Computational efficiency of staggered Wilson fermions in Lattice QCD: An exploratory study

David Adams

Division of Mathematical Sciences
Nanyang Technological University, Singapore

Based on joint work with Daniel Nogradi, Andriy Petrashyk and Christian Zielinski (to appear)

Outline

Introduction

Introduction

Will discuss staggered versions of Wilson fermions. These can also be used as kernel in the construction of domain wall fermions and overlap fermions on the lattice.

They are theoretically novel, and the hope is that they will also be computationally more efficient than the usual Wilson-based fermions, and perhaps have other advantageous features, e.g. improved chirality.

Main topic of this talk:
Will report on an exploratory numerical investigation of the computational efficiency of staggered Wilson fermions compared to usual Wilson fermions in quenched Lattice QCD.

But first some general background...

What are staggered Wilson fermions?

Usual Wilson fermion:

Naive fermion (16 species) + Wilson term

$\rightarrow$ 1 physical species, 15 doublers

Staggered Wilson fermion (the idea):

Staggered fermion (4 species) + "staggered Wilson term"

$\rightarrow$ 1 or 2 physical species, 3 or 2 doublers

Background

- Staggered Wilson fermions originated from an attempt to identify and understand the would-be zero-modes and index of the staggered Dirac operator, and construct overlap fermions from staggered fermions [D.A., PRL (2010), PLB (2011)]
  - Would-be zero modes identified from spectral flow of hermitian version of staggered Dirac operator
    $\rightarrow$ get staggered version of overlap Dirac operator
    $\rightarrow$ get staggered Wilson operator as kernel of staggered overlap operator
  - Number of fermion flavors reduced from 4 to 2 by "staggered Wilson term"
- 1-flavor version proposed later based on a different "staggered Wilson term" [C. Hoelbling, PLB (2011)].

Staggered Wilson terms in general

Can construct from combinations of the "flavored mass terms" for staggered fermions introduced in [Golterman & Smit, NPB (1984)].
– lifts the degeneracy of the 4 staggered fermion flavors

$\Rightarrow$ Can tune the usual mass to make one of the staggered fermion flavors massless.

Put in an overall factor $1/a \Rightarrow$ get one massless flavor and 3 flavors with masses $\sim 1/a$.

Thus 4 species (flavors) described by the original staggered fermion is reduced to 1 physical species and 3 doublers.

(If the flavored mass term has residual degeneracy, the number of physical species is only be reduced to 2.)

Two serious problems for this construction:

(1) The Golterman–Smit flavored mass terms break rotation symmetry and various staggered fermion symmetries.

$\rightarrow$ expect new fermionic and gluonic counterterms

$\rightarrow$ fine-tuning required to cancel the new counterterms and to keep the theory in a 1-flavor (or 2-flavor) phase.

(2) chirality problem: $\gamma_5^2 = \mathbf{1}$ is not exactly satisfied by the staggered version of γ_5

$\Rightarrow$ E.g. can't use staggered Wilson fermion with usual staggered version of γ_5 to make staggered versions of overlap and domain wall fermions.

Two serious problems for this construction:

(1) The Golterman–Smit flavored mass terms break rotation symmetry and various staggered fermion symmetries.

$\rightarrow$ expect new fermionic and gluonic counterterms

$\rightarrow$ fine-tuning required to cancel the new counterterms and to keep the theory in a 1-flavor (or 2-flavor) phase.

(2) chirality problem: $\gamma_5^2 = \mathbf{1}$ is not exactly satisfied by the staggered version of γ_5

$\Rightarrow$ E.g. can't use staggered Wilson fermion with usual staggered version of γ_5 to make staggered versions of overlap and domain wall fermions.

Recent development: both of these problems have been solved.

Solution to (1) – the problem of broken symmetries:

One of the flavored mass terms preserves rotation symmetry.
$\rightarrow$ make staggered Wilson term solely from this.

But it has 2-fold mass degeneracy
$\rightarrow$ get 2 (not 1) physical fermion species.
(This was the original 2-flavor construction of D.A.)

Still breaks a subset of the staggered fermion symmetries (the "shift symmetries").

Allows one new dimension 4 operator...

...fortunately its only effect on the 2 physical fermion species is a wavefunction renormalization!

The issue of broken symmetries in the 1-flavor case

The flavored mass term for the 1-flavor construction of staggered Wilson fermions in [Hoelbling, PLB (2011)] breaks rotation symmetry.

A residual subgroup of rotation symmetry survives though.

Turns out there are no new fermionic counterterms (besides the already mentioned one) [S. Sharpe, Kyoto workshop 2012].

But there is a new gluonic counterterm.

It shows up in the 1-loop contributions to the gluonic 2-, 3- and 4-point functions [D.A. (to appear)].

$\rightarrow$ would need to be included in the bare action and fine-tuned...

Solution to (2) – the chirality problem:

Recall staggered fermions have an exact *flavored* chiral symmetry.

The corresponding flavored γ_5, which we denote Γ_{55}, exactly satisfies $\Gamma_{55}^2 = \mathbf{1}$, and has spin$\otimes$flavor interpretation

$$\Gamma_{55} = \gamma_5 \otimes \xi_5$$

(ξ_μ = rep of Dirac algebra in flavor space).

$\rightarrow$ If staggered Wilson term is chosen such that $\xi_5 = \mathbf{1}$ on the physical fermion species, then

$$\Gamma_{55} = \gamma_5 \otimes \mathbf{1} \quad \text{on physical species}$$

$\rightarrow$ Can use Γ_{55} for the unflavored γ_5 in the staggered Wilson fermion theory!

Works for the constructions of D.A and Hoelbling.

(Also relies on a **technical miracle** – see [D.A., Lattice 2010 Proceedings].)

The 2-flavor staggered Wilson Dirac operator

$$D_{sW} = D_s + \frac{r}{2a}\Big(1 - \Gamma_{55}\Gamma_5\Big) + m$$

$D_s = i\eta_\mu \nabla_\mu$ = usual staggered Dirac operator

$\Gamma_5 = \gamma_5 \otimes \mathbf{1} + O(a^2)$ = usual unflavored staggered γ_5

$\Gamma_{55} = \gamma_5 \otimes \xi_5$ = flavored staggered γ_5

Note that in the "Wilson term" we have

$$1 - \Gamma_{55}\Gamma_5 = \mathbf{1} \otimes \mathbf{1} - \mathbf{1} \otimes \xi_{\mathbf{5}} + O(a^2)$$

Thus the 2 physical flavors are the ones on which $\xi_5 = \mathbf{1}$, while the 2 doubler flavors are the ones on which $\xi_5 = -\mathbf{1}$.

$\rightarrow$ The requirement of $\xi_5 = \mathbf{1}$ on the physical species is satisfied, so can take Γ_{55} as the flavor singlet γ_5 in for the staggered Wilson fermion theory in this case.

Motivation for 2-flavor staggered Wilson fermions

- Computationally more efficient than usual Wilson (hopefully)
- Genuinely new $\rightarrow$ use for universality checks in Lattice QCD

Drawback: SU(2) symmetry of the 2 physical flavors is broken by lattice effects (just like SU(4) symmetry of the usual staggered fermion flavors is broken by lattice effects).

However, situation for *flavor singlet* physics with 2-flavor staggered Wilson is just as good as for usual Wilson.

$\rightarrow$ Arenas where it may be useful if computational efficiency is confirmed:

- Precision computation of η' mass (and other flavor-singlet physics).
- Finite temperature QCD (esp. computation of bulk quantities and axial anomaly issue).

Numerical study

Investigated computational efficiency of 2-flavor staggered Wilson fermions compared to usual Wilson fermions in quenched lattice QCD simulation.

$16^3 \times 32$ lattice, $\beta = 6$, 200 configurations.

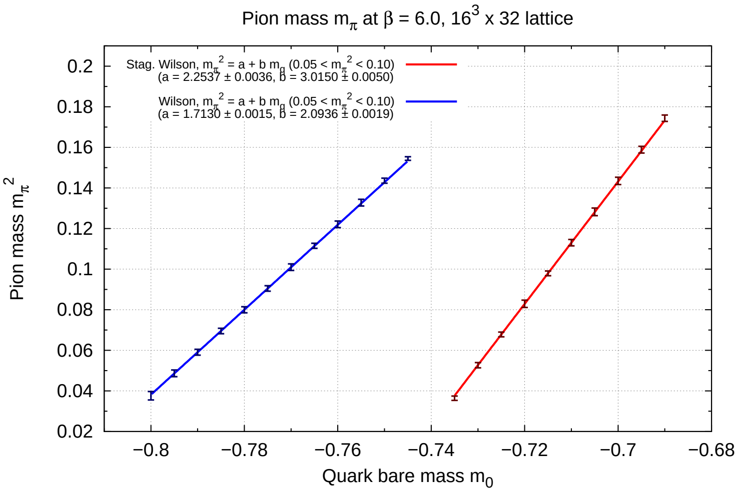
Used the Chroma/QDP software for Lattice QCD
($\rightarrow$ implementation of usual Wilson fermion is beyond reproach)

Goal:
Determine ratio of CPU times for computing the quark propagator as a function of the pion mass

To be meaningful, the comparison must be done at fixed values of a physical quantity! We use the **pion mass**.

$\rightarrow$ First need to determine the pion mass as a function of the bare quark mass for usual Wilson and staggered Wilson.

Pion mass as function of bare quark mass



Quark propagator computation time

Can be decomposed as

$$\text{time} = (\text{number of CG iterations}) \times (\text{time per iteration})$$

$\rightarrow$ Ratio of computation times for usual Wilson (W) and staggered Wilson (sW) decomposes as

$$\frac{\text{time}_W}{\text{time}_{sW}} = 4 \times \left(\frac{\text{iters}_W}{\text{iters}_{sW}}\right) \times \left(\frac{\text{time per iter}_W}{\text{time per iter}_{sW}}\right).$$

Factor 4 because 12 sources for usual Wilson quark propagator vs. 3 sources for staggered Wilson.

This allows a *theoretical estimate* of the computation time ratio.

Both ratios on the right-hand side can be estimated individually

Ratio estimates

(1)

$$\frac{\text{iters}_W}{\text{iters}_{sW}} \approx \frac{\sqrt{\kappa_W}}{\sqrt{\kappa_{sW}}}.$$

κ_W and κ_{sW} are the condition numbers of the lattice fermion matrix $D^\dagger D$ in the CG inversions for usual Wilson and staggered Wilson cases.

This approximation is expected to get better for smaller residue ϵ in the CG inversion.

(2)

$$\frac{\text{time per iter}_W}{\text{time per iter}_{sW}} = \frac{\text{flops}_W}{\text{flops}_{sW}} = \frac{1392}{1743} = 0.799$$

'flops' denotes the number of FLOPs per lattice site for the lattice Dirac operator (usual Wilson or staggered Wilson).

From direct measurement of computation time we find

$$\frac{\text{time per iter}_W}{\text{time per iter}_{sW}} \approx 0.60.$$

Explanation for why this is less than the flop ratio 0.799:

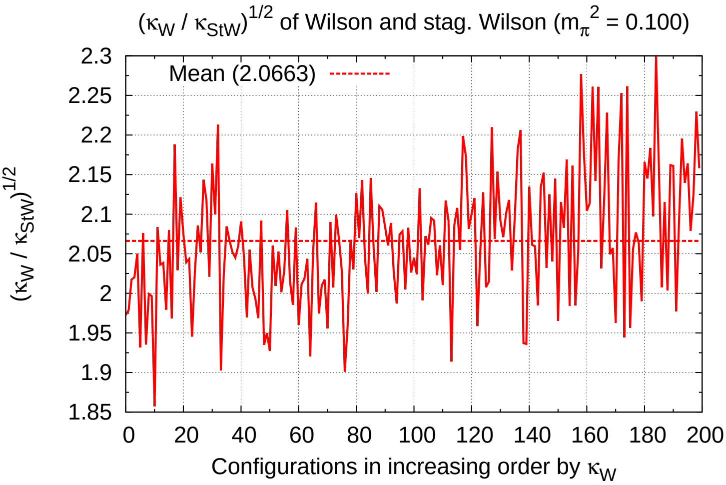
The staggered Wilson term couples opposite corners of lattice hypercubes
$\rightarrow$ Need to implement this in the Chroma/QDP software as a sequence of shifts of one step along each lattice direction.

These shift operations add nothing to the flop count, but they are computationally relatively expensive!

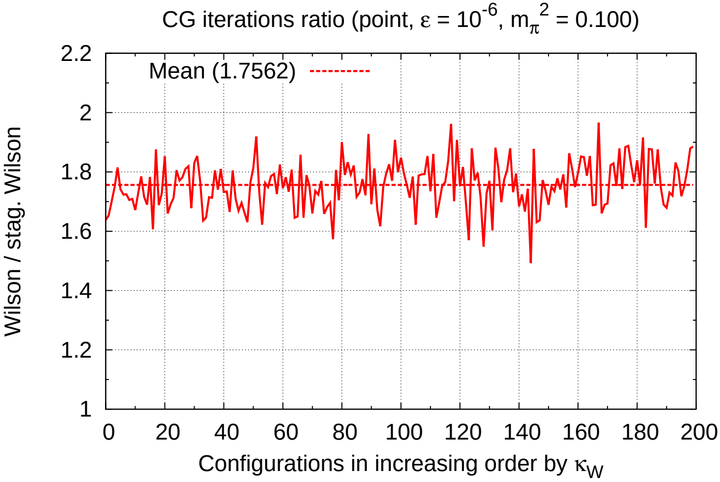
In principle it should be possible to extend the QDP code to allow shifts between opposite corners of hypercubes in a single operation.

– this would significantly reduce the computation time per iteration for staggered Wilson.

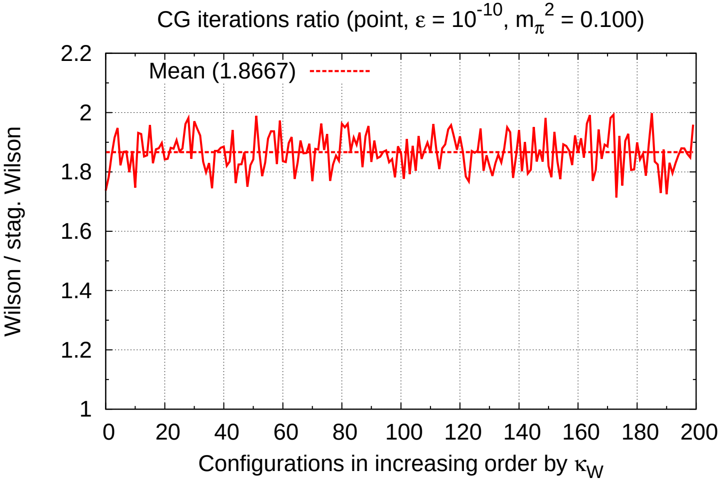
Condition number ratio at $m_\pi^2 = 0.10$



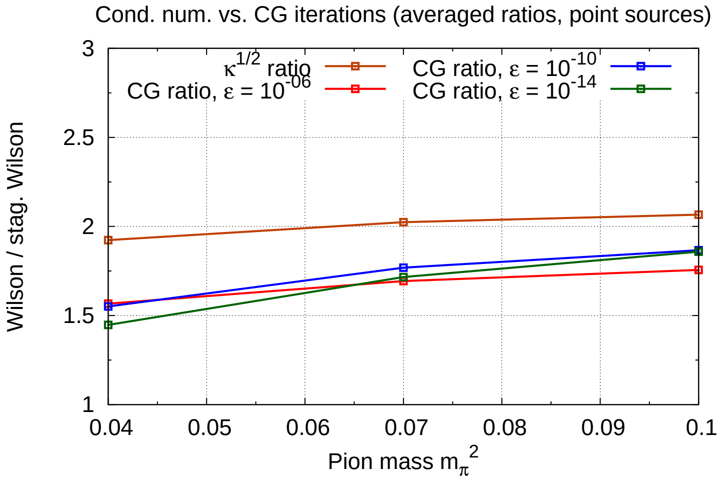
CG iterations ratio at $m_\pi^2 = 0.10$ with $\epsilon = 10^{-6}$



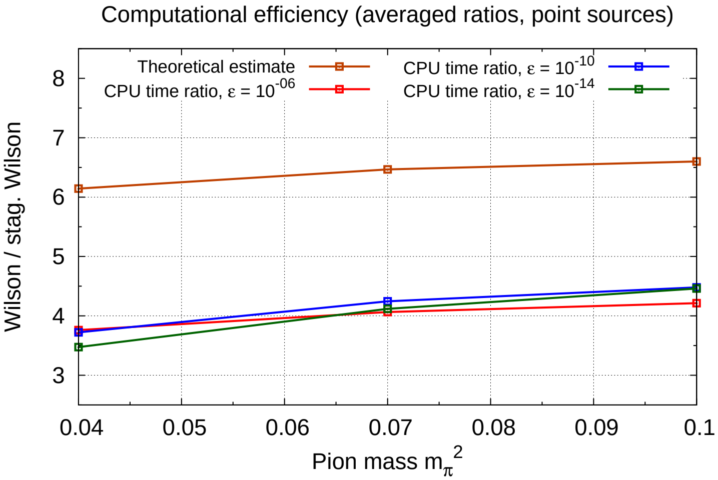
CG iterations ratio at $m_\pi^2 = 0.10$ with $\epsilon = 10^{-10}$



CG iterations ratio (average) as a function of pion mass



Computation time ratio (average) as a function of pion mass



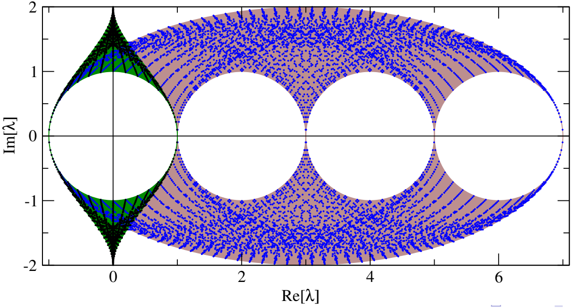
Staggered Wilson expected to be even more efficient for overlap and domain wall constructions

Reason: the spectrum of the staggered Wilson operator is closer to the Ginsparg–Wilson circle.

Free field spectrum [P. de Forcrand, Lattice 2010 conf.]:

Green: Eigenvalues of free field D_{sW} (staggered Wilson)
Blue: Eigenvalues of free field D_W (usual Wilson)

.



Staggered Wilson expected to be even more efficient for overlap and domain wall constructions

Need to compare the efficiency of staggered overlap vs usual Wilson overlap in a meaningful way:

- Compare at fixed values of a physical quantity (e.g. pion mass)
- The negative mass in the kernel operator $D - m_0$ needs to be chosen appropriately. Don't use same value for staggered Wilson and usual Wilson!
- Canonical value (free field case) is $m_0 = 1$ (in lattice units.

  But for Lattice QCD simulation should use $m_0 = 1 + m_c$, where m_c is the critical quark mass at which the pion mass vanishes (i.e. chiral limit)

  Note: m_c is different for staggered Wilson and usual Wilson!