Normalising flows for all-orders QED corrections in lattice quantum field theory

Nils Hermansson-Truedsson

Based on [hep-lat/2605.22444] in collaboration with G. Kanwar

September 9, 2026



THE UNIVERSITY *of* EDINBURGH



Motivation

- Precision tests of the Standard Model against experiments: Systematic control of the strong interactions $\longrightarrow$ Lattice QCD
- Isospin-breaking effects: Quarks are charged and $m_u \neq m_d$

$$\alpha_{\mathrm{EM}} = \frac{e^2}{4\pi} \sim \frac{m_u - m_d}{\Lambda_{\mathrm{QCD}}} \sim 1\% \qquad \longleftarrow \text{ Needed for precision!}$$

- Many lattice calculations where this is included in QCD+QED, e.g.
  - Hadron masses [BMW 13/15; RM123 13/22; Horsley et al. 16; RBC/UKQCD 17; FNAL/MILC 19; Clark et al. 22; Mainz 24; RC* 22/26]
  - Virtual corrections to leptonic decays: CKM elements [RM123S 15/18/19; RBC/UKQCD 22; Christ et al. 23]
  - Muon $g-2$ [White Paper 20/25 + updates by collaborations]
- Several observables still need improvement for novel precision tests e.g. CKM elements, g_A, $K \to \pi\pi$, ...

Including isospin-breaking corrections

- General observable in full theory QCD+QED

$$\langle \mathcal{O} \rangle_{\mathrm{QCD+QED}} = \frac{1}{Z_{\mathrm{QCD+QED}}} \int_{U,A,\psi,\overline{\psi}} \mathcal{O}\, e^{-S_{\mathrm{QCD+QED}}[U,A,\psi,\overline{\psi}]}$$

⋆ Dynamical approach: Fully sample QCD+QED ensembles

⋆ RM123 approach: Perturbative expansion in parameters around iso-symmetric point [RM123 13]

$$\langle \mathcal{O} \rangle_{\mathrm{QCD+QED}} = \langle \mathcal{O} \rangle_0 + \underbrace{\langle \mathcal{O}\delta S \rangle_0 + \langle \delta \mathcal{O} \rangle_0}_{e^2,(m_u - m_d)/\Lambda_{\mathrm{QCD}}} + \underbrace{\ldots}_{\mathrm{h.o.}}$$

⋆ Automatic differentiation: QCD isospin breaking at aribitrary order [Albandea et al. 26]

⋆ Our work: Normalising flows for including QED corrections

Scalar QED

- First step: Simple theory compared to QCD+QED
- We consider scalar QED: charged scalars ϕ, photons A_μ
- Goal: Introduce QED to the uncoupled theory through flows

$$S(\phi, A; e) = S_0(\phi) + S_A(A) + S_{\text{int}}(\phi, A; e)$$

- The free photon action takes the form [Davoudi et al. 18]

$$S_A(A) = -\frac{1}{2} A_\mu(x) \delta^2 A_\mu(x) \quad \longleftarrow \text{Feynman gauge}$$

$$\delta^2 = \frac{1}{a^2} \sum_\mu (\tau_\mu + \tau_{-\mu} - 2)$$

- For finite-volume QED, we use QED_L but methodology generalises

Scalar QED

- The scalar and interaction part is [Davoudi et al. 18]

$$S_0(\phi) + S_{\text{int}}(\phi, A; e) = \sum_x \phi^\dagger(x)\Delta\phi(x)$$

$$\Delta = m_0^2 + \frac{1}{a^2}\sum_\mu \left(2 - e^{ieaA_\mu}\tau_\mu - \tau_{-\mu}e^{-ieaA_\mu}\right) = \Delta_0 + \sum_{n=1}^{\infty} e^n \, \Delta^{(n)}$$

- The uncoupled Δ_0 defines the scalar free-theory action $S_0(\phi)$
- The interaction part then takes the form

$$S_{\text{int}}(\phi, A; e) = \sum_{n=1}^{\infty} e^n \, S_{\text{int}}^{(n)}(\phi, A)$$

Normalising flows

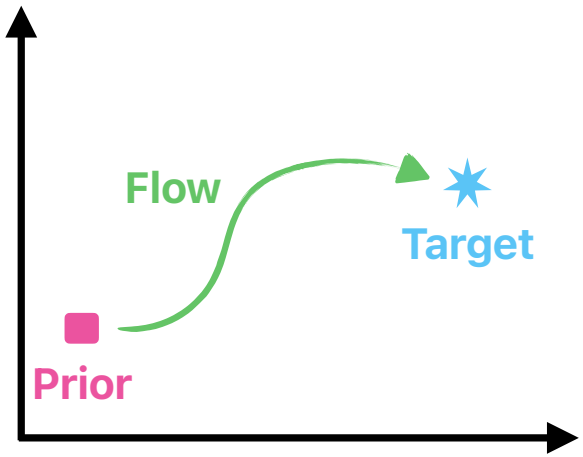
- Normalising flows discussed in many talks here
- We have the uncoupled prior theory

$$r(\phi, A) = \frac{1}{Z_0} e^{-S_0(\phi) - S_A(A)} \qquad \longleftarrow \text{ Prior}$$

- Factorised in ϕ/A, and easy to sample from (Gaussian)
- The target distribution for the flow is the full theory

$$p(\phi, A) = \frac{1}{Z_e} e^{-S_0(\phi) - S_A(A) - S_{\text{int}}(\phi, A; e)} \qquad \longleftarrow \text{ Target}$$

- Sample configurations ϕ, A according to prior
- Employ the flow on them to reach target
- Note: Target is gauge fixed



Normalising flows

- Define the normalising flow f, a diffemorphism such that

$$f : (\phi, A) \mapsto (\phi', A')$$

- The new fields follow the distribution

$$q(\phi', A') = r(\phi, A) \left| \det J_f \right|^{-1}$$

- Goal: Optimise flow f to get q close to the target p
- In practice: no flow is perfect $\longrightarrow$ reweighting for unbiased estimators

$$\hat{w}(\phi', A') = \frac{e^{-S_0(\phi') - S_A(A') - S_{\text{int}}(\phi', A'; e)}}{e^{-S_0(\phi) - S_A(A)}} \left| \det J_f \right|$$

- Observables are then given by

$$\langle \mathcal{O} \rangle_e = \frac{Z_e}{Z_0} \left\langle \mathcal{O}(\phi', A')\, \hat{w} \right\rangle_0 = \left\langle \mathcal{O}(\phi', A')\, w \right\rangle_0$$

Normalising flows: Removing order-e noise

- For many observables we have

$$\langle \mathcal{O} \rangle_e = \langle \mathcal{O} \rangle_0 + e^2 \, \langle \mathcal{O} \delta_{e^2} S_{\text{int}} \rangle_0 + e^4 \, \langle \mathcal{O} \delta_{e^4} S_{\text{int}} \rangle_0 + \dots$$

- At fixed statistics, there is residual noise at odd orders: $e, e^3, e^5, \dots$
- All-orders methods: Usually simulated at several values of e

  [BMW 15; RBC/UKQCD 18; RC* 22, 25]

- Flows: Residual odd-order noise in $\langle \mathcal{O} \rangle_e$ automatically improves when the flow learns the target action
- Use the flow f to define f_- with $-e$: An improved estimator

$$\langle \mathcal{O} \rangle_e = \frac{1}{2} \big\langle \mathcal{O}(\phi', A') w(\phi', A') + \mathcal{O}(\phi'_-, A'_-) w_-(\phi'_-, A'_-) \big\rangle$$

- Perfect-flow limit: Removes all residual effects

Normalising flows: Machine learned

- We consider two kinds of flows: Machine learned and linear (analytic)
- Residual-flow architecture [Chen et al. 19]

$$f: \quad \phi' = \phi + \varepsilon\, b_\phi(\phi, A) \qquad A' = A + \varepsilon\, P_{\mathrm{QED_L}}[b_A(\phi, A)]$$

- b_ϕ / b_A: UNet, convolutional networks, 1 hidden layer, 32 channels
- UNet: to capture short- and long-range features
- Training: Kullback-Leibler divergence + path gradients [Vaitl et al. 22]
- Flow is observable independent
- We will explore the optimal flow choice in the future

Normalising flows: Linear

- The scalar QED Lagrangian is very simple at order e

$$S_{\text{int}}^{(1)}(\phi, A) = \sum_{x,\mu} A_\mu(x)\, j_\mu(x) = \sum_{x,\mu} A_\mu(x)\, \frac{2}{a}\, \text{Im}[\phi^\dagger(x)\tau_\mu\phi(x)]$$

- Can define an analytic linear flow matching the target at leading order

$$\phi'^{\text{Lin.}}(x) = \phi(x), \qquad A'^{\text{Lin.}}_\mu(x) \equiv A_\mu(x) + eP_{\text{QED}_\text{L}}\left[(\delta^2)^{-1} j_\mu(x)\right]$$

- Jacobian of the linear flow has

$$\det J^{\text{Lin.}} = 1$$

- This would not directly generalise to fermionic theories, but it lets us test a cheap but imperfect flow

Observable

- We compute the scalar 2-point function at $\mathbf{p} = \mathbf{0}$

$$C_e(t) = \frac{1}{\hat{L}^{N_d-1}} \sum_{\mathbf{x}} \langle \phi(t, \mathbf{0})\, \phi^\dagger(0, \mathbf{0}) \rangle_e$$

- N_d is the number of dimensions
- This allows to extract the scalar mass shift through the correlator ratio

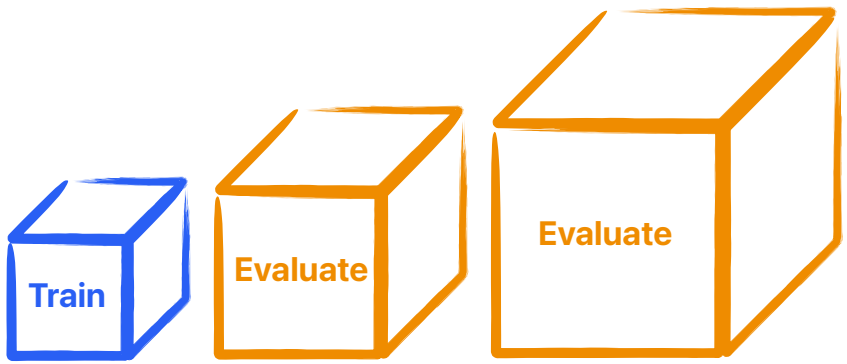
$$\frac{C_e(t)}{C_0(t)} \approx \frac{A_e}{A_0} e^{-\Delta m_e t} + \dots$$

- This can be studied in the lattice data as

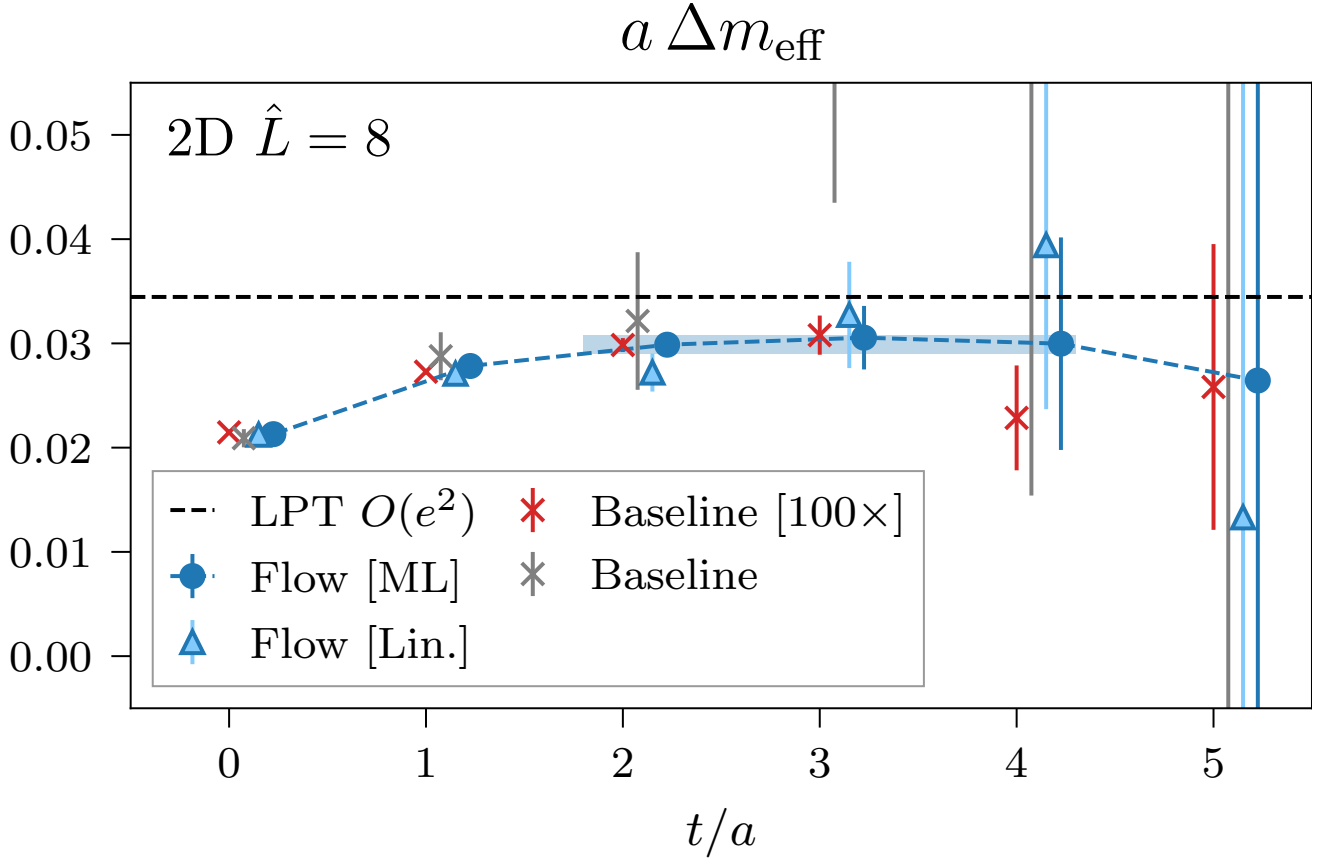
$$-\partial_t \log \left[\frac{C_e(t)}{C_0(t)} \right] = \Delta m_{\text{eff}} + \dots$$

Numerical setup

- We study $N_d = 2, 3, 4$ dimensions
- Parameters $m_0 = 1$ and $e = 0.3$ everywhere
- ML training: $\hat{L}_{\text{train}} = 8$ ($N_d = 2, 3$) or $\hat{L}_{\text{train}} = 4$ ($N_d = 4$)
- Evaluate models on range of larger volumes with $\hat{L} \leq 4 \times \hat{L}_{\text{train}}$
- Methods: Flows (ML/Linear); Baseline (all-orders stoch. reweighting)
- We also compare to lattice perturbation theory at order e^2

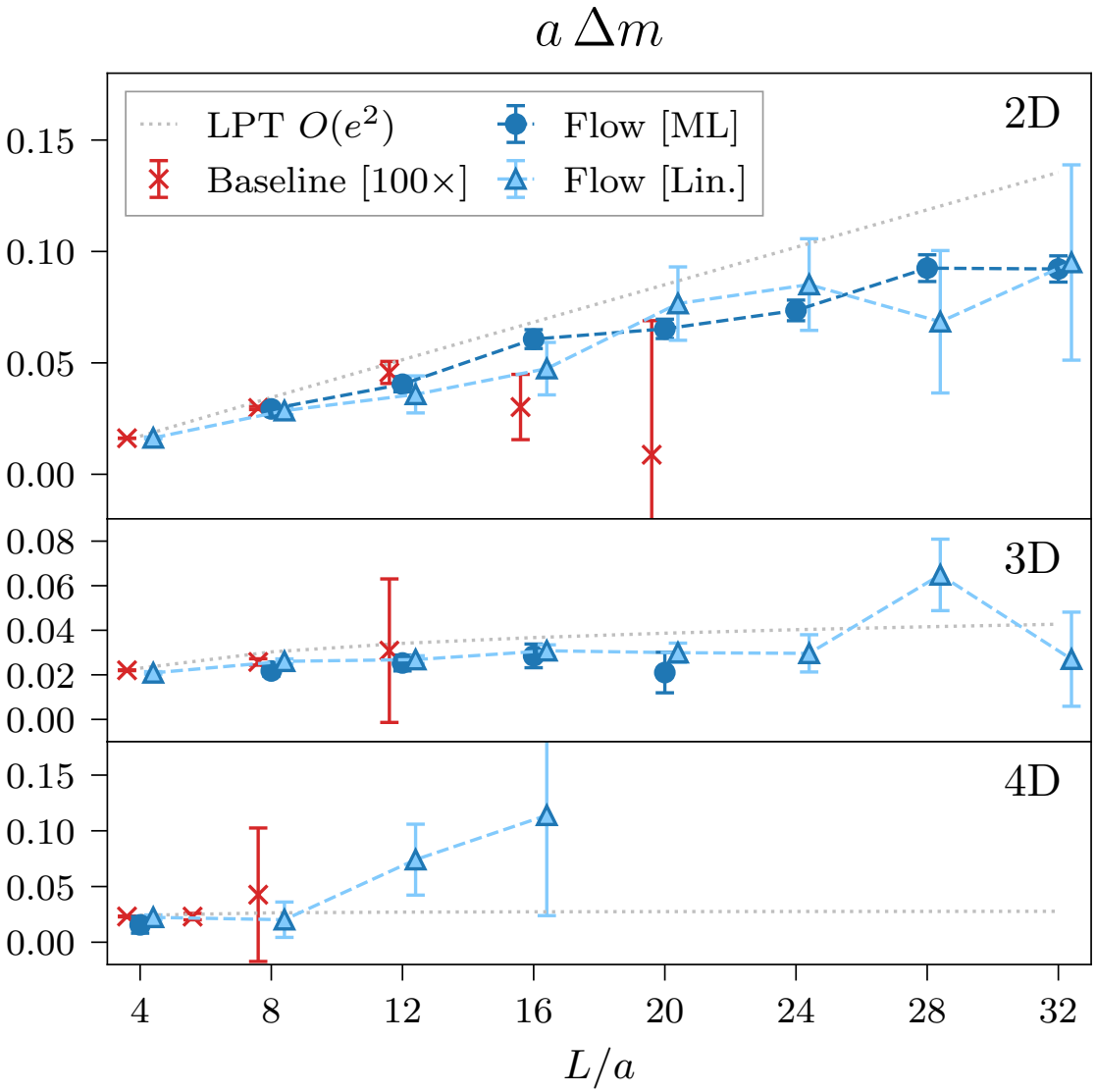


Numerical results: Effective mass shift



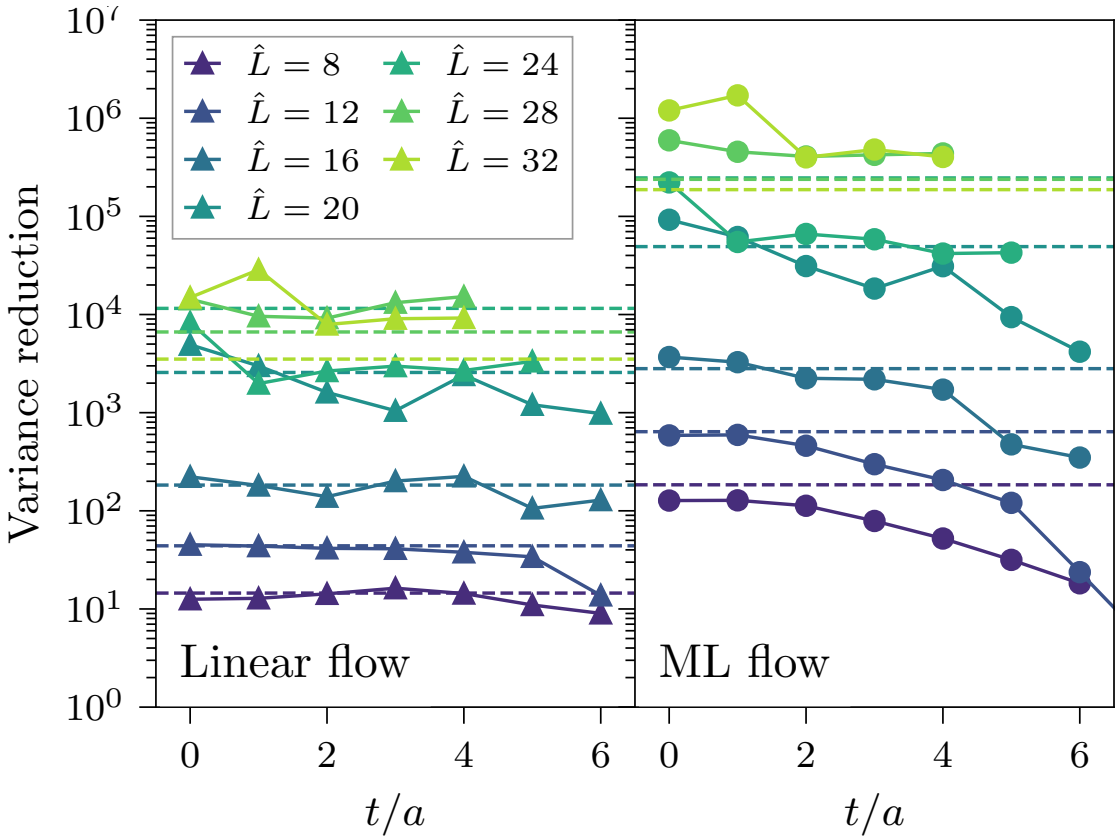
- Compare flows and baseline: 10^5 samples
- StN problem at large t (expected), but clear hierarchy in variance
- At 100 times more statistics, baseline becomes comparable to flows
- In comparison to LPT: Resolves e^4 dependence

Numerical results: Volume transfer



- Volume transfer of trained models
- Fitted mass shift. Agreement with other methods
- Linear flows: Perform well where we do not have ML flows

Numerical results: Variance improvement

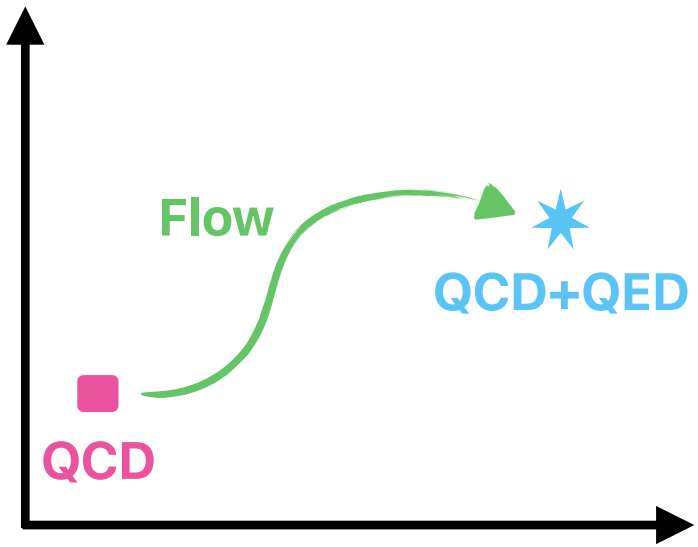


- Variance in C_e/C_0 in 2D relative to baseline
- Improves with the volume
- Dashed horizontal lines: Improvement in $\log Z_e/Z_0 = \log\langle \hat{w} \rangle_0$
- Before StN: Governs the variance improvement

$$\langle \mathcal{O} \rangle_e = \langle \mathcal{O}(\phi', A')\, w \rangle_0 = \frac{Z_e}{Z_0} \langle \mathcal{O}(\phi', A')\, \hat{w} \rangle_0$$

Conclusions and outlook

- Lattice scalar QED with normalising flows
- Uncoupled ($e = 0$) $\xrightarrow{\text{flow}}$ all-orders in e
- Observable-independent flow: Volume transfer and variance reduction
- Future questions:
  - ∗ Fixed simulation parameters here: Can we learn the dependence?
  - ∗ QCD/Fermion theories: How well do flows perform?
  - ∗ Many applications in QCD+QED, e.g. neutron physics, QED-disconnected diagrams, ...



Backup

Finite-volume QED

- Prescriptions to define QED in a finite volume:

⋆ $\mathrm{QED_M}$: Photon mass m_γ

[Endres et al. 16; Bussone et al. 18]

⋆ $\mathrm{QED_{SF}}$: Bound zero-mode fluctuations

[Göckeler et al. 90/97]

⋆ $\mathrm{QED_\infty}$: QED part in infinite volume

[Asmussen et al. 16; Feng, Jin 18; Christ et al. 23]

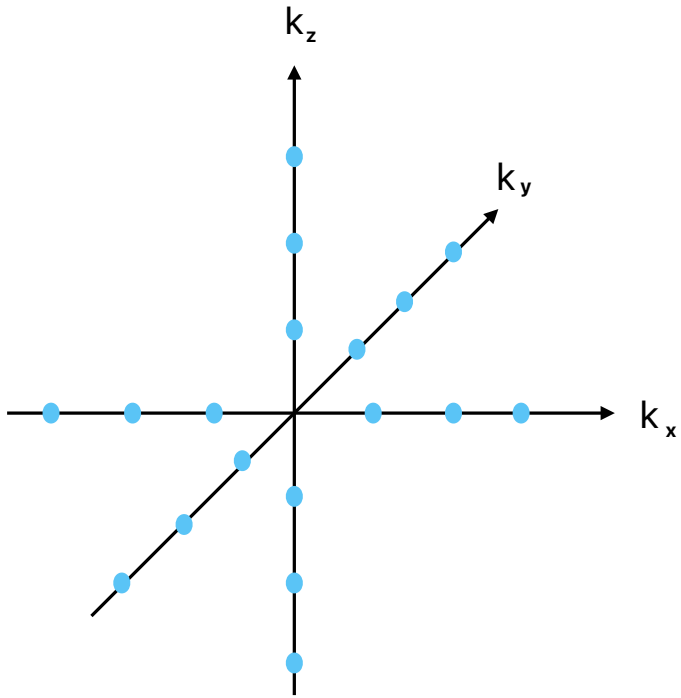
⋆ $\mathrm{QED_C}$: C-conjugated boundary conditions

[Kronfeld, Wiese 91–93; Lucini et al. 15; RC* 19]

⋆ $\mathrm{QED_{L/r}}$: Exclude/redistribute zero mode

[Davoudi et al. 19]

  - $\mathrm{QED_L}$: Exclude [Hayakawa, Uno 08]
  - $\mathrm{QED_r}$: Redistribute [Di Carlo, Hansen, NHT, Portelli 25]



- We consider $\mathrm{QED_L}$ here: would be interesting to explore others

Normalising flows: Machine learned

- Our flow f transforms to the distribution q

$$\log q(\phi', A') = -S_0(\phi) - S_A(A) - \log\det J_f - \log Z_0$$

- The target density is

$$\log p(\phi, A) = -S_0(\phi) - S_A(A) - S_{\text{int}}(\phi, A; e) - \log Z_e$$

- Optimised with Kullback-Leibler divergence + path gradients [Vaitl et al. 22]

$$\mathcal{L}_{kl} = \left\langle \log q(\phi', A') - \log p_{e^2}(\phi', A') - \log(Z_e/Z_0) \right\rangle_0$$

- We train models on a set of different lattices, and then evaluate them on others with larger volume $\longrightarrow$ volume transfer

- Jacobian evaluated through

$$\log\det J = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \operatorname{Tr}\left[J^k\right]$$

Lattice perturbation theory

- Compare to lattice perturbation theory at fixed order e^2

  [Davoudi et al. 19; Bijnens, NHT, Jüttner, Portelli et al. 19]

$$\delta_{e^2} C_e(t) = \frac{1}{T} \sum_{p_0} \frac{\Sigma(p_0, \mathbf{0})}{(\hat{p}_0^2 + m^2)^2} e^{ip_0 t}$$

- Calculate as an FFT
- One-loop self energy

$$\Sigma(p) = \frac{e^2}{TL^{N_d-1}} \sum_{k_0} \sum_{\mathbf{k}\neq\mathbf{0}} \left\{ \frac{N_d - 1 + \cos(ap_0)}{\hat{k}^2} - \frac{\widehat{2p-k}^2}{\hat{k}^2 \left(\widehat{p-k}^2 + m_0^2 \right)} \right\}$$



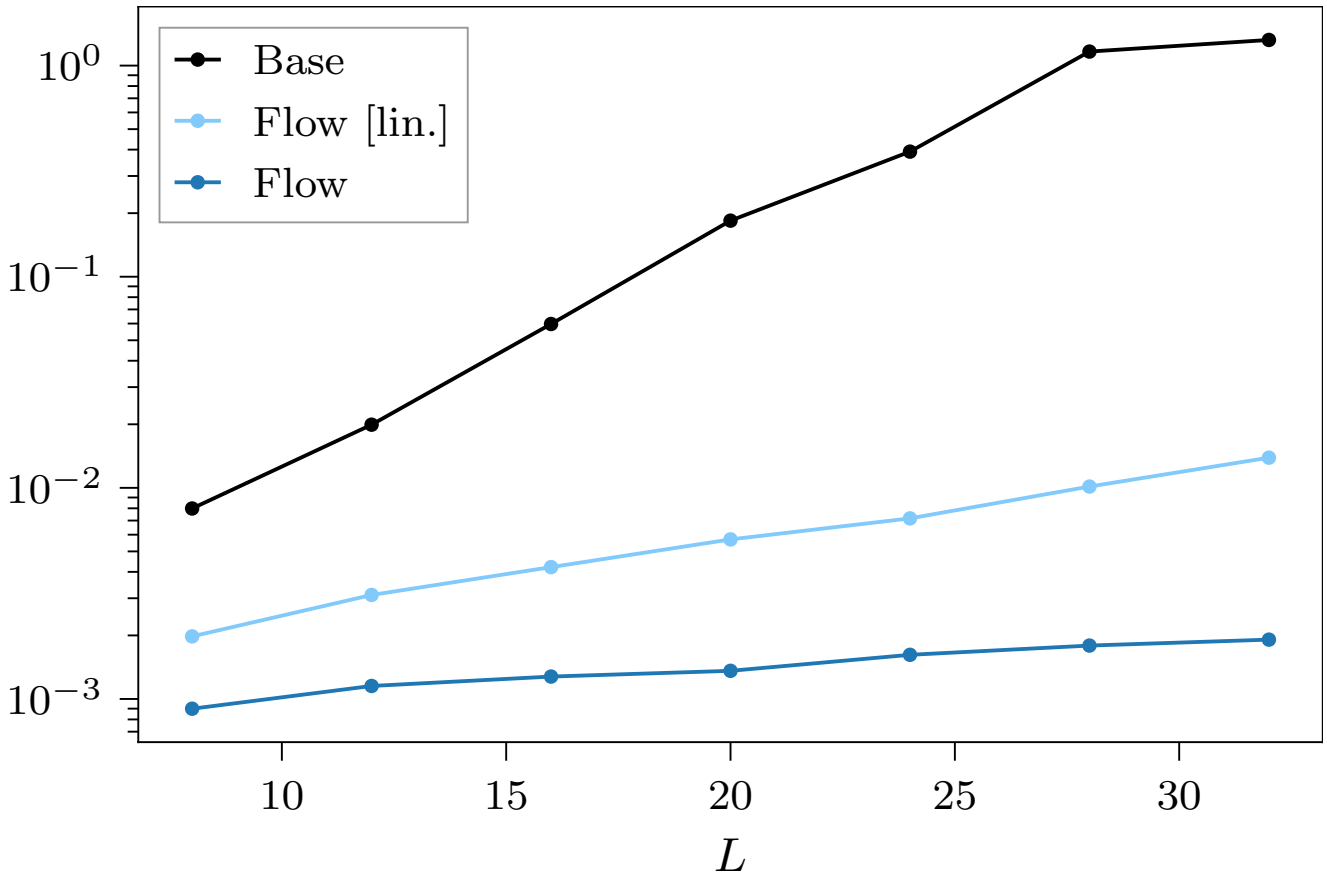
Lattice perturbation theory

- Works well for low N_d and one loop
- Number of terms are $\sim \hat{T}\,\hat{L}^{N_d-1}$: Difficult for higher N_d/loops
- Numerical approach for the sums [Bijnens, NHT, Jüttner, Portelli et al. 19]

$$\frac{1}{TL^{N_d-1}} \sum_{k_0} \sum_{\mathbf{k}\neq 0} f(k_0, \mathbf{k}) = \int_{-1}^{1} d^{N_d}x \, \frac{\mathcal{R}_{\mathrm{L}}(\lfloor \hat{L}\mathbf{x}/2 \rfloor)}{(2a)^{N_d}} \, f\left(\frac{\pi}{a}\frac{2}{\hat{T}} \left\lfloor \frac{\hat{T}}{2} x_0 \right\rfloor, \frac{\pi}{a}\frac{2}{\hat{L}} \left\lfloor \frac{\hat{L}}{2} \mathbf{x} \right\rfloor \right)$$

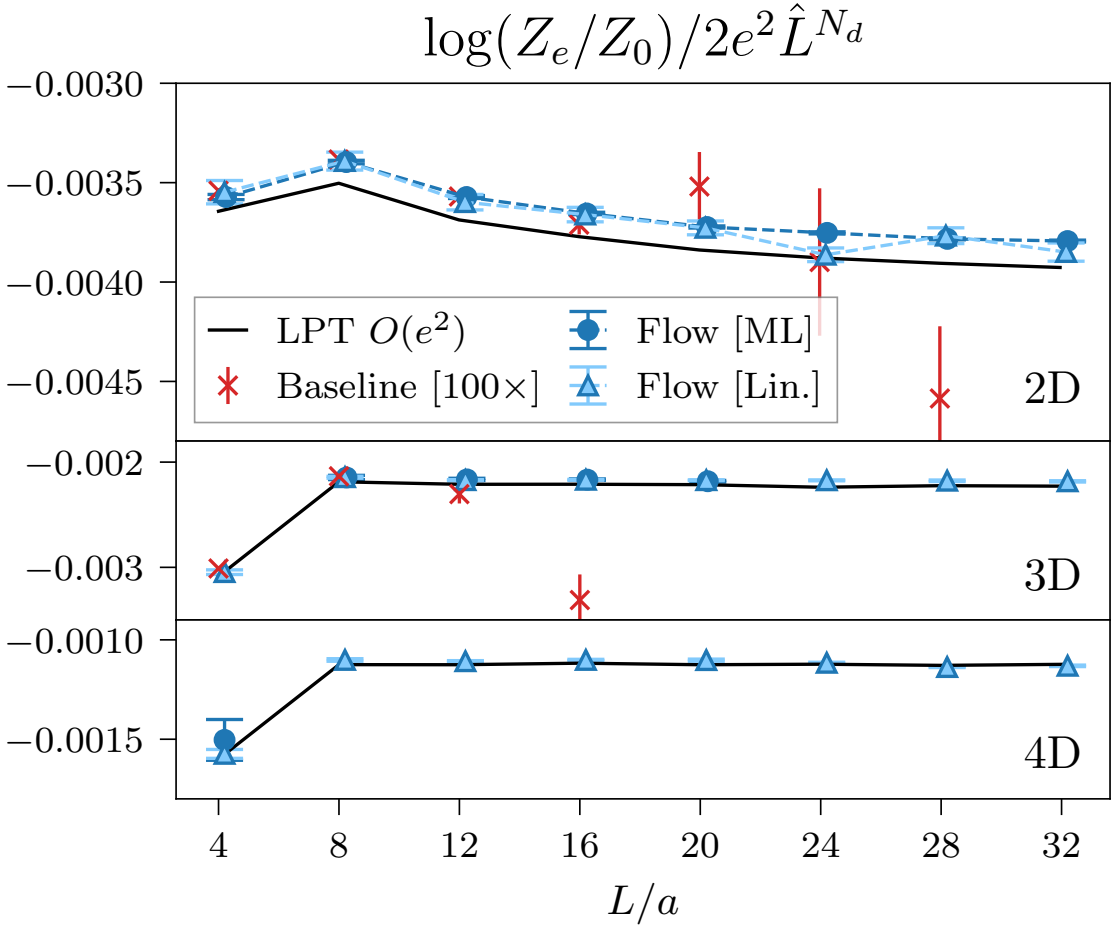
- Do integral with VEGAS Monte Carlo [Lepage 77/20]
- 10^6 Monte Carlo samples: uncertainties negligible here

Numerical results



- Error in ratio $C_e(t)/C_0(t)$

Numerical results



- Comparison of the log partition function ratio, $\log Z_e/Z_0$