enhanced for U(1) 2D than Hubbard and ϕ^4 (2D): non-linear constraint?

-loss gradient $P(c-S)$, at $P=1000$, needs extremely good overlap with prior*

New Protocol (still under testing)

Step 1: make the NF conditional with $s/t(\theta_{\text{frozen}}) \rightarrow s/t(\theta_{\text{frozen}}, c, \log P)$

Step 2: anneal in P, at each step sample uniformly from $\log P \in [\log 10, \log P_{\text{cieling}}]$

Step 3: "stitch" intervals via max likelihood across $c \in [0,2]$, using $q(U \mid c, \log P)$ samples*

Computing the DoS: Annealing in P + stitching

Main difficulties:

-learning the constrained target $\frac{P}{2}(c-S)^2$, more enhanced for U(1) 2D than Hubbard and ϕ^4 (2D): non-linear constraint?

-loss gradient $P(c-S)$, at $P=1000$, needs extremely good overlap with prior**

New Protocol (still under testing)

Step 1: make the NF conditional with $s/t(\theta_{\text{frozen}}) \rightarrow s/t(\theta_{\text{frozen}}, c, \log P)$

Step 2: anneal in P, at each step sample uniformly from $\log P \in [\log 10, \log P_{\text{cieling}}]$

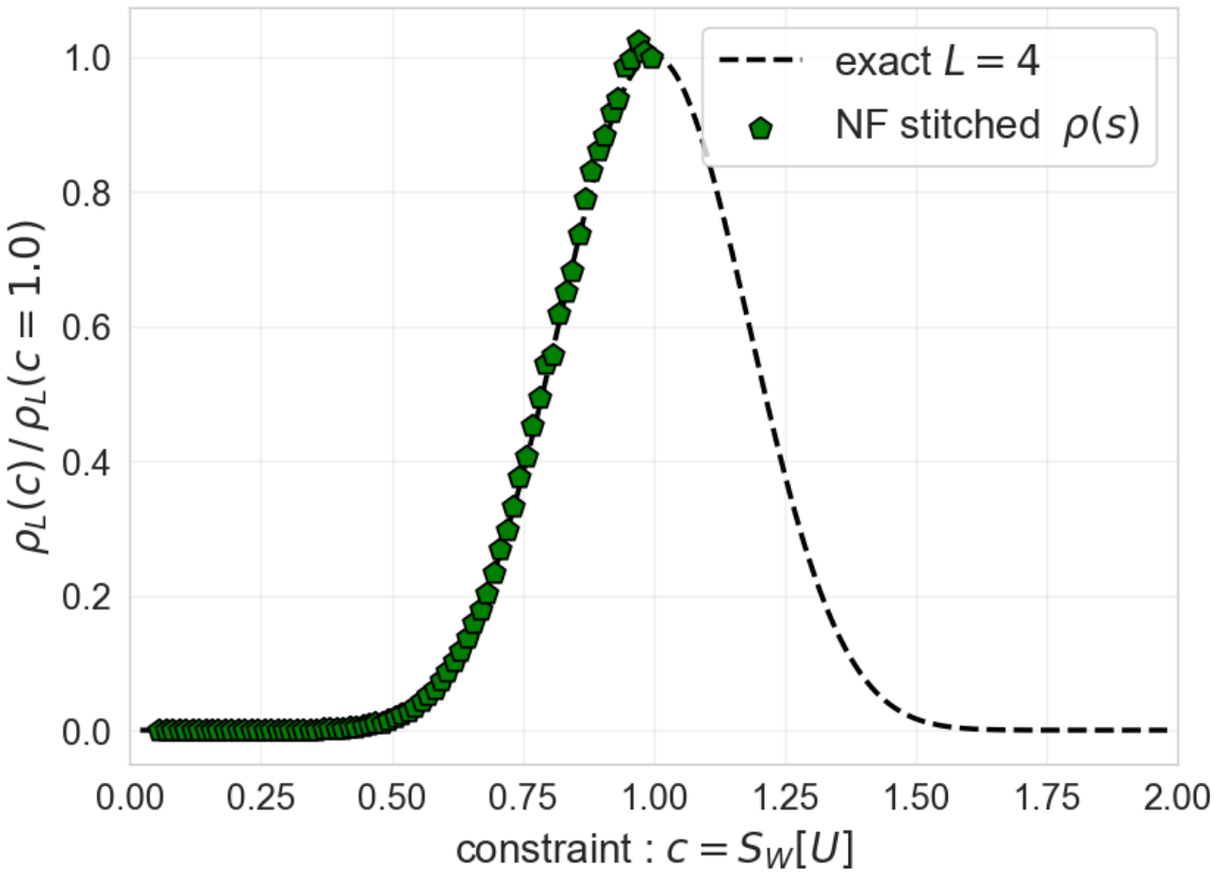
Step 3: "stitch" intervals via max likelihood across $c \in [0,2]$, using $q(U \mid c, \log P)$ samples*

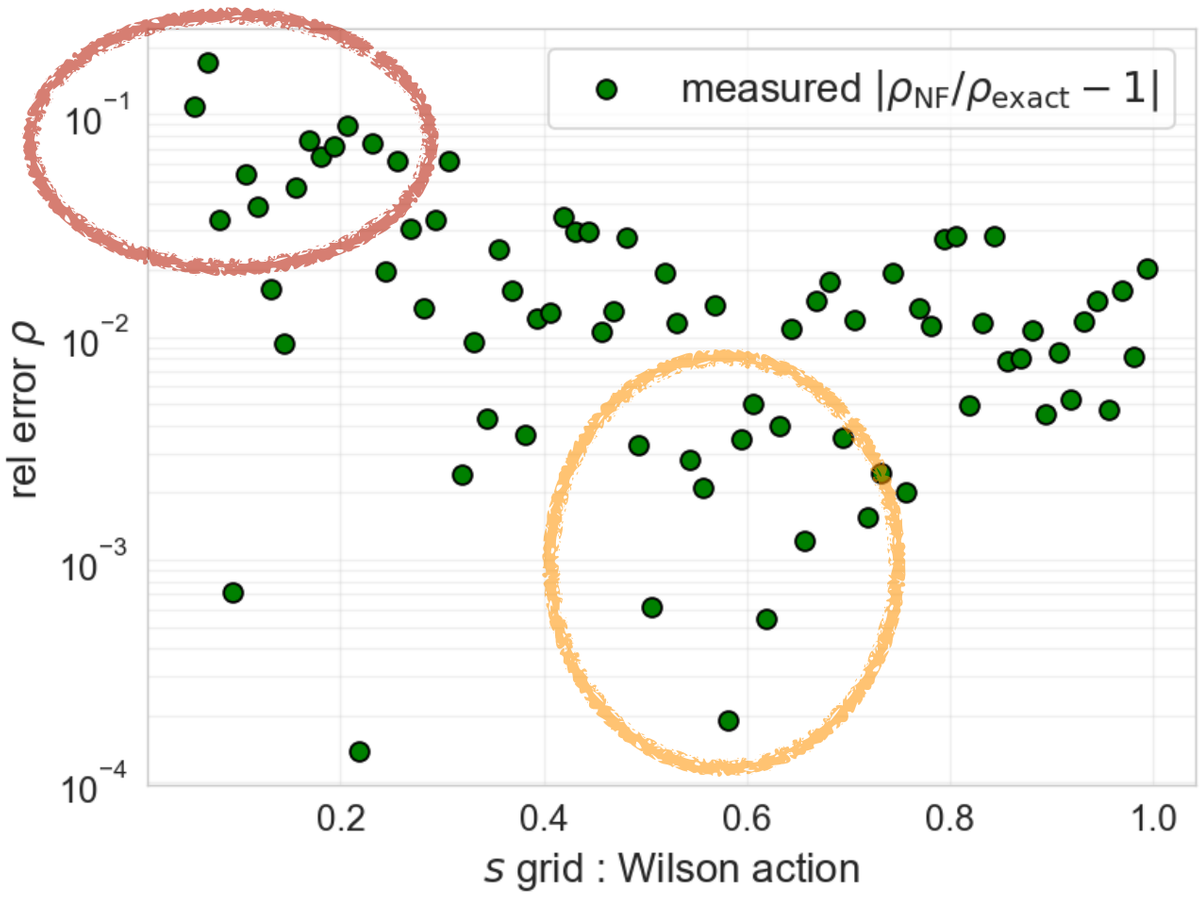
**annealing in c at fixed P=1000 does not work

*implemented using Claude AI

Computing the DoS: Annealing in P + stitching

1 conditional NF model trained at different c : $q(U \mid c, \log P)$





trained on $c \in [0.05, 1.0]$

ESS: ~27% at c = 0.894,
~1.8% at c = 0.05

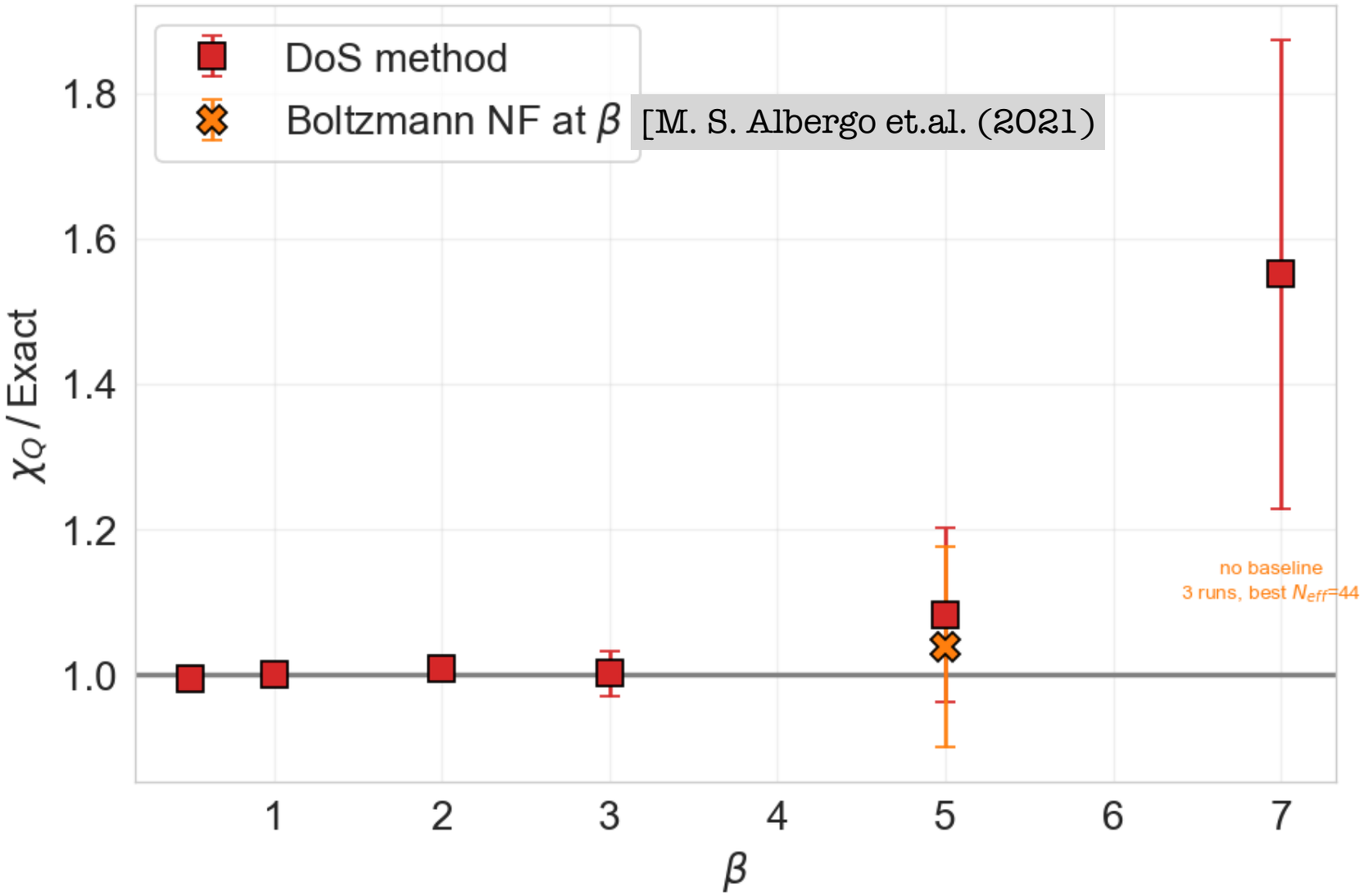
New results !

Can we infer anything meaningful from this?

Computing the DoS: Annealing in P + stitching

DoS route: One (cond) NF for all β compared to NF at fixed β keeping NN fixed

DoS route: Need $\rho(c)$ at even lower than $c < 0.05$ to get $\beta \geq 7$

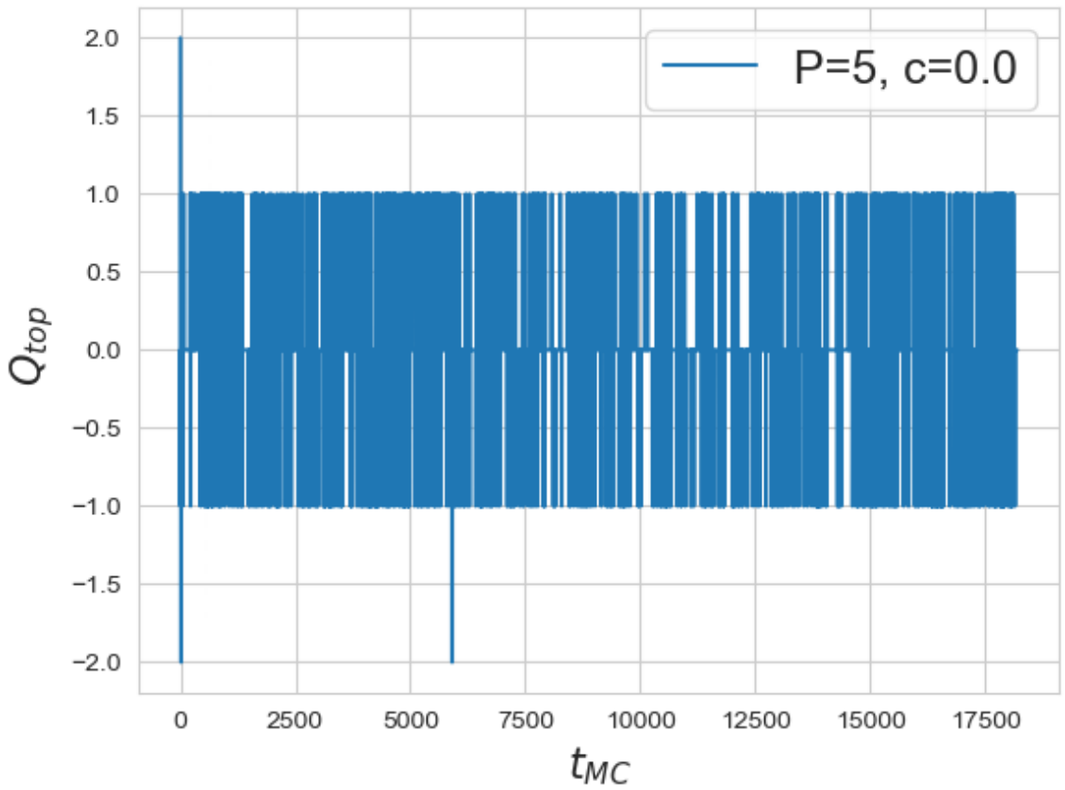


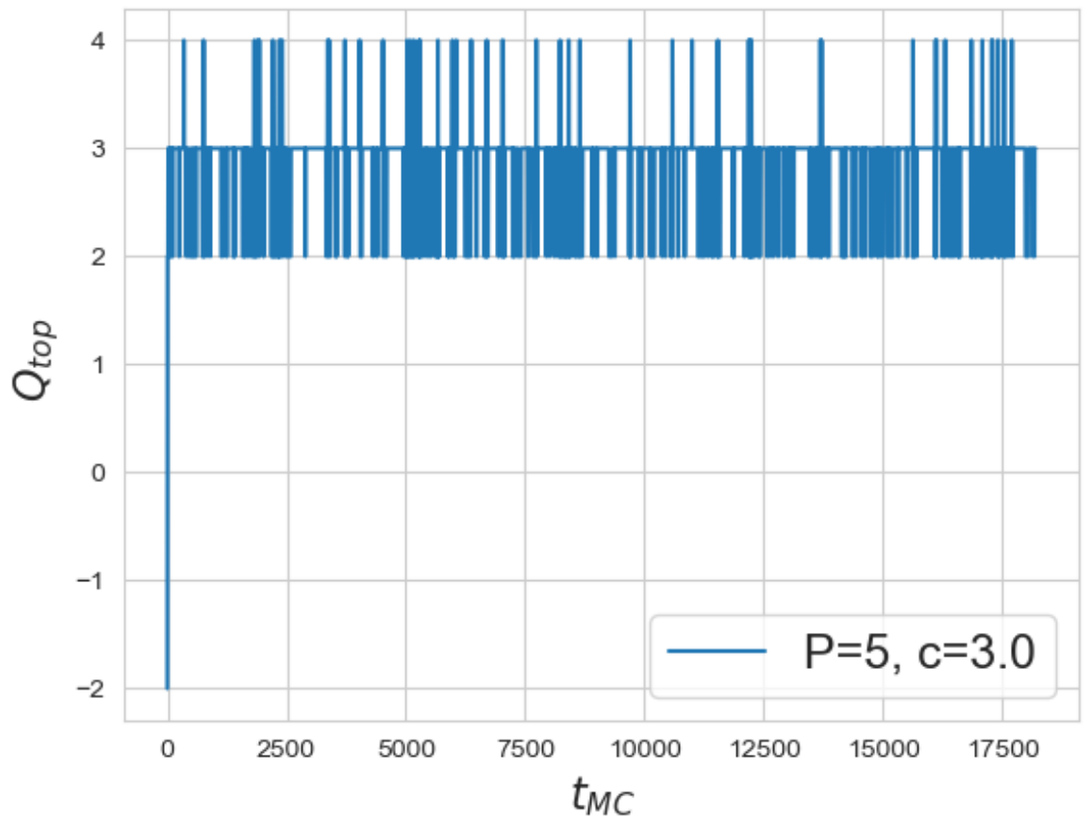
$L = 4$

New results !

Learning the constraints: 2D U(1) with θ–term

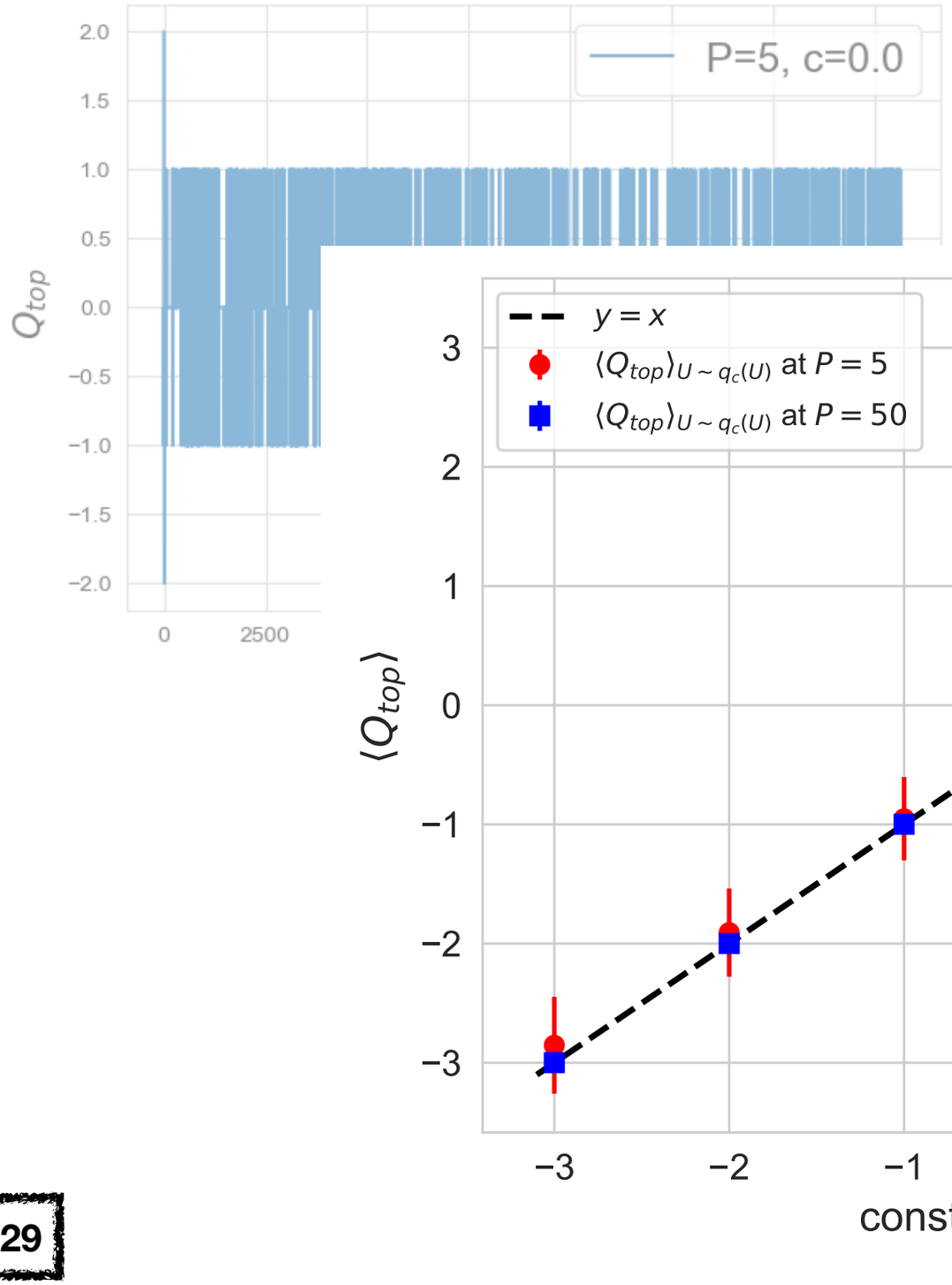
$$S_e[U] = S_{Wilson}[U] + i\theta Q_{top}[U] \qquad S_{c,P}[U] = S_{Wilson}[U] + \frac{P}{2}\left(c - Q_{top}[U]\right)^2 + \log N$$

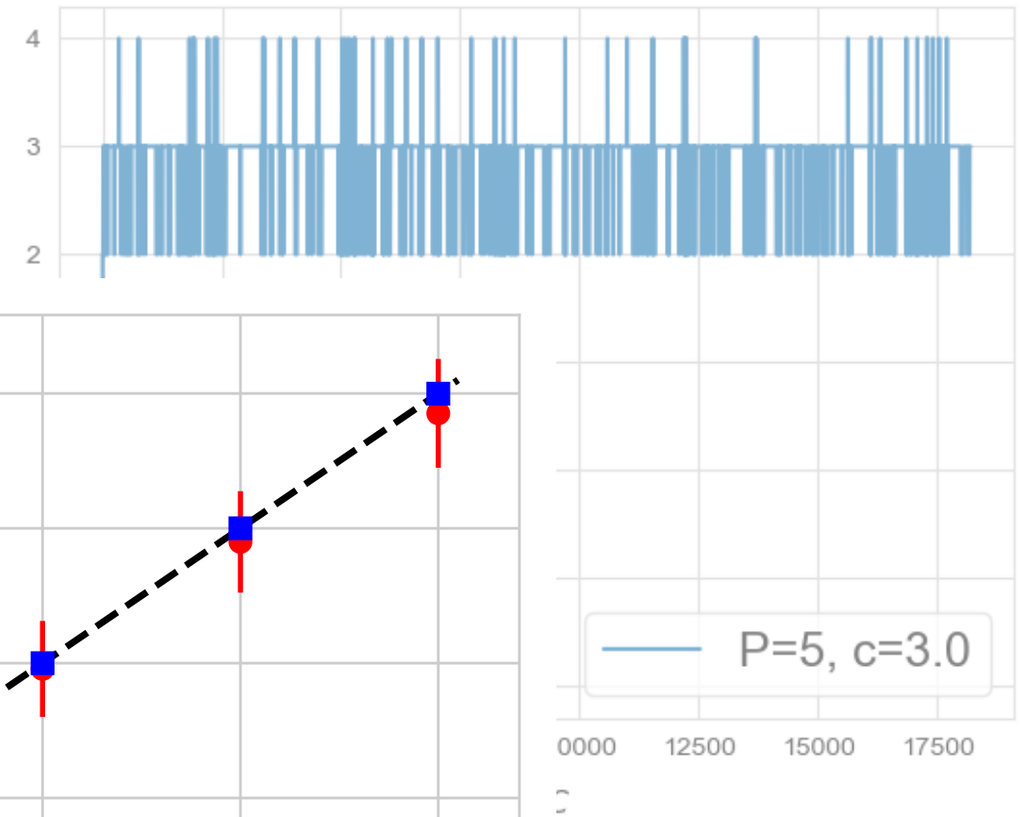


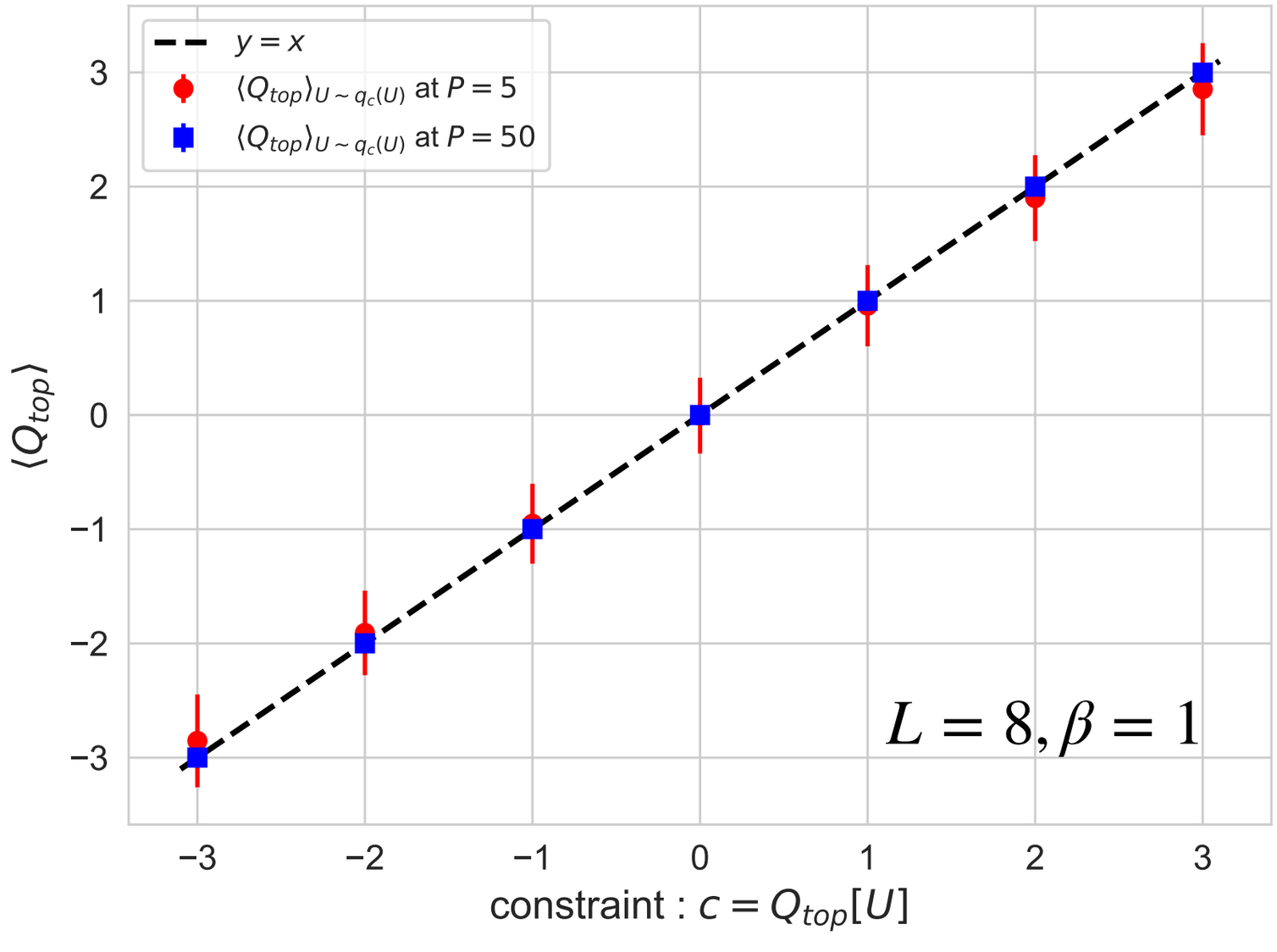


Learning the constraints: 2D U(1) with θ-term

$$S_e[U] = S_{Wilson}[U] + i\theta Q_{top}[U] \qquad S_{c,P}[U] = S_{Wilson}[U] + \frac{P}{2}\left(c - Q_{top}[U]\right)^2 + \log N$$







Summary & outlook

- NF-based DoS method applied to 2D U(1) LGT
  - w/o θ DoS reconstructed with
  - Learned configurations with fixed topological charge
- NF-based DoS to the doped Hubbard model
  - Full scan of partition function, number density in μ, up to $\beta = 100, N_t = 400$ - but $N_x = 2$

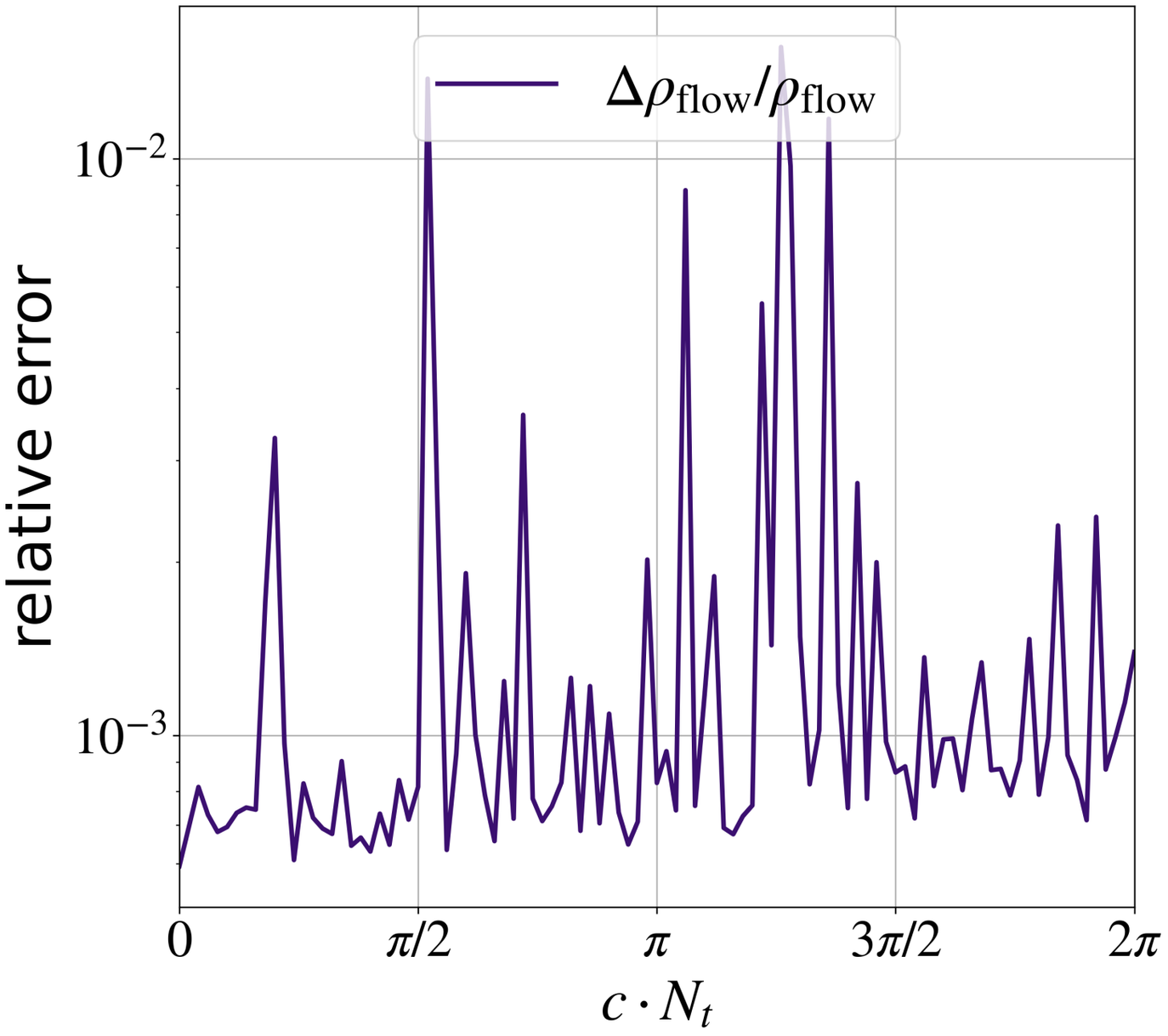
In progress

- Method development for 2D U(1): Technical (architecture based) + conceptual (Gaussian constraint, 2D DoS for θ)
- Beyond two-site Hubbard, beyond real NVP, annealing in β

Thank you for your attention !

BACKUP SLIDES

STATISTICAL ERRORS ON $\beta = 6, N_t = 24$



plot credit: F. Freche

Annealing in P, loss, ess, schedule, stitching

