An explicit formulation of the multiplicative Schwarz preconditionner

Guy Atenekeng Kahou, Emmanuel Kamgnia, Bernard Philippe

INSTITUT DE RECHERCHE EN INFORMATIQUE ET SYSTÈMES ALÉATOIRES

PUBLICATION INTERNE

N° 1738

AN EXPLICIT FORMULATION OF THE MULTIPLICATIVE SCHWARZ PRECONDITIONNER

GUY ATENEKENG KAHOU , EMMANUEL KAMGNIA , BERNARD PHILIPPE

ISSN 1166-8687

IRISA
CAMPUS UNIVERSITAIRE DE BEAULIEU - 35042 RENNES CEDEX - FRANCE

Institut de Recherche en Informatique et Systèmes Aléatoires
Campus de Beaulieu – 35042 Rennes Cedex – France
Tél. : (33) 02 99 84 71 00 – Fax : (33) 02 99 84 71 71
http://www.irisa.fr

An explicit formulation of the multiplicative Schwarz preconditionner

Guy Atenekeng Kahou * , Emmanuel Kamgnia ** , Bernard Philippe ***

Thème —
Projet SAGE

Publication interne n ° 1738 — Juillet 2005 — 25 pages

Abstract: We provide an explicit formulation of the splitting associated with the Multiplicative Schwarz iteration. We show the advantage of considering the explicit formulation, when the iteration is used as a preconditioner of a Krylov method.

Key-words: Domain decomposition, Multiplicative Schwarz, preconditionner, Krylov methods, Red-Black coloring, iterative methods.

(Résumé : tsvp)

This work was supported by AUF (Agence Universitaire de la Francophonie) for visiting LAMSIN/ENIT at Tunis and by the French ministry of Foreign Affairs under SARIMA program for visiting INRIA/IRISA at Rennes.

* gaatenek@irisa.fr
** kamgnia@uycdc.uninet.cm
*** philippe@irisa.fr

Centre National de la Recherche Scientifique
(UPRESSA 6074) Université de Rennes 1 – Insa de Rennes

Institut National de Recherche en Informatique
et en Automatique – unité de recherche de Rennes

Formulation explicite du préconditionnement Schwarz multiplicatif

Résumé : A partir d'une expression explicite du splitting défini par l'itération multiplicative de Schwarz, nous étudions son utilisation comme précondtionnement d'une méthode de Krylov.

Mots clés : Décomposition de domaine, Schwarz multiplicatif, préconditionnement, méthodes de Krylov, coloriage Rouge-Noir, méthodes itératives

@float

1 Introduction

Domain decomposition provides a class of divide-and-conquer methods suitable for the solution of linear or nonlinear systems of equations arising from the discretization of partial differential equations. For linear systems, domain decomposition methods can be viewed as preconditioners for Krylov subspace techniques.

As mentioned in [8], the term domain decomposition has slightly different meanings to specialists depending on their discipline : in parallel computing, it often means the process of distributing data from a computational model among the processors in a distributed memory computer. In numerical analysis, it means the separation of the physical domain into regions that can be modeled with different equations, with interfaces between the domains handled by various conditions. In preconditioning methods, which is our interest in this article, domain decomposition refers to the process of subdividing the solution of a large linear system into smaller problems whose solutions can be used to produce a preconditioner (or solver) for the system of equations that results from discretizing the PDE on the entire domain or more generally from any sparse matrix. In our work, we consider the latter and we suppose that the domain decomposition is with overlapping.

Traditionally, there are three classes of iterative methods which derive from domain decomposition : Additive and Multiplicative Schwarz for overlappings subdomains and Schur complement methods for non-overlapping subdomains. When using the Schwarz methods as solvers, the convergence rates are very slow and the convergence is mainly guarantied for symmetric positive definite matrices and M-matrices [3]. For that reason, the particular interest of Schwarz methods is as preconditioner of Krylov subspace methods since they can be efficient even when they would not converge as a full method.

When used as preconditioners, one is interested in deriving an explicit and useful expression of the preconditioner. For the Additive Schwarz method such an expression exists. To our knowledge, no explicit expression of the preconditionner is known for the Multiplicative Schwarz method. In this paper we derive such an expression.

In section 2 we suppose that one graph partitioner is applied resulting in subdomains with overlaps. If the domain decomposition were without overlap, it is known [4] that the multiplicative Schwarz algorithm would be equivalent to a Block Gauss Seidel iteration. In section 3, we derive an explicit formulation of the Multiplicative Schwarz preconditioner. In section 4, we discuss the use of such a preconditionner with a Krylov method and in section 5, we illustrate the behaviour of the preconditionner on some numerical tests.

2 Domain decomposition of a sparse matrix and notations

Let us consider a sparse matrix $A \in \mathbb{R}^{n\times n}$. The pattern of A is the set $\mathbb{P} = \{(k,l) | a_{k,l} \neq 0\}$ which is the set of the edges of the graph $G = (W, \mathbb{P})$ where $W = \{1, ..., n\} = [1 : n]$ is the set of vertices.

Definition 2.1 *A domain decomposition of matrix A into p subdomains is defined by a collection of sets of integers $W_i \subset W = [1 : n]$, $i = 1, ..., p$ such that :*

$$\left\{ \begin{array}{l} |i-j| > 1 \Longrightarrow W_i \cap W_j = \emptyset, \\ \mathbb{P} \subset \bigcup_{i=1}^{p} (W_i \times W_i). \end{array} \right.$$

Following this definition, a domain decomposition can be considered as resulting from a graph partitioner but with potential overlap between domains. It can be noticed that such a decomposition does not necessarily exist (e.g when A is a dense matrix in which case there is only one subdomain). For the rest of our discussion, we shall suppose that a graph partitioner has been applied and has resulted in p sets W_i whose union is W, $W = [1 : n]$. The submatrix of A corresponding to $W_i \times W_i$ is denoted by A_i
We shall denote by $L_i = \oplus_{j \in W_i}(e_j)$ the vector space of $\mathbb{R}^n$ of all the vectors with zero components for every index $j \notin W_i$. Let m_i be the dimension of L_i. The orthogonal projector onto L_i is defined by the sub-identity matrix I_i of order $n \times n$ whose diagonal elements are set to one if the corresponding node belongs to W_i and to zero otherwise. We also denote by A_i the extension of block A_i to the whole space, therefore:

$$A_i \quad = \quad I_i A I_i. \tag{1}$$

Finally, we define the complement sub-identity matrix $\bar{I}_i = I - I_i$ and the matrix,

$$\bar{A}_i \quad = \quad A_i + \bar{I}_i. \tag{2}$$

We assume thereafter that all the matrices $\bar{A}_i$, for $i = 1, \cdots, p$ are non singular. The generalized inverse A_i^+ of A_i satisfies $A_i^+ = I_i \bar{A}_i^{-1} = \bar{A}_i^{-1} I_i$.

Proposition 2.1 *For any domain decomposition as defined in Definition 2.1 the following property is true.*
$\forall\ i, j \in \{1, ..., p\}$,

$$|i-j| > 2 \Rightarrow I_i A I_j = 0.$$

Proof. Let $(k,l) \in W_i \times W_j$ such that $a_{k,l} \neq 0$. Since $(k,l) \in \mathbb{P}$, there exists $m \in \{1...n\}$ such that $k \in W_m$ and $l \in W_m$; therefore $W_i \cap W_m \neq \emptyset$ and $W_j \cap W_m \neq \emptyset$. Consequently, from Definition 2.1, $|i-m| \leq 1$ and $|j-m| \leq 1$, which implies $|i-j| \leq 2$. $\diamond$

Let us introduce a special situation which is often satisfied and which brings some simplification in the sequel.

Definition 2.2 *The domain decomposition is with* weak overlap *if and only if, for any* $i,\ j \in \{1, ..., p\}$ *the following is true*

$$|i - j| > 1 \Rightarrow I_i A I_j = 0.$$

The set of unknowns which represents the overlap is defined by the set of integers $J_i = W_i \cap W_{i+1}, i = 1, ..., p-1$, and let s_i be the dimension of the overlap. Similarly to (1) and (2), we define

$$C_i = O_i A O_i, \tag{3}$$

and

$$\bar{C}_i = C_i + \bar{O}_i, \tag{4}$$

where $O_i \in \mathbb{R}^{n \times n}$ is sub-identity matrix whose diagonal elements are set to one if the corresponding node belongs to J_i and to zero otherwise, and $\bar{O}_i = I - O_i$.

Example 2.1 *Figure 1 displays an example of a domain decomposition for a matrix in the case where all* W_i *are intervals of integers (*A_{ij} *denotes a block).*

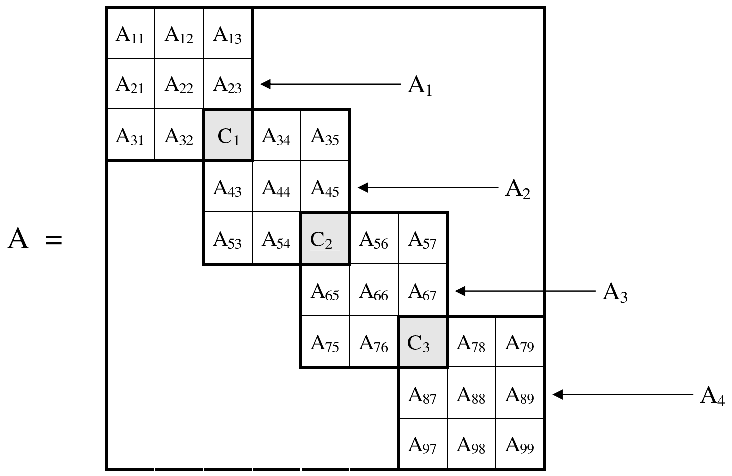

Figure 1: A matrix domain decomposition with block overlaps. $W_1 = w_1 \cup w_2 \cup w_3$, $W_2 = w_3 \cup w_4 \cup w_5$, $W_3 = w_5 \cup w_6 \cup w_7$, $W_4 = w_7 \cup w_8 \cup w_9$, where, for $i = 1, \cdots, 9$, w_i is the set of the row indices of A_{ii} with $C_1 = A_{33}$, $C_2 = A_{55}$ and $C_3 = A_{77}$.

There is a close connection between a block tridiagonal structure and a domain decomposition. For instance, in the previous example, if, in order to transform A into a block tridiagonal matrix, we assume that all the blocks $A_{i,i+2}$ and $A_{i+2,i}$ are zeros, the domain decomposition is obtained by defining the domains with three consecutive blocks and the overlaps correspond to only one block. One can easily verify that such an overlap is a weak overlap. If the domains were defined with only two blocks and the domain overlaps on one block, the overlap would not be weak.

In the definitions, the set of integers defining the subdomains are not necessarily intervals although this is often the situation as in the example. However, when considering other situations like a red-black block ordering, it is important to include the general case. Nevertheless, it is always possible to recover the special case by renumbering the unknowns.

3 Multiplicative Schwarz

The goal of Multiplicative Schwarz methods is to iteratively solve a linear system

$$Ax = b \tag{5}$$

where matrix A is decomposed into overlapping subdomains as described in the previous section. The iteration consists of solving the original equation in sequence on each subdomain. This is a well-known method ; for more details, see for instance [3, 4, 6, 7, 8, 10]). In this section, we present the main properties of the iteration and derive an explicit formulation of the corresponding matrix splitting.

3.1 Classical formulation

Let x^k be the current iterate and $r^k = b - Ax^k$ the corresponding residual. The classical formulation of multiplicative Schwarz proceeds as follow.

Algorithm 1 *: One iteration of the Multiplicative Schwarz Preconditioner builds p sub-iterates and their corresponding residuals by the following recursion :*

$$\begin{array}{ll} \textit{input} \; : & x := x^k \; ; \; r := r^k \; ; \\ \textit{for} \; i = 1 : p & \\ & x := x + A_i^+ r \; ; \\ & r := r - AA_i^+ r \; ; \\ \textit{end} & \\ \textit{output} \; : & x^{k+1} := x \; ; \; r^{k+1} := r \; ; \end{array}$$

It follows that :

$$r^{k+1} = (I - AA_p^+)...(I - AA_1^+)r^k. \tag{6}$$

This method corresponds to a relaxation iteration defined by some splitting $A = M - N$ such that the iteration matrices for the residual and the error are respectively

$$NM^{-1} = (I - AA_p^+) \cdots (I - AA_1^+) \text{ and} \tag{7}$$

$$M^{-1}N = A^{-1}NM^{-1}A = (I - A_p^+ A) \cdots (I - A_1^+ A). \tag{8}$$

The convergence of this iteration is proven for M-matrices and s.p.d. matrices (eg. see [3]).

Now, let us suppose that the goal is to consider another iterative method but preconditioned by one step of the Multiplicative Schwarz method. For that purpose, it is necessary to define

$$y = M^{-1}Ax \text{ or } y = AM^{-1}x$$

depending on the side of the preconditionning, for any vector x and where M is the matrix characterized by the previous splitting. From the expression of $M^{-1}N$ and NM^{-1} we can derive an expression of $M^{-1}A$ or AM^{-1} as follows :

$$M^{-1}A = I - (I - A_p^+ A) \cdots (I - A_1^+ A), \tag{9}$$

or

$$AM^{-1} = I - (I - AA_p^+) \cdots (I - AA_1^+). \tag{10}$$

3.2 Embedding in a system of larger dimension

If the subdomains do not overlap, it can be shown [4] that the Multiplicative Schwarz is equivalent to a Block Gauss-Seidel method applied on an extended system. In this section, following [9], we present an extended system which embeds the original system (5) into a larger one with no overlapping between subdomains.

For that purpose, we define the prolongation mapping and the restriction mapping. We assume for the whole section that the set of indices defining the domains are intervals. As mentioned before, this does not limit the scope of the study since a preliminary symmetric permutation of the matrix, corresponding to the same renumbering of the unknowns and the equations, can always end up with such a system.

Definition 3.1 *The prolongation mapping which injects $\mathbb{R}^n$ into a space $\mathbb{R}^m$ where $m = \sum_{i=1}^{p} m_i = n + \sum_{i=1}^{p-1} s_i$ is defined as follows :*

$$\begin{array}{rccc} \mathcal{D}: & \mathbb{R}^n & \rightarrow & \mathbb{R}^m \\ & x & \mapsto & \widetilde{x}, \end{array}$$

where $\widetilde{x}$ is obtained from vector x by duplicating all the blocks of entries corresponding to overlapping blocks.

The restriction mapping consists of projecting a vector $\widetilde{x} \in \mathbb{R}^m$ onto $\mathbb{R}^n$, which consists of deleting the subvectors corresponding to the first appearence of each overlapping blocks

$$\begin{array}{rccc} \mathcal{P}: & \mathbb{R}^m & \rightarrow & \mathbb{R}^n \\ & \widetilde{x} & \mapsto & x. \end{array}$$

Embedding the original system in a larger one is done for instance in [4, 9]. We present here a special case. In order to avoid a tedious formal presentation of the augmented system, we present its construction on an example which is generic enough to understand the definition of $\widetilde{A}$. In Figure 2, is displayed an example with four domains. Mapping $\mathcal{D}$ builds $\tilde{x}$ by duplicating some entries in vector x : mapping $x \rightarrow \widetilde{x} = \mathcal{D}x$ expands vector x to include subvectors y_3, y_5 and y_7. The equalities $y_3 = x_3$, $y_5 = x_5$ and $y_7 = x_7$ define a subspace $\mathcal{J}$ of $\mathbb{R}^m$. This subspace is the range of mapping $\mathcal{D}$. These equalities combined with the definition of matrix $\widetilde{A}$ show that $\mathcal{J}$ is an invariant subspace of $\widetilde{A}$: $\widetilde{A}\mathcal{J} \subset \mathcal{J}$. Therefore solving system $Ax = b$ is equivalent to solving system $\widetilde{A}\widetilde{x} = \widetilde{b}$ where $\widetilde{b} = \mathcal{D}b$. Operator $\mathcal{P}$ deletes entries x_3, x_5 and x_7 from vector $\tilde{x}$.

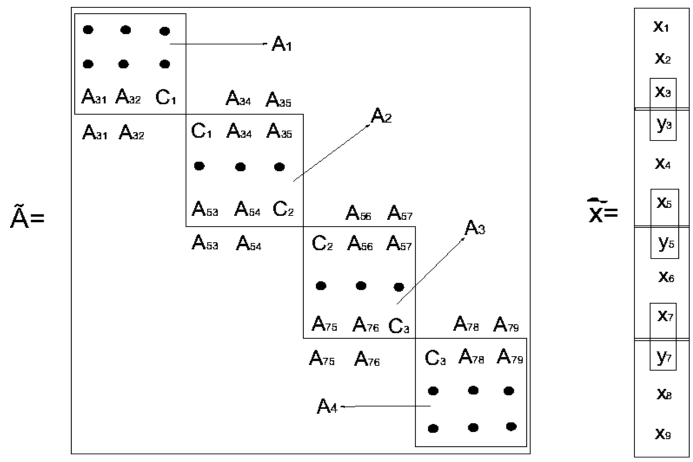

Figure 2: Definition of $\widetilde{A}$ (Extension of Matrix displayed in Fig 1).

Remark 3.1 *The following properties are straightforward consequences of the previous definitions :*

1. $Ax = \mathcal{P}\widetilde{A}\mathcal{D}x$,
2. $\mathcal{J} = \mathcal{R}(\mathcal{D}) \subset \mathbb{R}^m$ *is an invariant subspace of* $\widetilde{A}$,
3. $\mathcal{P}\mathcal{D} = I_n$ *and* $\mathcal{D}\mathcal{P}$ *is a projection onto* $\mathcal{J}$,
4. $\forall x,\ y \in \mathbb{R}^n$, $(y = Ax \Leftrightarrow \mathcal{D}y = \widetilde{A}\mathcal{D}x)$.

This can be illustrated by diagram (11) :

$$\begin{array}{ccccc} & \mathbb{R}^n & \stackrel{A}{\rightarrow} & \mathbb{R}^n & \\ \mathcal{D} & \downarrow & & \uparrow & \mathcal{P} \\ & \mathbb{R}^m & \stackrel{\widetilde{A}}{\rightarrow} & \mathbb{R}^m & . \end{array} \tag{11}$$

One iteration of the Multiplicative Schwarz method on the original system (5) corresponds to one Block-Seidel iteration on the enhanced system

$$\widetilde{A}\widetilde{x} \quad = \quad \mathcal{D}b, \tag{12}$$

where the diagonal blocks are the blocks defined by the p subdomains. More precisely, denoting by $\widetilde{M}$ the block lower triangular part of $\widetilde{A}$, the iteration defined in Algorithm 1 can be expressed as follows :

$$\begin{cases} \widetilde{x}^k = \mathcal{D}x^k, \\ \widetilde{r}^k = \mathcal{D}r^k, \\ \widetilde{x}^{k+1} = \widetilde{x}^k + \widetilde{M}^{-1}\widetilde{r}^k, \\ x^{k+1} = \mathcal{P}\widetilde{x}^{k+1}. \end{cases} \tag{13}$$

To prove it, let us partition $\widetilde{A} = \widetilde{M} - \widetilde{N}$, where $\widetilde{N}$ is the strictly upper block triangular part of $(-\widetilde{A})$. Matrices $\widetilde{M}$ and $\widetilde{N}$ are partitioned by blocks accordingly to the domain definition. One iteration of the Block Gauss-Seidel method can then be expressed by

$$\widetilde{x}^{k+1} = \widetilde{x}^k + \widetilde{M}^{-1}\widetilde{r}^k.$$

The resulting block triangular system is solved successively for each diagonal block. To derive the iteration, we partition $\widetilde{x}^k$ and $\widetilde{x}^{k+1}$ accordingly. At the first step, we obtain

$$\widetilde{x}_1^{k+1} = \widetilde{x}_1^k + A_1^{-1}\widetilde{r}_1^k, \tag{14}$$

which is identical to the first step of the Multiplicative Schwarz $x^{k+1/p} = x^k + A_1^+ r^k$. The i-th step $(i = 2, \cdots, p)$

$$\widetilde{x}_i^{k+1} = \widetilde{x}_i^k + A_i^{-1}\left(\widetilde{b}_i - \widetilde{M}_{i,1:i-1}\ \widetilde{x}_{1:i-1}^{k+1} - A_i\widetilde{x}_i^k + \widetilde{N}_{i,i+1:p}\ \widetilde{x}_{i+1:p}^k\right), \tag{15}$$

is equivalent to its counterpart $x^{k+(i+1)/p} = x^{k+i/p} + A_i^+ r^{k+i/p}$ in the Multiplicative Schwarz algorithm.

Therefore, we have the following diagram

$$\begin{array}{ccccc} & \mathbb{R}^n & \stackrel{M^{-1}}{\rightarrow} & \mathbb{R}^n & \\ \mathcal{D} & \downarrow & & \uparrow & \mathcal{P} \\ & \mathbb{R}^m & \stackrel{\widetilde{M}^{-1}}{\rightarrow} & \mathbb{R}^m & \end{array}$$

and we conclude that $M^{-1} = \mathcal{P}\widetilde{M}^{-1}\mathcal{D}$.

We must remark that there is an abuse in the notation, in the sense that the matrix denoted by M^{-1} can be singular even when $\widetilde{M}^{-1}$ is non singular.

We shall prove in the following theorem that this happens when one overlapping block is singular. Nevertheless, we keep the notation for its meaning in the general case.

3.3 Explicit formulation of the Multiplicative Schwarz preconditionner

Theorem 3.1 *Let $A \in \mathbb{R}^{n\times n}$ be decomposed into p subdomains as described in section 2 such that all the matrices $\bar{A}_i$, for $i = 1, \cdots, p$, and all the matrices C_i, for $i = 1, \cdots, p-1$,* *are non singular. The Multiplicative Schwarz preconditionner matrix M^{-1} can be explicitly expressed by :*

$$M^{-1} = \bar{A_p}^{-1}\bar{C}_{p-1}\bar{A}_{p-1}^{-1}\bar{C}_{p-2}\cdots\bar{A_2}^{-1}\bar{C}_1\bar{A_1}^{-1} \tag{16}$$

where matrices $\bar{A}_i$ and $\bar{C}_i$ are defined in section 2.

Proof. The Richardson iteration corresponding to the Multiplicative Schwarz preconditionner is expressed by the relation $x_{k+1} = x_k + M^{-1}r_k$. When we inject it in the augmented dimension we have the following iteration:
$\widetilde{x}_{k+1} = \widetilde{x}_k + \widetilde{M}^{-1}\widetilde{r}_k$, where $\widetilde{x}_k = \mathcal{D}x_k$, $\widetilde{r}_k = \mathcal{D}r_k$ and $\widetilde{M}^{-1}$ represents the Block Gauss-Seidel preconditionner built from $\widetilde{A}$. Let us set $\widetilde{M}^{-1}\widetilde{r}_k = \widetilde{t}_k$; then $\widetilde{x}_{k+1} = \widetilde{x}_k + \widetilde{t}_k$ and therefore, $x_{k+1} = \mathcal{P}\widetilde{x}_{k+1} = x_k + \mathcal{P}\widetilde{t}_k$.

In order to compute $\mathcal{P}\widetilde{t}_k$, we eliminate, within the system $\widetilde{r}_k = \widetilde{M}\widetilde{t}_k$, the unknowns which must be discarded by the projection $\mathcal{P}$. For that purpose, we partition the diagonal block (A_i) of $\tilde{A}$ as follows:
$A_i = \begin{pmatrix} B_i & F_i \\ E_i & C_i \end{pmatrix}$ and accordingly the vectors $\widetilde{t}^i_k = \begin{pmatrix} f^i_k \\ g^i_k \end{pmatrix}$ and $\widetilde{r}^i_k = \begin{pmatrix} u^i_k \\ v^i_k \end{pmatrix}$ for $i = 1, ..., p-1$ and $\tilde{t_k}^p = f^p_k$, $\tilde{r_k}^p = u^p_k$. We can observe from this representation that

$$r_k = \begin{bmatrix} u^1_k \\ \vdots \\ u^{p-1}_k \\ u^p_k \end{bmatrix} \quad \text{and} \quad t_k = \begin{bmatrix} f^1_k \\ \vdots \\ f^{p-1}_k \\ f^p_k \end{bmatrix}.$$

We now form the reduced system by only keeping the components of t_k:

$$(B_1 - F_1C_1^{-1}E_1)f^1_k = u^1_k - F_1C_1^{-1}v^1_k, \tag{17}$$

$$\begin{pmatrix} E_{i-1}f^{i-1}_k \\ 0 \end{pmatrix} + (B_i - F_iC_i^{-1}E_i)f^i_k = u^i_k - F_iC_i^{-1}v^i_k,\ i = 2, ..., p-1 \tag{18}$$

$$A_p\widetilde{t}^p_k + \begin{pmatrix} E_{p-1}f^{p-1}_k \\ 0 \end{pmatrix} = \widetilde{r}^p_k. \tag{19}$$

Note that, since $\widetilde{r}_k = \mathcal{D}r_k$, we have $v^i_k = R_{i+1}u^{i+1}_k$ where $R_{i+1}u^{i+1}_k$ consists of selecting the first s_i components of block vector u^{i+1}. Remember that s_i is the dimension of the overlap C_i.

Let $S_i = (B_i - F_iC_i^{-1}E_i)$ for $i = 1, ..., p-1$ be the local Schur complements associated to variable f^i_k and $S_p = A_p$. The reduced system becomes $Bt_k = z_k$ where the structure of B is defined in Figure 3.

The right hand side of the reduced system z_k can be written as $z_k = Tr_k$ where :

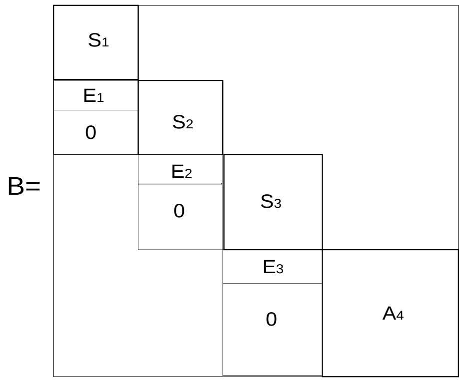

Figure 3: Matrix of reduced system (for p=4)

$$z_k = \begin{pmatrix} u_k^1 - F_1 C_1^{-1} R_2 u_k^2 \\ u_k^2 - F_2 C_2^{-1} R_3 u_k^3 \\ \vdots \\ u_k^{p-1} - F_{p-1} C_{p-1}^{-1} R_p \tilde{r}_k^p \\ \tilde{r}_k^p \end{pmatrix}$$

and therefore $z_k = Tr_k$ where

$$Tr_k = \begin{pmatrix} I_1 & \begin{pmatrix} -F_1 C_1^{-1} & 0 \end{pmatrix} & 0 & 0 \\ 0 & I_2 & \begin{pmatrix} -F_2 C_2^{-1} & 0 \end{pmatrix} & 0 \\ 0 & 0 & I_3 & \begin{pmatrix} -F_3 C_3^{-1} & 0 \end{pmatrix} \\ 0 & 0 & 0 & I_4 \end{pmatrix} \begin{pmatrix} u_k^1 \\ u_k^2 \\ u_k^3 \\ \tilde{r}_k^4 \end{pmatrix} \quad (20)$$

(displayed for $p = 4$). Matrix M, defined by the Multiplicative Schwarz splitting, is therefore characterized by the relation

$$TM = B.$$

We now prove that $M = \bar{A}_1\bar{C}_1^{-1}...\bar{C}_{p-1}^{-1}\bar{A}_p$ satisfies that relation. We first express the structure of the following matrix with the block structure defined by B and T:

$$\bar{A}_i\bar{C}_i^{-1} = \left(\begin{array}{c|c|c} \begin{pmatrix} I & & \\ & \ddots & \\ & & I \end{pmatrix} & & \\ \hline & \begin{pmatrix} B_i & G_i \\ \begin{bmatrix} E_i \\ 0 \end{bmatrix} & I \end{pmatrix} & \\ \hline & & \begin{pmatrix} I & & \\ & \ddots & \\ & & I \end{pmatrix} \end{array}\right) \begin{array}{l} \\ i^{th} \text{ row block} \\ (i+1)^{th} \text{ row block} \\ \\ \end{array}$$

for $i = 1, \cdots, p-1$. It is easy to prove by induction that for $i = 1, ..., p-2$

$$\begin{aligned} \Xi_i &= T(\bar{A}_1\bar{C}_1^{-1})(\bar{A}_2\bar{C}_2^{-1})...(\bar{A}_1\bar{C}_1^{-1}) \\ &= \begin{pmatrix} S_1 & & & & & & \\ \begin{bmatrix} E_1 \\ 0 \end{bmatrix} & S_2 & & & & & \\ & \ddots & \ddots & \ddots & & & \\ & & & S_i & & & \\ & & & \begin{bmatrix} E_i \\ 0 \end{bmatrix} & I & -G_{i+1} & \\ & & & & 0 & & \\ & & & & \ddots & \ddots & \ddots \\ & & & & & & -G_{p-1} \\ & & & & & 0 & I \end{pmatrix} \end{aligned}$$

and therefore

$$\begin{aligned} \Xi_{p-1} &= T(\bar{A}_1\bar{C}_1^{-1})(\bar{A}_2\bar{C}_2^{-1})...(\bar{A}_{p-1}\bar{C}_{p-1}^{-1}), \\ &= \begin{pmatrix} S_1 & & & \\ \begin{bmatrix} E_1 \\ 0 \end{bmatrix} & S_2 & & \\ & \ddots & \ddots & \\ & & S_{p-1} & \\ & & \begin{bmatrix} E_{p-1} \\ 0 \end{bmatrix} & \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix} \end{pmatrix}. \end{aligned}$$

Clearly we have

$$\Xi_{p-1}\bar{A}_p = B$$

which ends the proof. ⋄

Remark 3.2 *Let us describe more precisely the situation when one of the overlapping blocks C_i is singular. With this assumption, matrix* $S = \bar{A_p}^{-1}\bar{C}_{p-1}\bar{A}_{p-1}^{-1}\bar{C}_{p-2}\cdots\bar{A_2}^{-1}\bar{C}_1\bar{A_1}^{-1}$ *is singular and therefore it cannot be considered as being the inverse of a matrix M. In that situation, S cannot be used as preconditionner to solve system $Ax = b$: a left application of the preconditionner would lead to solve $SAx = SB$ which does not have a unique solution and a right application is impossible. It can be noticed that the singularity of one of the blocks C_i implies the singularity of the mapping $\widetilde{A}$ since the spectrum of $\widetilde{A}$ is equal to the union of the spectrum of A and the spectra of all the blocks C_i $(i = 1, \cdots, p-1)$ [9].*

Proposition 3.1 *The matrix N defined by the multiplicative Schwarz splitting $A = M - N$ can be expressed as follows:*

$$\begin{cases} N_{ij} & = G_i \cdots G_{j-1}B_j, \ when\ j > i+1 \\ N_{ii+1} & = G_iB_{i+1} - [F_i\ 0], \\ N_{ij} & = 0\ otherwise, \end{cases} \tag{21}$$

where $G_i = (F_iC_i^{-1}\ 0)$ for $i\ ,\ j = 1, ..., p-1$.

When the domain decomposition is with weak overlap, expression (21) becomes:

$$\begin{cases} N_{ii+1} & = G_iB_{i+1} - [F_i\ 0],\ for\ i = 1, ..., p-1, \\ N_{ij} & = 0\ otherwise. \end{cases} \tag{22}$$

Proof. It follows from equation 20 that the inverse of matrix $T = I - U$ is:

$$T^{-1} = \sum_{i=0}^{p-1} U^i$$

where :

$$U = \begin{pmatrix} 0 & G_1 & & 0 \\ & \ddots & \ddots & \\ & & & G_{p-1} \\ & & & 0 \end{pmatrix}.$$

Therefore, the matrix M can be expressed by:

$$\begin{cases} M_{ii} & = B_i, & \\ M_{ii+1} & = G_iB_{i+1}, & \\ M_{i+1i} & = \begin{bmatrix} E_i \\ 0 \end{bmatrix}, & for\ i\ ,\ j\ =\ 1\ \cdots\ p-1 \\ M_{ij} & = G_i \cdots G_{j-1}B_j, & if\ j > i+1 \\ M_{ij} & = 0, & otherwise. \end{cases}$$

By expressing matrix A in the same structure as the structure of M, we deduce that, $N = M - A$ satisfies relation (21). When the decomposition is of weak overlap, relation (22) follows from the fact that $G_i G_{i+1} = 0$. ⋄

Corollary 3.1 *In the splitting $A = M - N$ associated with the Multiplicative Schwarz method, matrix N is of rank $r \leq \sum_{i=1}^{p-1} s_i$.*

Proof. The proof is obvious when the decomposition is with weak overlap. For the general case, we have to prove that the rank of row block i of matrix N is less than s_i. The structure of row block N_i, for $i = 1, ..., p-1$, of matrix N is:

$$N_i = \left[\begin{array}{cccccccc} 0 & , & \cdots & 0 & , & [F_i \ 0] \begin{bmatrix} 0 & C_i^{-1} B_{i+1}(1,2) \\ 0 & -I \end{bmatrix}, & G_i G_{i+1} B_{i+2} & , \ \cdots \ , & G_i \cdots G_{p-1} B_p \end{array} \right].$$

Therefore, the rank of row block of N_i is limited by the rank of factor $[F_i \ 0]$ which cannot exceed s_i. This implies $r \leq \sum_{i=1}^{p-1} s_i$. ⋄

3.4 Symmetrization of M

Even when matrix A is symmetric, the preconditioned conjugate gradient method cannot be used directly since the Multiplicative Schwarz preconditioner is not symmetric. However, it can easily be symmetrized by including a second sweep corresponding to apply M^{-T} to current residual r^{k+1}. It can be written as follows :

$$\left\{ \begin{array}{l} x^{k+1} = x^k + M^{-1} r^k, \\ x^{k+2} = x^{k+1} + M^{-T} r^{k+1}. \end{array} \right.$$

Using that $r^{k+1} = N M^{-1} r^k$, we have

$$\begin{aligned} x^{k+2} &= x^k + M^{-1} r^k + M^{-T} N M^{-1} r^k, \\ &= x^k + M^{-T}(M^T + N) M^{-1} r^k, \\ &= x^k + M^{-T}(M^T + M - A) M^{-1} r^k. \end{aligned}$$

We deduce that the new preconditioning matrix (the one corresponding to two sweeps up and down) is such that:

$$\mathcal{M}^{-1} = M^{-T}(M^T + N) M^{-1} = M^{-T}(M^T + M - A) M^{-1}.$$

It can be shown [3] that when A is s.p.d., the preconditionner $\mathcal{M}$ is s.p.d. as well.

3.5 Red-Black ordering (M and N)

The Multiplicative Schwarz method, as described in Algorithm 1, clearly is not suitable for parallelism since the recursion between blocks prevents independent calculations. The classical way to overpass the drawback is to relax part of the recursion by a Red-Black coloring.

Recalling our assumption that overlap only occurs between consecutive blocks, like in Figure 1, all the blocks of odd numbers can be used in parallel and then the update is performed with all the blocks of even numbers. If we consider a reordering of the unknowns which labels first the components corresponding to the odd subdomains and then the even ones, it can easily be shown that the new preconditioner is still a Multiplicative Schwarz method but with only two subdomains ; the method becomes an Alternative Schwarz method. In such a situation the overlap is the union of all the elementary overlaps and therefore the total dimension of the overlap does not change.

4 Multiplicative Schwarz as preconditionner of Krylov methods

4.1 Early termination

When preconditioning a Krylov method, we consider in this section the advantage to consider a splitting $A = M - N$ in which N is rank deficient which is the case for the Multiplicative Schwarz preconditioner.

For solving the original system $Ax = b$, we define a Krylov method, as being an iterative method which builds, from an initial guess x^0, a sequence of iterates $x^k = x^0 + y^k$ such that $y^k \in \mathcal{K}_k(A, r^0)$ where $\mathcal{K}_k(A, r^0)$ is the Krylov subspace of degree k, built from the residual r^0 of the initial guess : $\mathcal{K}_k(A, r^0) = \mathbb{P}_{k-1}(A)r^0$ where $\mathbb{P}_{k-1}(\mathbb{R})$ is the set of polynomials of degree $k-1$ or less. The vector y^k is obtained by a characteristic property which depends on the method ; this property may be minimizing the error $x^k - x$ for a given norm, or projecting the initial error onto the subspace $\mathcal{K}_k(A, r^0)$ in a given direction. Nevertheless, we consider that, for a given k, when the property $x \in x^0 + \mathcal{K}_k(A, r^0)$ holds, it implies that $x^k = x$. This last property is satisfied when the Krylov subspace sequence becomes stationary :

$$(\mathcal{K}_k(A, r^0) = \mathcal{K}_{k+1}(A, r^0)) \Rightarrow (x \in \mathcal{K}_k(A, r^0)).$$

When a Krylov method is left preconditioned by the operator M, the Krylov subspaces to consider are : $\mathcal{K}_k(M^{-1}A, M^{-1}r^0)$. For a right preconditioning with the same operator, the subspace of interest becomes : $\mathcal{K}_k(AM^{-1}, r^0)$.

Proposition 4.1 *When $rank(N) = r < n$, then any Krylov method reaches the exact solution in at most $r+1$ iterations.*

Proof. In $M^{-1}A = I - M^{-1}N$ the matrix $M^{-1}N$ is of rank r. For any degree k, the following inclusion $\mathcal{K}_k(M^{-1}A, M^{-1}r^0) \subset (r^0) + \mathcal{R}(M^{-1}N)$ guarantees that the dimension of $\mathcal{K}_k(M^{-1}A, M^{-1}r^0)$ is at most $r+1$. Therefore, the method is stationnary from $k = r+1$ at the latest. The proof is identical for the right preconditioning. ◇

For a general non singular matrix, this result is applicable to the methods BiCG and QMR, preconditioned by the Multiplicative Schwarz method. In exact arithmetic, the number of iterations cannot exceed the total dimension s of the overlap by more than 1. The same result applies to GMRES(m) when m is greater than s.

For a symmetric positive definite matrix, the relevant method is PCG and it requires a s.p.d preconditionner. It is therefore necessary to symmetrize the basic multiplicative Schwarz preconditioner as done in section 3.4. When, the matrix is symmetric non definite, the symmetric preconditioner might also be non definite, and the method to consider is SQMR. However in these situations, the rank deficiency property is much less attractive than with the non symmetric basic case.

The previous remarks hold in exact arithmetic, but, as we shall see in the numerical tests, roundoff errors darken the picture especially for methods which rely on non orthonormal basis and for ill conditioned matrices.

4.2 Advantages of the explicit formulation

In the classical expression of the Multiplicative Schwarz iteration (Algorithm 1) the computation of the two steps

$$x^{k+1} \quad = \quad x^k + M^{-1}r^k \text{ and } r^{k+1} = b - Ax^{k+1},$$

is carried out recursively through the domains whereas the explicit formulation decouples the two computations. The computation of the residual is therefore more easily parallelized since it is withdrawn from the recursion. Another advantage of the explicit expression arises when it is used as a preconditioner of a method already coded in a library. In such a case, the user is supposed to provide a code for the procedure $x \to M^{-1}x$. Since the method computes the residual, the classical algorithm implies a double calculation of the residual.

We now show that the number of operations involved in both approaches remains roughly the same, although with a slight advantage to the explicit formulation.

Let us denote by $\mathcal{C}(p)_{cla}$ the cost of one iteration of Algorithm 1. One can verify that

$$\mathcal{C}(p)_{cla} = \sum_{i=1}^{p}(t_i + p_i), \tag{23}$$

where

$$\begin{cases} t_i &= \# \text{ of flops for processing on subdomain } i : x := x + A_i^+ r, \\ p_i &= \# \text{ of flops for processing on subdomain } i : r := r - A(A_i^+ r). \end{cases}$$

Let us also denote by $\mathcal{C}(p)_{exp}$ the number of operations for computing $x \to M^{-1}x$, by using the explicit form of the Multiplicative Schwarz preconditionner :

$$\mathcal{C}(p)_{exp} = \sum_{i=1}^{p-1}(t_i + \tau_i) + t_p, \tag{24}$$

where τ_i is the number of flops for multiplying a vector in subdomain i by C_i. The number of operations $\mathcal{C}(A)$ involved in the computation of a residual, which involves the multiplication by matrix A, satisfies the relation $\mathcal{C}(A) = \sum_{i=1}^{p} q_i - \sum_{i=1}^{p-1} \tau_i$ where q_i is the number of operations involved in the multiplication by block A_i. Since $q_i < p_i$ (they are usually close numbers), we obtain that

$$\mathcal{C}(p)_{cla} > \mathcal{C}(p)_{exp} + \mathcal{C}(A). \tag{25}$$

5 Numerical experiments

In this section we illustrate the numerical behaviour of the Multiplicative Schwarz preconditionner for Krylov subspace methods. The test matrices are taken from the Matrix Market suite [1]. The tests were carried out in MATLAB, and they were chosen to illustrate the property of early termination.

We denode by MS the Multiplicative Schwarz preconditioner and by SMS its Symmetrized Multiplicative Schwarz counterpart. For each method, the right-hand side is a fixed random vector, the initial guess is the nul vector and the tolerance for convergence is set to 10^{-8}.

Test1: Matrix: S3RMT3M3 shifted as $A := A + 10^{-3}\|A\|I$

Symmetric Reverse Cuthill Mackee reordering is applied on the matrix to reduce the bandwidth.
Source of S3RMT3M3 : Finite element analysis of Cylindrical Shells

- Order: 5357
- Type: Real Symmetric positive definite
- Condition number: 60.99

Domain decomposition

Block number	1	2	3	4	5	6
1st row index	1	875	1830	2810	3860	4700
last row index	1000	2000	3000	4000	4800	5357

- $rank(N) \leq \sum(rank(F_i)) = 92 + 144 + 164 + 131 + 89 = 620$
- Spectral radius of the iteration matrix: $\rho(M^{-1}N) = 0.4583$

Table 1 and Figure 4 show a nice convergence of methods since the spectral radius of the iteration matrix $M^{-1}N$ is small.

Test2: Matrix: BCCSTK20

Source of BCCSTK20 : Structural Engineering.

- Order: 485
- Type: Real Symmetric indefinite
- Condition number: 7.5×10^{12}

Domain decomposition

Block number:	1	2	3	4
1st row index	1	45	285	390
last row index	50	300	400	485

- $rank(N) \leq \sum(rank(F_i)) = 2 + 14 + 8 = 24$
- Spectral radius of the iteration matrix: $\rho(M^{-1}N) = 1$

That matrix is close to being singular and this is a difficult problem (see convergence in Table 2 and figure 5) when compared to the first one. Only GMRES succeeds. In BICG and QMR, a near-breakdown occurs [2].

Test3: Matrix: SHERMAN5

Symmetric Reverse Cuthill Mackee reordering on the matrix is applied to reduce the bandwidth.
Source of SHERMAN5 : Oil Reservoir simulation challenge matrices.

- Order: 3312
- Type: Real Unsymmetric
- Condition number: 3.9×10^5

Domain decomposition

Block number	1	2	3	4
1st row index	1	450	900	2495
last row index	500	970	2500	3312

- $rank(N) \leq \sum(rank(F_i)) = 31 + 61 + 0 = 92$
- Spectral radius of the iteration matrix: $\rho(M^{-1}N) = 0.8769$.

This is an unsymmetric problem and therefore, we only consider the methods GMRES, QMR and BiCG. Table 3 and Figure 6 show a good convergence for all the three methods. QMR and BiCG performs identically, which is the case when BiCG works well.

Test4: Matrix: GRE_1107

Source of GRE_1107 : Simulation of Computer System.

- Order: 1107
- Type: Real Unsymmetric
- Condition number: 9.7×10^7

Domain decomposition

Block number	1	2	3	4
1st row index	1	130	400	875
last row index	200	510	950	1107

- $rank(N) \leq \sum(rank(F_i)) = 38 + 73 + 49 + 0 = 160$
- Spectral radius of the iteration matrix: $\rho(M^{-1}N) = 5.3028 \times 10^4$.

This is another difficult problem. The matrix is almost singular. Table 4 and Figure 5 show the convergence of GMRES. It is interesting to note that the MS method would diverge since $\rho(M^{-1}N) > 1$. The last point in the residual of the graph reported in Figure 5 is surprizing since the 2-norm of the residual should define a non increasing sequence. It can easily be explained by the fact that during the inner iterations, the norm of the residual is computed by a formula which is not robust with respect to the loss of orthogonality of the basis whereas the residual is effectively computed at the basis restart. In BICG and QMR, near breakdown occurs which prevents convergence.

We can conclude from the sequence of tests that early termination property is not sufficient for obtaining convergence in floating point arithmetic. However GMRES, which appears to be much more robust, is clearly superior. In most of the cases, convergence was obtained even much earlier than what could be expected. However, GMRES suffers for the limitation on the size of the basis since a too large basis would imply a too high level of storage and a too large number of operations. Moreover, the loss of orthogonality within the basis may end up with a singular Hessenberg matrix which provokes a restart.

	Preconditionner	Number of Iteration	$\|r_k\|/\|r_0\|$
GMRES	MS	7	2.9095e-09
PCG	SMS	10	1.0679e-09
MINRES	SMS	11	9.8125e-09
QMR	MS	12	1.7088e-09
BICG	MS	9	1.0584e-09

Table 1: Convergence of the iterative methods on matrix S3RMT3M shifted and permuted

	Preconditionner	Number of Iteration	$\|r_k\|/\|r_0\|$
GMRES	MS	8	9.3112e-09
QMR	MS	no convergence	1
BICG	MS	no convergence	1

Table 2: Convergence of the iterative methods on matrix BCCSTK20

	Preconditionner	Number of Iteration	$\|r_k\|/\|r_0\|$
GMRES	MS	8	9.3112e-09
QMR	MS	12	1.1845e-09
BICG	MS	12	0 1.1910e-09

Table 3: Convergence of the iterative methods on matrix SHERMAN5 permuted

	Preconditionner	Number of Iteration	$\|r_k\|/\|r_0\|$
GMRES	MS	166	6.2627e-07
QMR	MS	no convergence	1
BICG	MS	no convergence	1

Table 4: Convergence of the iterative methods on matrix GRE_1107

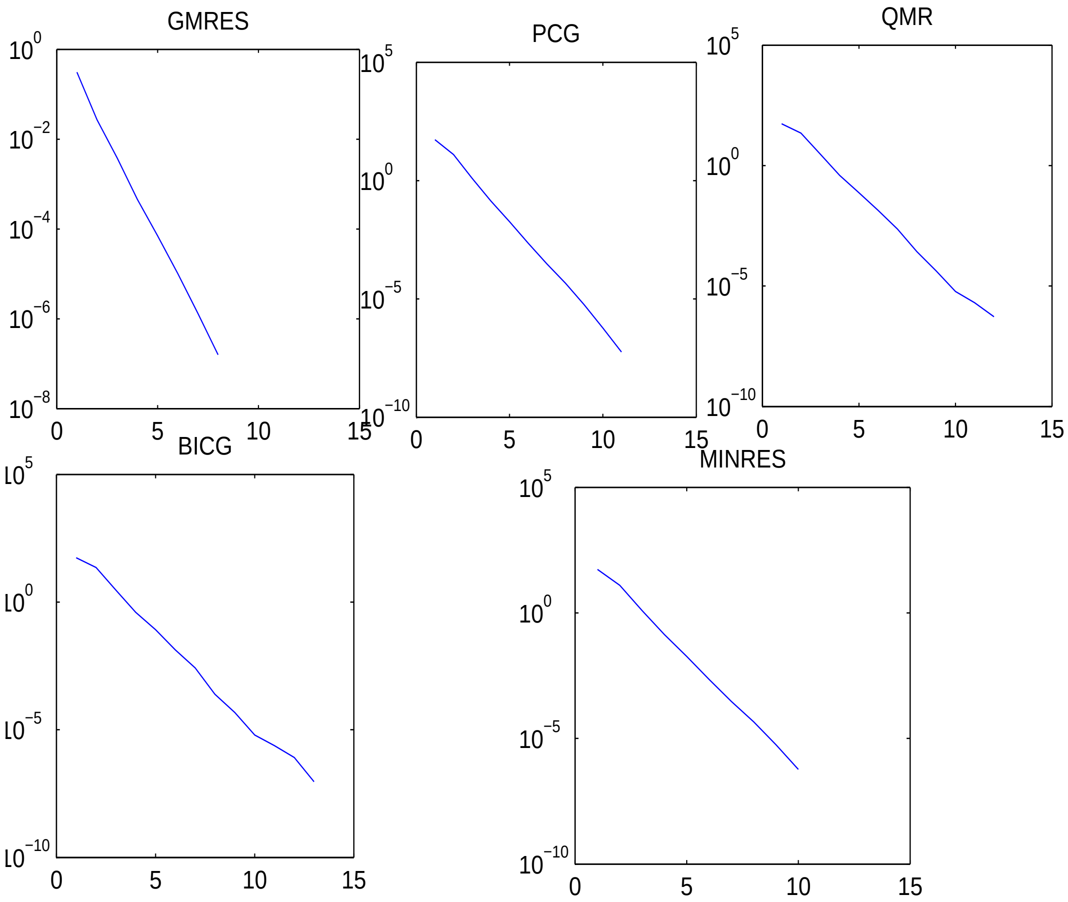

Figure 4: Convergence of the iterative methods on matrix S3RMT3M shifted and permuted

6 Conclusion

The Multiplicative Schwarz is a very efficient preconditionner especially for Krylov methods. We have established its early termination property which can reduce the number of iterations, depending on the amount of overlap.

In this work we have exhibited an explicit formulation of the Multiplicative Schwarz preconditionner. By decoupling the application of the preconditioner and the computation of the residual, we expect to be able to parallelize successive iterations. Such an approach is presently being developped on the GMRES method. A first basic parallel version of the codes is studied in [5]. Although the Additive Schwarz preconditioner is often prefered for its ability to be parallelized, it has a slower convergence rate. An efficient parallelized multiplicative version might change conclusion.

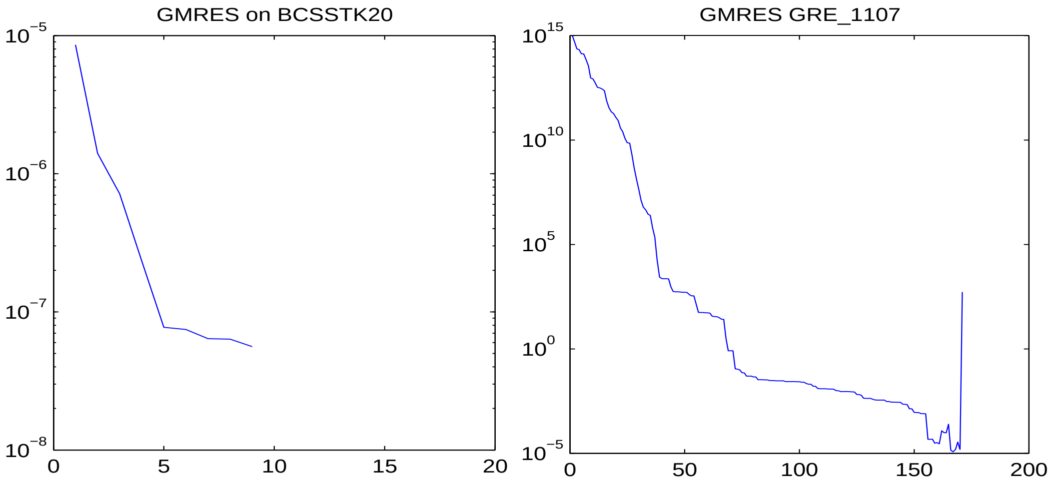

Figure 5: Convergence of the iterative methods on matrix GRE_1107

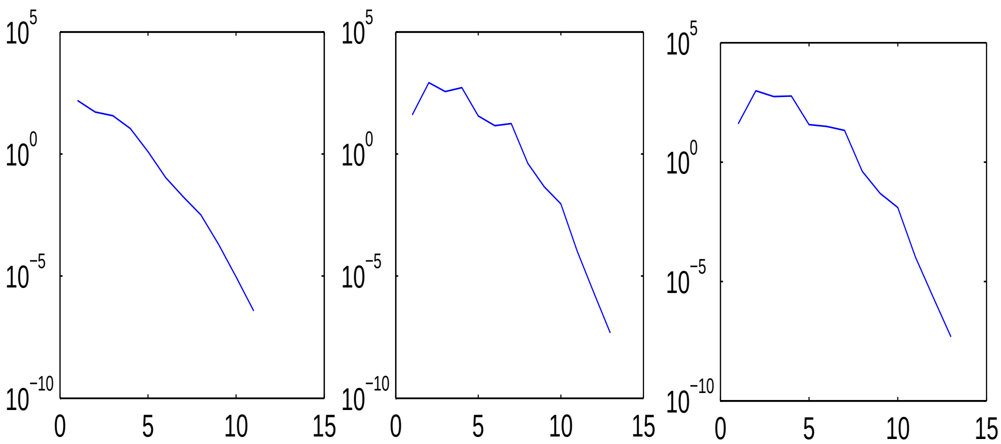

Figure 6: Convergence of the iterative methods on matrix Sherman5 permuted

7 Appendix: Alternative proof of Theorem 3.1

Notation: For $1 \leq i \leq j \leq p$, $I_{i:j}$ is the identity on the union of the domains W_k, $(i \leq k \leq j)$ and $\bar{I}_{i:j} = I - I_{i:j}$.

Lemma 7.1 *For any* $i \in \{1, \cdots, p-1\}$,

$$\bar{A}_{i+1} I_i + I_{i+1} \bar{A}_i - I_{i+1} A I_i = \bar{C}_i I_{i:i+1},$$

and for any $i \in \{1, \cdots, p\}$,

$$A_i^+ = \bar{A}_i^{-1} I_i = I_i \bar{A}_i^{-1}.$$

Straightforward (See figure 7)

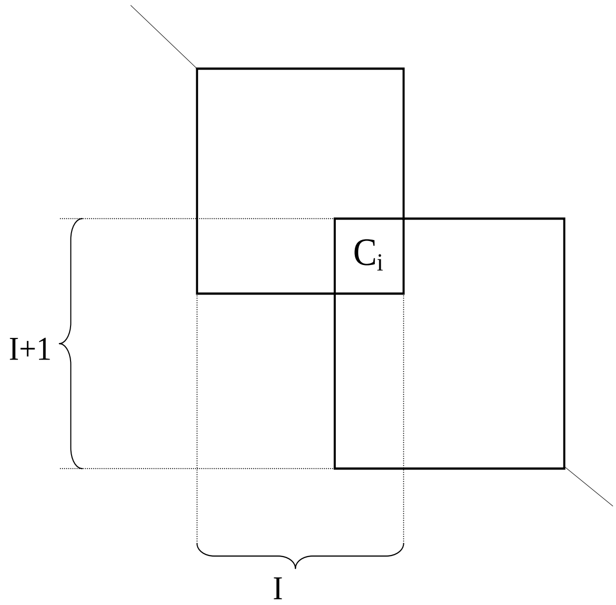

Figure 7: Subdomains of A

Proof of theorem 3.1 : Let us prove by induction that:

$$M^{-1} = \bar{A_p}^{-1} \bar{C}_{p-1} ... \bar{C}_1 \bar{A_1}^{-1}$$

Since $x^{k+1} = x^k + M^{-1} r^k$ and since x^{k+1} is obtained as result of successive steps, for $i = 1...p$: $x_{i+1} = x_i + A_i^+ r_i$ where x_i denotes $x^{k+i/p}$ and r_i denotes $r^{k+i/p}$, we shall prove by reverse induction on $i = p-1, \cdots, 0$ that:

$$x_p = x_i + \bar{A_p}^{-1} \bar{C}_{p-1} ... \bar{C}_{i+1} \bar{A}_{i+1}^{-1} I_{i+1:p} \; r_i \tag{26}$$

For $i = p-1$, the relation $x_p = x_{p-1} + \bar{A}_p^{-1} I_p r_p$ is obviously true.

Let us assume that (26) is valid for i and let us prove it for $i-1$.

$$\begin{aligned} x_p &= x_{i-1} + A_i^+ r_{i-1} + A_p^{-1} + ... + \bar{C}_{i+1} \bar{A}_{i+1}^{-1} I_{i+1:p} (I - A A_i^+) r_{i-1} \\ &= x_{i-1} + \bar{A}_p^{-1} ... \bar{C}_{i+1} (A_i^+ + A_{i+1}^+ I_{i+1:p} - \bar{A}_{i+1}^{-1} I_{i+1:p} A A_i^+) r_{i-1}. \end{aligned}$$

The last transformation was possible since the supports of $A_p, C_{p-1}, C_{p-1}, ..., C_{i+1}$ are disjoint from domain i. Let us transform the matrix expression:

$$\begin{aligned} B &= A_i^+ + A_{i+1}^+ I_{i+1:p} - \bar{A}_{i+1}^{-1} I_{i+1:p} A A_i^+ \\ &= \bar{A}_{i+1}^{-1} (A_{i+1}^+ I_i + I_{i+1:p} \bar{A}_i - I_{i+1:p} A I_i) \bar{A}_i^{-1} \end{aligned}$$

Lemma 7.1 and elementary calculations imply that:

$$\begin{aligned} B &= \bar{A}_{i+1}^{-1}(\bar{C}_i I_{i:i+1} + I_{i+2:p} - O_{i+1})\bar{A}_i^{-1} \\ &= \bar{A}_{i+1}^{-1}\bar{C}_i\bar{A}_i^{-1}I_{i:p} \end{aligned}$$

which proves that relation (26) is valid for $i-1$. This ends the proof.