SSOR Preconditioning of Improved Actions

N. Eicker\textsuperscript{a}, W. Bietenholz\textsuperscript{a}, A. Frommer\textsuperscript{b}, H. Hoeber\textsuperscript{c}, Th. Lippert\textsuperscript{a}, K. Schilling\textsuperscript{a,c}

\textsuperscript{a}HLRZ, c/o Research Center Jülich, D-52425 Jülich, Germany
\textsuperscript{b}Department of Mathematics, University of Wuppertal, D-42097 Wuppertal, Germany
\textsuperscript{c}Department of Physics, University of Wuppertal, D-42097 Wuppertal, Germany

We generalize local lexicographic SSOR preconditioning for the Sheikholeslami-Wohlert improved Wilson fermion action and the truncated perfect free fermion action. In our test implementation we achieve performance gains as known from SSOR preconditioning of the standard Wilson fermion action.

1. INTRODUCTION

The standard Wilson fermion action of lattice QCD leads to discretization-errors of $O(a)$ in lattice spacing, requiring prohibitively fine lattice resolutions in the approach to the chiral and continuum limits. The present trend to tackle this problem goes in two directions: (i) One approach is based on Symanzik’s on-shell improvement program, where irrelevant $O(a)$ counter terms are added to both, lattice action (Sheikholeslami-Wohlert-Wilson action SWA) and composite operators. The hope is to reach the continuum limit for a specific observable $O(a) = O_{\text{cont.}} + c_2 a^2 + \ldots$ without $O(a)$ contamination. (ii) Another promising ansatz is based on so-called perfect actions that are located on renormalized trajectories intersecting the critical surface in a fixed point of a renormalization group transformation. Perfect actions are in principle free of cutoff effects. However, they can only be realized approximatively as truncated perfect actions (TPA).

Simulations of dynamical fermions within these schemes meet the problem of the compute intensive solutions of the fermionic linear system $Mx = \phi$, well known from traditional actions. In the last three years, a considerable acceleration of the inversion of the standard Wilson fermion matrix has been achieved by introduction of the BiCGStab algorithm and novel parallel local-lexicographic SSOR preconditioning techniques. Obviously, the efforts should be combined, i.e. II-SSOR generalized for SWAs and TPAs in order to gain their full pay-off.

In general, both SWA and TPA can be written in the form

$$M = A + B + C + \cdots,$$

where $A$ represents diagonal blocks (containing $12 \times 12$ sub-blocks), $B$ is a nearest-neighbor hopping term, $C$ contains next-to-nearest-neighbor couplings. Usually next-next-nearest-neighbor couplings are truncated.

Our key observation is that one can include into the II-SSOR process (i) the internal degrees of freedom of the block diagonal term $A$ as arising in SWA and (ii) next-to-nearest-neighbor terms $C$ as present in TPA.

2. PRECONDITIONING SWA

Preconditioning amounts to the replacement of $M$, $x$ and $\phi$ by preconditioned quantities $\tilde{M}$, $\tilde{x}$ and $\tilde{\phi}$. The aim is to transform the matrix such that the spectrum becomes narrower, increasing the efficiency of the inversion. The matrix-vector multiplication is replaced by

$$v_i = Mp_i \Rightarrow \begin{cases} \text{solve } Pz_i = p_i \quad \text{solve } v_i = Mz_i \end{cases}.$$
\( P \) represents the preconditioning matrix. It can be decomposed into a product of three regular matrices \( P = RST \). This allows to apply the 'Eisenstat-trick' \(^6\) using the identity \( M = R + T - K \), with a fourth regular matrix \( K \).

Let us recall the essentials of \( ll \)-SSOR \(^5\): The matrix \( M \) is decomposed into a (block-) diagonal part \( D \), and two strictly upper and lower triangular parts \( U \) and \( L \), \( M = D - L - U \). This can be achieved by local lexicographic ordering, described in \(^5\). The choices for \( R \), \( S \) and \( T \) are \( R = \frac{1}{\omega}D - L \), \( S = \left( \frac{2}{\omega}D \right)^{-1} \) and \( T = \frac{1}{\omega}D - U \). Here \( \omega \) is an over-relaxation parameter to be chosen appropriately.

These special choices of \( R \), \( S \) and \( T \) simplify the task of solving the linear equation in (3). They lead to the following replacement of the matrix-vector multiplication:

\[
v_i = Mp_i \Rightarrow \begin{cases} \text{multiply by } D \\ \text{solve backward} \\ \text{solve forward} \end{cases}
\]

For standard Wilson fermions, the block-diagonal term is given by \( A \propto 1 \), which implies the natural choice of \( D \propto 1 \) in the SSOR scheme. Therefore, in (3) the multiplication by \( D \) and the multiplication by \( D^{-1} \) in the forward-/backward-solve is readily carried out. In the multiplication with a diagonal block, the 12 color-spin elements in the vector \( x \) are decoupled and can be treated simultaneously.

The situation changes if \( A \) is not a strict diagonal. We have the freedom to choose the splitting of \( A \) into the diagonal term \( D \) and the upper and lower terms \( U \) and \( L \). As efficient implementations require to store \( D^{-1} \), we thus can control the memory overhead. However, depending on the choice of \( D \) the elements of \( x \) are interleaved in the multiplication with a diagonal block. Therefore they can only be treated simultaneously, if we choose the diagonal part as \( D = A \), the choice with the largest memory overhead. For any other choice, SSOR subprocesses on the diagonal blocks have to be introduced.

As a test we have implemented the \( ll \)-SSOR preconditioning scheme within BiCGStab for SWA. The diagonal part of the related quark matrix contains four complex \( 3 \times 3 \) matrices \( F_i \):

\[
\begin{pmatrix}
1 + F_1 & F_2 & F_3 & F_4 \\
F_2^T & 1 - F_1 & F_3^T & -F_3 \\
F_3 & F_4 & 1 + F_1 & F_2 \\
F_4^T & -F_3 & F_2^T & 1 - F_1
\end{pmatrix}.
\]

This structure reduces the storage requirements by a factor of 4 and is well suited to a QCD optimized machine.

We tested the inverter in a 16\(^4\) pure gauge background at \( \beta = 6.0 \) for two choices of \( D \): (i) the true diagonal (true) with twelve \( 1 \times 1 \) blocks and (ii) the \( 1 \pm F_i \) blocks (block) as shown in (4). The \( c_{SW} \)-parameter was chosen as 1.0 and 1.6. We tested the algorithm for different values of \( \kappa \) on 4 field-configurations. The tests were done on the 32-node APE100/Quadrics Q4 in Wuppertal. We compare to unpreconditioned BiCGStab rescaled by a factor 2 to mimic odd-even preconditioning as a reference.

![Figure 1. Number of iterations at \( c_{SW} = 1.0 \).](image)

Fig. 1 shows iteration numbers at \( c_{SW} = 1.0 \) for both implementations. BiCGStab/2 represents the estimate for the odd-even inverter, \( \omega = 1.0 \) stands for the case without over-relaxation, the optimal value is \( \omega = 1.5 \). The gain is about a factor of 2 against BiCGStab/2 at \( \omega = 1 \); over-relaxation yields another 10 to 20% gain. The difference between the two choices of \( D \) is not significant. The results shown in Fig. 2 for \( c_{SW} = 1.6 \) are qualitatively identical to the \( c_{SW} = 1.0 \) case.
3. PRECONDITIONING TPA

Next we consider a perfect free lattice fermion action for arbitrary mass \( m \). As the couplings decay exponentially, a practical truncation scheme confines the couplings to a unit hypercube \([9]\).

The matrix for this “hypercube fermion” (HF), albeit with \( A \propto 1 \), seems considerably more complicated than the Wilson fermion matrix, due to contributions of type \( C \) and beyond. But \( ll\)-SSOR preconditioning and the Eisenstat-trick remain applicable. We will present a detailed treatment elsewhere \([10]\).

At this stage we discuss the effect of preconditioning by recourse to the multi-color approach, the extension of the “red-black” scheme. This leads us to \( 2^d \) non-interacting sub-lattices. We obtain many off-diagonal blocks that are fortunately largely suppressed. Denoting the maximal magnitude of the elements in \( L \) and \( U \) as \( O(\varepsilon) \), we apply the analog to the odd-even transformation and get

\[
M' = 1 - \left( \sum_{i \geq 1} L_i \right) \left( \sum_{j \geq 1} U_j \right) = 1 - LU - O(\varepsilon^3) .
\]

We expect the spectrum to be much closer to 1 since the eigenvalues of \( M' \) are all \( 1 - O(\varepsilon^2) \) (for \( M \) they are \( 1 - O(\varepsilon) \)).

Fig. 2 shows that the parameter \( \varepsilon \) obtained is smaller for the HF than that for the Wilson fermion, since the lattice derivative is somehow “smeared over the hypercube”. Thus we expect multi-color (and also SSOR) preconditioning to work very well.

4. CONCLUSIONS

We demonstrated that the application of the \( ll\)-SSOR preconditioning scheme leads to the most efficient preconditioning known for improved actions. Our method saves a large factor in memory compared to odd-even preconditioning.

REFERENCES

1. G. P. Lepage, Nucl. Phys. B (Proc. Suppl.) 47 (1996) 3.
2. M. Lüscher et al., Nucl. Phys. B 478 (1996) 365.
3. B. Sheikholeslami and R. Wohlert, Nucl. Phys. B 259 (1985) 572.
4. P. Hasenfratz and F. Niedermayer, Nucl. Phys. B 414 (1994) 785.
5. A. Frommer et al., IJMP C5 (1994) 1073.
6. S. Fischer et al., Comput. Phys. Commun. 98 (1996) 20.
7. S. Eisenstat, SIAM J. Sc. Stat. Comp. 2 (1981) 1.
8. K. Jansen and C. Liu, Comput. Phys. Commun. 99 (1997) 221.
9. W. Bietenholz and U.-J. Wiese, Nucl. Phys. B464 (1996) 319.
10. N. Eicker et al., in preparation.