A Multilevel Schwarz Preconditioner Based on a Hierarchy of Robust Coarse Spaces
Hussam Al Daas, Laura Grigori, Pierre Jolivet, Pierre-Henri Tourner

To cite this version:
Hussam Al Daas, Laura Grigori, Pierre Jolivet, Pierre-Henri Tourner. A Multilevel Schwarz Preconditioner Based on a Hierarchy of Robust Coarse Spaces. 2019. hal-02151184

HAL Id: hal-02151184
https://hal.archives-ouvertes.fr/hal-02151184
Submitted on 7 Jun 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.
A MULTILEVEL SCHWARZ PRECONDITIONER BASED ON A HIERARCHY OF ROBUST COARSE SPACES

HUSSAM AL DAAS†, LAURA GRIGORI‡, PIERRE JOLIVET§, AND PIERRE-HENRI TOURNIER¶

Abstract. In this paper we present a multilevel preconditioner based on overlapping Schwarz methods for symmetric positive definite (SPD) matrices. Robust two-level Schwarz preconditioners exist in the literature to guarantee fast convergence of Krylov methods. As long as the dimension of the coarse space is reasonable, that is, exact solvers can be used efficiently, two-level methods scale well on parallel architectures. However, the factorization of the coarse space matrix may become costly at scale. An alternative is then to use an iterative method on the second level, combined with an algebraic preconditioner, such as a one-level additive Schwarz preconditioner. Nevertheless, the condition number of the resulting preconditioned coarse space matrix may still be large. One of the difficulties of using more advanced methods, like algebraic multigrid or even two-level overlapping Schwarz methods, to solve the coarse problem is that the matrix does not arise from a partial differential equation (PDE) anymore. We introduce in this paper a robust multilevel additive Schwarz preconditioner where at each level the condition number is bounded, ensuring a fast convergence for each nested solver. Furthermore, our construction does not require any additional information than for building a two-level method, and may thus be seen as an algebraic extension.

Key words. domain decomposition, multilevel, elliptic problems, subspace correction

AMS subject classifications. 65F08, 65F10, 65N55

1. Introduction. We consider the solution of a linear system of equations

\begin{equation}
Ax = b,
\end{equation}

where \( A \in \mathbb{R}^{n \times n} \) is a symmetric positive definite (SPD) matrix, \( b \in \mathbb{R}^n \) is the right-hand side, and \( x \in \mathbb{R}^n \) is the vector of unknowns. To enhance convergence, it is common to solve the preconditioned system

\[ M^{-1}Ax = M^{-1}b. \]

Standard domain decomposition preconditioners such as block Jacobi, additive Schwarz, and restricted additive Schwarz methods are widely used [31, 9, 8]. In a parallel framework, such preconditioners have the advantage of relatively low communication costs. However, their role in lowering the condition number of the system typically deteriorates when the number of subdomains increases. Multilevel approaches have shown a large impact on enhancing the convergence of Krylov methods [32, 12, 7, 25, 20, 10, 21, 1, 15, 23]. In multigrid and domain decomposition communities, multilevel methods have proven their capacity of scaling up to large numbers of processors and tackling ill-conditioned systems [35, 4, 19]. While some preconditioners are purely algebraic [7, 20, 10, 26, 29, 16, 1], several multilevel methods are based on hierarchical meshing in both multigrid and domain decomposition communities [33, 9, 25, 15, 23]. Mesh coarsening depends on the geometry of the problem. One
Our preconditioner is based on a hierarchy of coarse spaces and is defined as following. At the first level, the set of unknowns is partitioned into $N_1$ subdomains and each subdomain has an associated matrix $A_{1,j} = R_{1,j}A_{1}R_{1,j}^{T}$ obtained by using appropriate restriction and prolongation operators $R_{1,j}$ and $R_{1,j}^{T}$ respectively, defined in the following section. The preconditioner is formed as an additive Schwarz preconditioner coupled with an additive coarse space correction, defined as,

$$M^{-1} = M_1^{-1} = V_1A_2^{-1}V_1^{T} + \sum_{j=1}^{N_1} R_{1,j}^{T}A_{1,j}^{-1}R_{1,j},$$

where $V_1$ is a tall-and-skinny matrix spanning a coarse space obtained by solving for each subdomain $j = 1$ to $N_1$ a generalized eigenvalue problem involving the matrix $A_{1,j}$ and the Neumann matrix associated with subdomain $j$. The coarse space matrix is $A_2 = V_1^{T}AV_1$. This is equivalent to the GenEO preconditioner, and is described in detail in [32] and recalled briefly in section 2. The dimension of the coarse space is proportional to the number of subdomains $N_1$. When it increases, factorizing $A_2$ by using a direct method becomes prohibitive, and hence the application of $A_2^{-1}$ to a vector should also be performed through an iterative method.

Our multilevel approach defines a hierarchy of coarse spaces $V_i$ and coarse space matrices $A_i$ for $i = 2$ to any depth $L + 1$, and defines a preconditioner $M_i^{-1}$ such that the condition number of $M_i^{-1}A_i$ is bounded. The depth $L + 1$ is chosen such that the coarse space matrix $A_{L+1}$ can be factorized efficiently by using a direct method. At each level $i$, the graph of the coarse space matrix $A_i$ is partitioned into $N_i$ subdomains, and each subdomain $j$ is associated with a local matrix $A_{i,j} = R_{i,j}A_iR_{i,j}^{T}$ obtained by
using appropriate restriction and prolongation operators $R_{i,j}$ and $R_{i,j}^T$, respectively. The preconditioner at level $i$ is defined as,

$$M_i^{-1} = V_i A_{i+1}^{-1} V_i^T + \sum_{j=1}^{N_i} R_{i,j}^T A_{i,j}^{-1} R_{i,j},$$

where the coarse space matrix is $A_{i+1} = V_i^T A_i V_i$.

One of the main contributions of the paper concerns the construction of the hierarchy of coarse spaces $V_i$ for levels $i$ going from 2 to $L$, that are built algebraically from the coarse space of the previous level $V_{i-1}$. This construction is based on the definition of local symmetric positive semi-definite (SPSD) matrices associated with each subdomain $j$ at each level $i$ that we introduce in this paper. These matrices are obtained by using the local SPSD matrices of the previous level $i-1$ and the previous coarse space $V_{i-1}$. They are then involved, with the local matrices $A_{i,j}$, in concurrent generalized eigenvalue problems solved for each subdomain $j$ that allows to compute the local eigenvectors contributing to the coarse space $V_i$.

We show in Theorem 5.3, section 5, that the condition number of $M_i^{-1} A_i$ is bounded and depends on the maximum number of subdomains at the first level that share an unknown, the number of distinct colors required to color the graph of $A_i$ so that $\{\text{span}\{R_{i,j}\}\}_{1 \leq j \leq N_i}$ of the same color are mutually $A_i$-orthogonal, and a user defined tolerance $\tau$. It is thus independent of the number of subdomains $N_i$.

The main contribution of this paper is based on the combination of two previous works on two-level additive Schwarz methods [3, 32]. The coarse space proposed in [32] guarantees an upper bound on the condition number that can be prescribed by the user. The SPSD splitting in the context of domain decomposition presented in [3] provides an algebraic view for the construction of coarse spaces. The combination of these two works leads to a robust multilevel additive Schwarz method. Here, robustness refers to the fact that at each level, an upper bound on the condition number of the associated matrix can be prescribed by the user a priori. The rest of the paper is organized as follows. In the next section, we present the notations used throughout the paper. In section 2, we present a brief review of the theory of one- and two-level additive Schwarz methods. We extend in section 3 the class of SPSD splitting matrices presented in [3] in order to make it suitable for multilevel methods. Afterwards, we define the coarse space at level $i$ based on the extended class of local SPSD splitting matrices associated with this level. Section 4 describes the partitioning of the domain at level $i + 1$ from the partitioning at level $i$. In Section 5, we explain the computation of the local SPSD matrices associated with each subdomain at level $i + 1$. We compute them using those associated with subdomains at level $i$. Section 6 presents numerical experiments on highly challenging diffusion and linear elasticity problems in two- and three-dimensional problems. We illustrate the theoretical robustness and practical usage of our proposed method by performing strong scalability tests up to 8,192 processes.

**Context and notation.** By convention, the finest level, on which (1.1) is defined, is the first level. A subscript index is used in order to specify which level an entity is defined on. In the case where additional subscripts are used, the first subscript always denotes the level. For the sake of clarity, we omit the subscript corresponding to level 1 when it is clear from context, e.g., matrix $A$. Furthermore, the subscripts $i$ and $j$ always refer to a specific level $i$ and its subdomain $j$, respectively. The number of levels is $L + 1$. Let $A_i \in \mathbb{R}^{n_i \times n_i}$ denote symmetric positive definite
matrices, each corresponding to level \( i = 1, \ldots, L + 1 \). We suppose that a direct solver can be used at level \( L + 1 \) to compute an exact factorization of \( A_{L+1} \).

Let \( B \in \mathbb{R}^{p \times q} \) be a matrix. Let \( P \subset [1:p] \) and \( Q \subset [1:q] \) be two sets of indices. The concatenation of \( P \) and \( Q \) is represented by \([P;Q]\). We note that the order of the concatenation is important. \( B(P,:) \) is the submatrix of \( B \) formed by the rows whose indices belong to \( P \). \( B(:,Q) \) is the submatrix of \( B \) formed by the columns whose indices belong to \( Q \). \( B(P,Q) = (B(P,:))(:,Q) \). The identity matrix of size \( p \) is denoted \( I_p \). We suppose that the graph of \( A_i \) is partitioned into \( N_i \) non-overlapping subdomains, where \( N_i \ll n_i \) and \( N_{i+1} \leq N_i \) for \( i = 1, \ldots, L \). We note that partitioning at level 1 can be performed by using a graph partitioning library such as ParMETIS \cite{22} or PT-SCOTCH \cite{11}. Partitioning at greater levels will be described later in section 4. In the following, we define for each level \( i = 1, \ldots, L \) notations for subsets and restriction operators that are associated with the partitioning. Let \( \Omega_i = [1:n_i] \) be the set of unknowns at level \( i \) and let \( \Omega_{i,j} \) for \( j = 1, \ldots, N_i \) be the subset of \( \Omega_i \) that represents the unknowns in subdomain \( j \). We refer to \( \Omega_{i,j} \) as the interior unknowns of subdomain \( j \). Let \( \Gamma_{i,j} \) for \( j = 1, \ldots, N_i \) be the subset of \( \Omega_i \) that represents the neighbor unknowns of subdomain \( j \), i.e., the unknowns at distance 1 from subdomain \( j \) through the graph of \( A_i \). We refer to \( \Gamma_{i,j} \) as the overlapping unknowns of subdomain \( j \). We denote \( \Omega_{i,j} = \{ \Omega_{i,j,\Gamma}, \Gamma_{i,j} \} \), for \( j = 1, \ldots, N_i \), the concatenation of interior and overlapping unknowns of subdomain \( j \). We denote \( \Delta_{i,j} \), for \( j = 1, \ldots, N_i \), the complementary of \( \Omega_{i,j} \) in \( \Omega_i \), i.e., \( \Delta_{i,j} = \Omega_i \setminus \Omega_{i,j} \). In Figure 1.1, a triangular mesh is used to discretize a square domain. The set of nodes of the mesh is partitioned into 16 disjoint subsets \( \Omega_{1,j,t} \), which represent a non-overlapping decomposition, for \( j = 1, \ldots, 16 \) (left). On the left, a matrix \( A_1 \) whose connectivity graph corresponds to the mesh is illustrated. The submatrix \( A_1(\Omega_{1,j,t},\Omega_{1,j,t}) \) is associated with the non-overlapping subdomain \( j \). Each submatrix \( A_1(\Omega_{1,j,t},\Omega_{1,j,t}) \) is colored with a distinct color. The same color is used to color the region that contains the nodes in the non-overlapping subdomain \( \Omega_{1,j,t} \). Note that if two subdomains \( j_1, j_2 \) are neighbors, the submatrix \( A_1(\Omega_{1,j_1,t},\Omega_{1,j_2,t}) \) has nonzero elements. For \( j = 1, \ldots, N_i \), we denote by \( n_{i,j,t}, \gamma_{i,j} \) and \( n_{i,j} \) the cardinality of \( \Omega_{i,j,t} \), \( \Gamma_{i,j} \) and \( \Omega_{i,j} \) respectively.

Let \( R_{i,j,t} \in \mathbb{R}^{n_{i,j,t} \times n_{i,j,t}} \) be defined as \( R_{i,j,t} = I_{n_{i,j,t}}(\Omega_{i,j,t},:) \). Let \( R_{i,j,\Gamma} \in \mathbb{R}^{n_{i,j,\Gamma} \times n_{i,j,\Gamma}} \) be defined as \( R_{i,j,\Gamma} = I_{n_{i,j,\Gamma}}(\Gamma_{i,j,\Gamma},:) \). Let \( R_{i,j} \in \mathbb{R}^{n_{i,j} \times n_{i,j}} \) be defined as \( R_{i,j} = I_{n_i}(\Omega_{i,j},:) \). Let \( \mathcal{P}_{i,j} = I_{n_i}(\Omega_{i,j,\Gamma},\Gamma_{i,j,\Delta},:) \in \mathbb{R}^{n_{i,j} \times n_{i,j,\Gamma}} \), be a permutation matrix associated with the subdomain \( j \). We denote \( D_{i,j} \in \mathbb{R}^{n_{i,j} \times n_{i,j}} \), \( j = 1, \ldots, N_i \), any set of non-negative diagonal matrices such that

\[
I_{n_i} = \sum_{j=1}^{N_i} R_{i,j}^T D_{i,j} R_{i,j}.
\]

We refer to \( \{D_{i,j} \}_{1 \leq j \leq N_i} \) as the algebraic partition of unity. Let \( V_i \in \mathbb{R}^{n_i \times n_i+1} \) be a tall-and-skinny matrix of full rank. We denote \( S_i \) the subspace spanned by the columns of \( V_i \). This subspace will stand for the coarse space associated with level \( i \). By convention, we refer to \( S_i \) as subdomain 0 at level \( i \). Thus, we have \( n_{i,0} = n_{i+1} \).
The interpolation operator at level $i$ is defined as:

$$\mathcal{S}_{i,2} : \prod_{j=0}^{N_i} \mathbb{R}^{n_{i,j}} \to \mathbb{R}^{n_i}$$

(1.2)

$$(u_j)_{0 \leq j \leq N_i} \mapsto \sum_{j=0}^{N_i} R_{i,j}^T u_j.$$  

Finally, we denote $\mathcal{V}_{i,j}$ the set of neighboring subdomains of each subdomain $j$ at level $i$ for $(i,j) \in \{1;L]\times\{1;N_i\}$.

$$\mathcal{V}_{i,j} = \{ k \in \{1;N_i\} : \Omega_{i,j} \cap \Omega_{i,k} \neq \emptyset \}.$$  

As previously mentioned, partitioning at level 1 can be performed by graph partitioning libraries such as ParMETIS [22] or PT-SCOTCH [11]. Partitioning at further levels will be defined later: the sets $\Omega_{i,j,1}$, $\Omega_{i,j,\Gamma}$, $\Omega_{i,j}$, and $\Delta_{i,j}$ for $i > 1$ are defined in subsection 4.2. The coarse spaces $S_i$ as well as the projection and prolongation operators $V_i^T$ and $V_i$ are defined in subsection 3.2. We suppose that the connectivity graph between the subdomains on each level is sparse. This assumption is not true in general, however, it is valid in structures based on locally constructed coarse spaces in domain decomposition as we show in this paper, see [18, Section 4.1 p.81] for the case of two levels.

2. Background. In this section, we review briefly several theoretical results related to additive Schwarz preconditioners. We introduce them for the sake of completeness.

**Lemma 2.1 (fictitious subspace lemma).** Let $A \in \mathbb{R}^{n_A \times n_A}$, $B \in \mathbb{R}^{n_B \times n_B}$ be two
symmetric positive definite matrices. Let $\mathcal{A}$ be an operator defined as
\[
\mathcal{A} : \mathbb{R}^{n_B} \to \mathbb{R}^{n_A}
\]

and let $\mathcal{A}^\top$ be its transpose. Suppose that the following conditions hold:

1. The operator $\mathcal{A}$ is surjective.
2. There exists $c_u > 0$ such that
   \[
   (\mathcal{A} v)^\top A (\mathcal{A} v) \leq c_u v^\top B v, \quad \forall v \in \mathbb{R}^{n_B}.
   \]
3. There exists $c_l > 0$ such that for all $v_{n_A} \in \mathbb{R}^{n_A}, \exists v_{n_B} \in \mathbb{R}^{n_B} | v_{n_A} = \mathcal{A} v_{n_B}$ and
   \[
   c_l v_{n_B}^\top B v_{n_B} \leq (\mathcal{A} v_{n_B})^\top A (\mathcal{A} v_{n_B}) = v_{n_A}^\top A v_{n_A}.
   \]

Then, the spectrum of the operator $\mathcal{A} B^{-1} \mathcal{A}^\top A$ is contained in the segment $[c_l, c_u]$.

Proof. We refer the reader to [12, Lemma 7.4 p.164] or [28, 27, 13] for a detailed proof.

**Lemma 2.2.** The operator $\mathcal{A}_{i,2}$ as defined in (1.2) is surjective.

**Proof.** The proof follows from the definition of $\mathcal{A}_{i,2}$ (1.2).

**Lemma 2.3.** Let $k_{i,c}$ for $i = 1, \ldots, L$ be the minimum number of distinct colors so that $\{\text{span}\{R_{i,j}\}\}_{1 \leq j \leq N_i}$ of the same color are mutually $A_i$-orthogonal. Then, we have
\[
(k_{i,c} + 1) \sum_{j=0}^{N_i} u_j^\top (R_{i,j} A_i R_{i,j}^\top) u_j, \quad \forall u_{B_i} = (u_j)_{0 \leq j \leq N_i} \in \prod_{j=0}^{N_i} \mathbb{R}^{n_{1,j}}.
\]

**Proof.** We refer the reader to [9, Theorem 12 p.93] for a detailed proof.

We note that at level $i$, the number $k_{i,c}$ is smaller than the maximum number of neighbors over the set of subdomains $[1; N_i]$

\[
k_{i,c} \leq \max_{1 \leq j \leq N_i} \# V_{i,j}.
\]

Due to the sparse structure of the connectivity graph between the subdomains at level $i$, the maximum number of neighbors over the set of subdomains $[1; N_i]$ is independent of the number of subdomains $N_i$. Then, so is $k_{i,c}$.

**Lemma 2.4.** Let $u_{A_i} \in \mathbb{R}^{n_{1,i}}$ and $u_{B_i} = \{u_j\}_{0 \leq j \leq N_i} \in \prod_{j=0}^{N_i} \mathbb{R}^{n_{1,j}}$ such that $u_{A_i} = \mathcal{A}_{i,2} u_{B_i}$. The additive Schwarz operator without any other restriction on the coarse space $S_i$ verifies the following inequality
\[
\sum_{j=0}^{N_i} u_j^\top (R_{i,j} A_i R_{i,j}^\top) u_j \leq 2 u_{A_i}^\top A_i u_{A_i} + (2k_{i,c} + 1) \sum_{j=1}^{N_i} u_j^\top R_{i,j} A_i R_{i,j}^\top u_j,
\]

where $k_{i,c}$ is defined in Lemma 2.3.
Lemma 2.5. Let $A, B \in \mathbb{R}^{m \times m}$ be two symmetric positive semi-definite matrices. Let $\ker(A)$, $\text{range}(A)$ denote the null space and the range of $A$ respectively. Let $\ker(B)$ denote the kernel of $B$. Let $L = \ker(A) \cap \ker(B)$, we denote $L^\perp_{\ker(A)}$ the orthogonal complementary of $L$ in $\ker(A)$. Let $P_0$ be an orthogonal projection on $\text{range}(A)$. Let $\tau$ be a positive real number. Consider the generalized eigenvalue problem,

\[ P_0 BP_0 u_k = \lambda_k A u_k, \]

\[ (u_k, \lambda_k) \in \text{range}(A) \times \mathbb{R}. \]

Let $P_\tau$ be an orthogonal projection on the subspace

\[ Z = L^\perp_{\ker(A)} \oplus \text{span} \{ u_k | \lambda_k > \tau \}, \]

then, the following inequality holds:

\[ (u - P_\tau u)^\top B (u - P_\tau u) \leq \tau u^\top A u, \quad \forall u \in \mathbb{R}^m. \]

Furthermore, $Z$ is the subspace of smallest dimension such that (2.1) holds.

**Proof.** We refer the reader to [3, Lemma 2.4] for a detailed proof.

2.1. GenEO coarse space. In [32, 12] the authors present the GenEO coarse space which relies on defining appropriate symmetric positive semi-definite (SPSD) matrices $\tilde{A}_j \in \mathbb{R}^{n \times n}$ for $j = 1, \ldots, N$. These are the unassembled Neumann matrices, corresponding to the integration on each subdomain of the operator defined in the variational form of the PDE. These matrices are local, i.e., $R_{j,\Delta} \tilde{A}_j = 0$. Furthermore, they verify the relations

\[ u^\top \tilde{A}_j u \leq u^\top A u, \quad \forall u \in \mathbb{R}^n, \]

\[ u^\top \sum_{j=1}^{N} \tilde{A}_j u \leq k_{\text{GenEO}} u^\top A u, \quad \forall u \in \mathbb{R}^n, \]

where $k_{\text{GenEO}} \leq N$ is the maximum number of subdomains that share an unknown.

2.2. Local SPSD splitting of an SPD matrix. In [3], the authors present the local SPSD splitting of an SPD matrix. Given the permutation matrix $P_j$, a local SPSD splitting matrix $\tilde{A}_j$ of $A$ associated with subdomain $j$ is defined as

\[ P_j \tilde{A}_j P_j^\top = \begin{pmatrix} R_{j,1} A_R_{j,1}^\top & R_{j,1} A_R_{j,1}^\top \\ R_{j,1} A_R_{j,1}^\top & A_{\tilde{L}}^\top \\ 0 & 0 \end{pmatrix}, \]

where $\tilde{A}_j^\top \in \mathbb{R}^{\gamma_j \times \gamma_j}$ satisfies the two following conditions: For all $u \in \mathbb{R}^{\gamma_j}$,

- $u^\top (R_{j,1} A R_{j,1}^\top) (R_{j,1} A R_{j,1}^\top)^{-1} (R_{j,1} A R_{j,1}^\top) u \leq u^\top \tilde{A}_j^\top u$

- $u^\top \tilde{A}_j^\top u \leq u^\top (R_{j,\Gamma} A R_{j,\Gamma}^\top) (R_{j,\Delta} A R_{j,\Delta}^\top)^{-1} (R_{j,\Delta} A R_{j,\Delta}^\top) u$. 

This manuscript is for review purposes only.
The authors prove that the matrices $\tilde{A}_j$ defined in such a way verify the following relations:

\begin{align}
R_{j,\Delta} \tilde{A}_j &= 0, \\
u^\top \tilde{A}_j u &\leq u^\top A u, \quad \forall u \in \mathbb{R}^n, \\
u^\top \sum_{j=1}^N \tilde{A}_j u &\leq k u^\top A u, \quad \forall u \in \mathbb{R}^n,
\end{align}

where $k$ is a number that depends on the local SPSD splitting matrices and can be at most equal to the number of subdomains $k \leq N$. The authors also show that the local matrices defined in GenEO [32, 12] can be seen as a local SPSD splitting.

In [3], the authors highlight that the key idea to construct a coarse space relies on the ability to identify the so-called local SPSD splitting matrices. They present a class of algebraically constructed coarse spaces based on the local SPSD splitting matrices. Moreover, this class can be extended to a larger variety of local SPSD matrices. This extension has the advantage of allowing to construct efficient coarse spaces for a multilevel structure in a practical way. This is discussed in the following section.

3. Extension of the class of coarse spaces. In this section we extend the class of coarse spaces presented in [3]. To do so, we present a class of matrices, that is larger than the class of local SPSD splitting matrices. This will be our main building block in the construction of efficient coarse spaces. Furthermore, this extension can lead to a straightforward construction of hierarchical coarse spaces in a multilevel Schwarz preconditioner setting.

3.1. Extension of the class of local SPSD splitting matrices. Regarding the two-level additive Schwarz method, the authors of [3] introduced the local SPSD splitting related to a subdomain as defined in (2.2). As it can be seen from the theory presented in that paper, it is not necessary to have the exact matrices $R_{j,\Gamma} A R_{j,\Gamma}^\top$, $R_{j,I} A R_{j,I}^\top$, and $R_{j,\Gamma} A R_{j,\Gamma}^\top$ in the definition of the local SPSD splitting in order to build an efficient coarse space. Indeed, the one and only necessary condition is to define for each subdomain $j$ an SPSD matrix $\tilde{A}_j$ for $j = 1, \ldots, N$ such that:

\begin{align}
R_{j,\Delta} \tilde{A}_j &= 0, \\
u^\top \sum_{j=1}^N \tilde{A}_j u &\leq k u^\top A u, \forall u \in \mathbb{R}^n,
\end{align}

where $k$ is a number that depends on the local SPSD matrices $\tilde{A}_j$ for $j = 1, \ldots, N$. The first condition means that $\tilde{A}_j$ has the local SPSD structure associated with subdomain $j$, i.e., it has the following form:

$$
p_j \tilde{A}_j p_j^\top = \begin{pmatrix} \tilde{A}_j^{I,\Gamma} & 0 \\ 0 & 0 \end{pmatrix},
$$

where $\tilde{A}_j^{I,\Gamma} \in \mathbb{R}^{n_j \times n_j}$. The second condition is associated with the stable decomposition property [34, 12]. Note that with regard to the local SPSD matrices, the authors in [32] only use these two conditions. That is to say, with matrices that verify conditions (3.1) the construction of the coarse space is straightforward through the
theory presented in either [32] or [3]. To this end, we define in the following the local SPSD (LSPSD) matrix associated with subdomain $j$ as well as the associated local filtering subspace that contributes to the coarse space.

**Definition 3.1 (local SPSD matrices).** An SPSD matrix $\tilde{A}_{i,j} \in \mathbb{R}^{n_i \times n_i}$ is called local SPSD (LSPSD) with respect to subdomain $j$ if

- $R_{i,j} A_{i,j} = 0$,
- $u^T \sum_{j=1}^{N_i} \tilde{A}_{i,j} u \leq k_i u^T A_i u$,

where $k_i > 0$.

We note that the local SPSD splitting matrices form a subset of the local SPSD matrices.

### 3.2. Multilevel coarse spaces.

This section summarizes the steps to be performed in order to construct the coarse space at level $i$ once we have the LSPSD matrices associated with each subdomain at that level.

**Definition 3.2 (coarse space based on LSPSD matrices).** Let $\tilde{A}_{i,j} \in \mathbb{R}^{n_i \times n_i}$ for $j = 1, \ldots, N_i$ be LSPSD matrices. Let $D_{i,j} \in \mathbb{R}^{n_i \times n_i}$ for $j = 1, \ldots, N_i$ be the partition of unity. Let $\tau_i > 0$ be a given number. For a subdomain $j \in [1; N_i]$, let

$$G_{i,j} = D_{i,j} (R_{i,j} A_{i,j} R_{i,j}^T) D_{i,j}.$$ 

Let $\tilde{P}_{i,j}$ be the projection on range($R_{i,j} \tilde{A}_{i,j} R_{i,j}^T$) parallel to ker($R_{i,j} \tilde{A}_{i,j} R_{i,j}^T$). Let $K_{i,j} = \ker(R_{i,j} \tilde{A}_{i,j} R_{i,j}^T)$, $L_{i,j} = \ker(G_{i,j}) \cap K_{i,j}$, and $L_{i,j}^{K_{i,j}}$ the orthogonal complementary of $L_{i,j}$ in $K_{i,j}$. Consider the generalized eigenvalue problem:

$$\tilde{P}_{i,j} G_{i,j} \tilde{P}_{i,j} u_{i,j} = \lambda_{i,j,k} R_{i,j} \tilde{A}_{i,j} R_{i,j}^T u_{i,j}.$$ 

(3.2)

Set

$$Z_{i,j} = L_{i,j}^{K_{i,j}} \oplus \text{span}\{u_{i,j} | \lambda_{i,j,k} > \tau_i\}.$$ 

(3.3)

Then, the coarse space associated with LSPSD matrices $\tilde{A}_{i,j}$ for $j = 1, \ldots, N_i$ at level $i$ is defined as:

$$S_i = \bigoplus_{j=1}^{N_i} R_{i,j}^T D_{i,j} Z_{i,j}.$$ 

(3.4)

Following notations from section 1, the columns of $V_i$ span the coarse space $S_i$. The matrix $A_{i+1}$ is defined as:

$$A_{i+1} = V_i^T A_i V_i.$$ 

(3.5)

The local SPSD splitting matrices at level 1 will play an important role in the construction of the LSPSD matrices at subsequent levels. In the following, we present an efficient approach for computing LSPSD matrices for levels greater than 1.

### 4. Partitioning for levels strictly greater than 1.

In this section, we explain how to obtain the partitioning sets $\Omega_{i,j}$ for $(i,j) \in [2; L] \times [1; N_i]$. Once the sets $\Omega_{i,j}$ for $j = 1, \ldots, N_i$ are defined at level $i$, the following elements are readily available: sets $\Omega_{i,j}$, $\Omega_{i,j}$, and $\Omega_{i,j}$; restriction operators $R_{i,j}$, $R_{i,j}$, $R_{i,j}$, $R_{i,j}$, and $R_{i,j}$; permutation matrices $\mathcal{P}_{i,j}$ for $j = 1, \ldots, N_i$. 

This manuscript is for review purposes only.
4.1. Superdomains as unions of several subdomains. In this section, we introduce the notion of superdomain. It refers to the union of several neighboring subdomains. Let \( \mathcal{G}_{i,1}, ... \), \( \mathcal{G}_{i,N_i}^{i+1} \) be disjoint subsets of \([1; N_i]\), where \( \bigcup_{j=1}^{N_i} \mathcal{G}_{i,j} = \bigcup_{j=1}^{N_i} \mathcal{G}_{i,j} \). We call the union of the subdomains \( \{ k \in [1; N_i] : k \in \mathcal{G}_{i,j} \} \) superdomain \( j \), for \( j = 1, ..., N_i^{i+1} \). Figure 4.1 gives an example of how to set superdomains. Though this definition of superdomains may look somehow related to the fine mesh, it is in practice done at the algebraic level, as explained later on. Note that the indices of columns and rows of \( A^{i+1} \) are associated with the vectors contributed by the subdomains at level \( i \) in order to build the coarse space \( S_i \), see Figure 4.2. Hence, defining subdomains on the structure of \( A^{i+1} \) is natural once we have the subsets \( \mathcal{G}_{i,j} \), for \( j = 1, ..., N_i^{i+1} \).

4.2. Heritage from superdomains. Let \( e_{i,j} \) be the set of indices of the vectors that span \( R_{i,j}^i D_{i,j} Z_{i,j} \) in the matrix \( V_i \) for some \( (i,j) \in [1; L - 1] \times [1; N_i] \), see Figure 4.2. We define \( \Omega_{i+1,j,I} = \bigcup_{k \in \mathcal{G}_{i,j}} e_{i,k} \), for \( j = 1, ..., N_i^{i+1} \). We denote \( \Omega_{i+1,j,I} \), the subset of \([1; n_i^{i+1}] \setminus \Omega_{i+1,j,I} \) whose elements are at distance 1 from \( \Omega_{i+1,j,I} \) through the graph of \( A^{i+1} \). We note that

\[
\Omega_{i+1,j,I} \subset \bigcup_{p \in \mathcal{V}_{i,j}} \bigcup_{k \in \mathcal{V}_{i,p}} e_{i,k},
\]

where \( \mathcal{V}_{i,j} \) represents the set of subdomains that are neighbors of subdomain \( j \) at level \( i \) for \( j = 1, ..., N_i \). The overlapping subdomain \( j \) is defined by the set \( \Omega_{i+1,j,I} = [\Omega_{i+1,j,I} \setminus \Omega_{i+1,j,I} \cup \bigcup_{p \in \mathcal{V}_{i,j}} \bigcup_{k \in \mathcal{V}_{i,p}} e_{i,k} \]. The rest of the sets, restriction, and prolongation operators can be defined as given in section 1.

5. LSPSD matrices for levels strictly greater than 1. In [32, 12, 3], different methods are suggested to obtain local SPSD splitting matrices at level 1. These matrices are used to construct efficient two-level additive Schwarz preconditioners. Here in this section, we do not discuss the construction of these matrices at level 1. We suppose that we have the local SPSD matrices \( A_{1,j} \in \mathbb{R}^{n_i \times n_i} \) for \( j = 1, ..., N_i \). We
focus on computing LSPSD matrices $\tilde{A}_{i,j} \in \mathbb{R}^{n_i \times n_j}$ for $(i, j) \in \{2; L\} \times \{1; N_i\}$. We also suppose that the coarse space $\mathcal{S}_i$ is available, i.e., the matrices $V_i$ and $A_2 = V_i^T A_1 V_i$ are known explicitly.

**Proposition 5.1.** Let $i$ be a fixed level index, and let $\tilde{A}_{i,j}$ be an LSPSD of $A_i$, (see Definition 3.1), associated with subdomain $j$, for $j = 1, \ldots, N_i$. Let $\mathcal{G}_{i,1}, \ldots, \mathcal{G}_{i,N_{i+1}}$ be a set of superdomains at level $i$ associated with the partitioning at level $i+1$, see subsection 4.1. Let $V_i^T$ be the restriction matrix to the coarse space at level $i$. Then, the matrix $\tilde{A}_{i+1,j}$ which is defined as:

$$\tilde{A}_{i+1,j} = \sum_{k \in \mathcal{U}_{i,j}} V_i^T \tilde{A}_{i,k} V_i,$$

satisfies the conditions in Definition 3.1. That is, $\tilde{A}_{i+1,j}$ is LSPSD of $A_{i+1}$ with respect to subdomain $j$ for $j = 1, \ldots, N_{i+1}$.

**Proof.** To prove that $\tilde{A}_{i+1,j}$ is LSPSD of $A_{i+1}$ with respect to subdomain $j$, we have to prove the following:

- $R_{i+1,j}^T \Delta \tilde{A}_{i+1,j} = 0$
- $u^T \sum_{j=1}^{N_{i+1}} \tilde{A}_{i+1,j} u \leq k_{i+1} u^T A_{i+1} u$ for all $u \in \mathbb{R}^{n_{i+1}}$.

First, note that $R_{i,k}^T \tilde{A}_{i,j} = 0$ for all non-neighboring subdomains $k$ of subdomain $j$. This yields $Z_{i,k}^T D_{i,k} R_{i,k} \tilde{A}_{i,j} = 0$ for these subdomains $k$.

Now, let $m \in \{1; n_{i+1}\} \setminus \Omega_{i+1,j}$. We will show that the $m$th row of $\tilde{A}_{i+1,j}$ is zero.

Following the partitioning of subdomains at level $i+1$, there exists a subdomain $\Omega_{p_0}$ such that the $m$th column of $V_i$ is part of $R_{i,p_0}^T D_{i,p_0} Z_{i,p_0}$. We denote this column vector by $v_m$. Furthermore, the subdomain $p_0$ is not a neighbor of any subdomain that is a part of the superdomain $\mathcal{G}_{i,j}$. Hence, $v_m^T \tilde{A}_{i,k} = 0$ for $k \in \mathcal{G}_{i,j}$. The $m$th row of $\tilde{A}_{i+1,j}$ is given as $v_m^T \sum_{k \in \mathcal{G}_{i,j}} \tilde{A}_{i,k} V_i$. Then, $v_m^T \sum_{k \in \mathcal{G}_{i,j}} \tilde{A}_{i,k} = 0$, and the $m$th row of $\tilde{A}_{i+1,j}$ is zero.
To prove the second condition, we have
\[ u^T \sum_{j=1}^{N_{i+1}} \tilde{A}_{i+1,j} u = u^T \sum_{j=1}^{N_{i+1}} \sum_{k \in \mathcal{G}_{i,j}} V_i^T \tilde{A}_{i,k} V_i u. \]

Since \( \{ \mathcal{G}_{i,j} \}_{1 \leq j \leq N_{i+1}} \) form a disjoint partitioning of \([1; N_i]\), we can write
\[ u^T \sum_{j=1}^{N_{i+1}} \tilde{A}_{i+1,j} u = u^T \sum_{k=1}^{N_i} V_i^T \tilde{A}_{i,k} V_i u, \]
\[ \leq u^T V_i^T \sum_{k=1}^{N_i} \tilde{A}_{i,k} V_i u. \]
\[ \tilde{A}_{i,k} \text{ is an LSPSD matrix of } A_i \text{ for } k = 1, \ldots, N_i. \] Hence, we have
\[ u^T \sum_{j=1}^{N_{i+1}} \tilde{A}_{i+1,j} u \leq k_i u^T V_i V_i A_i u, \]
\[ \leq k_i u^T A_{i+1} u. \]

We finish the proof by setting \( k_{i+1} = k_i. \) \( \square \)

Figure 5.1 gives an illustration of the LSPSD construction provided by Proposition 5.1. Figure 5.1 (top left) represents the matrix \( A_1 \). The graph of \( A_1 \) is partitioned into 16 subdomains. Each subdomain is represented by a different color. Figure 5.1 (top right) represents the matrix \( V_1 \) whose column vectors form a basis of the coarse space \( S_1 \). Colors of columns of \( V_1 \) correspond to those of subdomains in \( A_1 \). Figure 5.1 (bottom left) represents the matrix \( A_2 = V_1^T A_1 V_1 \). Note that column and row indices of \( A_2 \) are associated with column indices of \( V_1 \). Four subdomains are used at level 2. The partitioning at level 2 is related to the superdomain \( \mathcal{G}_{1,j} = [4(j-1)+1; 4(j-1)+4] \) for \( j = 1, \ldots, 4 \). Figure 5.1 (bottom right) represents an LSPSD matrix of \( A_2 \) with respect to subdomain 1 at level 2.

Theorem 5.2 shows that the third condition of the fictitious subspace lemma Lemma 2.1 holds at level \( i \) for \( i = 1, \ldots, L. \)

**Theorem 5.2.** Let \( \tilde{A}_{i,j} \) be an LSPSD of \( A_i \) associated with subdomain \( j \), for \( (i, j) \in [1; L] \times [1; N_i] \). Let \( \tau_i > 0, Z_{i,j} \) be the subspace associated with \( \tilde{A}_{i,j}, \) and \( P_{i,j} \) be the projection on \( Z_{i,j} \) as defined in Lemma 2.5. Let \( u_i \in \mathbb{R}^{n_i} \) and let \( u_{i,j} = (D_{i,j} (I_{n_{i,j}} - P_{i,j}) R_{i,j} u_i) \) for \( (i, j) \in [1; L] \times [1; N_i] \). Let \( u_{i,0} \) be defined as,
\[ u_{i,0} = (V_i^T V_i)^{-1} V_i^T \left( \sum_{j=1}^{N_i} R_{i,j}^T D_{i,j} P_{i,j} R_{i,j} u_i \right). \]

Let \( m_i = (2 + (2k_{i,e} + 1)k_i \tau_i)^{-1} \). Then,
\[ u_i = \sum_{j=0}^{N_i} R_{i,j}^T u_{i,j}, \]
and
\[ m_i \sum_{j=0}^{N_i} u_{i,j} R_{i,j} A_i R_{i,j}^T u_{i,j} \leq u_i^T A_i u_i. \]
Fig. 5.1. Illustration of the LSPSD construction provided by Proposition 5.1. Top left: the matrix $A_1$, top right: $V_1$, bottom left: the matrix $A_2 = V_1^T A_1 V_1$, bottom right: $A_{2,1} = \sum_{j \in \mathcal{G}_{1,1}} V_1^T A_{1,j} V_1$, where $\mathcal{G}_{1,1} = 1, \ldots, 4$

Proof. We have

$$\sum_{j=0}^{N_i} R_{i,j}^T u_{i,j} = V_i (V_i^T V_i)^{-1} V_i^T \left( \sum_{j=1}^{N_i} R_{i,j}^T D_{i,j} P_{i,j} R_{i,j} u_i \right) + \sum_{j=1}^{N_i} R_{i,j}^T u_{i,j}$$

Since for all $y \in S_i$, $V_i (V_i^T V_i)^{-1} V_i^T y = y$, we have

$$\sum_{j=0}^{N_i} R_{i,j}^T u_{i,j} = \sum_{j=1}^{N_i} R_{i,j}^T D_{i,j} P_{i,j} R_{i,j} u_i + \sum_{j=1}^{N_i} R_{i,j}^T \left( D_{i,j} \left( I_{n_{i,j}} - P_{i,j} \right) \right) R_{i,j} u_i$$

$$= \sum_{j=1}^{N_i} R_{i,j}^T D_{i,j} R_{i,j} u_i,$$

$$= u_i.$$ 

To prove the inequality (5.1), we start with the inequality from Lemma 2.4. We
have
\[
\sum_{j=0}^{N_i} u_{i,j}^T R_{i,j} A_i R_{i,j}^T u_{i,j} \leq 2 u_i^T A_i u_i + (2k_{i,c} + 1) \sum_{j=1}^{N_i} u_{i,j}^T R_{i,j} A_i R_{i,j}^T u_{i,j},
\]
where we chose \( u_{i,j} \) in Lemma 2.4 to be \( (u_{i,j})_{j=0,\ldots,N_i} \) and \( u_{A_i} = u_i \). In Definition 3.2, we defined \( Z_{i,j} \), such that for all \( w \in \mathbb{R}^{n_{i,j}} \) we have
\[
((I_{n_{i,j}} - P_{i,j})w)^T (D_{i,j} R_{i,j} A_i R_{i,j}^T D_{i,j}) ((I_{n_{i,j}} - P_{i,j})w) \leq \tau_i w^T (R_{i,j} \tilde{A}_{i,j} R_{i,j}^T) w.
\]
Hence, in the special case \( w = R_{i,j} u_i \), we can write
\[
((I_{n_{i,j}} - P_{i,j}) R_{i,j} u_i)^T (D_{i,j} R_{i,j} A_i R_{i,j}^T D_{i,j}) ((I_{n_{i,j}} - P_{i,j}) R_{i,j} u_i)
\]
\[
\leq \tau_i (R_{i,j} u_i)^T (R_{i,j} \tilde{A}_{i,j} R_{i,j}^T) (R_{i,j} u_i).
\]
Equivalently,
\[
u_{i,j}^T R_{i,j} A_i R_{i,j}^T u_{i,j} \leq \tau_i (R_{i,j} u_i)^T R_{i,j} \tilde{A}_{i,j} R_{i,j}^T (R_{i,j} u_i).
\]
Plugging this inequality in (5.2) gives
\[
\sum_{j=0}^{N_i} u_{i,j}^T R_{i,j} A_i R_{i,j}^T u_{i,j} \leq 2 u_i^T A_i u_i + (2k_{i,c} + 1) \tau_i \sum_{j=1}^{N_i} (R_{i,j} u_i)^T R_{i,j} \tilde{A}_{i,j} R_{i,j}^T (R_{i,j} u_i).
\]
Since \( \tilde{A}_{i,j} \) is local, we have
\[
(R_{i,j} u_i)^T R_{i,j} \tilde{A}_{i,j} R_{i,j}^T (R_{i,j} u_i) = u_i^T \tilde{A}_{i,j} u_i, \text{ for } j = 1, \ldots, N_i.
\]
By using the fact that \( \tilde{A}_{i,j} \) is LSPSD of \( A_i \) for \( j = 1, \ldots, N_i \), we obtain the following:
\[
\sum_{j=0}^{N_i} u_{i,j}^T R_{i,j} A_i R_{i,j}^T u_{i,j} \leq 2 u_i^T A_i u_i + (2k_{i,c} + 1) k_i \tau_i u_i^T A_i u_i.
\]
Multiplying both sides with \( m_i \) ends the proof, i.e.,
\[
m_i \sum_{j=0}^{N_i} u_{i,j}^T R_{i,j} A_i R_{i,j}^T u_{i,j} \leq u_i^T A_i u_i.
\]

Theorem 5.3 provides an upper bound on the condition number of the preconditioned matrix \( M_i^{-1} A_i \) for \( i = 1, \ldots, L \).

**THEOREM 5.3.** Let \( M_i \) be the additive Schwarz preconditioner at level \( i \) combined with the coarse space correction induced by \( S_i \) defined in (3.4). The following inequality holds,
\[
\kappa (M_i^{-1} A_i) \leq (k_{i,c} + 1) (2 + (2k_{i,c} + 1) k_i \tau_i).
\]

**Proof.** Lemma 2.2, Lemma 2.3, and Theorem 5.2 prove that the multilevel preconditioner verifies the conditions in Lemma 2.1 at each level \( i \). Hence, the spectrum of the preconditioned matrix \( M_i^{-1} A_i \) is contained in the interval \([2 + (2k_{i,c} + 1) k_i \tau_i]^{-1}, k_{i,c} + 1\] . Equivalently, the condition number of the preconditioned matrix at level \( i \) verifies the following inequality
\[
\kappa (M_i^{-1} A_i) \leq (k_{i,c} + 1) (2 + (2k_{i,c} + 1) k_i \tau_i).
\]
Proposition 5.1 shows that the constant $k_i$ associated with the LSPSD matrices at level $i$ is independent of the number of levels and bounded by the number of subdomains at level 1. Indeed,

$$k_1 \geq k_i \text{ for } i = 2, \ldots, L.$$  

Furthermore, in the case where the LSPSD matrices at the first level are the Neumann matrices, $k_1$ is bounded by the maximum number of subdomains at level 1 and an unknown.

The constant $k_{i,c}$ for $i = 1, \ldots, L$ is the minimum number of distinct colors so that $\{\text{span}\{R_{i,j}^T\}\}_{1 \leq j \leq N_i}$ of the same color are mutually $A_i$-orthogonal. Both constants $k_i$ and $k_{i,c}$ are independent of the number of subdomains for each level $i$.

The constant $\tau_i$ can be chosen such that the condition number of the preconditioned system at level $i$ is upper bounded by a prescribed value. Hence, this allows to have a robust convergence of the preconditioned Krylov solver at each level.

Algorithm 5.1 presents the construction of the multilevel additive Schwarz method by using GenEO. The algorithm iterates over the levels. At each level, three main operations are performed. First, the construction of the LSPSD matrices. At level 1, the LSPSD matrices are the Neumann matrices, otherwise, Proposition 5.1 is used to compute them. Once the LSPSD matrix is available, the generalized eigenvalue problem in (3.2) has to be solved concurrently. Given the prescribed upper bound on the condition number, $Z_{i,j}$ can be set. Finally, the coarse space is available and the coarse matrix is assembled.

### Algorithm 5.1 Multilevel GenEO

**Require:** $A_1 = A \in \mathbb{R}^{n \times n}$ SPD, $L + 1$ number of levels, $N_i$ number of subdomains at each level, $G_{i,j}$ sets of superdomains

**Ensure:** preconditioner at each level $i$, $M_i^{-1}$ with bounded condition number of $M_i A_i$

1: for $i = 1, \ldots, L$ do
2:   for each subdomain $j = 1, \ldots, N_i$ do
3:     $A_{i,j} = R_{i,j}^T A_i R_{i,j}^T$ (local matrix associated with subdomain $j$)
4:     if $i = 1$ then
5:       local SPSD $\tilde{A}_{i,j}$ is Neumann matrix of subdomain $j$
6:     else
7:       compute local SPSD matrix as
8:         $$\tilde{A}_{i,j} = \sum_{k \in G_{i,j}} V_{i-1}^T \tilde{A}_{i-1,k} V_{i-1}$$
9:     end if
10: end for
11: $S_i = \bigoplus_{j=1}^{N_i} D_{i,j} R_{i,j}^T Z_{i,j}$, $V_i$ basis of $S_i$
12: coarse matrix $A_{i+1} = V_i^T A_i V_i$, $A_{i+1} \in \mathbb{R}^{n_{i+1} \times n_{i+1}}$
13: end for
14: $M_i^{-1} = V_i A_i^{-1} V_i^T + \sum_{j=1}^{N_i} R_{i,j}^T A_{i,j}^{-1} R_{i,j}$

6. **Numerical experiments.** In this section, the developed theory is validated numerically with FreeFEM [14] for finite element discretizations and HPDDDM [19]
We present numerical experiments on two highly challenging problems illustrating the efficiency and practical usage of the proposed method. For both problems, we use \( N_1 = 2,048 \) MPI processes (equal to the number of subdomains at level 1). We compare the two-level GenEO preconditioner and its multilevel extension by varying \( N_2 \) between 4 and 256. For the two-level method, \( N_2 \) corresponds to the number of MPI processes that solve the coarse problem in a distributed fashion using MKL CPARDISO \cite{makinen2011}. For the multilevel method, \( N_3 \) is set to 1, i.e., a three-level method is used. The goal of these numerical experiments is to show that when one switches from a two-level method with an exact coarse solver, to our proposed multilevel method, the number of outer iterations is not impacted. Thus, three levels are sufficient. As an outer solver, since all levels but the coarsest are solved approximately, the flexible GMRES \cite{saad1986flexible} is used. It is stopped when relative unpreconditioned residuals are lower than \( 10^{-6} \). Subdomain matrices \( \{A_{i,j}\}_{1 \leq i \leq 2, 1 \leq j \leq N_i} \) are factorized concurrently using MKL PARDISO, and eigenvalue problems are solved using ARPACK \cite{lehoucq1998arpack}. In both, two- and three-level GenEO, we factorize the local matrices \( A_{1,j} \) for \( j \in [1; N_1] \) and solve the generalized eigenvalue problems concurrently at the first level. For this reason, we do not take into account the time needed for these two steps which are performed without any communication between MPI processes. We compare the time needed to assemble and factorize \( A_2 \) in the two-level approach against the time needed to assemble \( A_2 \) and local SPSD matrices \( A_{2,j} \) for \( j \in [1; N_2] \), solve the generalized eigenvalue problems concurrently on the second level, assemble, and factorize the matrix \( A_3 \) in the three-level approach. We also compare the time spent in the outer Krylov solver during the solution phase.

Readers interested by a comparison of the efficiency of GenEO and multigrid methods such as GAMG \cite{barszcz2014} are referred to \cite{daas2019}. FreeFEM scripts used to produce the following results are available at the following URL: https://github.com/prj-/aldaas2019multi\footnote{note to reviewers: the repository is not yet public but will be once the paper is accepted}.

6.1. Diffusion test cases. The scalar diffusion equation with highly heterogeneous coefficient \( \kappa \) is solved in \([0,1]^d \) \((d = 2 \ or \ 3)\). The strong formulation of the equation is:

\[
-\nabla \cdot (\kappa \nabla u) = 1 \quad \text{in } \Omega, \\
u = 0 \quad \text{on } \Gamma_D, \\
\frac{\partial u}{\partial n} = 0 \quad \text{on } \Gamma_N.
\]

The exterior normal vector to the boundary of \( \Omega \) is denoted \( n \). \( \Gamma_D \) is the subset of the boundary of \( \Omega \) corresponding to \( x = 0 \) in 2D and 3D. \( \Gamma_N \) is defined as the complementary of \( \Gamma_D \) with respect to the boundary of \( \Omega \). We discretize the equation using \( P_2 \) and \( P_4 \) finite elements in the 3D and 2D test cases, respectively. The number of unknowns is \( 441 \times 10^6 \) and \( 784 \times 10^6 \), with approximately 28 and 24 nonzero elements per row in the 3D and 2D cases, respectively. The heterogeneity is due to the jumps in the diffusion coefficient \( \kappa \), see Figure 6.1, which is modeled with the following \( P_0 \) function:

\[
\kappa = \begin{cases} 
10^5(\lfloor 9y \rfloor) & \text{if } \lfloor 9x \rfloor \equiv \lfloor 9y \rfloor \equiv 0 \pmod{2}, \\
1 & \text{elsewhere}.
\end{cases}
\]

The results in two dimensions are reported in Table 6.1. The number of outer iterations for both two- and three-level GenEO is 32. The size of the level 2 operator is
$n_2 = 25 \times 2,048 = 51,200$. It is striking that the multilevel method does not deteriorate the numerical performance of the outer solver. For the two-level method, the first column corresponds to the time needed to assemble the Galerkin operator $A_2$ from (3.5) (assuming $V_1$ has already been computed by ARPACK), and to factorize it using $N_2$ MPI processes. For the three-level method, the first column corresponds to the time needed to assemble level 2 local subdomain matrices $\{A_{2,j}\}_{1 \leq j \leq N_2}$, level 2 local SPSD matrices, solve the generalized eigenvalue problem (3.2) concurrently, assemble the Galerkin operator $A_3$ and factorize it on a single process. The size of the level 3 operator is $n_3 = 20 \times N_2$. For both two- and three-level methods, the second column is the time spent in the outer Krylov solver once the preconditioner has been set up. In the last column of the three-level method, the number of inner iterations for solving systems involving $A_2$, which is not inverted exactly anymore, is reported. Another important numerical property of our method is that, thanks to fully controlled bounds at each level, the number of inner iterations is low, independently of the number of superdomains $N_2$. Because this problem is not large enough, it is still tractable by a two-level method, for which HPDDM was highly optimized for. Thus, there is no performance gain to be expected at this scale. However, one can notice that the construction of the coarse operator(s) scales nicely with $N_2$ for the three-level method, whereas the performance of the direct solver MKL CPARDISO quickly plateaus because of the finer and finer parallel workload granularity.

The results in three dimensions are reported in Table 6.2. The number of outer
two-level GenEO || three-level GenEO
\begin{tabular}{|c|cc||c|cc|}
\hline
\textbf{N}_2 & \textbf{CS solve} & \textbf{CS solve inner iter} \\
\hline
4 & 7.0 & 20.9 & 16.9 & 43.6 & 17 \\
16 & 5.0 & 19.8 & 7.7 & 26.7 & 17 \\
64 & 5.1 & 20.1 & 5.8 & 32.7 & 15 \\
256 & 5.2 & 24.1 & 5.3 & 22.6 & 14 \\
\hline
\end{tabular}

\textbf{Table 6.2}

\textit{Diffusion 3D test case, comparison between two- and three-level GenEO}

iterations for both the two- and three-level GenEO is 19. The observations made in two dimensions still hold, and the dimensions of \( A_2 \) and \( A_3 \) are the same. Once again, it is important to note that the number of outer iterations is the same for both methods.

6.2. \textbf{Linear elasticity test cases}. The system of linear elasticity with highly heterogeneous elastic moduli is solved in 2D and 3D. The strong formulation of the equation is given as:

\[
\begin{align*}
\text{div } \sigma(u) + f &= 0 \quad \text{in } \Omega, \\
u &= 0 \quad \text{on } \Gamma_D, \\
\sigma(u) \cdot n &= 0 \quad \text{on } \Gamma_N.
\end{align*}
\]

The physical domain \( \Omega \) is a beam of dimensions \([0, 10] \times [0, 1]\), extruded for \( z \in [0, 1] \) in 3D. The Cauchy stress tensor \( \sigma(\cdot) \) is given by Hooke’s law: it can be expressed in terms of Young’s modulus \( E \) and Poisson’s ratio \( \nu \).

\[
\sigma_{ij}(u) = \begin{cases} 
2\mu \varepsilon_{ij}(u) & \text{if } i \neq j, \\
2\mu \varepsilon_{ii}(u) + \lambda \text{div}(u) & \text{if } i = j,
\end{cases}
\]

where

\[
\varepsilon_{ij}(u) = \frac{1}{2} \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} \right), \mu = \frac{E}{2(1 + \nu)}, \text{and } \lambda = \frac{E\nu}{1 - 2\nu}.
\]

The exterior normal vector to the boundary of \( \Omega \) is denoted \( n \). \( \Gamma_D \) is the subset of the boundary of \( \Omega \) corresponding to \( x = 0 \) in 2D and 3D. \( \Gamma_N \) is defined as the complementary of \( \Gamma_D \) with respect to the boundary of \( \Omega \). We discretize (6.1) using the following vectorial finite elements: \( (P_2, P_2, P_2) \) in 3D and \( (P_3, P_3) \) in 2D. The number of unknowns is \( 146 \times 10^6 \) and \( 847 \times 10^6 \), with approximately 82 and 34 nonzero elements per row in the 3D and 2D cases, respectively. The heterogeneity is due to the jumps in \( E \) and \( \nu \). We consider discontinuous piecewise constant values for \( E \) and \( \nu \): \( (E_1, \nu_1) = (2 \times 10^{11}, 0.25), \) \( (E_2, \nu_2) = (10^7, 0.45) \), see Figure 6.2.

Results in two (resp. three) dimensions are reported in Table 6.3 (resp. Table 6.4). The number of outer iterations are 73 and 45 respectively. For these test cases, we slightly relaxed the criterion for selecting eigenvectors in coarse spaces, which explains why the iteration counts increase. However, the same observations as for the diffusion test cases still hold. The dimension of the level 2 matrix is \( n_2 = 50 \times 2,048 = 102,050 \), while for the level 3 matrix it is \( n_3 = 20 \times N_2 \). We observe that the number of iterations of the inner solver increases slowly when increasing the number of subdomains from 4 to 256 in the 2D case and remains almost constant in the 3D case. In terms of
To show the potential of our method at larger scales, a three-dimensional linear elasticity problem of size $616 \times 10^6$ is now solved on $N_1 = 8,192$ processes and $N_2 = 256$ superdomains. With the two-level method, $A_2$ is assembled and factorized in 27.5 seconds. With the three-level method, this step now takes 13.2 seconds, see Table 6.5. Once again, as before, the number of outer iterations remains constant and equal to 53 for both methods. Not taking into account the preconditioner setup, the problem is solved in 52 seconds in both cases. At this regime, it is clear that there are important gains for the setup phase. At even greater scales, gains for the solution phase are also expected. Moreover, another interesting fact to note regarding computation time is that the generalized eigenvalue problems solved concurrently at the first level to obtain $V_1$ actually represents a significant part of the total time of 152.8 seconds (resp. 138.5 seconds) with the two- (resp. three-) level method: 113.3 seconds. This cost can be reduced by taking a larger number of (smaller) subdomains, with the drawback of increasing the size of $V_1$ and thus $A_2$. This drawback represents a clear bottleneck for the two-level method but is alleviated by using the three-level method, making it a good candidate for problems at greater scales.

7. Conclusion. In this paper, we reviewed generalities of overlapping Schwarz preconditioners and presented a framework for its multilevel extension. We generalized the local SPSD splitting presented in [3] to cover a larger set of matrices leading to more flexibility for building robust coarse spaces. Based on local SPSD matrices on the first level, we presented how to compute local SPSD matrices for coarser levels. The multilevel solver based on hierarchical local SPSD matrices is robust and guarantees a bound on the condition number of the preconditioned matrix at each level depending on predefined values. Numerical experiments illustrate the theory and prove the efficiency of the method on challenging problems of large size arising from heterogeneous linear elasticity and diffusion problems with jumps in the coefficients of multiple orders of magnitude.
Table 6.4
Elasticity 3D test case, comparison between two- and three-level GenEO

<table>
<thead>
<tr>
<th>$N^2$</th>
<th>two-level GenEO</th>
<th>three-level GenEO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS solve</td>
<td>inner iter</td>
</tr>
<tr>
<td>4</td>
<td>28.5</td>
<td>46.9</td>
</tr>
<tr>
<td>16</td>
<td>17.3</td>
<td>35.4</td>
</tr>
<tr>
<td>64</td>
<td>15.0</td>
<td>33.2</td>
</tr>
<tr>
<td>256</td>
<td>13.6</td>
<td>40.7</td>
</tr>
</tbody>
</table>

Table 6.5
Elasticity 3D test case, comparison between two- and three-level GenEO

<table>
<thead>
<tr>
<th>$N^2$</th>
<th>two-level GenEO</th>
<th>three-level GenEO</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CS solve</td>
<td>inner iter</td>
</tr>
<tr>
<td>256</td>
<td>27.5</td>
<td>52.0</td>
</tr>
</tbody>
</table>

8. Acknowledgements. This work was granted access to the HPC resources of TGCC@CEA under the allocation A0050607519 made by GENCI. The work of the second author was supported by the NLAFFET project as part of European Union’s Horizon 2020 research and innovation program under grant 671633.

REFERENCES


