Published by Institute of Physics Publishing for SISSA

Received: *July 9, 2005*
Accepted: *August 24, 2005*
Published: *September 12, 2005*

# Filtered overlap: speedup, locality, kernel non-normality and $Z_A \simeq 1$

**Stephan Dürr**

*Institut für theoretische Physik, Universität Bern*
*Sidlerstr. 5, CH-3012 Bern, Switzerland*
*E-mail:* durr@itp.unibe.ch

**Christian Hoelbling**

*Bergische Universität Wuppertal*
*Gaussstr. 20, D-42119 Wuppertal, Germany*
*E-mail:* christian.holbling@cern.ch

**Urs Wenger**

*NIC/DESY Zeuthen*
*Platanenallee 6, D-15738 Zeuthen, Germany*
*E-mail:* urs.wenger@desy.de

Abstract: We investigate the overlap operator with a UV filtered Wilson kernel. The filtering leads to a better localization of the operator even on coarse lattices and with the untuned choice $\rho = 1$. Furthermore, the axial-vector renormalization constant $Z_A$ is much closer to 1, reducing the mismatch with perturbation theory. We show that all these features persist over a wide range of couplings and that the details of filtering prove immaterial. We investigate the properties of the kernel spectrum and find that the kernel non-normality is reduced. As a side effect we observe that for certain applications of the filtered overlap a speed-up factor of 2-4 can be achieved.

Keywords: Lattice QCD.



 

## Contents



## 1. Introduction

From a theoretical viewpoint the ascent of "overlap" fermions [1−3], i.e. fermions which at zero quark mass satisfy the Ginsparg Wilson (GW) relation [4] ($\rho$ is a parameter that will be specified later)

$$\gamma_5 D + D\hat{\gamma}_5 = 0\,, \qquad \hat{\gamma}_5 = \gamma_5 \left(1 - \frac{1}{\rho} D\right) \tag{1.1}$$

and thus realize a lattice version of the continuum chiral symmetry [5]

$$\delta\psi = \hat{\gamma}_5 \psi\,, \qquad \delta\overline{\psi} = \overline{\psi}\gamma_5 \tag{1.2}$$

together with an index theorem [6, 7], represents a major breakthrough in the field of non-perturbative studies of QCD. We know how to discretize fermions in a way that preserves the relevant symmetries: (*i*) gauge invariance, (*ii*) flavor symmetry, and (*iii*) chiral invariance. Unfortunately, from a practical viewpoint the usefulness of this concept is limited by the fact that the overlap tends to be one to two orders of magnitude more expensive, in terms of CPU time, than a standard Wilson Dirac operator.

In this paper we study a variant of the overlap operator which makes use of a UV filtered Wilson kernel. Here, the "filtering" refers to replacing the original ("thin") links of the gauge configuration in the standard definition of the Wilson kernel by "thick" links obtained through APE [8] or HYP [9] smearing. This is a legal change of discretization as long as one keeps the iteration level and smearing parameters fixed all the way down to the continuum, since the "thick" links transform under a local gauge transformation



JHEP09(2005)030

in the same way as the "thin" links; it should be seen as a modification of the operator and not of the gauge background. Such filtering has been used in the context of staggered quarks, where it has been found to reduce UV fluctuations, in particular taste changing interactions due to highly virtual gluons [10]. In ref. [11] filtered staggered quarks were compared against overlap quarks (where the filtered version was merely considered for completeness), and it was observed that a single filtering step may speed up the forward application of the overlap operator $D_{\mathrm{ov}}$ on a source vector by a factor 2-4, depending on the gauge background. This was seen to come through a reduction of the degree of the Chebychev polynomial needed to approximate the inverse square root or sign function in the definition of the massless overlap [3]

$$aD_{\mathrm{ov}} = \rho \left[1 + D_{\mathrm{W},-\rho}(D^{\dagger}_{\mathrm{W},-\rho} D_{\mathrm{W},-\rho})^{-1/2}\right] = \rho\,[1 + \gamma_5\,\mathrm{sign}(a\gamma_5 D_{\mathrm{W},-\rho})] \tag{1.3}$$

with $D_{\mathrm{W},-\rho} = D_{\mathrm{W}} - \rho/a$ the Wilson operator at negative mass $-\rho/a$. However, what matters in view of most phenomenological applications is the performance of the massive operator (bare quark mass $m$)

$$D_{\mathrm{ov},m} = \left(1 - \frac{am}{2\rho}\right) D_{\mathrm{ov}} + m \tag{1.4}$$

in the process of calculating a given physical observable to a pre-defined accuracy. In other words the total CPU time spent depends on:

1. The number of forward applications of the shifted Wilson operator $D_{\mathrm{W},-\rho}$ (or, generally speaking, of the kernel) needed to construct the massless overlap operator (1.3).
2. The number of iterations spent on inverting the so-constructed massive operator (1.4) for a given renormalized quark mass (or a given $M_\pi^2$).
3. The number of gauge backgrounds needed to reach a pre-defined statistical accuracy of the desired observable at a given lattice spacing $a$.
4. The lattice spacing needed to enter the scaling window.

The main emphasis of this paper will be on point 1; in particular we attempt to give an understanding of the observed speedup in terms of the spectral properties of the underlying hermitean (shifted) Wilson operator $H_{\mathrm{W}} = \gamma_5 D_{\mathrm{W},-\rho}$. At first sight it might seem that point 2 does not need to be considered at all. At fixed bare mass $m$ and fixed $\rho$ the filtered and the unfiltered overlap do not differ on this point, since the number of forward applications of $D_{\mathrm{ov},m}$ to get a column of the inverse depends only on its condition number, and that is $2\rho/m$ for either variety. As we shall see, the optimum $\rho$ (w.r.t. locality) gets reduced through filtering whereas $Z_m = Z_S^{-1} = Z_P^{-1}$ increases and this means that in the filtered case one has to use a smaller bare mass to work at a fixed physical $m^{\mathrm{ren}} = Z_m m$. These two aspects tend to compensate, and as a result there is little net effect on point 2 from filtering. Whether in points 3 and 4 filtering brings further savings is not clear, but we plan to address this issue in the future.

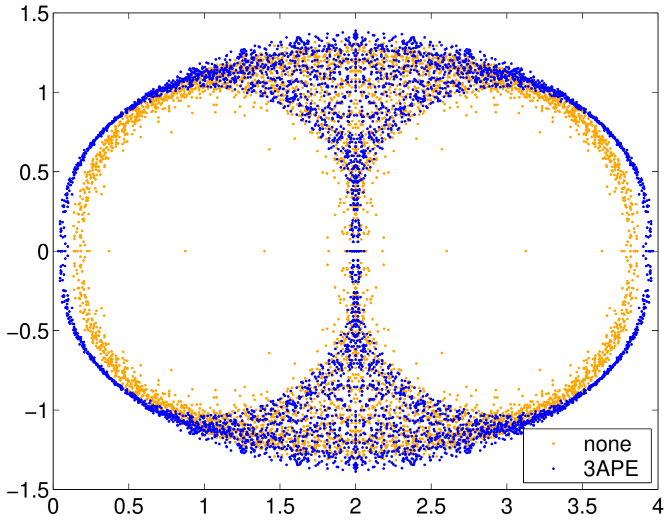


**Figure 1:** Eigenvalue spectra of $D_{\rm W}$ with $\kappa = \kappa_{\rm crit}^{\rm tree} = 0.25$ in the quenched Schwinger model ($16^2$, $\beta = 3.2$, 10 configurations) without filtering and after 3 steps with $\alpha = 0.5$ (i.e. equal weight to the original link and the staple, see [12] for details; in 2D APE involves already the full hypercube). Filtering depletes the "bellies", makes the physical (leftmost) branch narrower and shifts it to the left.

Let us try to obtain a first understanding of the effect of filtering in terms of the spectrum of the underlying (non-hermitean) Wilson operator. We are going to compute all eigenvalues, and to avoid spending too much CPU time on this illustration, we shall do this in 2D, but it is clear that the conceptual issue is — mutatis mutandis — the same as in 4D. In $d$ dimensions the Wilson Dirac operator has $d+1$ branches, and the respective flavor multiplicities are

$$\binom{d}{0}, \binom{d}{1}, \ldots, \binom{d}{d}. \tag{1.5}$$

Thus in 2D the Wilson operator has 3 branches with multiplicities 1,2,1, while in 4D it has 5 branches with multiplicities 1,4,6,4,1, respectively. Figure 1 shows the complete spectrum of $D_{\rm W}$ with the hopping parameter fixed at its tree-level critical value, $\kappa = 0.25$, on 10 configurations of size $16^2$ at $\beta = 3.2$ in the quenched Schwinger model. Besides the "thin" link operator also its UV filtered descendent is shown. In terms of the kernel spectrum the filtering is seen to have the following effects:

(a) The two/four "bellies" are depleted — in particular exactly real modes which cannot be assigned in a unique way to one of the three/five branches are severely suppressed.

(b) The horizontal scatter of any of the three/five branches diminishes.

(c) The additive mass renormalization of the physical (leftmost) branch is substantially reduced.

If one were to ignore the kernel non-normality (we shall come back to this point), the spectrum of $D_{\rm W}$ could be linked, on a mode-by-mode basis, to the one of $D_{{\rm W},-\rho}^\dagger D_{{\rm W},-\rho}$, $H_{\rm W} = \gamma_5 D_{{\rm W},-\rho}$ and $D_{\rm ov}$. Then the first observation above (the reluctance of the filtered

| $\beta$ | 5.66 | 5.84 | 6.00 | 6.26 |
|---|---|---|---|---|
| geometry | $8^4$ | $12^4$ | $16^4$ | $24^4$ |
| geometry | $8^3\times 16$ | $12^3\times 24$ | $16^3\times 32$ | — |

**Table 1:** Survey of matched 4D couplings and geometries with fixed $L/r_0 \approx 3$, according to the interpolation formula of ref. [14]. The first coupling is slightly out of bound (see discussion in [14]).

eigenvalues to show up near the projection point $\rho$) simply means that the effect of filtering on the spectrum of $H_\mathrm{W}$ is to deplete the vicinity of the origin by pushing the eigenvalues further towards the ends of the interval $[-2d+1, 2d-1]$. In spite of the caveat mentioned, the thinning effect that (any kind of) smearing has on the spectrum of $H_\mathrm{W}$ near zero is indeed the reason for the speedup in point 1 above. A bigger interval $[0, \epsilon^2[$ or $]-\epsilon, \epsilon[$ that does not need to be covered by the polynomial/rational approximation to the $1/\sqrt{.}$ or sign(.) function translates into a lower degree and thus into fewer forward applications of the kernel operator.

In the remainder of this article we shall address the spectral properties of $H_\mathrm{W}$ in more detail (section 2), and show that (a reasonable amount of) filtering does not degrade the locality properties of $D_\mathrm{ov}$, but rather makes the overlap operator *more local* (section 3). We continue with an explicit demonstration that the kernel non-normality gets reduced by filtering (section 4). We add some observations relevant to phenomenological applications of the filtered overlap; in particular $Z_A$ is shown to be much closer to the tree-level value 1 than for the unfiltered variety (section 5). We rate this as a sign that perturbation theory might work far better for the filtered overlap. We make an attempt to compare our simple filtering recipe against other approaches (section 6). Finally, the appendix contains spectral data which suggest that the spectral density of $H_\mathrm{W}$ at the origin is non-zero for any $\beta$ and any filtering level.

We shall use pure gauge backgrounds and set the scale through the Sommer parameter $r_0$ [13]. We choose the Wilson gauge action, and since $r_0(\beta)$ is known [14] it is easy to select $\beta$ values such that the resulting lattices are matched, i.e. have fixed spatial size $L \simeq 1.5\,\mathrm{fm}$, with the resolution varying by a factor 3 from the coarsest to the finest lattice — see table 1 for details. Henceforth we set $a = 1$.

## 2. Speedup and kernel spectrum

In a quenched simulation the overhead, in terms of CPU time, of overlap versus Wilson quarks comes in the first place from the polynomial or rational approximation to the $1/\sqrt{.}$ or sign(.) function in (1.3). Let us assume[1] that the lowest eigenvalue of the unfiltered $|H_\mathrm{W}|$ is 0.14 while the highest eigenvalue takes the free field value 7. This leads to the task to construct a polynomial/rational approximation of the inverse square root over the range $[\epsilon^2, 1]$ with $\epsilon = 0.02$ the inverse condition number of $|H_\mathrm{W}|$. Modest filtering will lift the lowest eigenvalue to something like 0.49, while the largest eigenvalue is almost invariant.

[1] In fact, these values are rather close to the actual situation at $\beta \simeq 6.0$, after projecting out the lowest 10-15 eigenvectors.



JHEP09(2005)030

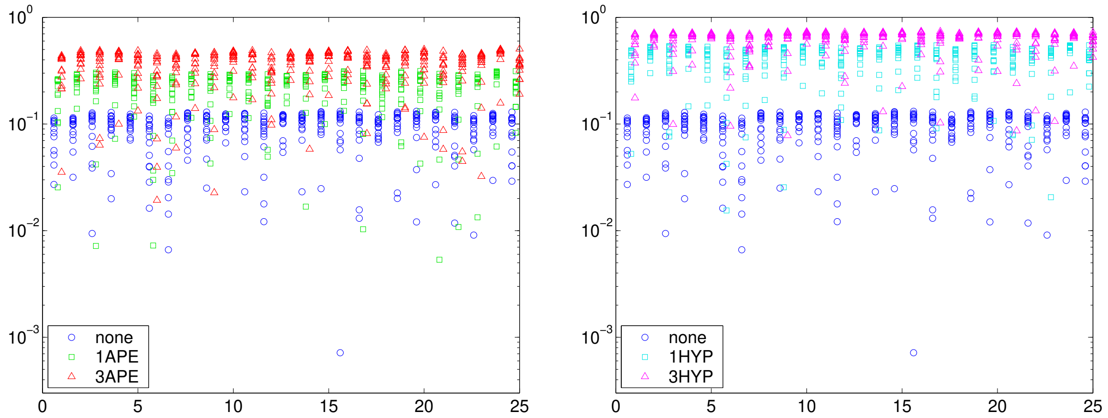


**Figure 2:** Sequence of the lowest 15 eigenvalues of $|H_\mathrm{W}|$ on 25 configurations at $\beta = 6.0$ without filtering and after 1,3 steps of APE (left) or HYP (right) filtering. Throughout $\rho = 1$.

Then the task is to construct the approximation over the range $[\tilde{\epsilon}^2, 1]$ with $\tilde{\epsilon} = 0.07$ the filtered inverse condition number. The lower bound increasing from 0.0004 to 0.0049 means that one gets away with a smaller overall polynomial degree or an increased minimum root of the denominator polynomial. Therefore, the filtered overlap requires fewer forward applications of $D^\dagger_{\mathrm{W},-\rho} D_{\mathrm{W},-\rho}$ and this is how the savings on CPU time in point 1 above come about. In the remainder of this section we will elaborate on this statement, replace the fictitious numbers by actual figures from real simulations and see that the conclusion remains unchanged.

Figure 2 shows, as an illustration, the 15 lowest eigenvalues [15] of $|H_\mathrm{W}|$ on 25 configurations at $\beta = 6.0$, without filtering and after 1,3 steps of APE or HYP smoothing. The filtering increases the upper end of the band of eigenvalues shown. In fact, just this upper end matters in terms of CPU time, since in practice one projects out the lowest few modes [16] and constructs the function in (1.3) over the relevant spectral range of $|H_\mathrm{W}|$ on the subspace orthogonal to these modes. Hence the sequence of the 15th eigenvalue represents the relevant quantity, if 14 modes are treated exactly, and this band gets lifted by filtering. Evidently, a single APE step is less efficient than a single HYP step, and adding two more steps lifts the 15th eigenvalue further, but the lifting factor is no more as large as it was in the first step. Here and below we use the parameters $\alpha_\mathrm{APE} = 0.5$, $\alpha_\mathrm{HYP} = (0.75, 0.6, 0.3)$ [9] (for details of the SU(3) projection see e.g. [17] or the appendix of [18]) and, unless stated otherwise, $\rho = 1$.

Figure 3 shows the mean and the standard deviation of the 15 lowest eigenvalues of $|H_\mathrm{W}|$, with our standard filtering options (none, 1 APE, 3 APE, 1 HYP, 3 HYP). In this logarithmic representation it is easy to see that (apart from the coarsest lattice which represents a special case discussed in appendix A) all 15 eigenvalues get lifted, at a given coupling, by virtually the same factor. Specifically, the 15th[2] eigenvalue gets multiplied by $\lambda_{1\,\mathrm{HYP}}/\lambda_\mathrm{none} = 4.8, 4.2, 3.2$ at $\beta = 5.84, 6.00, 6.26$. Thus the lifting effect that filtering has

[2]This number needs to be scaled with the physical box volume; working, for any given $\beta$, in a $(2.0\,\mathrm{fm})^4$ box instead of $(1.5\,\mathrm{fm})^4$, our statement would most likely be adequate for the 47th mode.

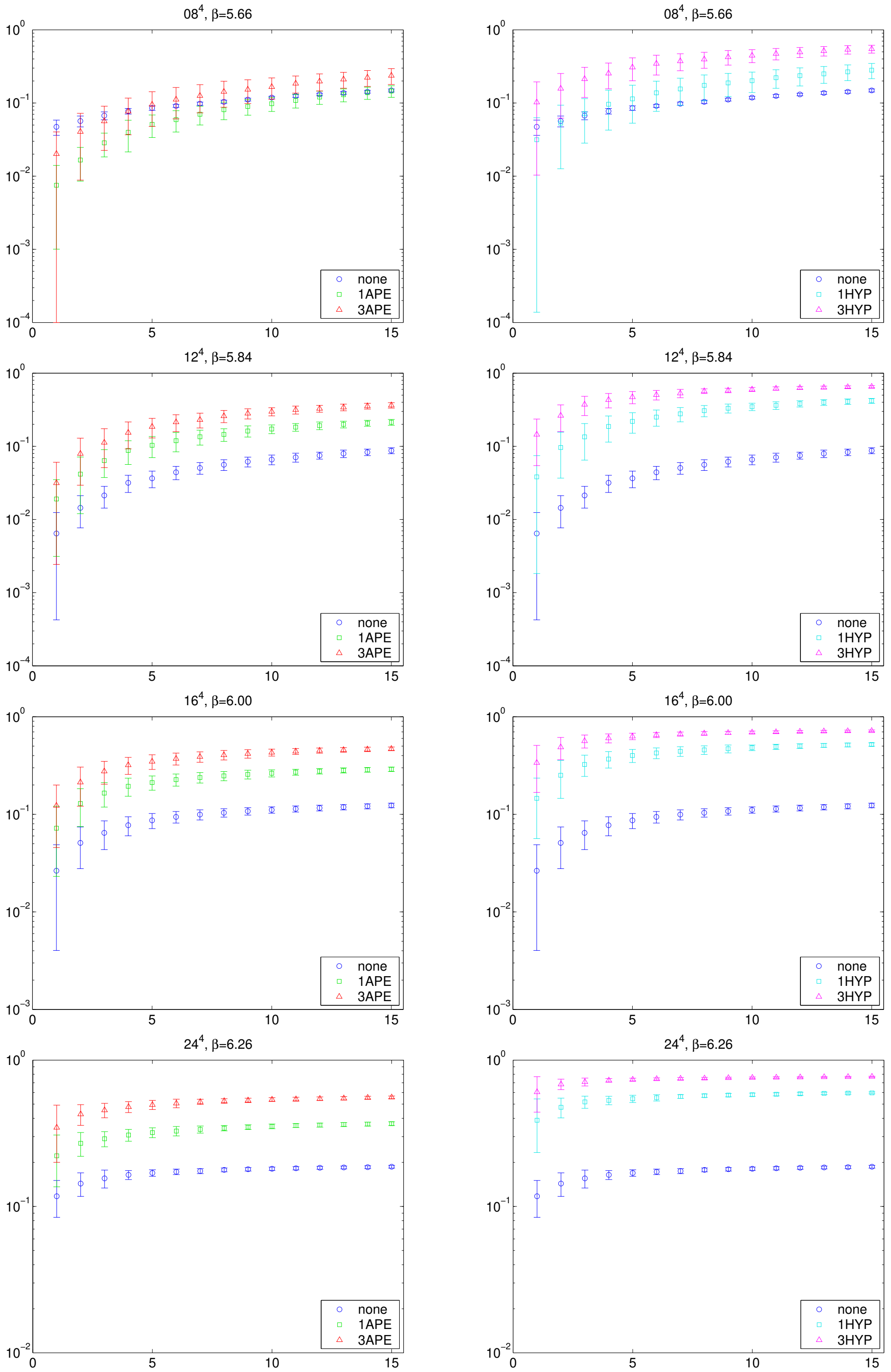


**Figure 3:** Mean and standard deviation ($\beta = 5.66, 5.84, 6.00, 6.26$, from top to bottom) of the 15 lowest eigenvalues of $|H_\mathrm{W}|$ at $\rho = 1$ in semi-logarithmic form with 0,1,3 steps of APE or HYP filtering.

| $\beta$ | 5.66 | 5.84 | 6.00 | 6.00 ($\rho = 1.4$) | 6.26 |
|---|---|---|---|---|---|
| none | 0.149(02) | 0.087(2) | 0.123(1) | 0.222(3) | 0.187(1) |
| 1 APE | 0.146(05) | 0.213(4) | 0.290(2) | 0.343(5) | 0.368(2) |
| 3 APE | 0.237(12) | 0.364(6) | 0.469(2) | 0.547(7) | 0.557(2) |
| 1 HYP | 0.281(14) | 0.419(7) | 0.519(2) | 0.618(5) | 0.597(2) |
| 3 HYP | 0.548(14) | 0.653(5) | 0.717(1) | 0.639(1) | 0.770(1) |

**Table 2:** Start of the "bulk" part of the eigenvalue spectrum of the shifted hermitean Wilson operator $|H_\mathrm{W}|$ without filtering and after one or three APE or HYP steps. We use the mean of the 15th-smallest eigenvalue to define the "bulk" edge. Unless indicated otherwise, the numbers refer to the case $\rho = 1$.

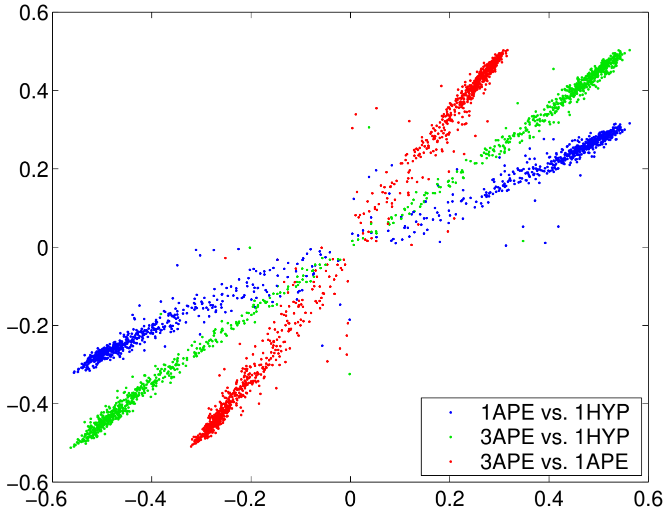


**Figure 4:** Correlation of the 15 eigenvalues closest to zero of $H_\mathrm{W}$ with various filtering options at $\beta = 6.0$.

on the "bulk" part of the $|H_\mathrm{W}|$ spectrum diminishes somewhat towards the continuum, but for accessible couplings it remains substantial. Details of the ensemble average of the 15th eigenvalue are collected in table 2. The second observation is that the bands become flatter at large $\beta$, hence the onset of the "bulk" becomes a less ambiguous concept at weaker coupling. Had we chosen the 10th or 20th mode to define the "bulk edge" instead of the 15th, this would cause a small change at $\beta = 6.26$, but it would make a substantial difference at the smallest $\beta$ shown.

A point of theoretical interest is whether the low-lying eigenvalues of the (shifted) hermitean Wilson operator $H_\mathrm{W}$ are correlated, between different smearing levels, just as the low-lying eigenvalues of the final $D_\mathrm{ov}$ were found to be correlated for large enough $\beta$ [11]. Figure 4 shows that this is almost true — the eigenvalues correlate if they are sufficiently large in absolute magnitude, but the correlation weakens closer to the origin. Here, a technical issue comes along. Ideally, one would pair the eigenvalues by considering a smooth interpolation between the two filtering recipes. Changes in topology (as seen by the overlap operator) would then be evident as stray points in quadrants 2 or 4. However, since

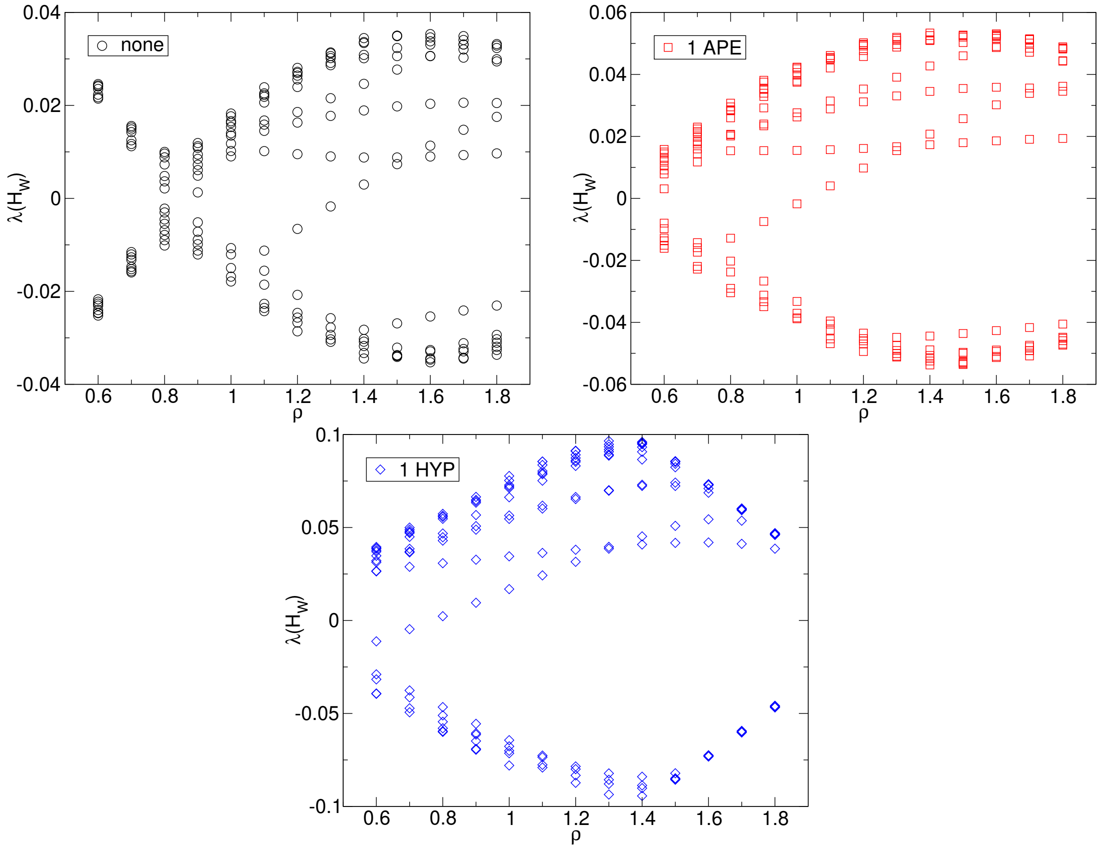


**Figure 5:** Eigenvalue flows of $H_\mathrm{W}$ with three filterings (none, 1 APE, 1 HYP) on one $16^4$ configuration.

we just know the eigenvalues shown we decided to pair them starting from 0. Now there are no points in quadrants 2 or 4 by definition and changes in topology manifest themselves through a reduced correlation of the few lowest eigenvalues in absolute magnitude. Such topology changes are expected to occur with an $O(a^2)$ re-definition of the overlap operator, e.g. by changing the filtering or $\rho$ [11].

A similar conclusion is drawn from the flow of eigenvalues $H_\mathrm{W}$ as shown in figure 5 for one $16^4$ configuration. One effect of filtering is to stretch the whole scenery in the vertical direction (note the vertical scale). Filtering also shifts the entire eigenvalue flow to the left which is consistent with the reduction of the additive mass renormalization of the kernel operator as discussed in the introduction. Note that there is, from a conceptual viewpoint, no reason to prefer one filtering level over any other one; what we see is just a manifestation of the $O(a^2)$ ambiguity of the overlap operator [11].

To assess the CPU time needed for the massless overlap, the behavior of the "bulk edge" of the $|H_\mathrm{W}|$ spectrum is one ingredient. What really matters is the condition number, thus we need to study the largest eigenvalue, too. From the naive discussion around figure 1 in the introduction one expects that filtering barely affects the largest eigenvalue of $|H_\mathrm{W}|$. It turns out that this is indeed true, for instance at $\beta = 6.0$ a single HYP filtering step lifts it from 6.55(1) to 6.88(1). Hence, filtering has an overall beneficial

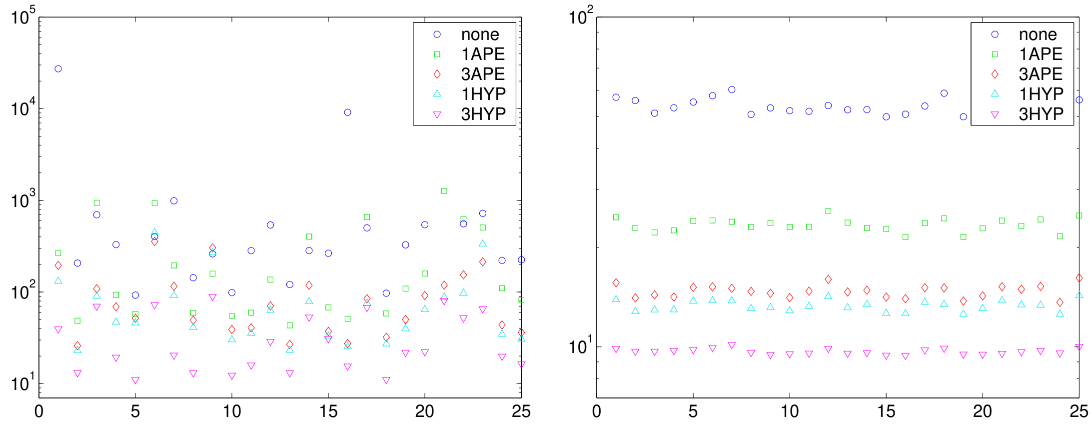


**Figure 6:** Condition number of $|H_{\rm W}|$ on 25 configurations at $\beta = 6.0$ without projection (left) and with 14 modes handled exactly (right), after 0,1,3 steps of APE or HYP filtering.

| $\beta$ | 5.66 | 5.84 | 6.00 | 6.00 ($\rho = 1.4$) | 6.26 |
|---|---|---|---|---|---|
| none | 43.5(04) | 75.2(15) | 53.2(4) | 27.9(4) | 35.3(2) |
| 1 APE | 47.2(18) | 31.8(06) | 23.4(2) | 18.6(3) | 18.5(1) |
| 3 APE | 30.5(17) | 19.0(03) | 14.7(1) | 11.9(2) | 12.4(1) |
| 1 HYP | 25.7(14) | 16.5(03) | 13.3(1) | 10.5(1) | 11.6(1) |
| 3 HYP | 12.8(04) | 10.6(01) | 9.69(3) | 10.2(1) | 9.04(2) |

**Table 3:** Mean condition number $1/\epsilon$ of $|H_{\rm W}|$ at $\rho = 1$ after projecting the 14 lowest eigenmodes.

effect on the condition number as illustrated in figure 6. Without projection the condition number fluctuates wildly and occasionally it may increase through filtering (i.e. the lowest eigenvalue decreases, cf. figure 2) but after projecting 14 eigenmodes this never occurs. The bottom line is that the combination of filtering and projection reduces the condition number much more vigorously than either one alone could do. Average condition numbers after projecting out 14 eigenmodes are collected in table 3 (regarding the first entry, cf. appendix A). As a side remark we note that the horizontal increase to the left explains why in a fixed physical volume simulating unfiltered overlap quarks on a coarse lattice is not so much cheaper than on a fine one; for the filtered version this penalty is reduced.

We have also studied the condition number of $|H_{\rm W}|$ as a function of the parameter $\rho$. With and without filtering the minimum is rather shallow and at a $\rho$ value above 1. Since in the free case

$$\epsilon = \begin{cases} \dfrac{\rho}{(8-\rho)} & \text{for} \;\; 0 < \rho \le 1 \\ \dfrac{(2-\rho)}{(8-\rho)} & \text{for} \;\; 1 \le \rho < 2 \end{cases} \tag{2.1}$$

we expect that larger $\beta$ values will further drive the minimum location towards $\rho = 1$.

The last step is to convert the reduced condition number, brought by the filtering, of $|H_{\rm W}|$ on the subspace orthogonal to the lowest 14 modes into a lower degree of the polynomial/rational approximation of the $1/\sqrt{.}$ function in (1.3) and thus into actual savings of CPU time in step 1 of the introduction.



JHEP09(2005)030

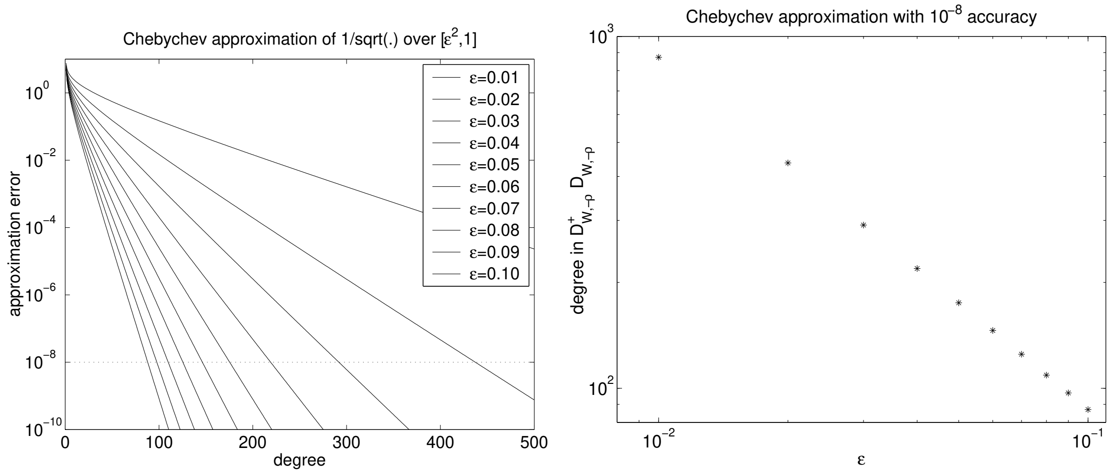


**Figure 7:** Accuracy of the Chebychev approximation of $(.)^{-1/2}$ over the interval $[\epsilon^2, 1]$ versus polynomial degree (left) and the degree (number of forward applications of $D^\dagger_{\mathrm{W},-\rho} D_{\mathrm{W},-\rho}$) needed to reach absolute accuracy $10^{-8}$ versus the inverse condition number $\epsilon$ of the shifted hermitean Wilson operator $|H_\mathrm{W}|$.

The precise speedup factor depends on the implementation of the massless overlap operator (1.3). For definiteness let us consider the approximation of the inverse square root over the range $[\epsilon^2, 1]$ through Chebychev polynomials [16]. Figure 7 shows on the l.h.s. for a few inverse condition numbers $\epsilon$ of $|H_\mathrm{W}|$ the well known exponential fall-off pattern of the truncation error of the Chebychev approximation versus the number of applications of $H^2_\mathrm{W} = D^\dagger_{\mathrm{W},-\rho} D_{\mathrm{W},-\rho}$. What matters for our purpose is the dependence of the polynomial degree required to reach a fixed minimax accuracy — say $\delta = 10^{-8}$ over the full approximation range — on $\epsilon$. As is evident from the r.h.s. of that figure, the relation

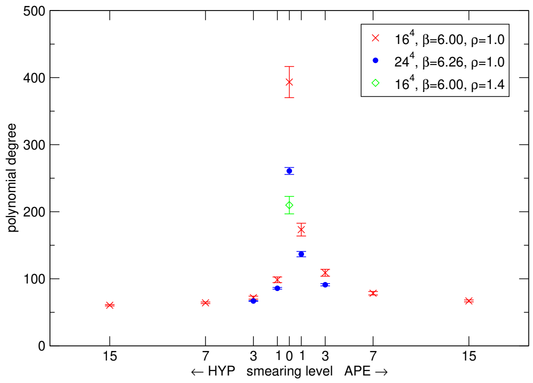


**Figure 8:** Mean and standard deviation of the Chebychev polynomial degree used to achieve a minimax accuracy of $10^{-8}$. At $\beta = 6.0$ and $\rho = 1$, a single HYP step results in a speedup by a factor $\sim 4$ for the massless overlap operator. Comparing to the situation with $\rho = 1.4$ and no filtering the factor is $\sim 2$.

$$\text{degree} \propto \epsilon^{-1} \tag{2.2}$$

holds in good approximation. Thus, from (2.2) and a look at table 3 one predicts that at $\beta = 6.0$ and $\rho = 1$ a single HYP step will speed up the construction of the overlap (on average) by a factor $53.2/13.3 = 4.00$, and this is in good agreement with what we find in actual runs (see figure 8). On a coarser lattice this factor would be somewhat larger (4.56 at $\beta = 5.84$) while on a finer lattice it tends to decrease (3.04 at $\beta = 6.26$), but it certainly remains substantial at all accessible couplings.

To approximate the inverse square root or sign function over the relevant range, two main strategies are found in the literature. Polynomial [19, 20, 16] and rational [20, 21] representations have been tried. We have concentrated on the Chebychev variant, since this one is efficient and easy to implement. It goes without saying that the lifting effect on the bulk of the $|\gamma_5 D_{\mathrm{kern}}|$ eigenvalues translates into similar savings on CPU time in step 1 of the introduction, if another representation is used. For instance, in the rational approach it is the increase of the smallest zero of the denominator polynomial that lets one get away with fewer iterations in the inner multishift CG.

For five dimensional variants of the overlap operator [22−26], in particular the domain wall formulation, the computational gain comes from the reduction of the extent of the fifth dimension needed to reach a given residual mass. What we would like to stress here is simply that our proposal to replace the "thin" links by "thick" links is generically useful for any kind of overlap variant.

## 3. Locality

It has been shown [27] that the overlap operator cannot be ultralocal, as opposed to the Wilson operator where $D_{\mathrm{W}}(x,y) = 0$ for $||x-y||_1 > 1$. To guarantee the universality of the underlying field theory and hence to obtain the correct continuum limit it is sufficient to have an operator with

$$D_{\mathrm{ov}}(x,y) \propto \exp(-\nu||x-y||) \qquad \text{for} \;\; ||x-y|| \gg 1 \tag{3.1}$$

where the localization $\nu$ is of the order of the cut-off, i.e. $\nu = O(1)$ [in lattice units]. In practice, for a given lattice spacing the condition (3.1) gives an upper bound on any physical mass that one can extract, and it is therefore crucial to have an operator as local as possible, i.e. with a maximal $\nu$. In [28] it has been demonstrated that the standard overlap operator indeed obeys (3.1). It is clear that their proof goes through for our filtered variant, but it is open in which way the localization $\nu$ is influenced. Naively, one might think that the locality will deteriorate, since the original links entering the covariant derivative of the filtered kernel spread over a larger volume. As first observed by Kovacs [29], the filtered overlap turns out to be even more local than the standard one and this is achieved without tuning $\rho$.

In figure 9 we plot the localization of $D_{\mathrm{ov}}$ at $\beta = 6.0$ with two projection parameters ($\rho = 1.0, 1.4$) and two filtering options (none, 1 HYP). The ordinate is the maximum over the 2-norm of $D_{\mathrm{ov}}\eta$ at $x$ with $\eta$ a normalized $\delta$-peak source vector at the point $y$ in the lattice, the abscissa is the "taxi driver" distance $d_1 = ||x-y||_1$ to the location of the $\delta$-peak, i.e. we plot the function

$$f(d_1) = \sup\left\{ ||(D_{\mathrm{ov}}\eta)(x)||_2 \;\middle|\; ||x-y||_1 = d_1 \right\} \tag{3.2}$$

versus $d_1$, as first studied in [28]. Comparing the two unfiltered operators (black/dark diamonds and crosses) one finds their result reproduced that (at this $\beta$) adjusting $\rho$ to a value around 1.4 lets $f(d_1)$ fall off steeper than with the value 1.0 which is the canonical choice in view of the spectrum of the Wilson operator sufficiently close to the continuum (cf. figure 1). The interesting observation is that a single HYP step together with $\rho = 1.0$ (red/light squares) results in an even steeper descent than the unfiltered version with $\rho = 1.4$ (which was chosen to nearly optimize the locality of the unfiltered operator). The last curve shown (red/light pluses) indicates that one should not attempt to combine the filtering with a $\rho$ value that would be optimal for the unfiltered operator.

An obvious question is whether filtering remains useful on fine lattices. Figure 10 shows the fall-off at four couplings with no smearing, 1 APE and 1 HYP step, with $\rho = 1$ fixed. On the coarsest lattice smearing alters the locality just modestly, on the two intermediate ones ($\beta = 5.84, 6.00$) the locality gets substantially improved, with HYP doing a better job than APE. On the finest lattice, the improvement is still sizable, but there is almost no difference among the two filtering recipes. At this coupling further smearing steps would then diminish the locality. The localization measured with the definition (B.1) [which we use for technical reasons discussed below] is summarized in table 4.

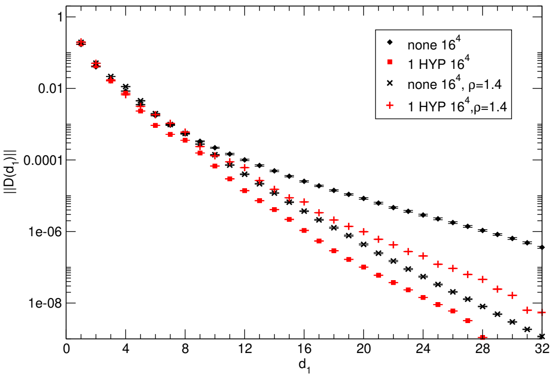


**Figure 9:** Localization of the overlap at $\beta = 6.0$ without filtering and after 1 HYP step, for $\rho = 1.0$ and $\rho = 1.4$. A single HYP step proves more efficient than optimizing $\rho$. Filtering and $\rho > 1$ should not be combined; for the filtered operator the untuned choice $\rho = 1$ is reasonable (but still not optimal).

There is a loose connection between the localization of $D_{\rm ov}$ and the spectrum of $H_{\rm W}$, for instance

$$||D_{\rm ov}(x,y)|| \leq \text{const} \times \exp\left(-\frac{\theta}{2}||x-y||_1\right) \tag{3.3}$$

is a bound found in [28], where $||.||$ is the matrix norm in Dirac and color space. The exponent $\theta/2$ in (3.3) is defined via the largest and smallest eigenvalue of $D^\dagger_{{\rm W},-\rho} D_{{\rm W},-\rho}$ through

$$\cosh(\theta) = \frac{\lambda_{\rm max}/\lambda_{\rm min} + 1}{\lambda_{\rm max}/\lambda_{\rm min} - 1} = \frac{1+\epsilon^2}{1-\epsilon^2} \tag{3.4}$$

where we like to express the r.h.s. in terms of the inverse condition number $\epsilon$ of $|H_{\rm W}|$. Expanding either side to first order one obtains the simple relation (after getting rid of the unphysical $\theta < 0$ solution)

$$\frac{\theta}{2} = \epsilon + O(\epsilon^2)\,. \tag{3.5}$$

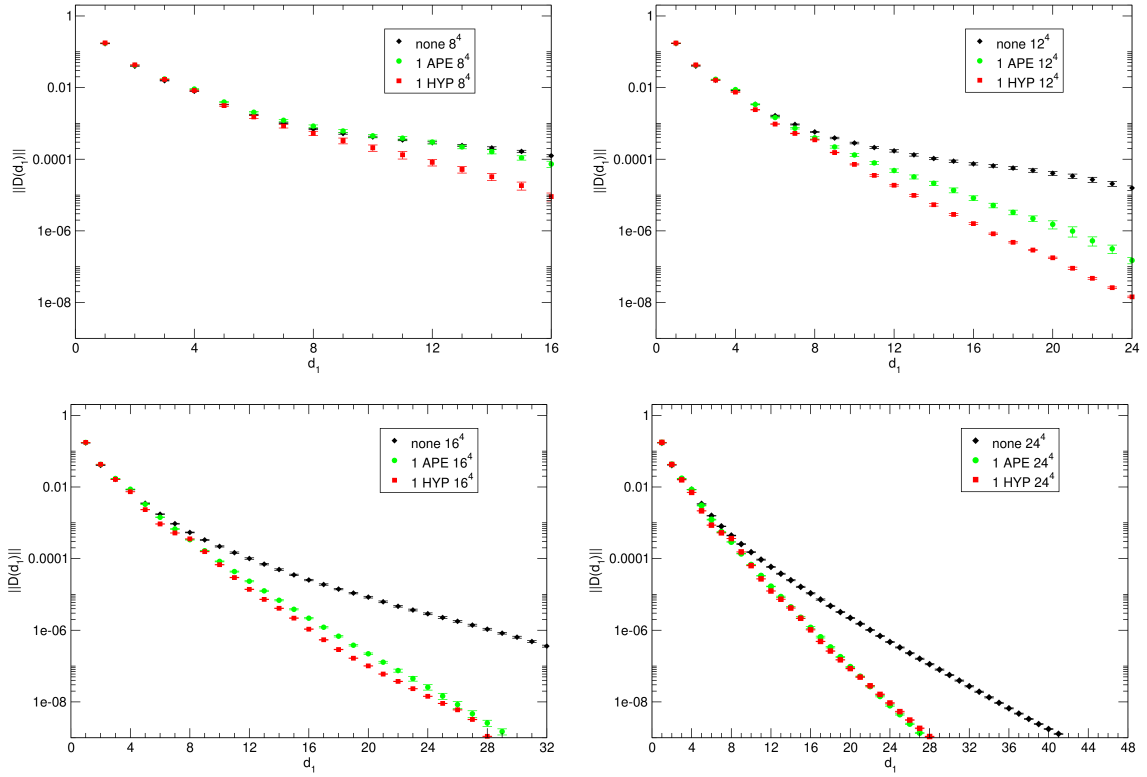


**Figure 10:** Localization of the unfiltered overlap and of the versions with one APE or HYP step. The data are for the matched ensembles ($8^4$ at $\beta = 5.66$, $12^4$ at $\beta = 5.84$, $16^4$ at $\beta = 6.00$, $24^4$ at $\beta = 6.26$) and $\rho = 1$ throughout. On sufficiently fine lattices the choice of smearing proves irrelevant.

As already mentioned in [28] the exponent $\theta/2$, defined via the spectral properties of the underlying $|H_{\mathrm{W}}|$, is a rather bad estimate for the actual localization $\nu$. The situation is not much better for the filtered variety, as a brief comparison of our tabulars 3 and 4 reveals.[3] Though quantitatively unsuccessful, this connection still gives a qualitative hint that the overlap operator with a filtered Wilson kernel might enjoy better localization properties due to the reduced condition number of $H_W$. There are more detailed bounds in the literature [30 − 34], but it seems fair to say that a quantitative understanding of the localization of $D_{\mathrm{ov}}$ in terms of the spectral properties of $H_W$ is a challenge.

[3]Table 3 contains the condition number on the subspace orthogonal to the 14 lowest modes, while $\epsilon$ in (3.3), (3.5) refers to the full operator. For two reasons we propose to re-interpret (3.5) as a prediction for the locality of $D_{\mathrm{ov}}$ with $\epsilon$ the ratio of the lower to the upper end of the *bulk* of eigenvalues of $|H_{\mathrm{W}}|$. A practical hint is that the unprojected condition number fluctuates wildly (see figure 6), whereas the localization is rather stable for all configurations in an ensemble. Furthermore, in [28] it is shown that an isolated near-zero mode of $|H_{\mathrm{W}}|$ does normally not affect the locality of $D_{\mathrm{ov}}$. Of course, one cannot repeat that argument indefinitely, but still a test whether a modified $\epsilon$ helps is interesting.

JHEP09(2005)030

| $\beta$ | 5.66 | 5.84 | 6.00 | 6.00 ($\rho = 1.4$) | 6.26 |
|---|---|---|---|---|---|
| none | 0.330(18) | 0.236(17) | 0.308(09) | 0.571(10) | 0.370(07) |
| 1 APE | 0.344(30) | 0.447(29) | 0.577(13) | 0.543(06) | 0.586(18) |
| 3 APE | 0.429(49) | 0.634(30) | 0.682(11) | 0.485(04) | 0.549(09) |
| 1 HYP | 0.469(48) | 0.642(32) | 0.695(10) | 0.480(05) | 0.554(05) |
| 3 HYP | 0.610(12) | 0.630(32) | 0.585(03) | 0.476(08) | 0.519(02) |

**Table 4:** Localization $\nu$ of the overlap operator with an unfiltered Wilson kernel and after 1 or 3 steps of APE or HYP filtering. At $\beta = 6.0$ we compare to $\rho = 1.4$ which is nearly optimal without filtering [28]. We use the definition (B.1); the error is only statistical.

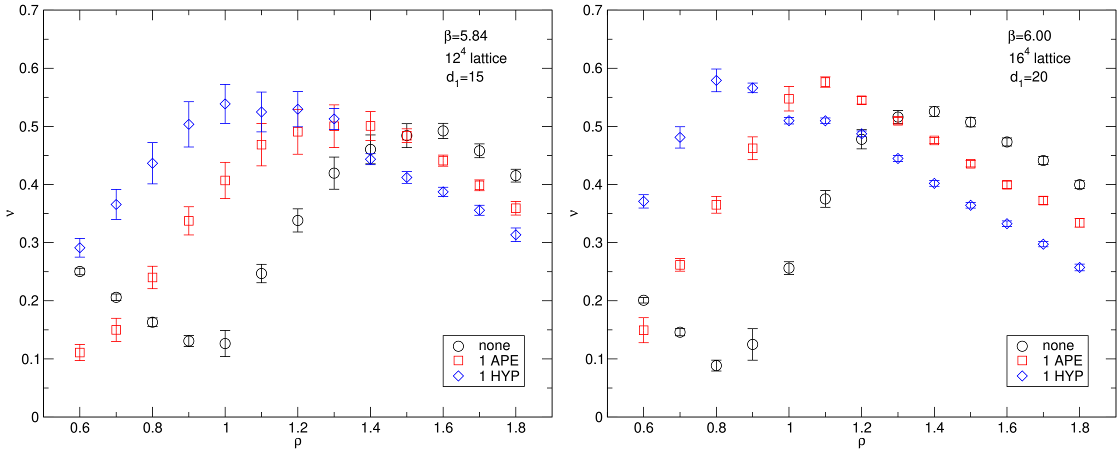


**Figure 11:** Localization $\nu$ vs. $\rho$ at $\beta = 5.84$ (left) and $\beta = 6.00$ (right) with no filtering, 1 APE or 1 HYP step. This is the only case where we deviate from our convention to define $\nu$ via (B.1) and use $d_1 = \frac{5}{4}L$.

The localization $\nu$ as a function of the projection parameter $\rho$ is presented in figure 11. For $\beta = 5.84, 6.00$ the optimum parameter for the unfiltered operator is around $\rho = 1.6, 1.4$, respectively. For the 1 HYP operator the localization at $\rho = 1.0$ does not fall short of the maximal one by a large amount; this is why we restrict much of our investigation with a filtered $D_{\mathrm{ov}}$ to the case $\rho = 1.0$. Still, the figure suggests that an optimal $\rho$ for the 1 HYP filtered operator may be *smaller* than 1, and it decreases with increasing $\beta$; at $\beta = 5.84$ we find $\rho_{\mathrm{opt}}^{1\,\mathrm{HYP}} \simeq 1.0$ and at $\beta = 6.00$ we find $\rho_{\mathrm{opt}}^{1\,\mathrm{HYP}} \simeq 0.8$.

After dealing with some technical issues to make sure that an $48^4$ lattice is large enough (see appendix B), we have studied $\nu$ defined via (B.1) as function of $\rho$ in the free case; the result is shown in figure 12. The pattern observed in figure 11 should thus not come as a surprise, filtering simply drives the locality properties of the overlap operator towards the free field case. In fact, figure 12 offers a simple explanation why it is so difficult to predict the localization $\nu$ from spectral properties of the underlying $|H_{\mathrm{W}}|$ operator — in the free case the inverse condition number (2.1) in the range $0 < \rho < 1$ is monotonic, while $\nu$ has a non-trivial extremum at $\rho_{\mathrm{opt}}^{\mathrm{free}} \simeq 0.54$.



JHEP09(2005)030

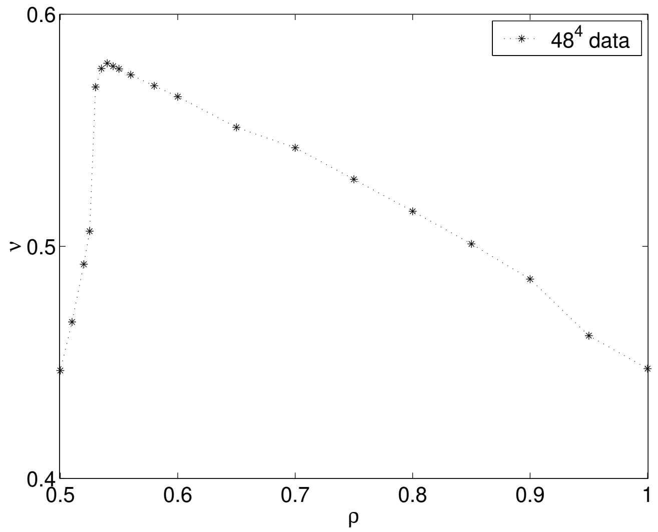


**Figure 12:** Localization (B.1) in the free field case with an extremum at $\rho_{\text{opt}}^{\text{free}} \simeq 0.54$ and a steep descent to the left. Note that in the $\rho$ range shown the inverse condition number $\epsilon = \rho/(8-\rho)$ is monotonic.

## 4. Kernel non-normality

An operator $A$ is called normal, if it commutes with its adjoint

$$[A, A^\dagger] = 0 \tag{4.1}$$

which implies that its left and right eigenbasis coincide. Normality has special implications for lattice Dirac operators. For a normal Dirac operator $D = \sum_k \lambda_k |k\rangle\langle k|$ which, in addition, is $\gamma_5$-hermitean

$$\gamma_5 D \gamma_5 = D^\dagger \tag{4.2}$$

we immediately obtain

$$D^\dagger = \sum_k \lambda_k^* |k\rangle\langle k| = \sum_k \lambda_k \gamma_5 |k\rangle\langle k| \gamma_5 \tag{4.3}$$

and this implies that eigenmodes with real $\lambda_k$ are chiral (or may be linearly combined to chiral modes in case of degeneracies). Furthermore, for such a $D$ the eigenvectors of the hermitean Dirac operator

$$H = \gamma_5 D = \sum_k \lambda_k \gamma_5 |k\rangle\langle k| \tag{4.4}$$

are given by $\sqrt{\lambda_k^*}|k\rangle \pm \sqrt{\lambda_k}\gamma_5|k\rangle$ with the corresponding eigenvalues $\pm|\lambda_k|$.

The continuum Dirac operator is normal, and so are the naive and staggered discretizations (but the latter two yield more than one flavor in the continuum limit). The GW relation (1.1) together with $\gamma_5$-hermiticity (4.2) also implies normality of the operator, hence $D_{\text{ov}}$ is normal. In fact, the overlap construction can be described as extracting the unique unitary part of $D_{\text{kern}}/\rho$ [35], and for a normal kernel it reduces to a simple radial projection of the $D_{\text{kern}}/\rho$ eigenvalues onto the unit circle.

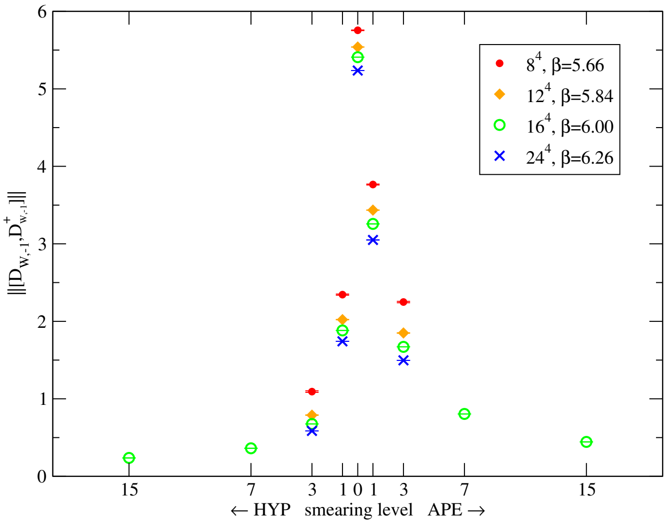


**Figure 13:** Non-normality (in lattice units) of the Wilson kernel, as defined in (4.5), as a function of smearing. The improvement of the kernel normality does not seem to degrade towards the continuum.

The shifted Wilson operator, which we use as a kernel, is not normal. Some consequences of this non-normality have been explored in other contexts [36]. Here, it suffices to point out that the relations between the eigenmodes of the overlap operator, its kernel and the hermitean Dirac operator are not as simple as above for the case of a normal operator. Typically, an eigenvector of the (hermitean) kernel will mix into every mode of the overlap operator, which we expect to have a detrimental effect on the efficiency of overlap construction algorithms. Thus, a practically relevant question is whether UV filtering can reduce the amount of non-normality of the overlap kernel.

To quantify the non-normality of $D_{\mathrm{W}}$ we measure the 2-norm of the commutator; technically

$$\left|\left|\left[D_{\mathrm{W},-1}, D_{\mathrm{W},-1}^{\dagger}\right]\left|\eta\right\rangle\right|\right| \tag{4.5}$$

is averaged over a number of normalized random vectors $|\eta\rangle$. In figure 13 the commutator (4.5) is shown for all $\beta$ and smearing levels (since this is not a physical observable, we use lattice units). Evidently, any kind of filtering reduces it — the filtered kernel is thus closer to normality and has left- and right-eigenvectors that are better aligned than for the unfiltered version. Whether "smart" overlap construction algorithms can be written which exploit this property is an open question.

## 5. Physics perspectives

To explore the physics potential of filtered overlap quarks a quenched spectroscopy study would be highly desirable. Physical results should reproduce — after a continuum extrapolation — results in the traditional "thin link" formulation. It would be interesting to see whether the speedup in point 1 of the introduction gets enhanced in points 2-4; in particular if scaling and/or asymptotic scaling set in earlier, this would make a real difference.



JHEP09(2005)030

Unfortunately, a detailed scaling study requires substantial computational resources, but as a first step in this direction we want to investigate the renormalization of the axial-vector current with filtered overlap quarks.

We follow the method of [37 − 39], where one starts from the usual (chirally rotated) densities

$$P(x) = \bar{\psi}_1(x)\gamma_5 \left[\left(1 - \frac{1}{2\rho}D_{\rm ov}\right)\psi_2\right](x) \tag{5.1}$$

$$A_\mu(x) = \bar{\psi}_1(x)\gamma_\mu\gamma_5 \left[\left(1 - \frac{1}{2\rho}D_{\rm ov}\right)\psi_2\right](x) \tag{5.2}$$

with $\psi_1 \neq \psi_2$ (flavor non-singlet) and defines the correlators $[x = (\mathbf{x}, t)]$

$$G_{PP}(t) = \sum_{\mathbf{x}} \langle P(\mathbf{x}, t) P^c(\mathbf{0}, 0)\rangle \tag{5.3}$$

$$G_{\nabla AP}(t) = \sum_{\mathbf{x}} \langle \bar{\nabla}_4 A_4(\mathbf{x}, t) P^c(\mathbf{0}, 0)\rangle \tag{5.4}$$

where $\bar{\nabla}_4$ is the symmetric derivative in the time direction and $P^c$ is the conjugate of (5.1), i.e. with the flavor indices $1 \leftrightarrow 2$ interchanged. With these correlators at hand one forms the ratio

$$\rho(t, m_1, m_2) = \frac{G_{\bar{\nabla}AP}(t)}{G_{PP}(t)} \tag{5.5}$$

where the second and third argument indicate that the spinors $\psi_1$ and $\psi_2$ in the densities (5.1), (5.2) are solutions to the massive operators $D_{{\rm ov},m_1}$ and $D_{{\rm ov},m_2}$, respectively. On account of the axial Ward identity (AWI) the ratio $\rho$ should be constant in time, and for light enough quarks (5.5) tends indeed to plateau rather nicely (see e.g. figure 1 in [39]). In a slightly sloppy but transparent notation the plateau value is $\rho(m_1, m_2)$. This quantity will — to the extent to which the AWI is respected at finite lattice spacing — only depend on the sum[4] of the quark masses, and thus defines the $m^{\rm AWI}$ quark masses

$$\rho(m_1, m_2) = \rho(m_1 + m_2) + O(a^2) = m_1^{\rm AWI} + m_2^{\rm AWI} + O(a^2)\,. \tag{5.6}$$

The actual data for our $Z_A$ determination for quenched filtered and unfiltered overlap quarks are generated with couplings and geometries as given in the last line of table 1. We restrict ourselves to the canonical choice $\rho = 1$. We plot $\rho(m_1, m_2)$ versus $m_1 + m_2$ for various quark mass combinations and filtering levels in figure 14 for $\beta = 5.66$ and in figure 15 for $\beta = 5.84, 6.00$, respectively. They form one universal band, i.e. different $m_1$ and $m_2$ combinations with a fixed sum $m_1 + m_2$ always give the same $\rho(m_1 + m_2)$ [within errors]. Furthermore, the relationship is in good approximation linear, but there is an anomaly without filtering at our strongest coupling (figure 14). Here, the slope is *negative*, and this

[4] In principle, we might use the covariant conserved current for overlap quarks (see [40] and the 2nd work in [44]) with the "thin" links replaced by "thick" links. Then the last term on the r.h.s. of (5.6), (5.7) would be absent, and the AWI would be an exact identity. However, there is a practical problem with APE or HYP filtering, due to the SU(3) projection involved. The solution via stout/EXP links is in exact analogy to the dynamical case discussed in the last section.

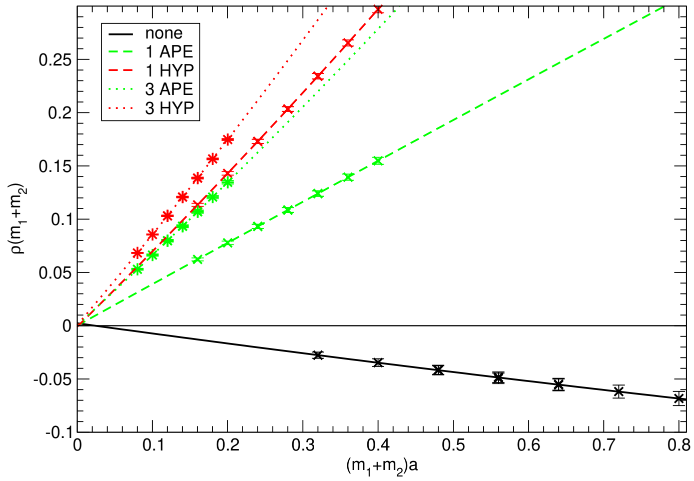


**Figure 14:** $m_1^{\mathrm{AWI}} + m_2^{\mathrm{AWI}}$ vs. $m_1^{\mathrm{bare}} + m_2^{\mathrm{bare}}$ [slope $= Z_A^{-1}$] at $\beta = 5.66$ . Without filtering $Z_A < 0$, while any filtering prescription gives $Z_A > 1$, with higher filtering levels resulting in a value closer to 1.

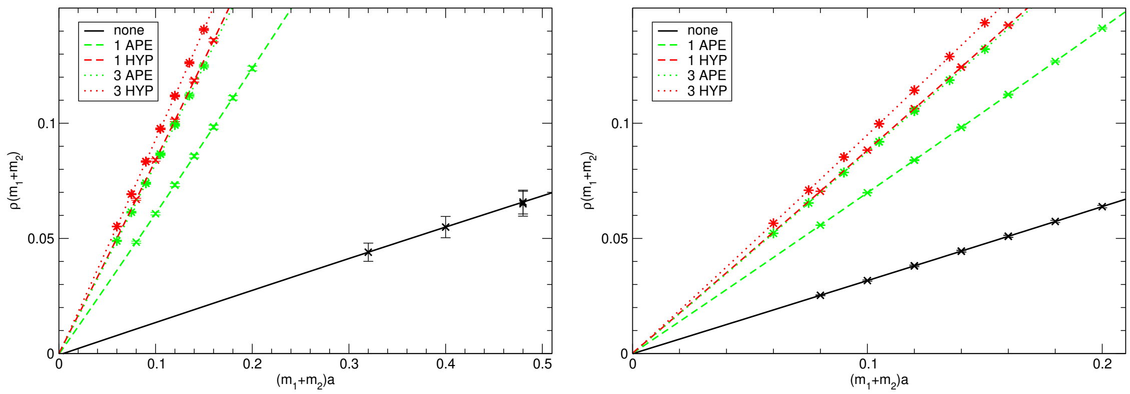


**Figure 15:** $m_1^{\mathrm{AWI}} + m_2^{\mathrm{AWI}}$ vs. $m_1^{\mathrm{bare}} + m_2^{\mathrm{bare}}$ at $\beta = 5.84, 6.00$. Higher filtering levels shift $Z_A$ closer to 1.

supports the view established in appendix A that with $\beta = 5.66$ and $\rho = 1.0$ the projection point is "in" or "to the left" of the physical branch of the underlying Wilson operator, and we effectively operate in the "zero fermion" sector. Also at $\beta = 5.84$ the unfiltered plateau was not very pronounced either, resulting in a large systematic uncertainty beyond the statistical error quoted below. We use the ansatz

$$\rho(m_1 + m_2) = \mathrm{const} + \frac{1}{Z_A}(m_1 + m_2) + \mathrm{const}\,(m_1 + m_2)^2 \tag{5.7}$$

and see whether we obtain acceptable fits and whether the first constant is consistent with zero. It turns out that this is the case, and the associate $Z_A$ values are summarized in table 5.

It is interesting to discuss both the general pattern of these $Z_A$ values and the relation to 1-loop perturbation theory. Evidently, at fixed $\beta$ and $\rho$ the filtered $Z_A$ is much closer to the tree-level value 1. We recover the relative strength ordering of section 2, i.e. one APE

JHEP09(2005)030

| $\beta$ | 5.66 | 5.84 | 6.00 | 6.00 ($\rho = 1.4$) |
|---|---|---|---|---|
| $Z_A^{\text{none}}$ | ill-def. | 7.06(73) | 3.145(94) | 1.554(1) [39] |
| $Z_A^{1\,\text{APE}}$ | 2.57(7) | 1.66(02) | 1.452(04) | — |
| $Z_A^{3\,\text{APE}}$ | 1.55(3) | 1.23(01) | 1.160(06) | — |
| $Z_A^{1\,\text{HYP}}$ | 1.44(2) | 1.22(01) | 1.153(03) | — |
| $Z_A^{3\,\text{HYP}}$ | 1.21(1) | 1.10(01) | 1.072(02) | — |
| $Z_A^{\text{none}}\vert_{1-\text{loop}}$ [41] | 1.280 | 1.272 | 1.264 | 1.120 |

**Table 5:** $Z_A$ determined via (5.7) with various filtering prescriptions. On the last line the 1-loop result [41] for the thin-link overlap $Z_{V,A} = 1 + C_F\, 0.198206 g_0^2 + O(g_0^4) = 1 + 1.585648/\beta + O(1/\beta^2)$ at $\rho = 1.0$ and $Z_{V,A} = 1 + C_F\, 0.090301 g_0^2 + O(g_0^4) = 1 + 0.722408/\beta + O(1/\beta^2)$ at $\rho = 1.4$ is added for comparison.

step is less efficient than 3 APE or a single HYP step, but the latter is topped by 3 HYP steps. At $\beta = 6.0$ we compare to $\rho = 1.4$ which is the the standard choice for the "thin link" overlap. Without filtering, $Z_A^{\text{none}}(\rho = 1.4) \simeq 1.554$ is about half of $Z_A^{\text{none}}(\rho = 1.0) \simeq 3.145$, and this means that the choice $\rho = 1.4$ is not just near-optimal w.r.t. locality, but also beneficial to tame (one particular) renormalization. Once the filtering recipe is specified, $Z_A$ seems to be monotonic in $6/\beta = g_0^2$, as expected from perturbation theory. In the unfiltered case the 1-loop value is included in the last line of table 5 for comparison. Assuming that in perturbation theory $1 < Z_A^{1\,\text{HYP}} < Z_A^{\text{none}}$ holds for $\rho = 1$, one may compare the deviation of the unfiltered $\beta = 6.0$ operator $3.145 - 1.264 = 1.881$ to $1.153 - 1 = 0.153$ which then amounts to an upper bound in the 1 HYP case. Evidently, the discrepancy is dramatically reduced, which in view of the perturbative results in [42, 43], should not come as a surprise. To get a slightly more quantitative view, we consider it useful to fit our data without filtering at $\rho = 1$ to a Pade-type ansatz of the form

$$Z_A^{\text{none}} = \frac{1 + c_1 x + c_2 x^2}{1 + (c_1 - 1.585648)x} \tag{5.8}$$

with $x = 1/\beta$, where the perturbative knowledge [41] (cf. caption of table 5) is built-in as a constraint. In the same spirit a Pade ansatz for any of the filtered operators reads

$$Z_A^{1\text{HYP}\,/\,3\text{HYP}} = \frac{1 + c_1 x + c_2 x^2}{1 + c_3 x} \tag{5.9}$$

with — as of now — no constraint on $c_1 - c_3$ yet. There is a problem with the functional forms (5.8), (5.9), since our data sets contain 2 and 3 entries, respectively, and there is zero degree of freedom. Still, for an illustration such a "fit" might be worth while, and the result is shown in figure 16. With sufficient data the curves would contain two pieces of information. The asymptotic slope for $x \to 0$ would predict the perturbative 1-loop coefficients for $Z_A^{1\,\text{HYP}}$, $Z_A^{3\,\text{HYP}}$. And the pole in (5.8), (5.9), i.e. the values $c_1 - 1.585648$ or $c_3$, respectively, would predict the coupling where the perturbative description breaks down. Hence, if the curves in figure 16 are indicative at all, it seems that filtering renders the perturbative 1-loop coefficient of $Z_A$ much smaller, but the perturbative range gets barely enhanced.

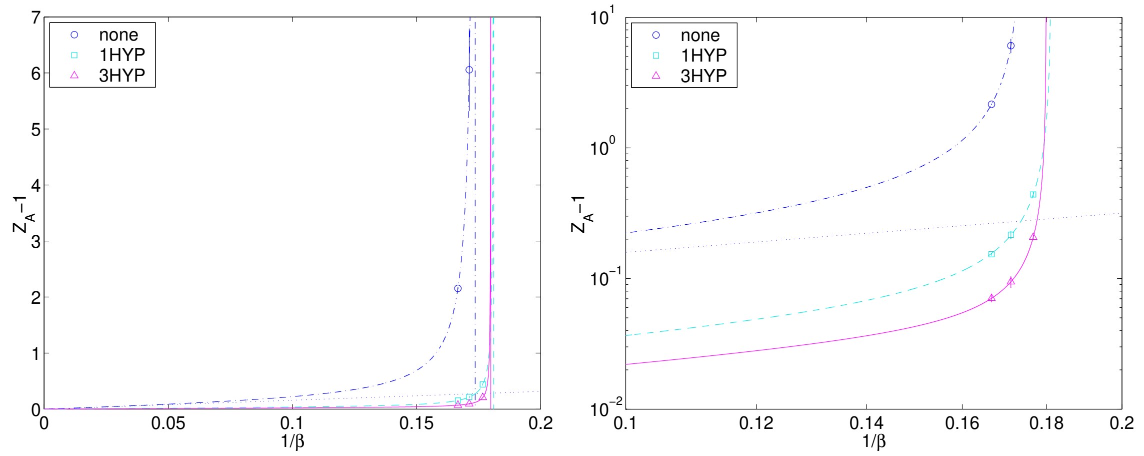


**Figure 16:** Pade functional forms through our $\rho = 1$ data in standard (left) or double logarithmic form (right). In the unfiltered case the constraint to reproduce the known 1-loop behavior is built in and the latter is indicated with a dotted line. It seems the 1-loop coefficients of the filtered $Z_A$ are dramatically reduced, while the perturbative range (blow-up point of the Pade ansatz) gets barely enlarged.

## 6. Discussion

In this paper we have studied the massless overlap operator constructed from a filtered Wilson kernel where the original "thin" links were replaced by "thick" links which behave in the same manner under local gauge transformations. This is a legal change of the fermion discretization as long as one particular filtering recipe [e.g. 1 HYP step with $\alpha_{\mathrm{HYP}} = (0.75, 0.6, 0.3)$] is maintained at all couplings. It amounts to an $O(a^2)$ re-definition of $D_{\mathrm{ov}}$ at fixed $\rho$, as does a change of $\rho$ at fixed filtering level.

Our key observations are the following. First, the onset of the "bulk" part of the spectrum of the underlying shifted hermitean Wilson operator $H_{\mathrm{W}} = \gamma_5(D_{\mathrm{W}} - \rho)$ gets lifted. This leads to an increased inverse condition number $\epsilon$ (after projection typically by a factor 2-4 through a single HYP step) and the latter reflects itself in a reduction (by the same factor) of the polynomial degree (and thus the number of forward applications of $H_{\mathrm{W}}^2$) needed to construct the inverse square root over the relevant range. What is the precise impact on CPU requirements to invert the massive operator is a topic for future research. Second, at standard couplings the filtered massless overlap is — even with the untuned canonical choice $\rho = 1$ — better localized than the unfiltered version with an optimally tuned $\rho$ could ever be. Our finding is backed by the observation that in the free case the optimum $\rho$ (w.r.t. locality) is around 0.54 and thus substantially smaller than the typical $\rho \simeq 1.4$ used in the past. Our third observation is that the filtered kernel is much closer to being a normal operator. In other words the left- and right-eigenvectors of $D_{\mathrm{W}}$ are better aligned with higher filtering level, and in this respect the effect of the "thick" links is the same as a shift much closer towards the continuum under which the overlap construction (1.3) tends to be a simple radial projection of the $D_{\mathrm{W}}$ eigenvalues. Finally, our fourth observation is that the renormalization constant of the axial-vector current is much closer to 1 with filtering than without. We rate this as a sign that lattice perturbation theory for the filtered overlap might work much better than for the unfiltered variety. If this is indeed so, and if it goes through for 4-fermion operators, it is likely to be the most important consequence of our work, since it offers the perspective of considerably reduced theoretical uncertainties in electroweak precision studies.

Let us finally discuss a variety of proposals in the literature that are similar in spirit to the one put forth in the present paper.

There is a top-level version deriving from "parametrized fixed-point fermions". The idea behind this approach pursued by Hasenfratz and Niedermayer is that true fixed-point fermions would satisfy the GW relation exactly [6], but a practical implementation is always ultralocal. Hence, sticking such an ansatz into the overlap formula (1.3) yields fermions with exact chiral symmetry and otherwise properties that are at least as good (but typically better) than the version with a plain Wilson kernel [44].

Bietenholz has considered a variety of actions, originally based on RG concepts [45]. The idea was that an action with a spectrum close to the GW circle could be iteratively improved in its chiral properties. Over time the focus has shifted towards using the overlap formula (1.3) to have exact chiral symmetry, but it is clear that the kernel of his "hypercubic overlap" benefits from a larger inverse condition number $\epsilon$ of $|\gamma_5 D_{\mathrm{kern}}|$ just as we do.

Gattringer and collaborators construct a "chirally improved" Dirac operator that involves the full Dirac Clifford algebra with links restricted to the hypercube. The coefficients are adjusted such that (for a given coupling) the violation of the GW relation is minimized [46]. The problem is the same as in the Bietenholz approach: a single forward application with such a kernel is so expensive that the improvement, if it is not "perfect", does not really pay off.

DeGrand has considered — both perturbatively and non-perturbatively — Wilson and clover action varieties that involve smeared gauge links [47, 42]. Based on this experience he went on to construct a "variant overlap" which starts from a kernel with only scalar/vector terms and smoothed links, and is thus sufficiently cheap as to allow for sticking it into the overlap formula [48, 43].

The closest to what we do is found in the work of Kovacs [29]. He uses a "fat-link clover" overlap in which all links are smeared, together with the tree-level value $c_{\mathrm{SW}} = 1$. As far as we know, he was the first author to notice that such a filtered kernel allows for the untuned choice $\rho = 1$, and still the resulting overlap shows good localization properties.

A related approach has been pursued by the Adelaide group [49]. Their "fat link irrelevant clover" overlap quarks are built from a clover action in which only the irrelevant pieces (i.e. the Wilson and the Sheikoleslami-Wohlert terms) use smeared links, but not the covariant derivative. They found a similar speedup factor in the construction of the overlap operator (cf. "step 1" in the introduction) and tied it to the reduced spectral density of $|\gamma_5 D_{\mathrm{kern}}|$ near the origin.

Finally, "overlap" quarks with smeared gauge links have been used by several lattice collaborations. RBC has found that the residual mass of domain-wall fermions at fixed $N_5$ gets reduced [50], though they miss out an important ingredient, the projection to SU(3).

UKQCD has used overlap valence quarks with 3-fold HYP smeared links on staggered sea as supplied by the MILC collaboration, finding a surprisingly good signal on as few as 10 configurations [51]. Similarly, LHP and NPLQCD have used filtered domain-wall valence quarks on staggered sea to compute the pion form factor [52] and the $I = 2$ $\pi\pi$-scattering length [53], respectively.

There is another idea that should not be confused with filtering. Using an improved gauge action has been found to reduce $\rho_{|H_\mathrm{W}|}(0)$ by up to an order of magnitude [54–56]. There is, however, an important practical difference to the filtering concept, which is a modification of the fermion action. As already discussed in [11], a better choice of the gauge action improves, in the first place, the very low end of the $|H_\mathrm{W}|$ eigenvalue distribution. After projecting out the lowest $O(15)$ eigenvectors (which nowadays is a standard thing to do [16]) much of the advantage is lost (in figure 3 of [56] the lifting factor diminishes to the right). By contrast, filtering lifts the complete low-energy end of the $|H_\mathrm{W}|$ eigenvalues (in our figure 3 one finds an almost-universal lifting factor) and the usefulness of filtering is not vitiated by the projection. Still, it might be interesting to see whether the two ideas can be fruitfully combined.

An extension of the filtering concept to full QCD is straightforward, albeit hampered by a technical problem. These days, most dynamical fermion simulations are set up with a HMC algorithm, and the latter requires the fermion action to be differentiable w.r.t. the gauge links. The kernel of our filtered overlap quarks is differentiable w.r.t. the "thick" links, but not w.r.t the elements of the original set, due to the projection involved in the APE [8] or HYP [9] procedure. A convenient way out is offered by the stout/EXP links introduced in [57], involving a differentiable mapping between the "thick" and "thin" links. In pure gauge observables the usefulness of this smearing recipe was found to be restricted to small parameter values [57, 18], and one may fear that this feature persists in stout/EXP overlap quarks, since in perturbation theory they are equivalent to APE filtered overlap fermions with $\alpha_\mathrm{APE} = 1/(1 - 6\alpha_\mathrm{EXP})$ [58]. Thus, due to the pole at $\alpha_\mathrm{EXP} = 1/6$ we expect them to have a "narrow therapeutic range" in parameter space, but it is clear that there is no conceptual issue in simulating full QCD with filtered overlap quarks beyond the difficulties met in the unfiltered case [59].

To summarize, our suggestion is to use the overlap recipe (1.3) with an unimproved ($c_\mathrm{SW} = 0$) Wilson kernel in which all links are replaced by some smeared descendents of the actual gauge background. We recommend to stay with a moderate amount of link "fattening", e.g. with a single step of standard HYP smearing [9]. The projection parameter $\rho$ may be fixed at its canonical value 1, and in this sense the filtered overlap involves *less tuning* than the unfiltered version.[5] An important restriction is that the choice of iteration level and smearing parameter must be the same for all couplings considered in a scaling study. This is one point on which our proposal differs from some of the attempts reviewed above which involve coefficients (e.g. in the extended $\gamma$-algebra) that are adjusted "by hand" to yield a GW-type spectrum at one standard value of the gauge coupling. The

[5]Of course, there are parameters in the filtering recipe, but our results show that they hardly matter. Thus filtering allows one to trade a parameter that needs to be tuned for parameters on which the lattice data show very little sensitivity.

other difference is that our kernel remains cheap and still requires fewer $D_{\text{kern}}^{\dagger}D_{\text{kern}}$ forward applications. From a practical viewpoint, a clear advantage is the ease of implementation of the "filtered overlap" — everyone with a running overlap code has it (in disguise).

## Acknowledgments

It is a pleasure to thank Ferenc Niedermayer and Tom DeGrand for useful conversation or correspondence. This paper was supported by the Swiss NSF.

## A. Cumulative eigenvalue distributions in 4D and 2D

In this appendix we discuss what can be learned from the cumulative eigenvalue distribution (CED). We consider both the eigenvalues in 4D generated for the main part of this paper and data from dedicated runs in the quenched Schwinger model (QED with massless fermions in 2D) to elucidate the effects that filtering and changing $\beta$ have on the spectral density of the hermitean Wilson operator $H_{\text{W}} = \gamma_5 D_{\text{W},-1}$.

Figure 17 presents the cumulative eigenvalue distribution (CED) of the 15 smallest eigenvalues of $|H_{\text{W}}|$ on the ensembles discussed before. We show it both in standard form and in double logarithmic form, and the scale on the ordinate follows from the requirement that it would extend up to 1, if all eigenvalues were calculated (cf. figure 18 below). For the two intermediate couplings ($\beta = 5.84, 6.00$) we see the expected linear rise of the CED near the origin, which soon gets complemented by a higher order piece. The coefficient of the linear part is a measure for the spectral density of the hermitean Wilson operator at the origin, $\rho_{|H_{\text{W}}|}(0)$. That density being non-zero means that there is a finite probability to encounter arbitrarily small eigenvalues. The main effect of smearing is to reduce this spectral density, as is evident from the double logarithmic plots — here the initial slope 1 piece gets shifted downwards, and this corresponds to a smaller coefficient in front of the linear piece in the standard representation. Our data at $\beta = 6.26$ are of lesser quality — here we definitely cannot identify a linearly dominated regime. The situation is far more favorable in figure 18 where quenched Schwinger model data are shown. Apart from the statistics, the main difference is that all eigenvalues (extending up to $\sim 3$ in 2D) are included. The higher the filtering level or $\beta$, the more pronounced is the "jump" in the CED at $\lambda \simeq 1$. Note that with a chiral kernel *all* eigenvalues of $|\gamma_5 D_{\text{kern}}|$ would be there, i.e. the CED would be a step function at $\lambda = 1$. Finally, to come back to figure 17, the situation at the strongest coupling ($\beta = 5.66$) is different, since here the linear piece in the unfiltered CED is *not* larger than in the filtered versions. This is, because our choice $\rho = 1$ lets us "loose" the fermion — at this coupling our projection point is somewhere "in" the physical branch or "to the left" of it, while for the filtered version $\rho = 1$ is still appropriate. One might avoid such a situation by choosing a larger $\rho$ with the unfiltered kernel, but an even safer option might be to refrain from simulating unfiltered overlap quarks on such coarse lattices. It looks like this is a situation where the filtered overlap may help a lot, since it allows simulations on coarser lattices than the unfiltered operator, but in order to really be useful such simulations should be in the scaling regime (and not just in the right universality class), and this is, of course, not yet clear.



JHEP09(2005)030

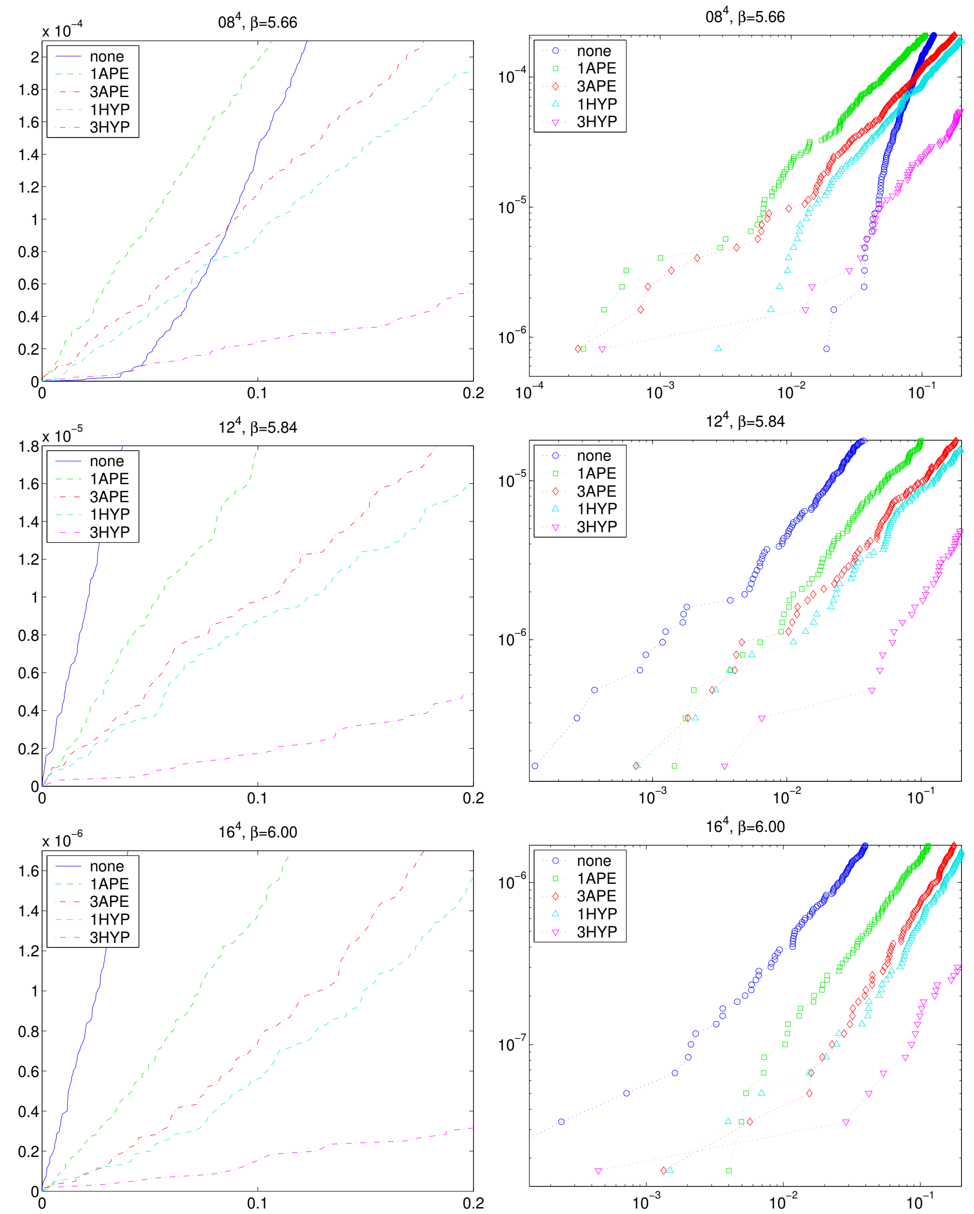


**Figure 17:** Cumulative eigenvalue distribution (CED) of $|H_{\mathrm{W}}|$ in standard (left) or double logarithmic (right) form for $\beta = 5.66, 5.84, 6.00$ (from top to bottom) with 0,1,3 steps of APE or HYP filtering. At each $\beta$ the upper cut in the vertical direction is the same on the left and on the right and the upper end in the horizontal direction is 0.2 throughout. Note the change in the ordinate scale between different couplings. At $\beta = 6.26$ we definitely lack the statistics needed to see a linearly dominated regime.

The spectral properties of $H_{\mathrm{W}}$ play a role in the context of the physical interpretation of the Aoki phase [60]. The latter is a conjectured phase, originally specific to $N_f = 2$ active Wilson fermions at negative mass, in which after switching off an external trigger term

$$S_{\mathrm{source}} = \pm h \overline{\psi} \gamma_5 \sigma_3 \psi \tag{A.1}$$

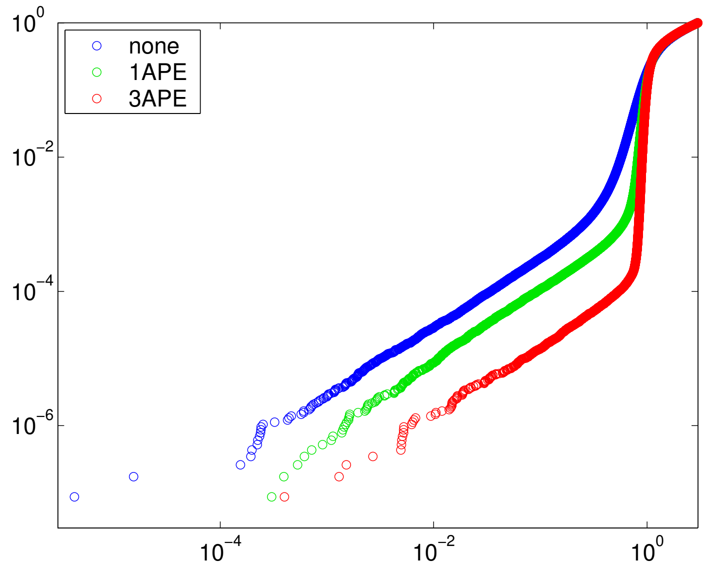


**Figure 18:** Log-log plot of the cumulative eigenvalue distribution (CED) of $|H_\mathrm{W}|$ in the Schwinger model ($N_f = 0$, $16^2$ geometry, $\beta = 3.2$, $\rho = 1$) without filtering (top curve) and after 1 or 3 filtering steps. In all cases the CED starts out linearly, raises sharply somewhere near $\lambda \simeq 1$ and reaches 1 at $\lambda \simeq 3$ (in 2D). This figure includes all eigenvalues of all three operators on 22,500 decorrelated configurations.

parity and flavor break spontaneously and a condensate (const $\neq 0$)

$$\lim_{h\to 0^\pm} \langle \overline{\psi} \gamma_5 \sigma_3 \psi \rangle = \pm\mathrm{const} \tag{A.2}$$

forms. Good numerical evidence for a non-zero condensate (A.2) in the (dynamical) 2-flavor case for an appropriate choice of the negative mass $-\rho(\beta)$ is found in [61]. Ref. [62] argues that in the massless limit of the continuum theory a condensate of the form (A.2) is simply an axial rotation of the usual (flavor diagonal) condensate and thus breaks neither parity nor flavor. They relate the spectral density of $D_\mathrm{W}$ to that of $H_\mathrm{W}$ and argue that the absence of a gap (around the origin) of the latter is indicative of chiral symmetry breaking and that $\rho_{|H_\mathrm{W}|}(0) > 0$ if and only if (A.2) is non-zero. This was later elucidated to be a continuum argument [63], which — in view of our section 4 — might be an important point.

The next issue is whether there is an Aoki phase in the quenched theory with 2 valence (but 0 sea) flavors [64, 65]. The simplest expectation is that qualitatively the picture with the 5 Aoki "fingers" goes through, though the phase boundary is somewhat shifted w.r.t. the $N_f = 2$ case.

Our 4D data in figure 17 clearly show the suppression of $\rho_{|H_\mathrm{W}|}(0)$ as one approaches the continuum, but we cannot see any sign that this distribution would vanish at some "critical" coupling. Given the uniform pattern in the figures (apart from the scale on the $y$-axis they seem qualitatively similar), it seems more likely to us that $\rho_{|H_\mathrm{W}|}(0)$ will stay non-zero for arbitrary couplings.

To test this view, we analyze the quenched Schwinger model where high statistics can be reached. The couplings and geometries are chosen such as to have a fixed physical volume, with a box size about 5 times larger than the Compton wavelength of the lightest



JHEP09(2005)030

| $\beta$ | 0.8 | 1.8 | 3.2 | 5.0 | 7.2 |
|---|---|---|---|---|---|
| geometry | 8×8 | 12×12 | 16×16 | 20×20 | 24×24 |
| statistics | 90,000 | 40,000 | 22,500 | 14,400 | 10,000 |
| statistics | — | — | — | 275,000 | 75,000 |

**Table 6:** Matched quenched 2D lattices; reweighting to $N_f = 1$ the lightest particle in the chiral limit would fit 5 times into the box. In the first set all $|H_\mathrm{W}|$ eigenvalues are determined with statistics such as to have an equal number of eigenvalues. In the second set only the lowest 10 eigenvalues are computed.

degree of freedom in the chiral limit of the $N_f = 1$ theory. A survey of the parameters is given in table 6 and for technical details we refer to [12].

Figure 19 provides an overview over the complete $|H_\mathrm{W}|$ eigenvalue distribution; one sees a "peak" at $\lambda = 1$ forming that gets more pronounced with higher $\beta$ and higher filtering level. This "peak" corresponds to the "jump" at $\lambda = 1$ in the CED of $|H_\mathrm{W}|$ in figure 18.

Figure 20 presents the distribution of the lowest eigenvalue of $|H_\mathrm{W}|$. At low $\beta$ this distribution accumulates at zero, at intermediate values of the coupling there is a horizontal band of eigenvalues connecting down to zero, and at the largest $\beta$ there are just scattered eigenvalues. Evidently, one cannot draw a final conclusion whether these scattered eigenvalues really make up for a non-zero $\rho(0)$, but it seems worth while to study this band in the region of $\beta$ values where it is clearly visible and see whether changing $\beta$ implies some structure, or whether it just stays flat, regardless of $\beta$.

Figure 21 presents the CED in the area of interest, the very-low $\lambda$ region. At $\beta = 5.0, 7.2$ we show the data from the high-statistics run with 10 eigenvalues per configuration (bottom line of table 6), but we checked that the results are consistent with what we get from the runs where all eigenvalues were determined. At a given coupling filtering clearly reduces $\rho_{|H_\mathrm{W}|}(0)$. The overall impression is that changing $\beta$ merely rescales the $y$-axis, in striking analogy with what we have seen in 4D (figure 17). If this is indeed true, the natural conclusion is that $\rho_{|H_\mathrm{W}|}(0)>0$ at any finite $\beta$ in the quenched theory.

Figure 22 contains a summary of our determinations of the spectral densities $\rho_{|H_\mathrm{W}|}(0)$ for various $\beta$ and filtering levels, extracted from the initial slopes in the CED shown in figure 21. It looks like eventually the density decreases exponentially in $\beta$ and changing the filtering level amounts to an overall rescaling factor which is, in good approximation, independent of $\beta$. Obviously, this is just numerical evidence, but the message seems to be as clear as one can possibly hope for from a numerical experiment. Note that for practical reasons we cannot take the infinite volume limit, but given our physical box size we expect finite volume effects to be exponentially small.

We remind the reader that (both in 2D and in 4D) we were working at fixed negative mass $m_0 = -1$ and pushed towards the continuum line. In other words, we were trying to stay as far outside the Aoki phase as one can, if one wants to be in the supercritical region with one overlap fermion. Of course, our data do not exclude the existence of a critical $\beta$, but they favor the view that there is no $\beta_\mathrm{crit}$ that makes the $|H_{\mathrm{W},-1}|$ spectral density strictly zero and therefore we conjecture that $\rho_{|H_\mathrm{W}|}(0) > 0$ throughout the supercritical region. Still, we do not see why this would create a problem for the localization of the overlap operator, since the two seem not one-to-one inversely connected.



JHEP09(2005)030

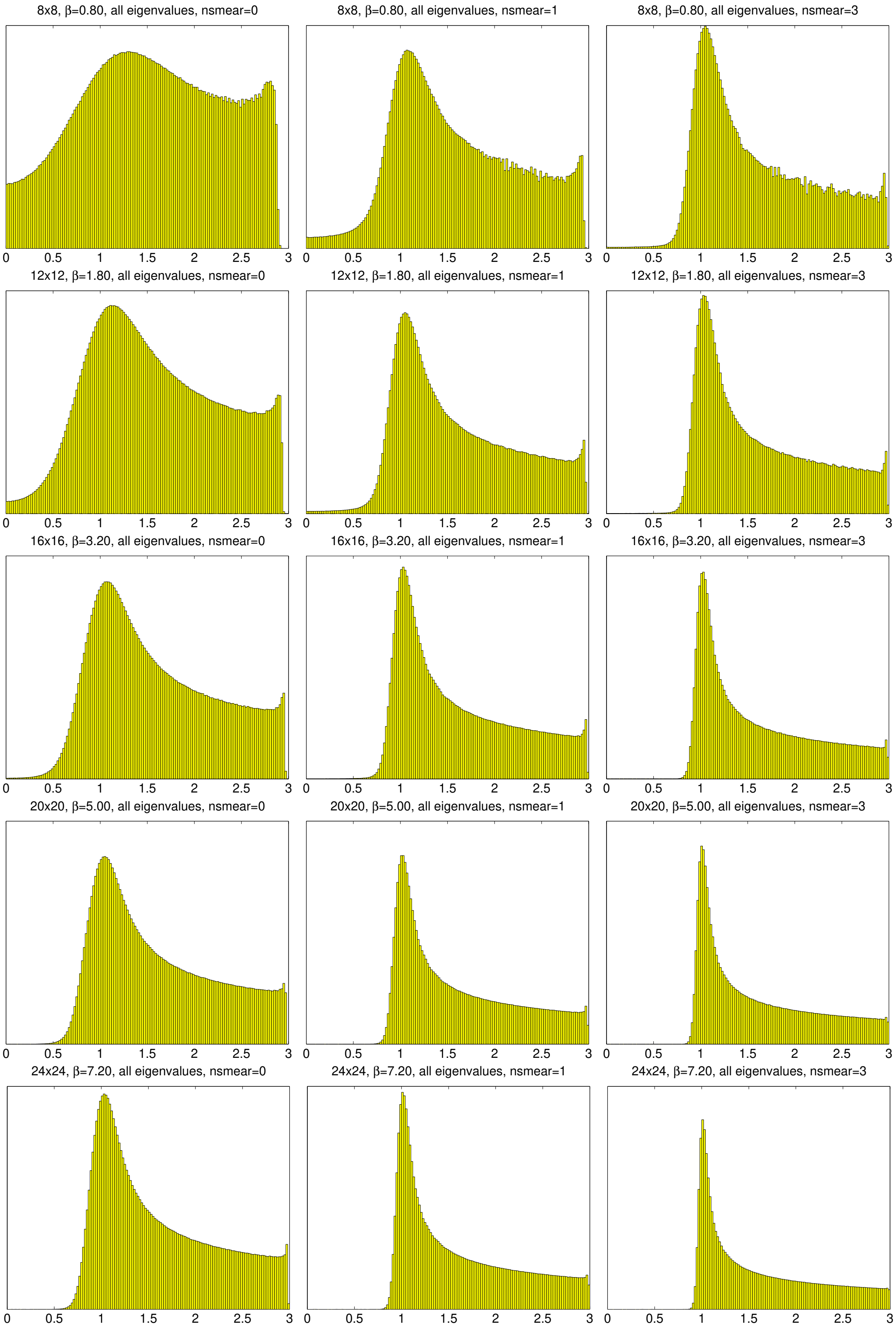


**Figure 19:** Distribution of all eigenvalues of $|H_\mathrm{W}| = |\gamma_5 D_{\mathrm{W},-1}|$ at five $\beta$ and 0,1,3 smearings.

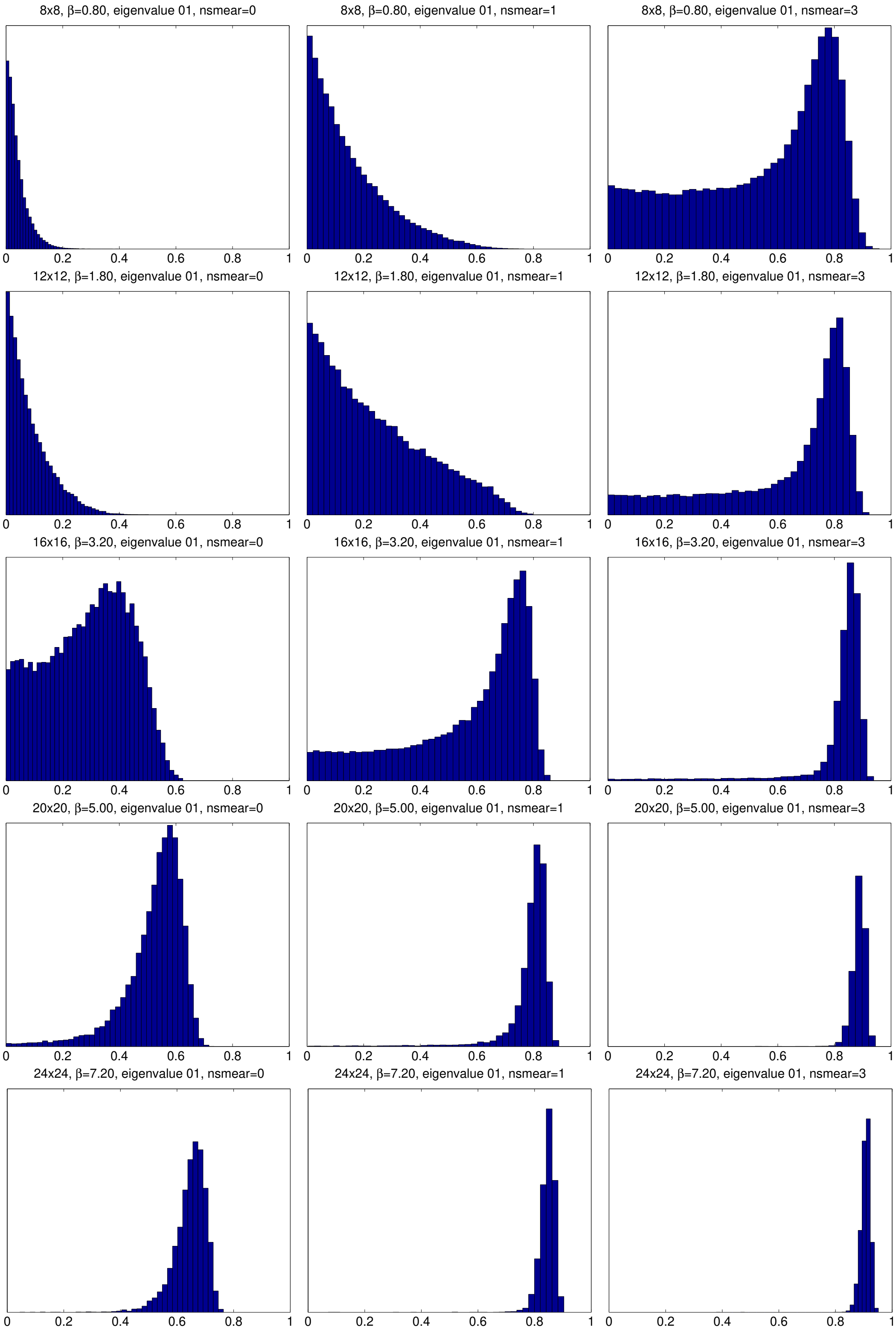


**Figure 20:** Distribution of the 1st eigenvalue of $|H_{\mathrm{W}}| = |\gamma_5 D_{\mathrm{W},-1}|$ at five $\beta$, with 0,1,3 smearings.

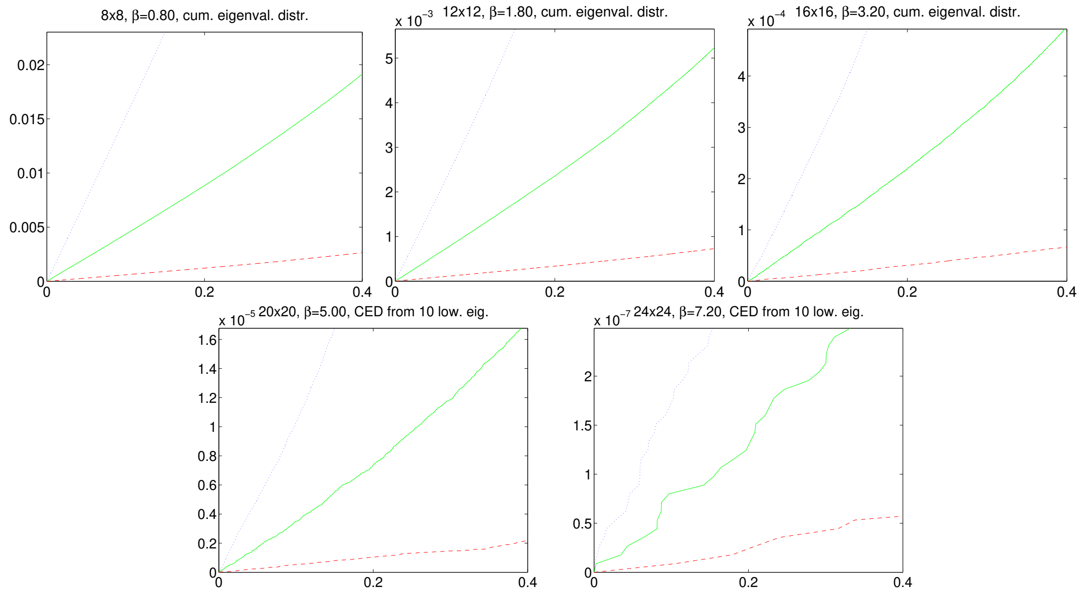


**Figure 21:** Cumulative eigenvalue distribution of $|H_{\mathrm{W}}| = |\gamma_5 D_{\mathrm{W},-1}|$ at five $\beta$, with 0,1,3 smearings.

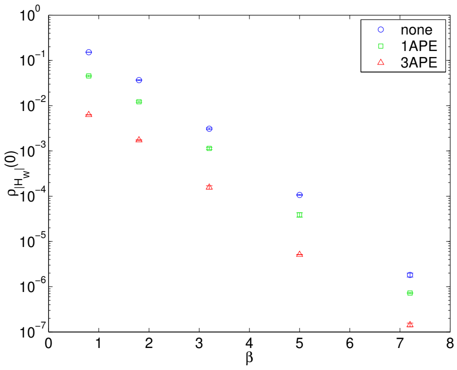


**Figure 22:** $\rho_{|H_{\mathrm{W}}|}(0)$ versus $\beta$ in the quenched Schwinger model. Filtering seems to reduce, for all $\beta$, the eigenvalue density by a near-universal factor. The data suggest an exponential fall-off at large $\beta$.

Finally, there is a simple argument that $\rho_{|H_{\mathrm{W}}|}(0) > 0$ holds in all quenched or unquenched theories with a massive overlap determinant at all couplings. Acquire infinite statistics at $\beta = 0$, $N_f = 0$. When integrating over the full configuration space the spectral density is certainly non-zero. Results for the case of interest at finite $\beta$ and maybe finite $N_f$ can be obtained through reweighting. As long as one can guarantee that there is no configuration where the reweighting factor vanishes, the spectral density will be modified, but it cannot be made strictly zero. This holds true in the quenched case and in the dynamical theory with an overlap determinant ($m \neq 0, -2\rho$), but it would not be true with Wilson fermions at a negative mass.

## B. Overlap operator locality in the free case

In this appendix we collect some technical points to make sure that a numerical investigation of the localization $\nu$ versus $\rho$ as shown in figure 12 for a $48^4$ lattice is not overwhelmed with finite size effects.

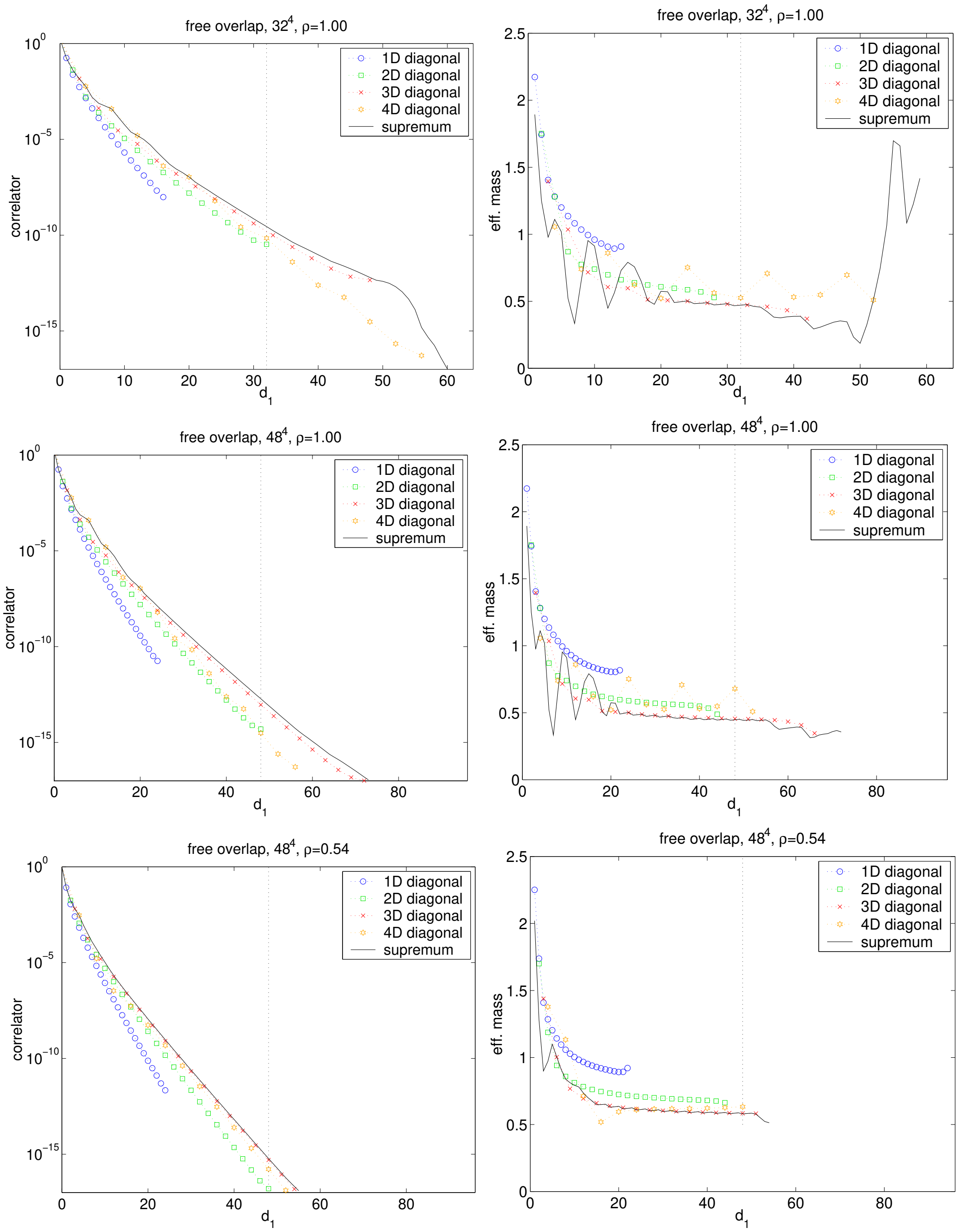


**Figure 23:** Free overlap couplings in four directions and the function (3.2) together with their "effective masses" for ($32^4$, $\rho = 1$) and ($48^4$, $\rho = 1$). Some effective masses are missing, since correlator values below $10^{-17}$ have been cut off (double precision limit). The good agreement of the $32^4$ and $48^4$ data with evaluation point $d_1 = L$ (dotted vertical lines) suggests that these data are much less affected by finite size effects than those near the maximal $d_1 = 2L$. The correlator with $\rho = 0.54$ is visibly steeper.

Considering $f(d_1)$ as given in (3.2) [28] on a finite lattice one encounters a technical problem that is evident in figure 23. The free field case is far from showing rotational symmetry and the supremum function (3.2) has a couple of initial bumps (in particular at $d_1 = 4, 8, 12$), and this means that one needs to go to sufficiently large distances to measure the slope in a logarithmic representation. On the other hand, the choice to measure the distance in the 1-norm leads to rather large finite size effects for $d_1 > L$, in particular the region near the maximal distance $d_1 = 2L$ is heavily contaminated. Therefore, we tried

$$\nu = \frac{1}{2}\log\left(\frac{f(L-1)}{f(L+1)}\right) \tag{B.1}$$

as a technical definition of the localization $\nu$ in (3.1). The comparison between the $32^4$ and $48^4$ geometries shows that our choice to evaluate the logarithmic derivative at $d_1 = L$ produces rather consistent $\nu$ values, and we take this as a sign that they cannot be far from the asymptotic exponent. For the (two-digit-precision) projection parameter that we find to be optimal w.r.t. locality in the free case, $\rho_{\mathrm{opt}}^{\mathrm{free}} \simeq 0.54$, the correlator is explicitly shown to be steeper than in the $\rho = 1$ case.

## References


[1] D.B. Kaplan, *A method for simulating chiral fermions on the lattice*, *Phys. Lett.* **B 288** (1992) 342 [hep-lat/9206013].

[2] Y. Shamir, *Chiral fermions from lattice boundaries*, *Nucl. Phys.* **B 406** (1993) 90 [hep-lat/9303005].

[3] R. Narayanan and H. Neuberger, *Chiral determinant as an overlap of two vacua*, *Nucl. Phys.* **B 412** (1994) 574 [hep-lat/9307006]; *A construction of lattice chiral gauge theories*, *Nucl. Phys.* **B 443** (1995) 305 [hep-th/9411108];
H. Neuberger, *Exactly massless quarks on the lattice*, *Phys. Lett.* **B 417** (1998) 141 [hep-lat/9707022]; *More about exactly massless quarks on the lattice*, *Phys. Lett.* **B 427** (1998) 353 [hep-lat/9801031].

[4] P.H. Ginsparg and K.G. Wilson, *A remnant of chiral symmetry on the lattice*, *Phys. Rev.* **D 25** (1982) 2649.

[5] M. Lüscher, *Exact chiral symmetry on the lattice and the Ginsparg- Wilson relation*, *Phys. Lett.* **B 428** (1998) 342 [hep-lat/9802011].

[6] P. Hasenfratz, V. Laliena and F. Niedermayer, *The index theorem in QCD with a finite cut-off*, *Phys. Lett.* **B 427** (1998) 125 [hep-lat/9801021].

[7] F. Niedermayer, *Exact chiral symmetry, topological charge and related topics*, *Nucl. Phys.* **73** *(Proc. Suppl.)* (1999) 105 [hep-lat/9810026].

[8] APE collaboration, M. Albanese et al., *Glueball masses and string tension in lattice QCD*, *Phys. Lett.* **B 192** (1987) 163.

[9] A. Hasenfratz and F. Knechtli, *Flavor symmetry and the static potential with hypercubic blocking*, *Phys. Rev.* **D 64** (2001) 034504 [hep-lat/0103029];
A. Hasenfratz, R. Hoffmann and F. Knechtli, *The static potential with hypercubic blocking*, *Nucl. Phys.* **106** *(Proc. Suppl.)* (2002) 418 [hep-lat/90110168].

[10] T. Blum et al., *Improving flavor symmetry in the Kogut-Susskind hadron spectrum*, *Phys. Rev.* **D 55** (1997) 1133 [hep-lat/9609036]; *Improving flavor symmetry in the Kogut-Susskind hadron spectrum*, *Phys. Rev.* **D 55** (1997) 1133 [hep-lat/9609036];
G.P. Lepage, *Flavor symmetry restoration and symanzik improvement for staggered quarks*, *Phys. Rev.* **D 59** (1999) 074502 [hep-lat/9809157];
K. Orginos, D. Toussaint and R.L. Sugar [MILC Collab.], *Phys. Rev.* **D 60** (1999) 054503 [hep-lat/9903032].

[11] S. Durr, C. Hoelbling and U. Wenger, *Staggered eigenvalue mimicry*, *Phys. Rev.* **D 70** (2004) 094502 [hep-lat/0406027].

[12] S. Dürr and C. Hoelbling, *Staggered versus overlap fermions: a study in the schwinger model with $N_f = 0, 1, 2$*, *Phys. Rev.* **D 69** (2004) 034503 [hep-lat/0311002]; *Scaling tests with dynamical overlap and rooted staggered fermions*, *Phys. Rev.* **D 71** (2005) 054501 [hep-lat/0411022].

[13] R. Sommer, *A new way to set the energy scale in lattice gauge theories an its applications to the static force and alpha-s in* SU(2) *Yang-Mills theory*, *Nucl. Phys.* **B 411** (1994) 839 [hep-lat/9310022].

[14] ALPHA collaboration, M. Guagnelli, R. Sommer and H. Wittig, *Precision computation of a low-energy reference scale in quenched lattice QCD*, *Nucl. Phys.* **B 535** (1998) 389 [hep-lat/9806005].

[15] T. Kalkreuter and H. Simma, *An accelerated conjugate gradient algorithm to compute low lying eigenvalues: a study for the dirac operator in* SU(2) *lattice QCD*, *Comput. Phys. Commun.* **93** (1996) 33 [hep-lat/9507023].

[16] P. Hernandez, K. Jansen and L. Lellouch, *A numerical treatment of neuberger's lattice Dirac operator*, hep-lat/0001008.

[17] J. Kiskis, R. Narayanan and H. Neuberger, *Does the crossover from perturbative to nonperturbative physics in QCD become a phase transition at infinite n?*, *Phys. Lett.* **B 574** (2003) 65 [hep-lat/0308033].

[18] S. Durr, *Gauge action improvement and smearing*, hep-lat/0409141.

[19] B. Bunk, *Fractional inversion in Krylov space*, *Nucl. Phys.* **63** *(Proc. Suppl.)* (1998) 952 [hep-lat/9805030].

[20] R.G. Edwards, U.M. Heller and R. Narayanan, *A study of practical implementations of the overlap-Dirac operator in four dimensions*, *Nucl. Phys.* **B 540** (1999) 457 [hep-lat/9807017].

[21] J. van den Eshof, A. Frommer, T. Lippert, K. Schilling and H.A. van der Vorst, *Numerical methods for the QCD overlap operator. I: sign- function and error bounds*, *Comput. Phys. Commun.* **146** (2002) 203 [hep-lat/0202025].

[22] H. Neuberger, *Vector like gauge theories with almost massless fermions on the lattice*, *Phys. Rev.* **D 57** (1998) 5417 [hep-lat/9710089].

[23] A. Borici, *Truncated overlap fermions*, *Nucl. Phys.* **83** *(Proc. Suppl.)* (2000) 771 [hep-lat/9909057].

[24] R. Narayanan and H. Neuberger, *An alternative to domain wall fermions*, *Phys. Rev.* **D 62** (2000) 074504 [hep-lat/0005004].

[25] A. Borici, A.D. Kennedy, B.J. Pendleton and U. Wenger, *The overlap operator as a continued fraction*, *Nucl. Phys.* **106** *(Proc. Suppl.)* (2002) 757 [hep-lat/0110070];
U. Wenger, *The overlap Dirac operator as a continued fraction*, hep-lat/0403003.

[26] R.C. Brower, H. Neff and K. Orginos, *Moebius fermions: improved domain wall chiral fermions*, hep-lat/0409118.

[27] I. Horvath, *Ginsparg-Wilson relation and ultralocality*, *Phys. Rev. Lett.* **81** (1998) 4063 [hep-lat/9808002]; *Ginsparg-Wilson-Luscher symmetry and ultralocality*, *Phys. Rev.* **D 60** (1999) 034510 [hep-lat/9901014]; *Ginsparg-Wilson relation and ultralocality*, *Phys. Rev. Lett.* **81** (1998) 4063 [hep-lat/9808002]. W. Bietenholz, *On the absence of ultralocal Ginsparg-Wilson fermions*, [hep-lat/9901005].

[28] P. Hernandez, K. Jansen and M. Lüscher, *Locality properties of Neuberger's lattice Dirac operator*, *Nucl. Phys.* **B 552** (1999) 363 [hep-lat/9808010].

[29] T.G. Kovacs, *The topological susceptibility with dynamical overlap fermions*, hep-lat/0111021; T.G. Kovacs, *Locality and topology with fat link overlap actions*, *Phys. Rev.* **D 67** (2003) 094501 [hep-lat/0209125].

[30] H. Neuberger, *Bounds on the Wilson Dirac operator*, *Phys. Rev.* **D 61** (2000) 085015 [hep-lat/9911004].

[31] Y. Kikukawa, *Locality bound for effective four-dimensional action of domain-wall fermion*, *Nucl. Phys.* **B 584** (2000) 511 [hep-lat/9912056].

[32] K. Fujikawa and M. Ishibashi, *Locality properties of a new class of lattice Dirac operators*, *Nucl. Phys.* **B 605** (2001) 365 [hep-lat/0102012].

[33] D.H. Adams, *General bounds on the wilson-dirac operator*, *Phys. Rev.* **D 68** (2003) 065009 [hep-lat/9907005]; *Improving the locality of the overlap Dirac operator via approximate solutions of the Ginsparg-Wilson relation*, *Nucl. Phys.* **129** *(Proc. Suppl.)* (2004) 513 [hep-lat/0309148];
D.H. Adams and W. Bietenholz, *Axial anomaly and index of the overlap hypercube operator*, *Eur. Phys. J.* **C 34** (2004) 245 [hep-lat/0307022].

[34] T.-W. Chiu, *Locality of optimal lattice domain-wall fermions*, *Phys. Lett.* **B 552** (2003) 97 [hep-lat/0211032].

[35] L. Giusti, C. Hoelbling and C. Rebbi, *Considerations on Neuberger's operator*, *Nucl. Phys.* **83** *(Proc. Suppl.)* (2000) 896 [hep-lat/9906004].

[36] W. Kerler, *Dirac operator normality and chiral fermions*, *Chin. J. Phys.* **38** (2000) 623 [hep-lat/9912022];
I. Hip, T. Lippert, H. Neff, K. Schilling and W. Schroers,*Instanton dominance of topological charge fluctuations in QCD?*, *Phys. Rev.* **D 65** (2002) 014506 [hep-lat/0105001]; *The consequences of non-normality*, *Nucl. Phys.* **106** *(Proc. Suppl.)* (2002) 1004 [hep-lat/0110155].

[37] M. Lüscher, S. Sint, R. Sommer and P. Weisz, *Chiral symmetry and O(a) improvement in lattice QCD*, *Nucl. Phys.* **B 478** (1996) 365 [hep-lat/9605038].

[38] L. Giusti, C. Hoelbling and C. Rebbi, *Light quark masses with overlap fermions in quenched QCD*, *Phys. Rev.* **D 64** (2001) 114508 [hep-lat/0108007]; *Light quark masses with overlap fermions in quenched QCD*, *Phys. Rev.* **D 64** (2001) 114508 [hep-lat/0108007].

[39] F. Berruto et al., *Preliminary results from a simulation of quenched QCD with overlap fermions on a large lattice*, *Nucl. Phys.* **129** *(Proc. Suppl.)* (2004) 471 [hep-lat/0310006].

[40] Y. Kikukawa and A. Yamada, *A note on the exact lattice chiral symmetry in the overlap formalism*, hep-lat/9810024.

[41] C. Alexandrou, E. Follana, H. Panagopoulos and E. Vicari, *One-loop renormalization of fermionic currents with the overlap-Dirac operator*, *Nucl. Phys.* **B 580** (2000) 394 [hep-lat/0002010].

[42] C.W. Bernard and T. DeGrand, *Perturbation theory for fat-link fermion actions*, *Nucl. Phys.* **83** *(Proc. Suppl.)* (2000) 845 [hep-lat/9909083].

[43] T. DeGrand, *One loop matching coefficients for a variant overlap action and some of its simpler relatives*, *Phys. Rev.* **D 67** (2003) 014507 [hep-lat/0210028].

[44] P. Hasenfratz et al., *The construction of generalized dirac operators on the lattice*, *Int. J. Mod. Phys.* C12 (2001) 691–708 [hep-lat/0003013]; *Testing the fixed point QCD action and the construction of chiral currents*, *Nucl. Phys.* **B 643** (2002) 280 [hep-lat/0205010].

[45] W. Bietenholz, *Solutions of the Ginsparg-Wilson relation and improved domain wall fermions*, *Eur. Phys. J.* **C 6** (1999) 537 [hep-lat/9803023]; *Solutions of the Ginsparg-Wilson relation and improved domain wall fermions*, *Eur. Phys. J.* **C 6** (1999) 537 [hep-lat/9803023]; *Convergence rate and locality of improved overlap fermions*, *Nucl. Phys.* **B 644** (2002) 223 [hep-lat/0204016];
W. Bietenholz and I. Hip, *The scaling of exact and approximate Ginsparg-Wilson fermions,Nucl. Phys.* **B 570** (2000) 423 [hep-lat/9902019];
W. Bietenholz and S. Shcheredin, *Relating chiral perturbation theory and qcd simulations with overlap hypercube fermions* [hep-lat/0502010]. .

[46] C. Gattringer, *Phys. Rev.* **D 63** (2001) 114501 *A new approach to Ginsparg-Wilson fermions*, [hep-lat/0003005];
C. Gattringer and I. Hip, *New approximate solutions of the Ginsparg-Wilson equation: tests in* $2-D$, *Phys. Lett.* **B 480** (2000) 112 [hep-lat/0002002];
C. Gattringer, I. Hip and C.B. Lang, *Approximate Ginsparg-Wilson fermions: a first test*, *Nucl. Phys.* **B 597** (2001) 451 [hep-lat/0007042].

[47] MILC collaboration, T. DeGrand, *Simple observables from fat link fermion actions*, *Phys. Rev.* **D 60** (1999) 094501 [hep-lat/9903006].

[48] MILC collaboration, T. DeGrand, *A variant approach to the overlap action*, *Phys. Rev.* **D 63** (2001) 034503 [hep-lat/0007046];
T. DeGrand, A. Hasenfratz and T.G. Kovacs, *Improving the chiral properties of lattice fermions*, *Phys. Rev.* **D 67** (2003) 054501 [hep-lat/0211006].

[49] W. Kamleh, D.H. Adams, D.B. Leinweber and A.G. Williams, *Accelerated overlap fermions*, *Phys. Rev.* **D 66** (2002) 014501 [hep-lat/0112041].

[50] RBC collaboration, M. Lin, *Rough gauge fields, smearing and domain wall fermions*, hep-lat/0409115.

[51] UKQCD collaboration, K.C. Bowler, B. Joo, R.D. Kenway, C.M. Maynard and R.J. Tweedie, *Lattice QCD with mixed actions*, hep-lat/0411005.

[52] Lattice Hadron Physics collaboration, F.D.R. Bonnet, R.G. Edwards, G.T. Fleming, R. Lewis and D.G. Richards, *Lattice computations of the pion form factor*, hep-lat/0411028.

[53] NPLQCD collaboration, S.R. Beane, P.F. Bedaque, K. Orginos and M.J. Savage, $I = 2\pi\pi$ *scattering from fully-dynamical mixed-action lattice QCD*, hep-lat/0506013.

[54] CP-PACS collaboration, A. Ali Khan et al., *Eigenvalues of the hermitian Wilson-Dirac operator and chiral properties of the domain-wall fermion*, *Nucl. Phys.* **94** *(Proc. Suppl.)* (2001) 725 [hep-lat/0011032].

[55] Y. Aoki et al., *Domain wall fermions with improved gauge actions*, *Phys. Rev.* **D 69** (2004) 074504 [hep-lat/0211023].

[56] K. Jansen and K.-i. Nagai, *Reducing residual-mass effects for domain-wall fermions*, *JHEP* **12** (2003) 038 [hep-lat/0305009].

[57] C. Morningstar and M.J. Peardon, *Analytic smearing of* SU(3) *link variables in lattice QCD*, *Phys. Rev.* **D 69** (2004) 054501 [hep-lat/0311018].

[58] T. DeGrand, private communication.

[59] Z. Fodor, S.D. Katz and K.K. Szabo, *Dynamical overlap fermions, results with hybrid Monte-Carlo algorithm*, *JHEP* **08** (2004) 003 [hep-lat/0311010]; *Dynamical overlap fermions, results with hybrid Monte-Carlo algorithm*, *JHEP* **08** (2004) 003 [hep-lat/0311010];
T. DeGrand and S. Schaefer, *Physics issues in simulations with dynamical overlap fermions*, *Numerical methods for the QCD overlap operator IV: hybrid Monte Carlo*, *Phys. Rev.* **D 71** (2005) 034507 [hep-lat/0412005];
N. Cundy, S. Krieg, G. Arnold, A. Frommer, T. Lippert and K. Schilling, *Algorithms for Lattice QCD with Dynamical Fermions*, [hep-lat/0502007];
A.D. Kennedy, *Nucl. Phys.* **140** *(Proc. Suppl.)* (2005) 190 [hep-lat/0409167].

[60] S. Aoki, *New phase structure for lattice QCD with Wilson fermions*, *Phys. Rev.* **D 30** (1984) 2653.

[61] E.-M. Ilgenfritz, W. Kerler, M. Muller-Preussker, A. Sternbeck and H. Stuben, *A numerical reinvestigation of the Aoki phase with $N_f = 2$ Wilson fermions at zero temperature*, *Phys. Rev.* **D 69** (2004) 074511 [hep-lat/0309057].

[62] K.M. Bitar, U.M. Heller and R. Narayanan, *A new method to measure the chiral condensate in lattice QCD using wilson fermions*, *Phys. Lett.* **B 418** (1998) 167 [hep-th/9710052].

[63] S.R. Sharpe and J. Singleton, Robert L., *Spontaneous flavor and parity breaking with Wilson fermions*, *Phys. Rev.* **D 58** (1998) 074501 [hep-lat/9804028].

[64] M. Golterman and Y. Shamir, *Localization in lattice QCD*, *Phys. Rev.* **D 68** (2003) 074501 [hep-lat/0306002].

[65] M. Golterman, S.R. Sharpe and J. Singleton, Robert L., *Effective theory for quenched lattice QCD and the Aoki phase*, *Phys. Rev.* **D 71** (2005) 094503 [hep-lat/0501015].


JHEP09(2005)030