Applying polynomial-filtering to mass-preconditioned HMC

Taylor Haar

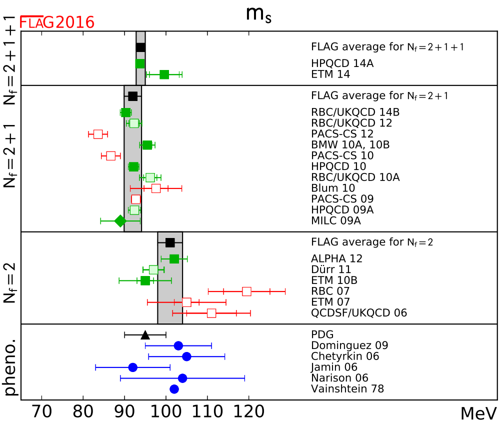
Collaborators: Waseem Kamleh, Yoshifumi Nakamura, James Zanotti

University of Adelaide

International Nuclear Physics Conference, 2016

Motivation

- Recent lattice QCD simulations are reaching the physical quark masses, and giving 2% errors for many quantities



Motivation

$$\mathcal{O}\left(\boxed{1}\right)\quad \mathcal{O}\left(\boxed{2}\right)\quad \mathcal{O}\left(\boxed{3}\right)$$

$$\mathcal{O}\left(\boxed{4}\right)\quad \mathcal{O}\left(\boxed{5}\right)\quad \mathcal{O}\left(\boxed{6}\right) \longrightarrow \boxed{\langle \mathcal{O} \rangle = \frac{1}{N}\sum_{i=1}^{N} \mathcal{O}[U_i, \psi_i, \overline{\psi}_i]}$$

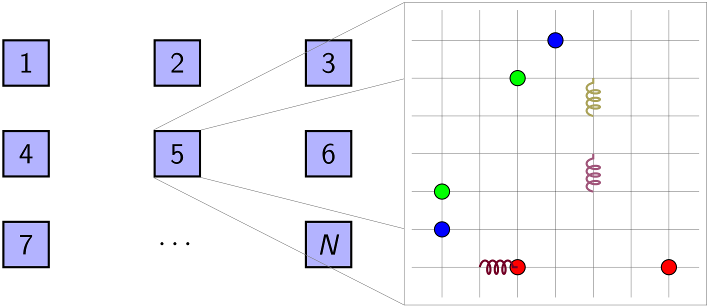
$$\mathcal{O}\left(\boxed{7}\right)\quad \cdots \quad \mathcal{O}\left(\boxed{N}\right)$$

- To measure observables on the lattice, we take an average of the observable over a large number of configurations distributed according to the Boltzmann distribution $\exp(-S[U, \psi^{(f)}, \overline{\psi}^{(f)}])$.
- Generating sufficiently many independent configurations takes several months or years $\implies$ Find algorithmic improvements

Outline

1. Lattice QCD
   - Hybrid Monte Carlo
   - Filtering methods
2. Results

Configurations



- Each configuration is given by the state of the lattice gauge field U and the quark fields $\psi^{(f)}$, which describe the gluons and quarks in the QCD vacuum.

Hybrid Monte Carlo

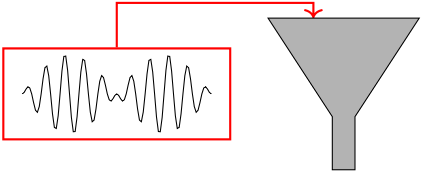
- Use Hybrid Monte Carlo (HMC) to generate the configurations $(U_i, \psi_i^{(f)})$
- Successive states are produced via a molecular dynamics integration that approximately conserves the Euclidean action, which, in the case of two degenerate quarks, is given by

$$S = S_G[U] + S_F[U, \psi^{(f)}, \overline{\psi}^{(f)}] = S_G[U] + \phi^\dagger (D^\dagger D)^{-1} \phi.$$

- The main cost of Hybrid Monte Carlo is in the repeated inversion of the fermion matrix $K[U, m] = D^\dagger D$, which describes the fermion interactions between lattice sites.

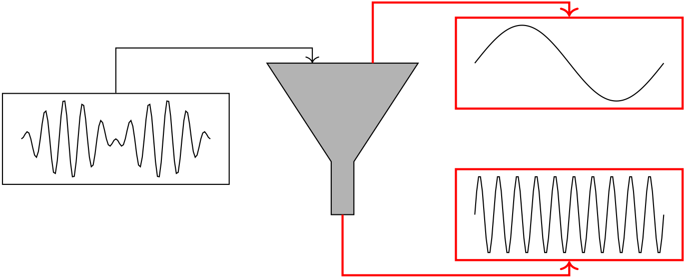
Filtering methods

- The issue with K is its wide range of energy/frequency scales.
  - Large energies = need a small integration step size to capture
  - Small energies = large condition number, so harder to invert
- Filtering methods separate these components, which can then be dealt with on separate integration scales and hence improve the computational cost.



Filtering methods

- The issue with K is its wide range of energy/frequency scales.
  - Large energies = need a small integration step size to capture
  - Small energies = large condition number, so harder to invert
- Filtering methods separate these components, which can then be dealt with on separate integration scales and hence improve the computational cost.



Filtering methods

Filtering methods filter the fermion action $S_F = \phi^\dagger K^{-1}\phi$:

Mass preconditioning (MP) (a.k.a. Hasenbusch preconditioning)

- Use a fermion matrix $J = K[U, m']$ with heavier quark mass $m' > m$

$$S_{MP} = \phi_1^\dagger J^{-1}\phi_1 + \phi_2^\dagger J K^{-1}\phi_2$$

- Parametrized by the mass parameter m'

Polynomial filtering (PF)

- Use a polynomial approximation $P(K)$ to the inverse

$$S_{PF} = \phi_1^\dagger P(K)\phi_1 + \phi_2^\dagger [P(K)K]^{-1}\phi_2$$

- Parametrized by the polynomial order p

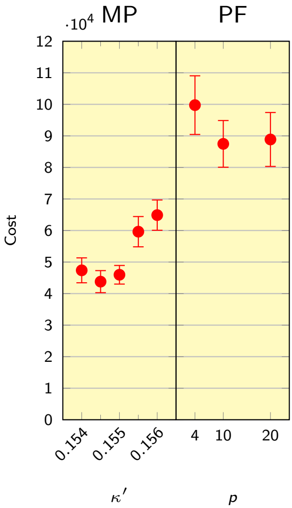
Investigating filtering methods

- We tested these methods out on a $16^3 \times 32$ lattice with two degenerate Wilson fermions, pion mass $m_\pi \sim 400$ MeV.
- Mass preconditioning is tuned using the hopping parameter κ', which is inversely related to the (single) mass parameter m'.

Goal

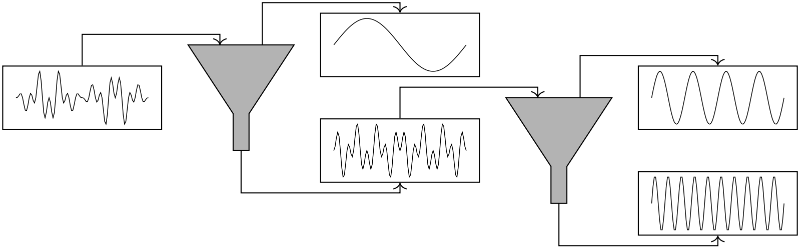
- Want to minimize the cost = number of K matrix-vector multiplications to generate an independent configuration.
- Can vary κ' and p to minimize the cost.

Cost of HMC with filtering: single filter



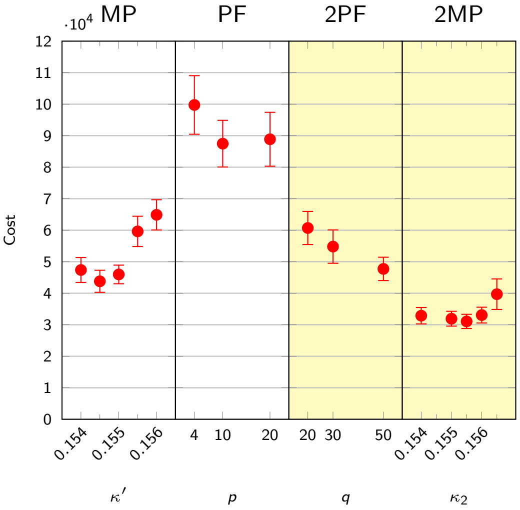
Multiple filters

- The obvious next step is to use multiple filters



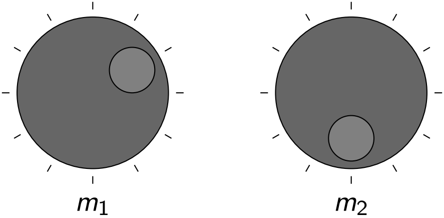
- This decomposes the energy distribution further, which can improve performance.

Cost of HMC with filtering: two filters



Mixed filters

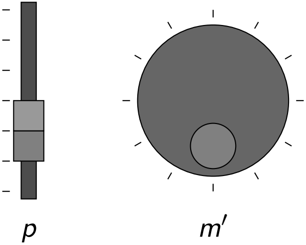
- Using multiple mass preconditioners, like 2MP, is one of the most common filtering methods.
- Problem: have to fine-tune the mass parameters m_1 and m_2 for optimal performance



- Have to generate several configurations to test each (m_1, m_2)

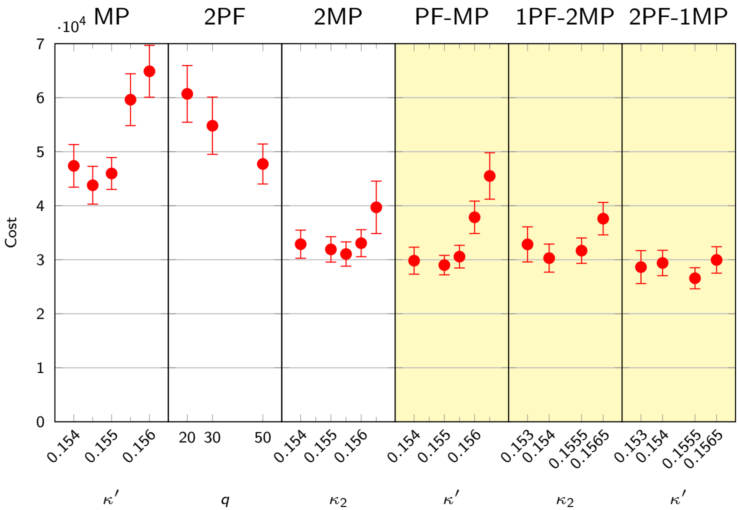
Mixed filters

- Idea: mix up with PF.



- PF does not require fine-tuning due to p's integral nature, and is well-suited to capture the high energy modes.

Cost of HMC with filtering: PF-MP



Summary

- The 2PF-1MP filtering scheme shows the most promise: it gives good performance with no fine-tuning.
- Further research would investigate its effectiveness at lighter quark masses.
- Read the paper at arXiv:1609.02652.

Lattice parameters

- We used two degenerate flavours of Wilson quarks with even-odd preconditioning on a $16^3 \times 32$ lattice
- Hopping parameter $\kappa = 0.15825$ implies a pion mass $m_\pi \sim 400$ MeV
- Beta parameter $\beta = 5.6$ implies lattice spacing $a \sim 0.08$ fm.

Lattice action

- The basic Euclidean action is given by $S = S_G[U] + S_F[U, \phi]$
- The gauge action S_G is relatively easy to calculate
- The fermion action S_F for two degenerate quarks can be written as

$$S_F = \phi^\dagger K^{-1} \phi \tag{1}$$

where $K = D^\dagger D$, D is the Dirac matrix

- The fermion action for polynomial filtered HMC is

$$S_{1PF} = \phi_1^\dagger P(K) \phi_1 + \phi_2^\dagger [P(K)K]^{-1} \phi_2 \tag{2}$$

- The fermion action for mass preconditioned HMC is

$$S_{1MP} = \phi_1^\dagger J^{-1} \phi_1 + \phi_2^\dagger J K^{-1} \phi_2 \tag{3}$$

Hybrid Monte Carlo

- Generate ϕ from the Gaussian distribution $\exp(-S_F[U, \phi])$
- Introduce a conjugate momentum field P to the gauge field U, and construct the Hamiltonian

$$H = \sum \mathrm{Tr} P^2 + S[U, \phi] \tag{4}$$

- Integrate a molecular dynamics trajectory based on Hamilton's equations to generate candidate state (P', U')
- Accept this state with probability $\min(1, \exp[H - H'])$

Filtering parameters

- The polynomial used is a Chebyshev approximation to the inverse, with real parameters μ and ν set to 1.2 and 0.9 in order to minimize the net force.
- 2PF, PF-MP runs: $p = p_1 = 4$
- 2MP runs: $\kappa_1 = 0.145$
- 1PF-2MP runs: $p = 4$, $\kappa_1 = 0.145$
- 2PF-1MP runs: $p_1 = 4$, $p_2 = 24$

Summary plot