

Numerical properties of staggered overlap fermions

Philippe de Forcrand*

Institute for Theoretical Physics, ETH Zürich, CH-8093 Zürich, Switzerland
and
CERN, Physics Department, TH Unit, CH-1211 Geneva 23, Switzerland
E-mail: forcrand@phys.ethz.ch

Aleksi Kurkela

Institute for Theoretical Physics, ETH Zürich, CH-8093 Zürich, Switzerland
and
Department of Physics, McGill University, 3600 rue University, Montréal, QC H3A 2T8, Canada
E-mail: aleksi.kurkela@mcgill.ca

Marco Panero

Institute for Theoretical Physics, ETH Zürich, CH-8093 Zürich, Switzerland
and
Department of Physics and Helsinki Institute of Physics, University of Helsinki,
FIN-00014 Helsinki, Finland
E-mail: marco.panero@helsinki.fi

We report the results of a numerical study of staggered overlap fermions, following the construction of Adams which reduces the number of tastes from 4 to 2 without fine-tuning. We study the sensitivity of the operator to the topology of the gauge field, its locality and its robustness to fluctuations of the gauge field. We make a first estimate of the computing cost of a quark propagator calculation, and compare with Neuberger's overlap.

The XXVIII International Symposium on Lattice Field Theory, Lattice2010
June 14-19, 2010
Villasimius, Italy

*Speaker.

1. Introduction

The staggered Dirac operator,

$$\mathcal{D}_s = \frac{1}{2} \sum_{\mu} \eta_{\mu} (V_{\mu} - V_{\mu}^{\dagger}) \quad \text{with: } (V_{\mu})_{xy} = U_{\mu}(x) \delta_{y, x+\hat{\mu}} \quad \text{and } \eta_{\mu}(x) = (-1)^{\sum_{\nu < \mu} x_{\nu}}, \quad (1.1)$$

is the most computationally efficient way to discretize the Dirac operator, and is commonly used in large-scale lattice QCD simulations. Yet, it leads to 4 degenerate quark ‘‘tastes’’ in the continuum limit, and the determinant is raised to the power $N_f/4$ in N_f -flavor simulations. The systematic error associated with this ‘‘rooting’’ is a subject of hot debate. In the literature, two different types of staggered-like operators which avoid rooting by representing only two tastes have been proposed: *i*) minimally doubled staggered fermions [3, 4], which represent 2 tastes with minimal fine-tuning; *ii*) staggered overlap fermions [1, 2], where the degeneracy of the spectrum is lifted by a taste-dependent mass term, and the resulting operator is used as the kernel in Neuberger’s overlap [6].

Here we study the numerical properties of *(ii)*. In spite of the additional complexity of the overlap, it may be simpler than the multiple fine-tuning required in *(i)* [5]. Note also that the overlap kernel entering *(ii)* can also be used without overlap, at the expense of fine-tuning. We first motivate the overlap kernel devised in [1] and study its topological properties, then consider the staggered overlap operator of [2].

2. Index of overlap kernel

The index of a gauge field configuration is obtained from the flow of the eigenvalues of $H(m) = \gamma_5(\mathcal{D} + m)$ as a function of m . To probe the topological properties of the gluon field, it is essential that the varying term $m\gamma_5$ be a taste singlet. This is not the case for staggered fermions, if one makes for γ_5 the customary choice $\gamma_5 \rightarrow \varepsilon(x, y) = (-1)^{\sum_{\mu} x_{\mu}} \delta_{x, y}$, which has the decomposition $\varepsilon = \gamma_5 \otimes \gamma_5$ in spin \otimes taste space. Instead, Adams proposes to trade ε for $\Gamma_5 = \eta_5 C$, where $C = \frac{1}{24} \sum_{ijkl=\text{perm}(1234)} C_i C_j C_k C_l$, $C_{\mu} = \frac{1}{2}(V_{\mu} + V_{\mu}^{\dagger})$, is the symmetrized sum of 4-link parallel transporters connecting a site to its opposite in an elementary hypercube, and $\eta_5(x, y) = \prod_{\mu=1}^4 \eta_{\mu} = (-1)^{x_1+x_3} \delta_{x, y}$ is the corresponding product of 4 phase factors. Γ_5 has the spin \otimes taste decomposition $\gamma_5 \otimes \mathbf{1}$: it is a taste singlet, which allows $m\Gamma_5$ to probe the topology of the gauge field.

In Figs. 1, we compare the flow of eigenvalues with m for Adams’ operator

$$H_A(m) = \varepsilon \mathcal{D}_s + m\Gamma_5 \quad (2.1)$$

and for Neuberger’s operator $H_W(m) = \gamma_5(\mathcal{D}_W + m)$ (where \mathcal{D}_s and \mathcal{D}_W are the zero bare mass staggered and Wilson Dirac operators, respectively), on the same $SU(3)$ gauge field configurations.

The top pair of figures corresponds to the free case (the solid lines are analytic results). In the second set of figures, the gauge field is that of a smooth, cooled instanton. The eigenvalue flow shows one crossing in Neuberger’s case, two nearly degenerate crossings in Adams’ case, reflecting the corresponding number of flavors. The third pair of figures corresponds to a thermalized configuration ($\beta = 6$ with Wilson gauge action). While the pattern of eigenvalue flow is less clear for both operators, it is remarkable that the eigenvalue gap in Adams’ case seems to persist for

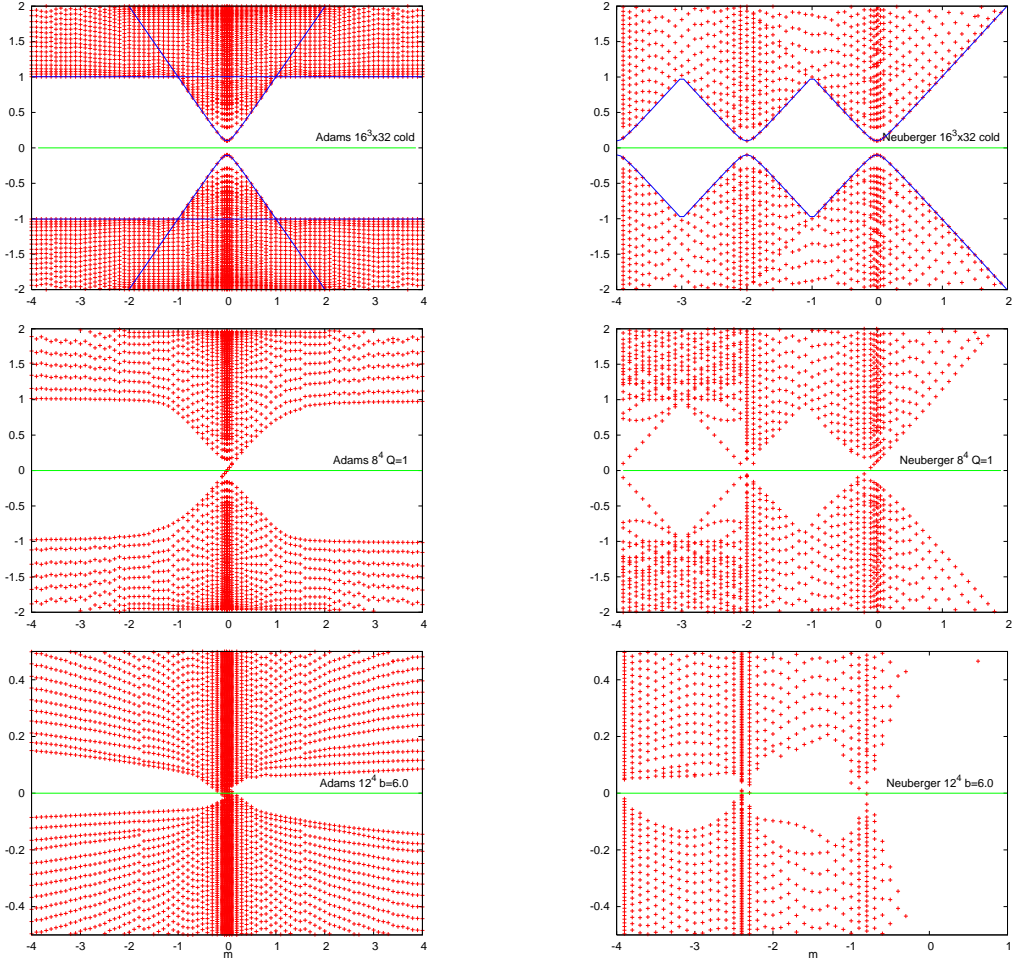


Figure 1: Flow of eigenvalues vs. m for Adams' operator (left) and Neuberger's operator (right).

arbitrarily large values of $|m|$ ¹, while in Neuberger's case additional eigenvalue crossings appear at negative values of m (corresponding to theories with a larger number of flavors), shifted from their free-field values by gauge field fluctuations.

3. Adams' staggered overlap operator

Having constructed a Hermitian kernel $H_A(m) = \varepsilon \not{D}_s + m\Gamma_5$ sensitive to the topology of the gauge field, one can plug this kernel into Neuberger's overlap: $\not{D}_{ov} = \mathbf{1} + \gamma_5 \text{sign}(H(-m_0)) = \mathbf{1} + D/\sqrt{D^\dagger D}$, where $D = \gamma_5 H$ and $D^\dagger = \gamma_5 D \gamma_5$. Thus, Adams' staggered overlap operator is

$$\not{D}_{sov} \equiv \mathbf{1} + \varepsilon \text{sign}(H_A(-m_0)) = \mathbf{1} + D_A/\sqrt{D_A^\dagger D_A}, \quad (3.1)$$

with $D_A = \varepsilon H_A(-m_0) = \not{D}_s - m_0 \varepsilon \Gamma_5$. Note that the mass term $-m_0 \varepsilon \Gamma_5$ has the spin \otimes taste decomposition $\mathbf{1} \otimes \gamma_5$: it is "flavored". If one considers low-momentum eigenstates $\tilde{\Psi}$ of \not{D}_s , satisfying

¹In Adams' case, eigenvalues come in pairs $\lambda(m) \leftrightarrow -\lambda(-m)$, because $\varepsilon H_A(m) \varepsilon = -H_A(-m)$.

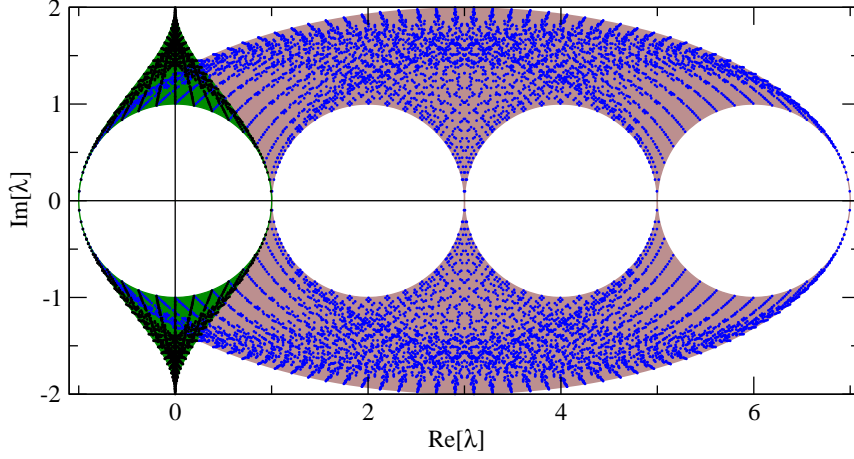


Figure 2: Spectrum of the free Wilson operator (blue) and of the free Adams' operator (green).

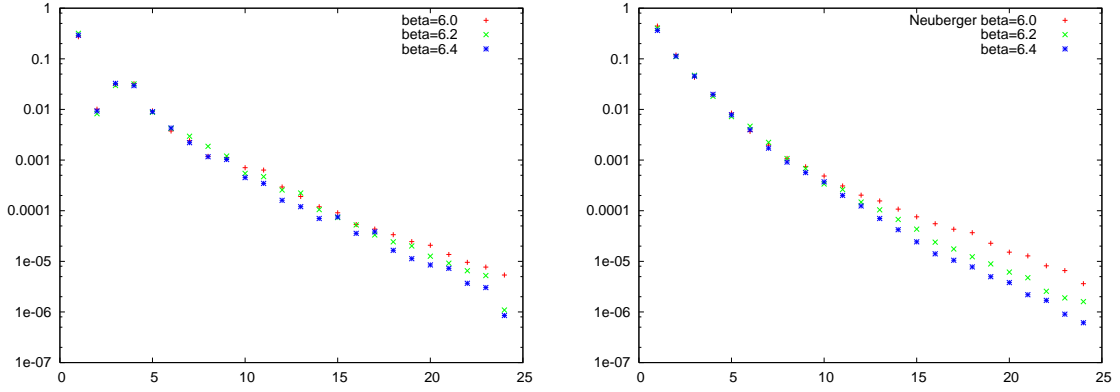


Figure 3: Maximum magnitude $\max_y |D_{ov}(x,y)|$ of the Dirac operator matrix elements versus Manhattan distance $|x-y|$, for Adams' operator (left) and for Wilson's operator (right).

$\not{D}_s \tilde{\Psi} \approx 0$, they will obey $\langle \tilde{\Psi}^\dagger (\epsilon \Gamma_5) \tilde{\Psi} \rangle \approx \pm 1$ depending on their taste content. Of the initial 4 tastes, two combinations will give $+1$ and become physical, light modes of the overlap operator; the other two combinations will give -1 and become heavy doublers.

This is clear in the free case: Fig. 2 shows the spectrum of D_A and of the Wilson operator D_W , for $m_0 = 1$ and a free field. The splitting of the 4 tastes into 2 pairs is achieved in a symmetric way, more elegant than the reduction from 16 flavors to 1 in the Wilson case. Moreover, since Adams' kernel has a spectrum already much closer to the unit circle than Wilson's operator, one may expect a smaller number of operations to achieve the unitary projection $D/\sqrt{D^\dagger D}$ of the kernel operator D in Adams' case than in Neuberger's case. This is investigated in Sec. 5.

4. Locality

First, we compare the locality of Adams' overlap operator with that of Neuberger's. In both cases, the matrix elements $D_{ov}(x,y)$ are non-zero for arbitrarily distant sites x and y , as the overlap operator is not ultra-local. What matters, however, is the decrease in magnitude of $|D_{ov}(x,y)|$ with

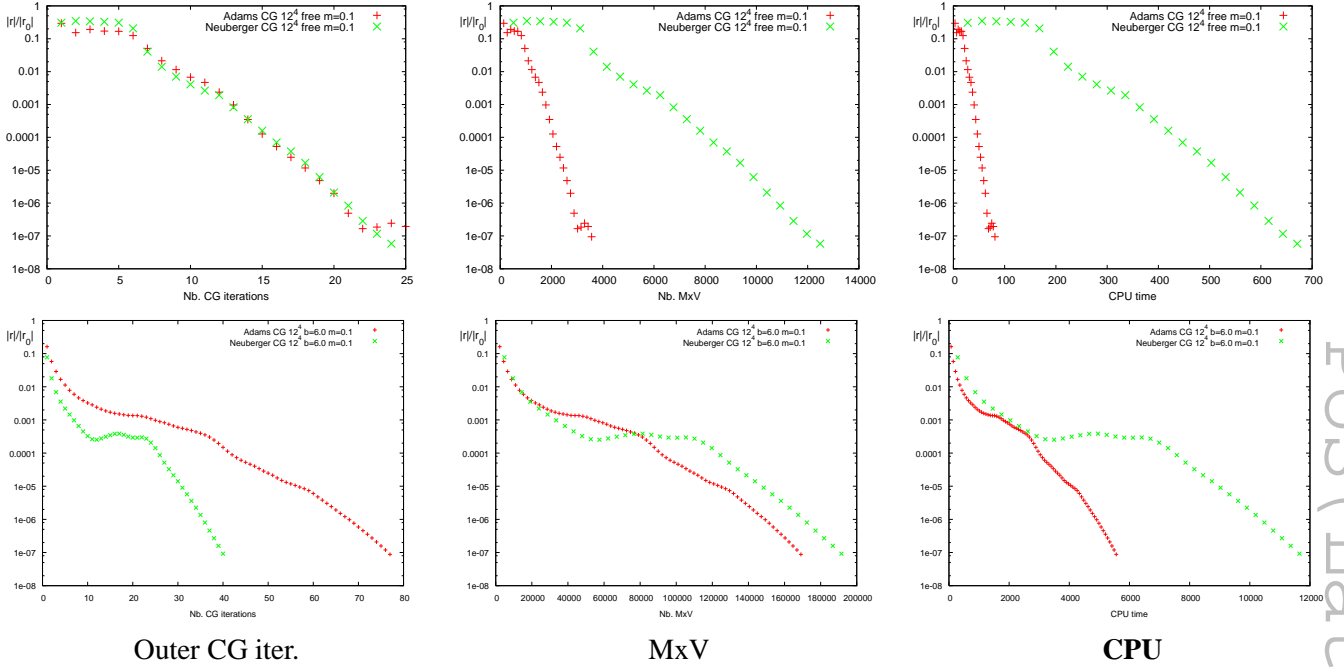


Figure 4: Computer cost of one overlap propagator, measured in outer CG iterations (left), matrix-vector multiplications (middle) and CPU time (right). The values for Adams' operator are shown in red, those for Neuberger's operator in green. The gauge field is the free field (top), or a $\beta = 6.0$ configuration (bottom).

the distance $|x - y|$, which should be bounded by $\exp(-|x - y|/(ca))$, where (ca) is a localization length proportional to the lattice spacing a , and thus shrinking to zero in the continuum limit.

Fig. 3 shows the maximum magnitude $\max_y |D_{ov}(x_0, y)|$ versus the Manhattan distance $|x_0 - y|$, chosen to follow the conventions of Ref. [7] for the Neuberger operator. The left figure corresponds to Adams' overlap operator, the right one to Neuberger's, on the same gauge configurations at 3 values of β . While Adams' operator behaves differently at short distance because of the 4-link transporters, at large distances the decay of the matrix elements is exponential as in Neuberger's case, with a similar localization length.²

5. Robustness to gauge fluctuations and efficiency

We have studied the computational cost of a quark propagator calculation with Adams' operator and compared it to Neuberger's propagator (for one component) on the same gauge field background, and with the same numerical approach. As the matrix to invert is 4 times smaller, and its spectrum is closer to the unit circle, at least in the free case (Fig. 2), Adams' operator may be computationally cheaper.

The propagator is obtained as the solution of $(D_{ov} + m)^\dagger (D_{ov} + m)x = (D_{ov} + m)^\dagger b$, using a conjugate gradient iterative solver, using the following simple and robust method [9]: at each iteration of this outer CG, $\text{sign}(H)$ is applied to a vector v through a Lanczos process, building a

²This happens even though the kernel of Adams' operator is much less local than that of Neuberger's: a less ultralocal kernel may lead to a more local overlap operator [8].

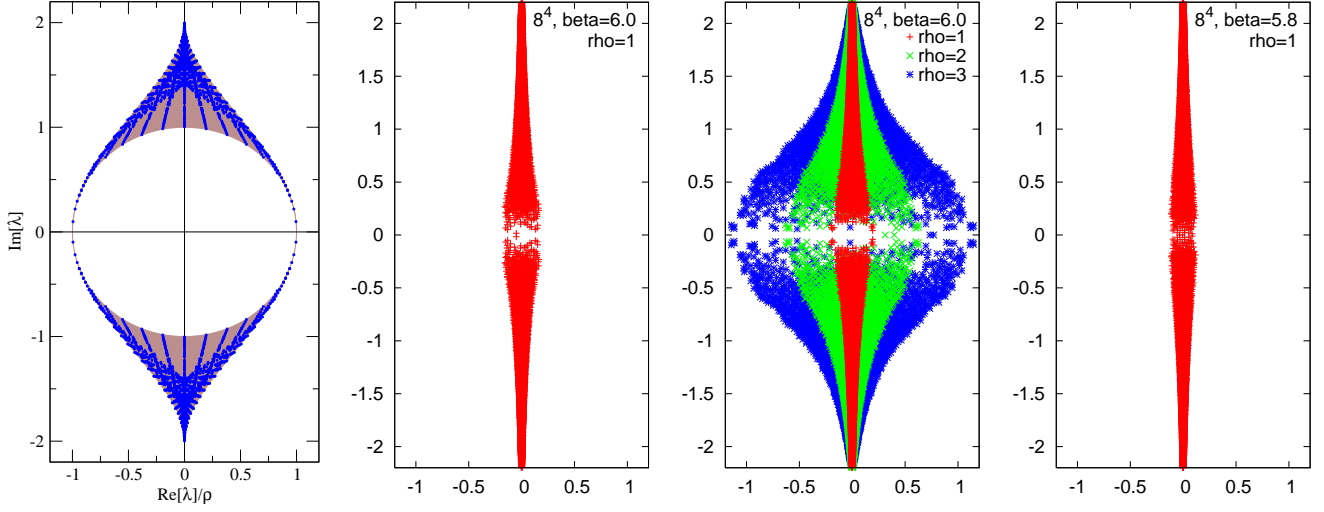


Figure 5: Spectra of Adams' operator. Left to right: free field, $\beta = 6.0$, $\beta = 6.0$ with larger ρ , $\beta = 5.8$.

tridiagonal matrix. Its eigenvalues are representative of those of H , and we replace them by their sign. The results are presented in Fig. 4, for a free field (top) and a $\beta = 6$ configuration (bottom). The 3 figures in each row show the relative norm of the residual, $|r|/|r_0|$, vs. outer CG iterations (left), number of matrix-vector multiplications (middle) and CPU time (right).

In the free field case, the CPU time to find the solution for Adams' operator is almost an order of magnitude smaller than for Neuberger's, thanks to a construction of the sign function requiring fewer matrix-vector multiplications, each with a smaller CPU cost. The inversion of the unitary operator converges at the same rate, reflecting the similar spectral properties in the infrared.

The situation changes on a $\beta = 6$ configuration. The outer CG now converges noticeably faster in Neuberger's case (left). This advantage is offset by the cost of the sign function, which still is cheaper in Adams' case (middle). Finally, the CPU time per matrix-vector multiplication is a factor $\mathcal{O}(2)$ smaller in Adams' case (Adam's matrix is one quarter the size of Neuberger's, with each site connected to $(8+16)$ neighbours, instead of 8 with 2 Dirac components). In total, the CPU time to find the solution is only $\mathcal{O}(2)$ times smaller in Adams' case.

This loss of efficiency can be traced to changes in the spectrum of Adams' kernel D_A in the presence of gauge field fluctuations, as illustrated in Fig. 5. While the free spectrum (left) is remarkably close to the unit circle, it is quite different at $\beta = 6$ (2nd panel). The splitting of the 4 tastes into 2 pairs is markedly reduced. The reason is that the taste-dependence of the mass operator $\varepsilon\Gamma_5$ is achieved via 4-link transporters: fluctuations in the gauge links are raised to the 4th power. It is the same reason for which the chirality $\langle\Psi^\dagger\Gamma_5\Psi\rangle$ of near-zero modes of the ordinary staggered operator is so small [10]. Here, one may attempt to restore the mass splitting of the pairs of tastes, by increasing the mass parameter m_0 , called ρ in Adams' Ref. [2]. The effect of such increase is shown Fig. 5 (3rd panel). The gap in the spectrum, which was the complete unit disk in the free case, and which shrank to a small but disk-like shape at $\beta = 6$ for $m_0 = 1$, now becomes a very narrow band. Unitary projection of the operator becomes more difficult, and after unitarization many modes are present near the origin, which makes inversion more difficult as well.

This figure also shows that the spectrum remains centered about the origin: changing m_0 is not the analogue of changing the mass in the Wilson operator, which shifts the whole spectrum. Rather, m_0 is the analogue of Wilson's hopping parameter r [2].

This is why the eigenvalue gap in the Hermitian operator $H_A(m)$ (Fig. 1) persists for large values of $|m|$. Shifting the whole spectrum of D_A by a taste-independent mass term is also possible, but will destroy the symmetry of the spectrum about the origin without any computational advantage. Finally, Fig. 5 (right) shows how the gap in the spectrum of D_A closes at $\beta = 5.8$.

6. Conclusion

Our study shows that Adams' staggered overlap operator works as advertised: the taste-dependent mass operator in its kernel yields 2 massless tastes without fine-tuning, and the topology and locality properties are similar to Neuberger's operator.

On very smooth gauge configurations, the computer cost of a quark propagator is nearly an order of magnitude less than in Neuberger's case, but the 4-link transporters in the flavored mass term reduce this advantage to a factor $\mathcal{O}(2)$ on $\beta = 6$ configurations. Another drawback of Adams' construction is that the continuous symmetry of the massless staggered overlap operator is $U(1)$, not $SU(2)$ as one would wish for a 2-flavor chiral symmetry.

The lack of robustness and of full chiral symmetry can both be addressed by modifying the mass operator, for example with 2-link transporters [11] that reduce the number of light tastes to 1. Preliminary results [12] confirm our expectations, but do not bring the cost of staggered overlap fermions near that of ordinary staggered fermions: avoiding rooting still has its price.

References

- [1] D. H. Adams, Phys. Rev. Lett. **104** (2010) 141602 [arXiv:0912.2850 [hep-lat]].
- [2] D. H. Adams, arXiv:1008.2833 [hep-lat].
- [3] L. H. Karsten, Phys. Lett. B **104** (1981) 315; F. Wilczek, Phys. Rev. Lett. **59** (1987) 2397.
- [4] M. Creutz, JHEP **0804** (2008) 017 [arXiv:0712.1201 [hep-lat]]; A. Borici, Phys. Rev. D **78** (2008) 074504 [arXiv:0712.4401 [hep-lat]].
- [5] S. Capitani, M. Creutz, J. Weber and H. Wittig, JHEP **1009** (2010) 027 [arXiv:1006.2009 [hep-lat]].
- [6] H. Neuberger, Phys. Lett. B **417** (1998) 141 [arXiv:hep-lat/9707022]; Phys. Lett. B **427** (1998) 353 [arXiv:hep-lat/9801031].
- [7] P. Hernandez, K. Jansen and M. Lüscher, Nucl. Phys. B **552** (1999) 363 [arXiv:hep-lat/9808010].
- [8] W. Bietenholz, Fortsch. Phys. **56** (2008) 107 [arXiv:hep-lat/0611030]; S. Dürr and G. Koutsou, arXiv:1012.3615 [hep-lat].
- [9] A. Borici, Phys. Lett. B **453** (1999) 46 [arXiv:hep-lat/9810064].
- [10] P. de Forcrand et al., Nucl. Phys. Proc. Suppl. **73** (1999) 578 [arXiv:hep-lat/9810033].
- [11] C. Hoelbling, arXiv:1009.5362 [hep-lat].
- [12] Ph. de Forcrand, talk at workshop "QCD and numerical analysis QCDNA VI", Boston Univ., Sept. 2010, see <http://super.bu.edu/brower/qcdna6/talks/deforcrand.pdf>; Ph. de Forcrand, A. Kurkela and M. Panero, in preparation.