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

Nils Hermansson-Truedsson

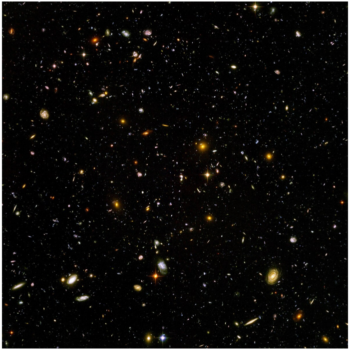
Based on [hep-lat/2605.22444] in collaboration with G. Kanwar

August 19, 2026

THE UNIVERSITY of EDINBURGH

Motivation

- Standard Model
- Best theory of particle physics
- Very successful
- Universe: Clear signs that we need something beyond
- ⋆ Dark matter
- ⋆ Matter/antimatter asymmetry

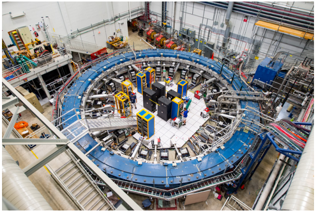


Hubble Ultra Deep Field

There is new physics to discover: How?

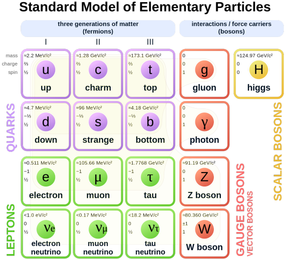
Precision tests

- Precision tests of the Standard Model: theory and experiment
- Low energy: Many interesting observables [Flavour Lattice Averaging Group (NHT)]
  My focus: decays of e.g. pions/kaons/neutrons, EDMs
- Control of strong force QCD + electromagnetism QED for % precision



Experiment Fermilab

=



Theory prediction

$\underbrace{+\quad ?}_{\text{Unknown}}$

Lattice QCD+QED

- Simulate quantum field theory with action $S_{\mathrm{QCD+QED}}$

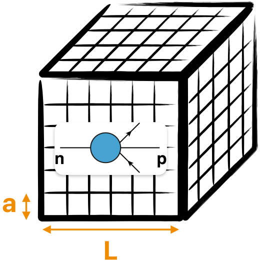
$$\underbrace{\langle \mathcal{O} \rangle_{\mathrm{QCD+QED}}}_{\text{Decay process, ...}} = \int_{\text{fields}} \mathcal{O}\, \frac{e^{-S_{\mathrm{QCD+QED}}}}{Z_{\mathrm{QCD+QED}}} = \int_{\text{fields}} \mathcal{O}\, p_{\mathrm{QCD+QED}}$$

- In essence: Sampling problem from $p_{\mathrm{QCD+QED}} \sim e^{-S_{\mathrm{QCD+QED}}}$
- Traditionally: Monte Carlo. Machine learning can help
- Discretise theory and spacetime on grid of finite extent

⋆ Discretisation: a

⋆ Finite volume: L

⋆ Physics: $L \to \infty$, $a \to 0$, ...



QCD+QED in practice

- Parametric dependence on e^2 in QCD+QED

$$\langle \mathcal{O} \rangle_{\mathrm{QCD+QED}} \sim e^2,\, e^4,\, e^6, \ldots$$

- How to determine the expectation value?

1. Fixed-order e^2: Expansion around QCD ($e^2 \sim 0.09$) [RM123 13]

$$\langle \mathcal{O} \rangle_{\mathrm{QCD+QED}} \approx \langle \mathcal{O} \rangle_{\mathrm{QCD}} + e^2 \langle \delta\mathcal{O} \rangle_{\mathrm{QCD}}$$

   Standard for most calculations but more challenging operator and cost: difficult for full e^2 answer [RC* 25] and neutron decays [Gorchtein et al. 23]

2. All-orders approach: Fully sample QCD+QED [BMW 15; RC* 25] but difficult because unphysical $e, e^3, \ldots$ noise to remove

3. Our work: normalising flows for all-orders QED at reduced variance

Normalising flows

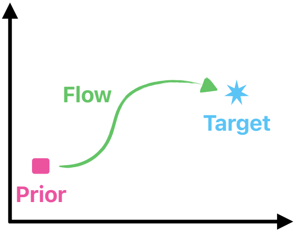
- Normalising flows have received much attention in the lattice community: Transform between distributions

  Lattice@CERN 26: Talks by Tomiya, Kanwar, Romero Lopez, NHT, Nada, Ramos, Shanahan

- First step: Consider scalar QED: charged scalar fields ϕ, photons A_μ
- Action in scalar QED: $S_0(\phi, A) + S_{\text{int}}(\phi, A; e)$
- Start: Uncoupled prior ($e = 0$)

$$p_0(\phi, A) = \frac{1}{Z_0} e^{-S_0(\phi, A)}$$

  Benefit: Gaussian, easy to sample



- Goal: Coupled target

$$p_e(\phi, A) = \frac{1}{Z_e} e^{-S_0(\phi, A) - S_{\text{int}}(\phi, A; e)} \longrightarrow \langle \mathcal{O} \rangle_e = \int \mathcal{O}\, p_e(\phi, A)$$

Normalising flows

- Define the normalising flow f, a diffemorphism:

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

- The new fields follow the distribution

$$q(\phi', A') = p_0(\phi, A)\,|\det J_f|^{-1}$$

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

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

- Target observables are then given by reweighting

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

Normalising flows

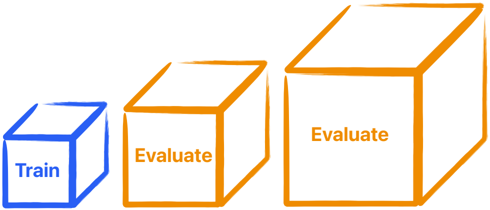
- We consider two kinds of flows: Machine learned and linear
- Residual-flow architecture [Chen et al. 19] (to be further explored)

$$f: \quad \phi' = \phi + \varepsilon\, b_\phi(\phi, A) \qquad A' = A + \varepsilon\, b_A^L(\phi, A)$$

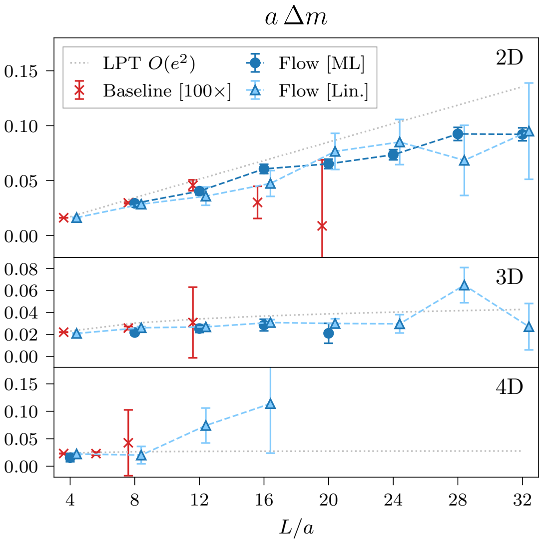
- b_ϕ / b_A^L: UNets to capture short- and long-range features
- Observable-independent training: Kullback-Leibler divergence to learn the full target action $\rightarrow$ Contains all orders in e
- Perfect-flow limit: Removes all unphysical order $e, e^3, e^5, \ldots$ in $\langle \mathcal{O} \rangle_e$
- Major benefit: Observable-independent flowed approach automatically reduces unwanted noise

Numerical setup

- We compute the mass shift due to QED in $N_d = 2, 3, 4$
- Methods: Flows (ML/Linear); Baseline (all-orders stoch. reweighting)
- ML training: $\hat{L}_{\rm train} = 8$ ($N_d = 2, 3$), 4 ($N_d = 4$)
- Major benefit: Volume transfer, evaluate models up to $4 \times \hat{L}_{\rm train}$

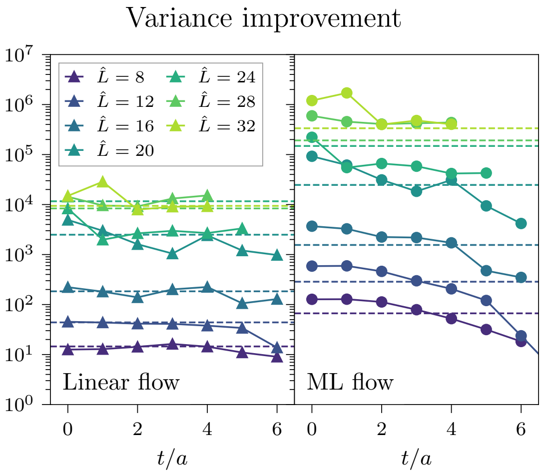


Numerical results



- Mass shift: Flows vs baseline
- Clear hierarchy in variance
- Great volume transfer of trained models: baseline too noisy
- Comparison to lattice perturbation theory: Flows resolve e^4

Numerical results

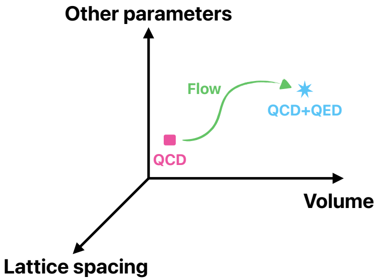


- Where does the improvement originate?
- Mass extracted from time-dependence in C_e/C_0
- Variance in C_e/C_0 in 2D relative to baseline
- Dashed horizontal lines: Improvement in $\log Z_e/Z_0 = \log\langle\hat{w}\rangle_0$
- This is what governs the variance improvement, before the dip

Conclusions and outlook

- Proof-of-concept study: 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? Reuse QCD distributions
  * Many applications, e.g. neutron physics at PSI, ESS, ...



Backup

Scalar QED

- We consider scalar QED: charged scalars ϕ, photons A_μ

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

- The actions take the form [Davoudi et al. 18]

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

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

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

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

- For finite-volume QED, we use $\mathrm{QED_L}$ but methodology generalises

Normalising flows: Linear

- The scalar QED Lagrangian is very simple at order e

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

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

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

- Jacobian of the linear flow has

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

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

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

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

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

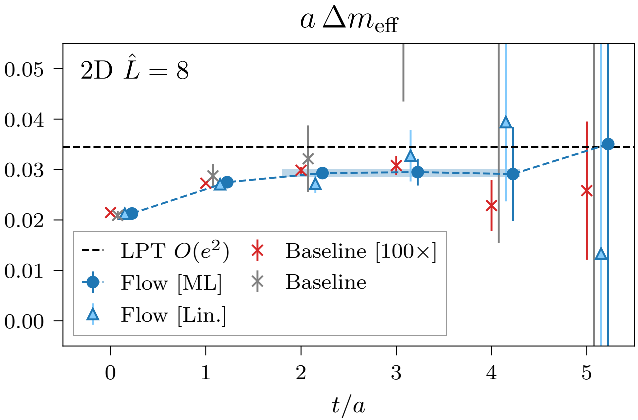
$$\mathcal{L}_{kl} = \left\langle \log q(\phi', A') - \log p_e(\phi', A') - \log(Z_e/Z_0) \right\rangle_0$$

- Jacobian evaluated through

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

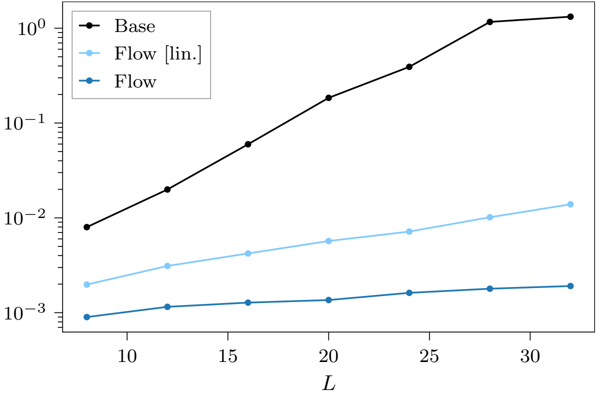
- We use UNets, convolutional networks, 1 hidden layer, 32 channels

Numerical results: Effective mass shift



- Compare flows and baseline: 10^5 samples
- Clear hierarchy in variance
- At 100 times more statistics, baseline becomes comparable to flows
- In comparison to LPT: Resolves e^4 dependence

Numerical results



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

$$C_e(t) = \frac{1}{\hat{L}^{N_d-1}} \sum_{\mathbf{x}} \langle \phi(t,\mathbf{0})\, \phi^\dagger(0,\mathbf{0}) \rangle_e\,, \quad C_0(t) = \frac{1}{\hat{L}^{N_d-1}} \sum_{\mathbf{x}} \langle \phi(t,\mathbf{0})\, \phi^\dagger(0,\mathbf{0}) \rangle_0$$

Lattice perturbation theory

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

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

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

- Calculate as an FFT
- One-loop self energy (exactly or with $\textsc{Vegas}$ Monte Carlo [Lepage 77/20])

$$\Sigma(p) = \frac{e^2}{T L^{N_d - 1}} \sum_{k_0} \sum_{\mathbf{k} \neq \mathbf{0}} \left\{ \frac{N_d - 1 + \cos(a p_0)}{\hat{k}^2} - \frac{\widehat{2p - k}^2}{\hat{k}^2 \left(\widehat{p - k}^2 + m_0^2 \right)} \right\}$$