GMRES acceleration of restricted Schwarz iterations

Pacull[1] and Aubert[1]

We present here an analysis of the Richardson iterations preconditioned by either the restricted additive [2] or multiplicative Schwarz [6] operators, and the associated GMRES Krylov sub-space acceleration. The framework of study is purely algebraic and general sparse unsymmetrical and indefinite matrices are considered. This paper can be seen as an extension of [1, 10], in which a block preconditioned system is downsized to an interface system. The following study is circumscribed to restricted Schwarz preconditioners.

At first, the equivalence between the primary and interface iterations is described. Then, the interface system operator is depicted as a Schur complement of the permuted preconditioned global matrix. Finally, the benefit of the Krylov sub-space acceleration of the interface iterations, over the primary ones, is exhibited. Note that exact solves of the sub-domain problems is assumed throughout.

The linear system to solve is :

$$A\,u = f \tag{1}$$

with $A \in \mathbb{R}^{n\times n}$, $u \in \mathbb{R}^n$ and $f \in \mathbb{R}^n$. We assume that A is close to structurally symmetric, which is a common property of matrices originating from PDE problems.

As a preparatory step, we start by introducing the vertex-based partitioning process and the notations used hereafter.

1 Introduction

1.1 Graph partitioning and overlap

We denote $\mathscr{G}$ the adjacency graph of matrix A, $\mathscr{V} = \{1,2,...,n\}$ the nodes of $\mathscr{G}$, and $\mathscr{E}$ the edges, which correspond to the non-zero off-diagonal elements of A. The graph $\mathscr{G}$ is considered to be undirected: given an unordered pair of distinct nodes $(v_1,v_2) \in \mathscr{V}^2$, we have $(v_1,v_2) \in \mathscr{E}$ if and only if $A(v_1,v_2) \neq 0$ or $A(v_2,v_1) \neq 0$.

A non-overlapping partition of $\mathscr{V}$ with p sub-domains corresponds to p non-empty sub-sets, $\{\mathscr{V}_i\}_{1\leq i\leq p}$, such that $\mathscr{V} = \cup_{i=1}^{p} \mathscr{V}_i$ and $\mathscr{V}_j \cap \mathscr{V}_k = \varnothing$ for $1 \leq j < k \leq p$. The usual goal when performing this graph-partitioning task is to minimize the overall edge cut, which is the total number of edges $(v_i,v_j) \in \mathscr{E}$ with v_i and v_j belonging to distinct sub-domains, while equilibrating the number of nodes per sub-domain to approximatively n/p. Dealing with p equal sub-sets aims at balancing the dis-

[1] Fluorem, 64 Chemin des Mouilles 69130 Ecully, France e-mail: {fpacull}{saubert}@fluorem.com

tributed computational and memory load per processor. Minimizing the number of edges crossing the partition boundaries results in a reduced communication volume between processors.

Increasing the δ-overlap is beneficial regarding the convergence rate of Schwarz methods (see [6] for example): starting from $\mathscr{V}_{i,0} \equiv \mathscr{V}_i$, this consists in growing recursively each sub-set $\mathscr{V}_{i,\delta}$ by adding some of the adjacent nodes, in order to form a larger set $\mathscr{V}_{i,\delta+1}$.

For each sub-domain and for each δ level, $n_{i,\delta} \equiv |\mathscr{V}_{i,\delta}|$ refers to the cardinality of the node sub-set.

1.2 Notations regarding restrictions operators

Similarly to what is done in [9], three different sub-sets of nodes are defined in association with a given sub-domain $\mathscr{V}_{i,\delta}$: $\mathscr{V}_{i,\delta}^{int}$, $\mathscr{V}_{i,\delta}^{loc}$ and $\mathscr{V}_{i,\delta}^{ext}$. The internal nodes $\mathscr{V}_{i,\delta}^{int}$ are the nodes of $\mathscr{V}_{i,\delta}$ that have their graph neighborhood fully included in $\mathscr{V}_{i,\delta}$. The local interface nodes $\mathscr{V}_{i,\delta}^{loc}$ are the nodes of $\mathscr{V}_{i,\delta}$ that have a least one of their neighbors outside of $\mathscr{V}_{i,\delta}$. Finally, the external interface nodes $\mathscr{V}_{i,\delta}^{ext}$ are the nodes that do not belong to $\mathscr{V}_{i,\delta}$, but which have at least one of their neighbors within $\mathscr{V}_{i,\delta}$.

Note that $\mathscr{V}_{i,\delta}^{ext}$ is the set of candidate nodes for growing the sub-set $\mathscr{V}_{i,\delta}$: $\mathscr{V}_{i,\delta+1} \subseteq \mathscr{V}_{i,\delta} \cup \mathscr{V}_{i,\delta}^{ext}$.

An important sub-set of nodes for our study is the global set of external interface node, simply called the *interface nodes* hereafter: $\mathscr{V}_{\delta}^{ext} \equiv \cup_{i=1}^{p} \mathscr{V}_{i,\delta}^{ext}$, with cardinality $n_{\delta}^{ext} \equiv |\mathscr{V}_{\delta}^{ext}|$. The complementary sub-set of $\mathscr{V}_{\delta}^{ext}$ is denoted by $\bar{\mathscr{V}}_{\delta}^{ext} \equiv \mathscr{V} \backslash \mathscr{V}_{\delta}^{ext}$.

In the following, notations from [7] are used to describe the different operators associated with the algebraic Schwarz preconditioners. For the i-th sub-domain, we denote $R_{i,\delta} \in \mathbb{R}^{n_{i,\delta} \times n}$ the restriction operator associated with $\mathscr{V}_{i,\delta}$. $R_{i,\delta}^{ext}$ is the restriction operator associated with $\mathscr{V}_{i,\delta}^{ext}$. The special restriction operator used in the restricted Schwarz iterations, is defined as follows: $\tilde{R}_{i,\delta} \equiv R_{i,\delta} R_{i,0}^{T} R_{i,0} \in \mathbb{R}^{n_{i,\delta} \times n}$.

The node sub-set $\bar{\mathscr{V}}_{i,\delta}$ refers to the following set difference: $\bar{\mathscr{V}}_{i,\delta} \equiv \mathscr{V} \setminus \mathscr{V}_{i,\delta}$, and $\bar{R}_{i,\delta}$ to the restriction operator associated with $\bar{\mathscr{V}}_{i,\delta}$. R_{δ}^{ext} and $\bar{R}_{\delta}^{ext}$ are the restriction operators associated with $\mathscr{V}_{\delta}^{ext}$ and $\bar{\mathscr{V}}_{\delta}^{ext}$ respectively.

The local parts of the operator A are the following ones: $A_{i,\delta} \equiv R_{i,\delta} A R_{i,\delta}^{T}$ for the inner coupling, and $A_{i,\delta}^{ext} \equiv R_{i,\delta} A R_{i,\delta}^{ext\,T}$ for the outer coupling.

Finally, the vector y stands for the vector of interface node unknowns

$$y = R_{\delta}^{ext} u \in \mathbb{R}^{n_{\delta}^{ext}} \tag{2}$$

while $x = \bar{R}_{\delta}^{ext} u \in \mathbb{R}^{n - n_{\delta}^{ext}}$ stands for the complementary unknowns, located at the non-interface nodes $\bar{\mathscr{V}}_{\delta}^{ext}$.

2 Richardson iterations with a restricted Schwarz preconditioner

The preconditioned Richardson iteration $u^{(k+1)} = u^{(k)} + M^{-1}(f - Au^{(k)})$, is expressed as the stationary iteration

$$u^{(k+1)} = F\,u^{(k)} + g \tag{3}$$

where $F = I - M^{-1}A$ and $g = M^{-1}f$ are the iteration matrix and vector. We only consider here the restricted additive (RAS) and multiplicative (RMS) Schwarz preconditioners, as defined for example in [6]:

$$F_{RAS,\delta} = I - \sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} A_{i,\delta}^{-1} R_{i,\delta} A \tag{4}$$

$$F_{RMS,\delta} = \prod_{i=p}^{1} \left(I - \tilde{R}_{i,\delta}^{T} A_{i,\delta}^{-1} R_{i,\delta} A \right) \tag{5}$$

As pointed out in [1, 10], under some specific conditions, the primary iteration (3) can be reduced to an equivalent interface iteration, in terms of the unknown y defined in (2):

$$y^{(k+1)} = Gy^{(k)} + h \tag{6}$$

In order to gain more insight into this interface system, let us derive the iteration (6) starting from (3). If the restriction R_{δ}^{ext} is applied to (3), we get the following iteration: $y^{(k+1)} = R_{\delta}^{ext} F x^{(k)} + h$, with $h \equiv R_{\delta}^{ext} g$. We now make use of the following relation:

$$\begin{aligned} R_{i,\delta} A &= R_{i,\delta} A (R_{i,\delta}^{T} R_{i,\delta} + \bar{R}_{i,\delta}^{T} \bar{R}_{i,\delta}) \\ &= A_{i,\delta} R_{i,\delta} + R_{i,\delta} A \bar{R}_{i,\delta}^{T} \bar{R}_{i,\delta} \\ &= A_{i,\delta} R_{i,\delta} + A_{i,\delta}^{ext} R_{i,\delta}^{ext} \end{aligned} \tag{7}$$

Thus, in the restricted additive Schwarz case, we have:

$$\begin{aligned} F_{RAS,\delta} &= I - \sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} A_{i,\delta}^{-1} (A_{i,\delta} R_{i,\delta} + A_{i,\delta}^{ext} R_{i,\delta}^{ext}) \\ &= I - \sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} R_{i,\delta} - \sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} A_{i,\delta}^{-1} A_{i,\delta}^{ext} R_{i,\delta}^{ext} \end{aligned} \tag{8}$$

Using the following equality, $\sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} R_{i,\delta} = \sum_{i=1}^{p} R_{i,0}^{T} R_{i,0} R_{i,\delta}^{T} R_{i,\delta} = \sum_{i=1}^{p} R_{i,0}^{T} R_{i,0} = I$, we get:

$$F_{RAS,\delta} = -\sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} A_{i,\delta}^{-1} A_{i,\delta}^{ext} R_{i,\delta}^{ext} \tag{9}$$

This shows that the iteration matrix $F_{RAS,\delta}$ only depends on the interface nodes.

For the multiplicative case, by using (7), we get:

$$F_{RMS,\delta} = \prod_{i=p}^{1} \left(I - \tilde{R}_{i,\delta}^T A_{i,\delta}^{-1} R_{i,\delta} A \right)$$
$$= \prod_{i=p}^{1} \left(I - R_{i,0}^T R_{i,0} - \tilde{R}_{i,\delta}^T A_{i,\delta}^{-1} A_{i,\delta}^{ext} R_{i,\delta}^{ext} \right) \quad (10)$$

For simplicity reasons, we call a_i the left term in the parentheses and b_i the right term: $a_i \equiv I - R_{i,0}^T R_{i,0}$, $b_i \equiv \tilde{R}_{i,\delta}^T A_{i,\delta}^{-1} A_{i,\delta}^{ext} R_{i,\delta}^{ext}$. By noticing that $\prod_{i=p}^{1} a_i = 0$ and that $a_i b_j = b_j$ if $i \neq j$, we get:

$$F_{RMS,\delta} = \sum_{k=1}^{p} \sum_{p \geq i_1 > \ldots > i_k \geq 1} (-1)^k b_{i_1} \ldots b_{i_k} \quad (11)$$

The important thing is that the b_i terms only depend on the interface nodes, and so does $F_{RMS,\delta}$ consequently.

Hence we have, in both restricted Schwarz cases:

$$F = F R_\delta^{ext\,T} R_\delta^{ext} \text{ and } F \bar{R}_\delta^{ext\,T} \bar{R}_\delta^{ext} = 0 \quad (12)$$

Indeed we know from [10] that k belongs to $\bar{\mathcal{V}}_\delta^{ext}$ (that is, k is not an interface node) if and only if the k-th column of F is null, and if and only if the k-th column of M is equal to the k-th column of A.

We can now state that with the coherent initial interface conditions $y^{(0)} = R_\delta^{ext} u^{(0)}$, the following relation between $u^{(k)}$ and $y^{(k)}$ holds:

$$y^{(k+1)} = R_\delta^{ext} u^{(k+1)} = R_\delta^{ext} \left[F u^{(k)} + g \right] = G y^{(k)} + h \text{ for } k \geq 1 \quad (13)$$

The iteration matrix G can be expressed as follows: $G = R_\delta^{ext} F R_\delta^{ext\,T}$. Note that this relation holds whatever the initial condition $x^{(0)} = \bar{R}_\delta^{ext} u^{(0)}$ is.

We now focus on the interface system: $(I - G) y^{(\infty)} = h$.

3 Restricted Schwarz and Schur

In [3, 8], it is shown that a multiplicative Schwarz iterate is identical to a block Gauss-Seidel sweep applied to the Schur complement system on the interface unknowns, provided that coherent initial conditions are used. Similar results also holds between the additive Schwarz iterate and a block Jacobi sweep of the Schur complement system. The considered Schur complement S is related to the interface nodes of the non-overlapping partition. In the overlapping case, it is possible to decompose the sub-domains into smaller disjoint parts and express the global matrix as a preconditioned version of S, thanks to block Gaussian elimination. As stated in [4]:

> the overlapping method is equivalent to a non-overlapping method with a specific interface preconditioner. One can think of the overlapping method implicitly computing the effect of this preconditioner by the extra operations performed on the overlapping region.

We observe that in our case, the interface unknowns may not correspond to the interface nodes of the non-overlapping partition. If we consider the permuted matrix $P_\delta M^{-1} A P_\delta^T$, with M being a restricted Schwarz preconditioner, and $P_\delta^T = \left[\bar{R}_\delta^{ext\,T} R_\delta^{ext\,T}\right]$, we get the following linear system:

$$P_\delta M^{-1} A P_\delta^T \begin{Bmatrix} x \\ y \end{Bmatrix} = \begin{Bmatrix} \bar{R}_\delta^{ext} g \\ h \end{Bmatrix} \tag{14}$$

We note that the matrix $P_\delta M^{-1} A P_\delta^T$ is a 2×2 block matrix:

$$P_\delta M^{-1} A P_\delta^T = \begin{bmatrix} \bar{R}_\delta^{ext} M^{-1} A \bar{R}_\delta^{ext\,T} & \bar{R}_\delta^{ext} M^{-1} A R_\delta^{ext\,T} \\ R_\delta^{ext} M^{-1} A \bar{R}_\delta^{ext\,T} & R_\delta^{ext} M^{-1} A R_\delta^{ext\,T} \end{bmatrix} \tag{15}$$

In the previous section, we saw that $F \bar{R}_\delta^{ext\,T} = 0$, which implies that

$$\bar{R}_\delta^{ext} M^{-1} A \bar{R}_\delta^{ext\,T} = I \tag{16}$$

$$R_\delta^{ext} M^{-1} A \bar{R}_\delta^{ext\,T} = 0 \tag{17}$$

We also have the following equalities:

$$\bar{R}_\delta^{ext} M^{-1} A R_\delta^{ext\,T} = -\bar{R}_\delta^{ext} F R_\delta^{ext\,T} \tag{18}$$

$$R_\delta^{ext} M^{-1} A R_\delta^{ext\,T} = I - R_\delta^{ext} F R_\delta^{ext\,T} = I - G \tag{19}$$

Plugging these equalities into (14), we get:

$$P_\delta M^{-1} A P_\delta^T \begin{Bmatrix} x \\ y \end{Bmatrix} = \begin{bmatrix} I & -\bar{R}_\delta^{ext} F R_\delta^{ext\,T} \\ 0 & I - G \end{bmatrix} \begin{Bmatrix} x \\ y \end{Bmatrix} = \begin{Bmatrix} \bar{R}_\delta^{ext} g \\ h \end{Bmatrix} \tag{20}$$

The matrix $I - G$ can be seen as a Schur complement of $P_\delta M^{-1} A P_\delta^T$ with respect to the identity operator applied to the non-interface nodes. The inverse of $P_\delta M^{-1} A P_\delta^T$ can be expressed in this way:

$$\left(P_\delta M^{-1} A P_\delta^T\right)^{-1} = \begin{bmatrix} I & \bar{R}_\delta^{ext} F R_\delta^{ext\,T} (I - G)^{-1} \\ 0 & (I - G)^{-1} \end{bmatrix} \tag{21}$$

Also, equation (20) gives us some information about the spectrum of $I - G$:

$$\sigma\left(M^{-1} A\right) = \sigma\left(P_\delta M^{-1} A P_\delta^T\right) = \sigma(I) \cup \sigma(I - G) \tag{22}$$

The spectrum of $I-G$ is the spectrum of $M^{-1}A$ augmented with the eigenvalue 1, which has a multiplicity of $n-n_\delta^{ext}$.

We remark that the cost of explicitly building the $I-G$ matrix is prohibitive, regarding the significant resources required. In the RAS case, the matrix G writes:

$$G = -R_\delta^{ext} \sum_{i=1}^{p} \tilde{R}_{i,\delta}^T A_{i,\delta}^{-1} A_{i,\delta}^{ext} R_{i,\delta}^{ext} R_\delta^{ext\,T} \tag{23}$$

This represents $|\mathscr{V}_{i,\delta}^{ext}|$ local systems to solve for each sub-domain, which solution is dense. This is why iterative methods are preferred.

4 Krylov acceleration

Since matrix A is assumed to be unsymmetrical and indefinite, the GMRES Krylov sub-space method [8] is used to accelerate the iteration (6), as proposed in [1]. The GMRES method is chosen over some other Krylov techniques for its monotonous convergence property. The algorithm used to solve the interface system is presented next, in a left-preconditioned version.

Algorithm 1 GMRES resolution of $(I-G)y = h$

$r_0 = R_\delta^{ext} M^{-1}(b - Ax_0)$, $\beta = \|r_0\|$, and $v_1 = r_0/\beta$
for $j = 1,\ldots,m$ **do**
  $w \leftarrow R_\delta^{ext} M^{-1} A R_\delta^{ext\,T} v_j$
  for $i = 1,\ldots,j$ **do**
    $h_{i,j} \leftarrow (w, v_i)$
    $w \leftarrow w - h_{i,j} v_i$
  end for
  ...
end for
...
Compute $z_m = \text{argmin}_z \|\beta e_1 - \bar{H}_m z\|$ and $y_m = R_\delta^{ext} x_0 + V_m z_m$
If satisfied $y^{(\infty)} \leftarrow y_m$ else restart with $x_0 = R_\delta^{ext\,T} y_m$

An important point is that **Algorithm 1** only differs from the usual one by the use of the restriction and prolongation operators R_δ^{ext} and $R_\delta^{ext\,T}$. Also, one extra step is required to solve the global solution from the interface solution $y^{(\infty)}$:

$$u^{(\infty)} = (I - M^{-1}A) R_\delta^{ext\,T} y^{(\infty)} + g \tag{24}$$

In this last step, the preconditioner M^{-1} can differ from the one used in the GMRES algorithm. For example, if $M_{RAS,\delta}^{-1}$ is chosen, we get:

$$u^{(\infty)} = \sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} A_{i,\delta}^{-1} R_{i,\delta} \left(b - A\bar{R}_{i,\delta}^{T}\bar{R}_{i,\delta} R_{\delta}^{ext\,T} y^{(\infty)} \right)$$
$$= \sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} A_{i,\delta}^{-1} \left(R_{i,\delta} b - A_{i,\delta}^{ext} R_{i,\delta}^{ext} R_{\delta}^{ext\,T} y^{(\infty)} \right) \tag{25}$$

Algorithm 1 represents less floating point operations and also requires less memory to store the Arnoldi vectors than when GMRES is applied to the primary unknowns, with almost no extra work regarding the implementation.

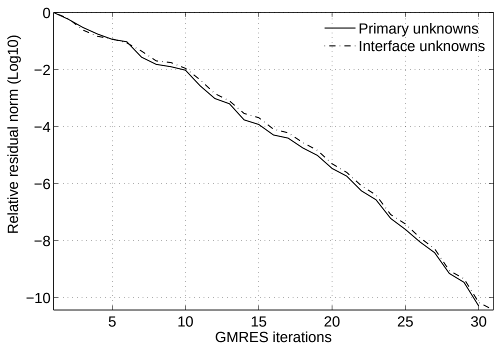

Fig. 1 Full GMRES convergence of the global and interface systems. GT01R matrix from the UF sparse matrix collection is used. Initial condition is $x^{(0)} = \{1, ..., 1\}^T$. The domain is divided into 2 parts ($p = 2$) with an overlap of $\delta = 1$ (all the adjacent nodes are included). The number of primary and interface unknowns is 7980 and 420 respectively.

Fig. 1 presents the GMRES convergence of both primary and interface systems. Matrix GT01R from the UF sparse matrix collection [5] is used. We observe that the convergence behaviors are similar, but slightly differ because of the non-interface nodes. The size of the global system is 7980, while it is 420 for the interface system.

Also, the new vector $w \leftarrow R_{\delta}^{ext} M^{-1} A R_{\delta}^{ext\,T} v_j$ in the outer loop of **Algorithm 1** is equivalent to this one: $w \leftarrow (I - R_{\delta}^{ext} F R_{\delta}^{ext\,T}) v_j$, in which only local "homogeneous" problems are solved. For example in the RAS case, we have:

$$w \leftarrow \left(I + R_{\delta}^{ext} \sum_{i=1}^{p} \tilde{R}_{i,\delta}^{T} A_{i,\delta}^{-1} A_{i,\delta}^{ext} R_{i,\delta}^{ext} R_{\delta}^{ext\,T} \right) v_j \tag{26}$$

The local operator $A_{i,\delta}^{-1}$ is applied to $A_{i,\delta}^{ext} R_{i,\delta}^{ext} R_{\delta}^{ext\,T} v_j$, which only concerns the local interface nodes of the sub-domain, $\mathscr{V}_{i,\delta}^{loc}$. This means that for the local problem in (26), the right-hand side is null for the internal nodes $\mathscr{V}_{i,\delta}^{int}$. Thus, a local Schur complement approach may be used to deal with each local problem, associated to an iterative local solver and the LU factorization of the two diagonal blocks of $A_{i,\delta}$ corresponding to the internal and the local interface nodes.

5 Conclusion

The restricted Schwarz iterations have been described in details. It appears that the restricted Schwarz operators benefit from the indirect preconditioning effect of the overlap, but also from the non-overlapping property of the restricted local operator images. We have seen that solving the interface system instead of the primary one, is advantageous regarding memory usage and floating point operation count. This represents only a slight modification of the global algorithm, but requires exact local solves. Another advantage is that the local problems can be treated as homogeneous problems.

References

1. Brakkee, E., Wilders, P.: A domain decomposition method for the advection-diffusion equation. Tech. rep., Faculty of Technical Mathematics and Informatics, Delft University of Technology, Delft (1994)
2. Cai, X.C., Sarkis, M.: A restricted additive Schwarz preconditioner for general sparse linear systems. SIAM J. Sci. Comput. **21**(2), 792–797 (electronic) (1999)
3. Chan, T.F., Goovaerts, D.: On the relationship between overlapping and nonoverlapping domain decomposition methods. SIAM J. Matrix Anal. Appl. **13**(2), 663–670 (1992)
4. Chan, T.F., Mathew, T.P.: Domain decomposition algorithms. In: Acta numerica, 1994, Acta Numer., pp. 61–143. Cambridge Univ. Press, Cambridge (1994)
5. Davis, T.A., Hu, Y.: The University of Florida sparse matrix collection. ACM Trans. Math. Software **38**(1), Art. 1, 25 (2011)
6. Frommer, A., Nabben, R., Szyld, D.B.: An algebraic convergence theory for restricted additive and multiplicative Schwarz methods. In: N. Debit, M. Garbey, R. Hoppe, D. Keyes, Y. Kuznetzov, J. Périaux (eds.) Domain Decomposition Methods in Science and Engineering, Thirteenth International Conference on Domain Decomposition Methods, Lyon, France, pp. 371–377. CIMNE, UPC, Barcelona (2002)
7. Frommer, A., Szyld, D.B.: An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms. SIAM J. Numer. Anal. **39**(2), 463–479 (electronic) (2001)
8. Saad, Y.: Iterative methods for sparse linear systems, second edn. Society for Industrial and Applied Mathematics, Philadelphia, PA (2003)
9. Saad, Y., Sosonkina, M.: Distributed Schur complement techniques for general sparse linear systems. SIAM J. Sci. Comput. **21**(4), 1337–1356 (electronic) (1999/00)
10. Wilders, P., Brakkee, E.: Schwarz and Schur: an algebraical note on equivalence properties. SIAM J. Sci. Comput. **20**(6), 2297–2303 (electronic) (1999)