Deflated solver for the Dirac equation in openQCD

Martin Lüscher August 2013; revised February 2022

1. Introduction

The DFL_SAP_GCR solver for the Dirac equation included in the openQCD distribution combines a Krylov-space solver with the Schwarz alternating procedure and local deflation [1,2]. A technical modification suggested by Frommer et al. [3] was added later, which led to a simplification of the structure of the algorithm and a slightly better performance. The various parts of the solver are briefly described in this note.

2. Preconditioned Krylov-space solver

The DFL_SAP_GCR solver is essentially a preconditioned GCR algorithm that solves the Dirac equation

$$D\psi = \eta \tag{2.1}$$

for any given source field η, while the Schwarz alternating procedure and low-mode deflation are elements of the preconditioner M. A flexible form, FGCR, of the GCR algorithm is used, where M does not need to have any properties other than being an approximate inverse of the Dirac operator D. In particular, M may be a non-linear operator and may involve approximate iterative procedures.

The FGCR solver recursively builds up the Krylov space

$$\mathcal{K}_n = \text{span}\{\eta, DM\eta, (DM)^2\eta, \ldots, (DM)^n\eta\} \tag{2.2}$$

and finds the field ψ_n in $M\mathcal{K}_{n-1}$, which minimizes the square norm of the residue $\rho_n = \eta - D\psi_n$. The algorithm stops if

$$\|\eta - D\psi_n\| \leq \omega\|\eta\| \tag{2.3}$$

for some specified tolerance ω and norm $\|\cdot\|$ (the supported norms are the standard square norm $\|\cdot\|_2$ and the uniform norm $\|\cdot\|_\infty$). Since ρ_n is contained in $\mathcal{K}_n$, the next Krylov space can be obtained by adding $DM\rho_n$ to the current one. The next approximate solution is then a linear combination of the fields $M\rho_0,\ldots,M\rho_n$.

In practice the FGCR algorithm is restarted after generating some maximal number of Krylov vectors and if the computed approximate solution of the Dirac equation is not sufficiently accurate. Before a restart, the latter is saved and the new Krylov space is generated starting from its residue, i.e. the Dirac equation is now solved with the source set to the residue. The new solution is then added to the saved one, the algorithm is restarted again, and so on, until the norm of the residue reaches the desired level.

3. Schwarz alternating procedure (SAP)

The use of the SAP in lattice QCD was proposed long ago in ref. [4] and its implementation in `openQCD` still follows the lines of that paper.

3.1 Block grids

A block grid is a division of the global lattice in non-overlapping rectangular blocks of lattice points. All blocks must have identical shape and there must be an even number of blocks along all four coordinate axes. Moreover, the block sizes l_0,l_1,l_2,l_3 must be even and at least 4.

The point z in a given block with the smallest coordinates $(z_0,\ldots,z_3)$ is referred to as the base point of the block. Since the number of blocks is even in all directions, the block grid can be chessboard-coloured, i.e. the blocks can be classified as even or odd depending on the parity of the sum $\sum_{\mu=0}^{3} z_\mu/l_\mu$ of the block coordinates.

Blocks of lattice points and block grids are important objects in `openQCD`. There are, for example, functions mapping fields from the global lattice to fields attached to the blocks. For technical reasons, each block must currently be fully contained in the local lattice that contains its base point.

3.2 Multiplicative Schwarz preconditioner

Let Λ be a block in a given block grid,

$$P_\Lambda\psi(x)=\begin{cases}\psi(x) & \text{if } x\in\Lambda,\\ 0 & \text{otherwise,}\end{cases} \tag{3.1}$$

the operator that restricts any quark field $\psi(x)$ to the block points and $D_\Lambda = P_\Lambda D P_\Lambda$ the restriction of the Dirac operator to the block. D_Λ acts on block fields like the full Dirac operator with Dirichlet boundary conditions along the exterior boundary of the block.

An approximate solution of the Dirac equation (2.1) may now be obtained starting from $\psi = 0$ by alternately solving the equation on all even and all odd blocks, in each step keeping the current solution on the other blocks fixed. On a given block Λ, this amounts to solving the equation

$$D_\Lambda \chi = \rho - P_\Lambda D(1 - P_\Lambda)\psi \tag{3.2}$$

and replacing $\psi(x)$ by $\chi(x)$ at all $x \in \Lambda$, where ρ is the current residue $\eta - D\psi$. Note that the field on the blocks in each set (even or odd) of blocks can be updated independently from one another.

After $n_{\rm cy}$ such (SAP) update cycles, a field

$$\psi = M_{\rm sap}\eta \tag{3.3}$$

is obtained, which solves the Dirac equation to some precision. The operator $M_{\rm sap}$ implicitly defined in this way may thus serve as preconditioner for the Dirac operator. $M_{\rm sap}$ is a linear operator, if the block equations (3.2) are solved exactly, but as previously mentioned, preconditioners may be non-linear and approximate solutions are therefore acceptable. In `openQCD`, the block equations are solved by applying $n_{\rm mr}$ steps of the minimal residual (MR) algorithm, where the latter may be even-odd preconditioned. $M_{\rm sap}$ then has 2 parameters, $n_{\rm mr}$ and $n_{\rm cy}$, in addition to the block size and the choice of preconditioned or ordinary block MR algorithm.

4. Local deflation

The SAP is a smoothing operation that damps the high-mode components of the residue of the calculated approximate solution of the Dirac equation, while local deflation is designed to reduce its low-mode components and thus complements the SAP.

4.1 Deflation subspace

The term "low mode" is used here for any quark field ψ satisfying

$$\|D\psi\|_2 \ll \|\psi\|_2. \tag{4.1}$$

Deflation subspaces can be constructed from sets $\{\psi_1, \ldots, \psi_{N_s}\}$ of N_s linearly independent low modes by projecting them to the blocks of a block grid. On each block Λ, the projected fields $P_\Lambda\psi_1, \ldots, P_\Lambda\psi_{N_s}$ span a linear space of dimension N_s. The dimension of these deflation subspaces is thus equal to N_s times the number n_b of blocks.

If the projected fields are orthonormalized, block by block, any quark field can be easily projected to the subspace by computing its scalar products with these basis fields. The subspace may thus be identified with the linear space of complex vectors of length $N_s n_b$.

4.2 Preconditioning

Let P_0 be the orthogonal projector to a deflation subspace, however constructed, and $A = P_0 D P_0$ the restriction of the Dirac operator to the subspace. The quark fields ψ may then be split into two components according to

$$\psi = \{\psi - D\zeta\} + D\zeta, \qquad \zeta = A^{-1}P_0\psi, \tag{4.2}$$

where it is taken for granted that the "little Dirac operator" A is a non-singular operator in the deflation subspace. Since

$$P_0\{\psi - D\zeta\} = 0, \qquad P_0 D\zeta = P_0\psi, \tag{4.3}$$

and if all low modes of the Dirac operator have large components along the deflation subspace, the splitting approximately separates the high- from the low-mode components of ψ.

As a consequence, the deflated preconditioner

$$M\psi = M_{\mathrm{sap}}\{\psi - D\zeta\} + \zeta \tag{4.4}$$

tends to be much more effective than M_{sap} alone. The little Dirac equation

$$A\zeta = P_0\psi, \qquad P_0\zeta = \zeta, \tag{4.5}$$

must be solved each time the preconditioner is applied, which is potentially costly in terms of computer time. An exact solution of the equation is however not required and approximate solutions up to relative tolerances as large as 10^{-2} are usually perfectly sufficient.

4.3 Low mode generation

The construction of the deflation subspace along the lines of subsect. 4.1 starts from a set $\{\psi_1, \ldots, \psi_{N_s}\}$ of low modes of the Dirac operator. In `openQCD` these fields are generated by initializing them to random values and applying $M_{\rm sap}$ and later the deflated preconditioner M a number of times, the deflation subspace required by the latter being constructed from the current set of fields.

The process (a form of approximate inverse iteration) systematically lowers the Rayleigh quotients $\|D\psi_k\|_2/\|\psi_k\|_2$, but tends to produce fields that are approximately linearly dependent, especially in its later phase, where the Rayleigh quotients are already small. This behaviour can be easily counteracted by orthonormalizing the fields before the deflated preconditioner is applied.

5. Solver for the little Dirac equation

Each application of the preconditioner (4.4) requires the little Dirac equation

$$A\zeta = \lambda \tag{5.1}$$

to be solved for a given source λ. As already mentioned, an accurate solution is not required, but since the little Dirac operator tends to have a large condition number, a careful choice of the solver algorithm is important.

5.1 Even-odd preconditioning

The deflation block grid is of the same kind as the one used in the case of the SAP (see subsect. 3.1). In particular, the block grid can be chessboard-coloured and the blocks accordingly divide into even and odd ones.

If the blocks are ordered such that the even ones come first, the little Dirac operator assumes the block-diagonal form

$$A = \begin{pmatrix} A_{\rm ee} & A_{\rm eo} \\ A_{\rm oe} & A_{\rm oo} \end{pmatrix}. \tag{5.2}$$

The little Dirac equation (5.1) may then be reduced to its even-odd preconditioned form,

$$\hat{A}\zeta_{\rm e} = A_{\rm ee}^{-1}\{\lambda_{\rm e} - A_{\rm eo}A_{\rm oo}^{-1}\lambda_{\rm o}\}, \tag{5.3}$$

$$\hat{A} = 1_{\rm e} - A_{\rm ee}^{-1}A_{\rm eo}A_{\rm oo}^{-1}A_{\rm oe}, \tag{5.4}$$

where $\lambda_{\rm e}$ and $\lambda_{\rm o}$ denote the components of the source λ on the even and odd blocks. Once eq. (5.3) is solved, the odd components of the solution are given by

$$\zeta_{\rm o} = A_{\rm oo}^{-1}\{\lambda_{\rm o} - A_{\rm oe}\zeta_{\rm e}\}. \tag{5.5}$$

Clearly, these equations assume that the diagonal parts $A_{\rm ee}$ and $A_{\rm oo}$ of the little Dirac operator are safely invertible (an error condition is raised if this is not so).

5.2 Exact global-mode deflation

The global modes $\psi_1,\ldots,\psi_{N_{\rm s}}$ that generated the deflation subspace through projection to the deflation block grid are, by construction, contained in the deflation subspace. They are low modes of the Dirac operator and thus of the little Dirac operator too.

The removal of these modes from the even-odd preconditioned little Dirac equation (5.3) through an oblique projection is, in general, profitable. If $v_1,\ldots,v_{N_{\rm s}}$ denote the orthonormalized even components of the global modes, the "little little Dirac operator"

$$B_{kl} = (v_k, \hat{A}v_l), \quad k,l = 1,\ldots,N_{\rm s}, \tag{5.6}$$

may be defined and the projection alluded to above is achieved using the projectors

$$P_L\chi_{\rm e} = \chi_{\rm e} - \sum_{k,l}\hat{A}v_k(B^{-1})_{kl}(v_l,\chi_{\rm e}), \tag{5.7}$$

$$P_R\chi_{\rm e} = \chi_{\rm e} - \sum_{k,l}v_k(B^{-1})_{kl}(v_l,\hat{A}\chi_{\rm e}), \tag{5.8}$$

where χ stands for any element of the deflation subspace.

The projected even-odd preconditioned little Dirac equation

$$P_L\hat{A}\bar{\zeta}_{\rm e} = P_L\hat{\lambda}_{\rm e}, \qquad \bar{\zeta}_{\rm e} = P_R\zeta_{\rm e}, \qquad \hat{\lambda}_{\rm e} = A_{\rm ee}^{-1}\{\lambda_{\rm e} - A_{\rm eo}A_{\rm oo}^{-1}\lambda_{\rm o}\}, \tag{5.9}$$

derives from eq. (5.3) by left-multiplication with the projector P_L and using the identity $P_L\hat{A} = \hat{A}P_R$. Now if $\bar{\zeta}_{\rm e}$ solves the projected equation (5.9) and satisfies the (consistent) constraint

$$P_R\bar{\zeta}_{\rm e} = \bar{\zeta}_{\rm e}, \tag{5.10}$$

the solution of eq. (5.3),

$$\zeta_{\rm e} = \bar{\zeta}_{\rm e} + \sum_{k,l} v_k(B^{-1})_{kl}(v_l, \hat{\lambda}_{\rm e}), \tag{5.11}$$

is obtained by adding a global-mode correction. An inexact solution of the projected equation moreover yields an approximate solution of eq. (5.3) with residue

$$\hat{\lambda}_{\rm e} - \hat{A}\zeta_{\rm e} = P_L\hat{\lambda}_{\rm e} - P_L\hat{A}\bar{\zeta}_{\rm e} \tag{5.12}$$

exactly given by the one of $\bar{\zeta}_{\rm e}$.

5.3 Solver algorithm

The projected even-odd preconditioned equation (5.9) is solved using the FGCR algorithm (the one briefly described in sect. 2) with a non-restarted GCR algorithm as preconditioner. Non-restarted means that the algorithm stops if the desired level of accuracy is reached or if the specified maximal number of Krylov vectors has been generated.

In order to accelerate the computations, the preconditioner is implemented using single-precision data and arithmetic, but the solution of eq. (5.9) is obtained with double precision (even if this may often be unnecessary).

References

[1] M. Lüscher, *Local coherence and deflation of the low quark modes in lattice QCD*, J. High Energy Phys. 0707 (2007) 081

[2] M. Lüscher, *Deflation acceleration of lattice QCD simulations*, J. High Energy Phys. 0712 (2007) 011

[3] A. Frommer, K. Kahl, S. Krieg, B. Leder, M. Rottmann, *Adaptive aggregation based domain decomposition multigrid for the lattice Wilson Dirac operator*, SIAM J. Sci. Comput. 36 (2014) A1581

[4] M. Lüscher, *Solution of the Dirac equation in lattice QCD using a domain decomposition method*, Comput. Phys. Commun. 156 (2004) 209