



Robust coarse spaces for Abstract Schwarz preconditioners via generalized eigenproblems

E. Agullo, L. Giraud, L. Poirel

**RESEARCH
REPORT**

N° 8978

November 2016

Project-Teams HiePACS



Robust coarse spaces for Abstract Schwarz preconditioners via generalized eigenproblems

E. Agullo*, L. Giraud*, L. Poirel*

Project-Teams HiePACS

Research Report n° 8978 — November 2016 — 19 pages

Abstract: The solution of large sparse linear systems is one of the most important kernels in many numerical simulations. The domain decomposition methods (DDM) community has developed many efficient and robust solvers in the last decades. While many of these solvers fall in Abstract Schwarz (AS) framework, their robustness has often been demonstrated on a case-by-case basis. In this paper, we propose a bound for the condition number of all deflated AS methods provided that the coarse grid consists of the assembly of local components that contain the kernel of some local operators. We show that classical results from the literature on particular instances of AS methods can be retrieved from this bound. We then show that such a coarse grid correction can be explicitly obtained algebraically via generalized eigenproblems, leading to a condition number independent of the number of domains. This result can be readily applied to retrieve the bounds previously obtained via generalized eigenproblems in the particular cases of Neumann-Neumann (NN), additive Schwarz (aS) and optimized Robin but also generalizes them when applied with approximate local solvers. Interestingly, the proposed methodology turns out to be a comparison of the considered particular AS method with generalized versions of both NN and aS for tackling the lower and upper part of the spectrum, respectively. We furthermore show that the application of the considered grid corrections in an additive fashion is robust in the aS case although it is not robust for AS methods in general. In particular, the proposed framework allows for ensuring the robustness of the aS method applied on the Schur complement (aS/S), either with deflation or additively, and with the freedom of relying on an approximate local Schur complement, leading to a new powerful and versatile substructuring method. Numerical experiments illustrate these statements.

Key-words: Abstract Schwarz framework, preconditioning, SPD linear systems, robustness, coarse space, generalized eigenvalue problems.

* Inria Bordeaux Sud-Ouest, 200 avenue de la Vieille Tour, F-33400 Talence, France

**RESEARCH CENTRE
BORDEAUX – SUD-OUEST**

200 avenue de la Vieille Tour
33405 Talence Cedex

Sur la construction d'espaces grossiers robustes pour des préconditionneurs dans le cadre Schwarz Abstrait via la résolution de problèmes aux valeurs propres généralisés

Résumé : La résolution de grands systèmes linéaires est l'un des noyaux de calcul critiques dans de nombreuses simulations. À cet effet des méthodes de décomposition de domaines robustes et efficaces ont été proposées durant ces dernières décades. Bien que nombre de ces méthodes rentrent dans le cadre théorique des méthodes de Schwarz Abstraites (AS), leur robustesse est souvent établie au cas par cas. Dans ce rapport, nous proposons une borne pour le conditionnement des méthodes AS avec déflation basée un espace grossier qui contient le noyau de certains opérateurs locaux. Nous montrons que cette borne permet sur cette forme d'AS de retrouver des résultats classiques de la littérature. Nous montrons ensuite que ces espaces grossiers peuvent être construites explicitement et algébriquement via la résolution de problèmes aux valeurs propres généralisés; ce qui conduit à des conditionnements indépendants du nombre de sous-domaines. Ce résultat peut être utilisé pour retrouver les bornes existantes construites en résolvant des problèmes aux valeurs propres généralisés dans les cas particuliers de Neumann-Neumann (NN), Schwarz additif (aS) et Robin optimisé mais permet également de les généraliser lorsque des solveurs locaux approchés sont utilisés. Par ailleurs, nous montrons que nos opérateurs grossiers peuvent être utilisés de façon additive de façon robust dans le cadre AS alors que ce n'est pas le cas dans le cas général AS. In particulier, le cadre que nous proposons permet d'assurer la robustess d'aS appliqué sur la matrice initiale ou sur son complément de Schur pour sa version avec déflation ou simplement additive avec la flexibilité additionnelle de pouvoir utiliser des compléments de Schur locaux approchés. Ceci conduit à une technique de sous-structuration puissante et versatile. Des résultats expérimentaux sont présentés qui valident l'ensemble de ces propositions.

Mots-clés : Cadre Schwarz abstrait, préconditionneur, système linéaire symétrique défini positif, robustesse, espace grossier, problème aux valeurs propres généralisé.

1 Introduction

Domain decomposition methods [Toselli and Widlund, 2005, Mathew, 2008, Dolean et al., 2015] are a natural source of parallelism in scientific computing and for instance, sparse linear systems can be solved using substructuring methods by decomposing a global domain in smaller nonoverlapping subdomains. A simultaneous factorization of the interior problems leads to the interface problem of finding $x_\Gamma \in V$ such that

$$\mathcal{S}x_\Gamma = f. \quad (1.1)$$

In this study, we consider the particular case where \mathcal{S} is a symmetric positive definite (SPD) matrix, in which case Equation (1.1) can be solved in parallel using the preconditioned conjugate gradient method (PCG). We suppose that the search space can be decomposed as $V = \sum_{i=1}^N \text{range}(\mathcal{R}_{\Gamma_i}^T)$, where \mathcal{R}_{Γ_i} is a full-rank restriction matrix. This sum does not need to be direct. In the substructuring context, \mathcal{R}_{Γ_i} is the canonical restriction from the global interface Γ to the local interface of subdomain i : $\Gamma_i \subset \Gamma$.

Note that all the work presented in the present paper is discussed in the context of substructuring methods but results can also be applied directly to any SPD matrix \mathcal{A} . In case of ambiguity, we will mention whether the considered method is applied *on* \mathcal{A} (with the $/\mathcal{A}$ notation) or *on* \mathcal{S} (with the $/\mathcal{S}$ notation).

A good preconditioner \mathcal{M} for Problem (1.1) should have the two following properties: (1) \mathcal{M} is SPD and *close* to \mathcal{S}^{-1} , namely, the condition number $\kappa(\mathcal{M}\mathcal{S})$ should be as small as possible; (2) it is easy to compute $\mathcal{M}u$ for any vector u (at least easier than $\mathcal{S}^{-1}u$). Domain decomposition methods are often used to build such preconditioners of the form

$$\mathcal{M}_{AS} = \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \hat{\mathcal{S}}_i^\dagger \mathcal{R}_{\Gamma_i} \quad (1.2)$$

where $\hat{\mathcal{S}}_i$ is a local problem associated with \mathcal{S} on subdomain i , and † represents a pseudo-inverse.

These preconditioners have been studied for long using the Abstract Schwarz (AS) theory (see, e.g., Toselli and Widlund [2005] for a recent overview). Two particular cases of preconditioners that fit this description are the Neumann-Neumann (NN) preconditioner [Mandel, 1993], with $\hat{\mathcal{S}}_i = D_i^{-1} \mathcal{S}_i D_i^{-1}$, and the additive Schwarz (aS) preconditioner, with $\hat{\mathcal{S}}_i = \mathcal{R}_{\Gamma_i} \mathcal{S} \mathcal{R}_{\Gamma_i}^T$:

$$\mathcal{M}_{NN} = \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T D_i \mathcal{S}_i^\dagger D_i \mathcal{R}_{\Gamma_i}, \quad \mathcal{M}_{aS} = \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T (\mathcal{R}_{\Gamma_i} \mathcal{S} \mathcal{R}_{\Gamma_i}^T)^{-1} \mathcal{R}_{\Gamma_i}, \quad (1.3)$$

where $(D_i)_{i=1}^N$ is a partition of unity such that $\sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T D_i \mathcal{R}_{\Gamma_i} = \mathcal{I}$.

Unless $\hat{\mathcal{S}}_i$ perfectly mimics the global action of \mathcal{S} in subdomain i , $\kappa(\mathcal{M}_1 \mathcal{S})$ may significantly increase with the number N of subdomains, leading to scalability issues for these methods. Furthermore, if $\hat{\mathcal{S}}_i$ is singular, the pseudo-inverse is only defined up to an element in $\ker \hat{\mathcal{S}}_i$. To solve these two problems, a coarse space V_0 such that $\mathcal{R}_{\Gamma_i}^T \ker \hat{\mathcal{S}}_i \subset V_0$ and including also the low-energy modes of $\mathcal{M}_1 \mathcal{S}$ can be introduced, leading to the deflated AS preconditioner

$$\mathcal{M}_{AS,D} = V_0 (V_0^T \mathcal{S} V_0)^\dagger V_0^T + (\mathcal{I} - \mathcal{P}_0) \left(\sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \hat{\mathcal{S}}_i^\dagger \mathcal{R}_{\Gamma_i} \right) (\mathcal{I} - \mathcal{P}_0)^T \quad (1.4)$$

where $\mathcal{P}_0 = V_0(V_0^T \mathcal{S} V_0)^\dagger V_0^T \mathcal{S}$ is the \mathcal{S} -orthogonal projection onto V_0 . A simpler additive two-level preconditioner can be obtained by just adding the coarse solve to the one-level preconditioner

$$\mathcal{M}_{AS,2} = V_0(V_0^T \mathcal{S} V_0)^\dagger V_0^T + \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \hat{\mathcal{S}}_i^\dagger \mathcal{R}_{\Gamma_i}. \quad (1.5)$$

Note that both these deflated and additive applications of the coarse correction lead to preconditioners that are themselves general AS preconditioners (see, *e.g.*, Le Tallec and Vidrascu [1998]).

While many studies have proposed bounds on the condition number $\kappa(\mathcal{M}\mathcal{S})$ on particular numerical cases, often relying on analytical assumptions and most of the time requiring to work on the exact Schur complement, Le Tallec and Vidrascu [1998] have introduced a new class of preconditioners that generalize NN and derived a bound on the condition number for this relatively wide class of preconditioners. The exact matrix \mathcal{S} is still used in PCG and the approximation $\hat{\mathcal{S}}$ is only used to build the preconditioner, guaranteeing a convergence towards the actual solution of Equation (1.1). These preconditioners are called *generalized NN* in the original article. Because the generalization consists of handling an approximate Schur matrix, we will instead refer to them as *approximate NN* preconditioners in the present article.

These approximate NN preconditioners generalize classical NN but are very special cases of AS, not including aS for instance. The first contribution (Section 2) of this paper is to extend the result from Le Tallec and Vidrascu [1998] by using a generic local preconditioner and cover a broader range of AS methods, which we name *approximate deflated AS* methods and consist of all deflated AS methods whose coarse grid consists of the assembly of local components that contain the kernel of some local operators (and are formally introduced below, in Definition 1). Interestingly, the bound we exhibit (Theorem 1) highlights the key position of NN and aS among other local preconditioners in the Schwarz framework: they provide two bounds on the spectrum of the preconditioned operator, and the convergence of any AS local preconditioner can be evaluated by comparing it to these two well-known methods.

This bound depends on a generalized Rayleigh quotient which is often estimated using function analysis. Alternatively, we propose to control this Rayleigh quotient by building the coarse space using eigenvectors of well chosen generalized eigenproblems (Theorem 2). For that, we follow the Generalized Eigenvalue in the Overlap (GenEO) procedure Spillane et al. [2013]. This second contribution (Section 3) results in an explicit procedure for building a robust coarse space of any approximate deflated AS method leading to a bound on the condition number (hence on the number of iterations of PCG) independent of the number of subdomains. This result can be readily applied to retrieve the bounds previously obtained via generalized eigenproblems in the particular cases of aS (applied on the original problem, aS/ \mathcal{A}) Spillane et al. [2013], NN Spillane and Rixen [2012] and optimized Robin (SORAS) Haferssas et al. [2015]. It also generalizes these results to the approximate case. It can besides be applied to define a coarse space for the aS method applied on the Schur complement (aS/ \mathcal{S}) Carvalho et al. [2001] as illustrated in the numerical experiments (Section 5).

The third contribution (Section 4) of this paper is that the application of the considered coarse grid corrections in an additive fashion is robust in the approximate aS case (although it is not robust for AS methods in general). The bound we obtain (Theorem 3) can also be applied for retrieving the bound obtained in Spillane et al. [2013] when the coarse correction is applied additively to the aS method on the original matrix (aS/ \mathcal{A}) and again generalizes it to the approximate case. Our bound can furthermore be readily applied for ensuring the robustness of the aS method applied on the Schur complement (aS/ \mathcal{S}), possibly relying on an approximate local Schur complement.

The rest of the paper is organized as follows. Section 2 introduces a new, wide class of approximate (deflated) AS preconditioners and provides a bound on their condition number, which depends on a generalized Rayleigh quotient. Applying the GenEO procedure on two well chosen generalized eigenproblems, Section 3 proposes a procedure to explicitly compute the coarse space while bounding this Rayleigh quotient leading to a bound on the condition number (hence on the number of iterations of PCG) independent of the number of subdomains. Section 4 shows that a similar result (and procedure) can be obtained when the coarse grid correction is additively applied, in the case of approximate aS problems. Numerical experiments illustrate our discussion in Section 5 before concluding in Section 6.

2 Approximate Abstract Schwarz preconditioners

In this section, we first define a class of approximate (deflated) AS preconditioners, which combine a local preconditioner $\hat{\mathcal{S}}_i$, an approximate matrix $\tilde{\mathcal{S}}$ and a coarse space V_0 in Section 2.1. We then provide a bound on the condition number of this class of methods in Section 2.2, whose proof is provided in Section 2.3.

2.1 Context

Definition 1. Approximate (deflated) Abstract Schwarz preconditioner $\tilde{\mathcal{M}}_{AS,D}$

In order to build such a preconditioner, we need the three following objects:

1. a symmetric positive semi-definite (SPSD) local preconditioner $\hat{\mathcal{S}}_i$,
2. an approximation $\tilde{\mathcal{S}}$ of \mathcal{S} such that

$$\exists (\tilde{\mathcal{S}}_i)_{i=1}^N, \quad \tilde{\mathcal{S}} = \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \tilde{\mathcal{S}}_i \mathcal{R}_{\Gamma_i} \text{ and } \tilde{\mathcal{S}}_i \text{ is SPSPD,} \quad (2.1)$$

$$\exists \omega_-, \omega_+ > 0, \quad \forall v \in V \quad \omega_- v^T \mathcal{S} v \leq v^T \tilde{\mathcal{S}} v \leq \omega_+ v^T \mathcal{S} v, \quad (2.2)$$

3. and a coarse space V_0 such that

$$\exists (V_i^0)_{i=1}^N, \quad V_0 = \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T V_i^0 \quad \text{with} \quad \ker \hat{\mathcal{S}}_i + \ker D_i^{-1} \tilde{\mathcal{S}}_i D_i^{-1} \subset V_i^0. \quad (2.3)$$

We can then define a coarse matrix $\tilde{\mathcal{S}}_0 = V_0^T \tilde{\mathcal{S}} V_0$, a coarse projection $\tilde{\mathcal{P}}_0 = V_0 \tilde{\mathcal{S}}_0^\dagger V_0^T \tilde{\mathcal{S}}$ and the *approximate (deflated) AS preconditioner* is then defined as

$$\tilde{\mathcal{M}}_{AS,D} = V_0 \tilde{\mathcal{S}}_0^\dagger V_0^T + (\mathcal{I} - \tilde{\mathcal{P}}_0) \left(\sum_{i=1}^N \mathcal{R}_i^T \hat{\mathcal{S}}_i^\dagger \mathcal{R}_i \right) (\mathcal{I} - \tilde{\mathcal{P}}_0)^T. \quad (2.4)$$

Remark. If we choose $\hat{\mathcal{S}}_i = \tilde{\mathcal{S}}_i^{(NN)} = D_i^{-1} \tilde{\mathcal{S}}_i D_i^{-1}$ we retrieve the approximate NN preconditioner $\tilde{\mathcal{M}}_{NN,D}$ as defined in Le Tallec and Vidrascu [1998]. We can similarly define an *approximate aS* preconditioner $\tilde{\mathcal{M}}_{aS,D}$ with $\hat{\mathcal{S}}_i = \tilde{\mathcal{S}}_i^{(aS)} = \mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T$. Furthermore, if no approximation is used, then $\tilde{\mathcal{S}} = \mathcal{S}$ and $\omega_- = \omega_+ = 1$.

2.2 Convergence result for $\tilde{\mathcal{M}}_{AS,D}$

In each subdomain, we note $N_i = \#\{j \neq i, \mathcal{R}_i \hat{\mathcal{S}} \mathcal{R}_j^T \neq 0\}$ the number of neighbors of a subdomain through the connectivity graph of \mathcal{S} . We also define two local subspaces \hat{V}_i^\perp and \tilde{V}_i^\perp as the orthogonal spaces of V_i^0 for the outer products inferred by $\hat{\mathcal{S}}_i$ in $\text{range}(\hat{\mathcal{S}}_i)$ and $\tilde{\mathcal{S}}_i^{(NN)}$ in $\text{range}(\tilde{\mathcal{S}}_i^{(NN)})$ respectively. Then,

$$\text{range}(\mathcal{R}_{\Gamma_i}) = \hat{V}_i^\perp \oplus V_i^0 = \tilde{V}_i^\perp \oplus V_i^0, \quad (2.5)$$

$$\forall u \in V_i^0 \quad \forall v \in \hat{V}_i^\perp \quad \forall w \in \tilde{V}_i^\perp \quad u^T \hat{\mathcal{S}}_i v = u^T \tilde{\mathcal{S}}_i^{(NN)} w = 0. \quad (2.6)$$

Finally, for any SPSD matrix \mathcal{B} and vector u , we note $|u|_{\mathcal{B}} = \sqrt{u^T \mathcal{B} u}$ the \mathcal{B} -seminorm of u ; if \mathcal{B} is SPD, we note it $\|u\|_{\mathcal{B}}$.

Theorem 1. Convergence results for approximate AS

The condition number of the preconditioned matrix $\tilde{\mathcal{M}}_{AS,D}\mathcal{S}$ is bounded by

$$\kappa(\tilde{\mathcal{M}}_{AS,D}\mathcal{S}) \leq \frac{\omega_+}{\omega_-} \left(1 + \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\hat{\mathcal{S}}_i}^2}{|v|_{\tilde{\mathcal{S}}_i^{(NN)}}^2} \right) \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \hat{V}_i^\perp} \frac{|v|_{\tilde{\mathcal{S}}_i^{(aS)}}^2}{|v|_{\hat{\mathcal{S}}_i}^2} \right),$$

where $\tilde{\mathcal{S}}_i^{(NN)} = D_i^{-1} \tilde{\mathcal{S}}_i D_i^{-1}$ and $\tilde{\mathcal{S}}_i^{(aS)} = \mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T$.

Remark. We see three factors in this bound:

- The first one, with ω_+ and ω_- , controls the quality of the approximation $\tilde{\mathcal{S}}$.
- The second one is a generalized Rayleigh quotient between the local preconditioner $\hat{\mathcal{S}}_i$ and the approximate NN preconditioner $\tilde{\mathcal{S}}_i^{(NN)} = D_i^{-1} \tilde{\mathcal{S}}_i D_i^{-1}$.
- The last one is a generalized Rayleigh quotient between the local preconditioner $\hat{\mathcal{S}}_i$ and the approximate aS preconditioner $\tilde{\mathcal{S}}_i^{(aS)} = \mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T$.

Proof. The proof of Theorem 1 is a direct consequence of lemmas 3 and 5, using the fact that

$$\kappa(\tilde{\mathcal{M}}_{AS,D}\mathcal{S}) = \frac{\lambda_{\max}(\tilde{\mathcal{M}}_{AS,D}\mathcal{S})}{\lambda_{\min}(\tilde{\mathcal{M}}_{AS,D}\mathcal{S})}. \quad \square$$

Corollary 1. Convergence results for aS and NN

We define the deflated aS and NN preconditioners $\tilde{\mathcal{M}}_{NN,D}$ and $\tilde{\mathcal{M}}_{aS,D}$ by replacing $\hat{\mathcal{S}}_i$ with $\tilde{\mathcal{S}}_i^{(aS)}$ or $\tilde{\mathcal{S}}_i^{(NN)}$ respectively in Equation (2.4). We also define $N_c = \max_{1 \leq i \leq N} (N_i + 1)$.

Then, the condition number of $\tilde{\mathcal{M}}_{NN,D}\mathcal{S}$ and $\tilde{\mathcal{M}}_{aS,D}\mathcal{S}$ are bounded by

$$\kappa(\tilde{\mathcal{M}}_{aS,D}\mathcal{S}) \leq \frac{\omega_+}{\omega_-} \left(1 + \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\tilde{\mathcal{S}}_i^{(aS)}}^2}{|v|_{\tilde{\mathcal{S}}_i^{(NN)}}^2} \right) N_c,$$

$$\kappa(\tilde{\mathcal{M}}_{NN,D}\mathcal{S}) \leq \frac{\omega_+}{\omega_-} \max \left(1, \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\tilde{\mathcal{S}}_i^{(aS)}}^2}{|v|_{\tilde{\mathcal{S}}_i^{(NN)}}^2} \right) N_c.$$

Proof. The proof of Corollary 1 is a consequence of lemmas 3 and 4 for aS, and lemmas 2 and 5 for NN. \square

2.3 Proof of Theorem 1

To estimate the condition number of $\tilde{\mathcal{M}}_{AS,D}\mathcal{S}$, we need to bound the spectrum of this operator from above and below. The lower bound is a consequence of the Stable Decomposition Lemma as stated in [Toselli and Widlund, 2005].

Lemma 1. Stable decomposition lemma

If there exists a constant C_0 , local matrices \mathcal{B}_i and extension operators \mathcal{I}_i , such that $\ker \mathcal{B}_i \subset \ker \mathcal{I}_i$ and every $u \in V$ admits a decomposition

$$u = \sum_{i=0}^N \mathcal{I}_i u_i, \quad \{u_i \in V_i, 0 \leq i \leq N\} \quad \text{that satisfies} \quad \sum_{i=0}^N |u_i|_{\mathcal{B}_i}^2 \leq C_0^2 \|u\|_{\mathcal{S}}^2.$$

Then

$$\lambda_{\min}(\mathcal{M}\mathcal{S}) \geq C_0^{-2}, \quad \text{where} \quad \mathcal{M} = \sum_{i=0}^N \mathcal{I}_i \mathcal{B}_i^\dagger \mathcal{I}_i^T.$$

Proof. see, e.g., Lemma 2.5 in Toselli and Widlund [2005]. \square

Then, although it is not directly used in the proof of Theorem 1, we first expose in Lemma 2 a lower bound for the spectrum of NN ($\hat{\mathcal{S}}_i = \tilde{\mathcal{S}}_i^{(NN)}$) as it provides a good insight on the reason behind the Rayleigh quotients in the bound presented in Lemma 3 for the general case.

Lemma 2. Lower bound for the approximate Neumann-Neumann preconditioner

$$\text{Let } \tilde{\mathcal{M}}_{NN,D} = V_0 \tilde{\mathcal{S}}_0^\dagger V_0^T + (\mathcal{I} - \tilde{\mathcal{P}}_0) \left(\sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \tilde{\mathcal{S}}_i^{(NN)\dagger} \mathcal{R}_{\Gamma_i} \right) (\mathcal{I} - \tilde{\mathcal{P}}_0)^T.$$

Then,

$$\lambda_{\min}(\tilde{\mathcal{M}}_{NN,D}\mathcal{S}) \geq \frac{1}{\omega_+}.$$

Proof. This is a consequence of Lemma 1 (see Theorem 1 in Le Tallec and Vidrascu [1998]). \square

In the more general case, the bound can be obtained by comparing the local preconditioner $\hat{\mathcal{S}}_i$ with the approximate NN preconditioner through a Rayleigh quotient.

Lemma 3. Lower bound for the approximate Abstract Schwarz preconditioner

$$\lambda_{\min}(\tilde{\mathcal{M}}_{AS,D}\mathcal{S}) \geq \frac{1}{\omega_+} \left(1 + \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\hat{\mathcal{S}}_i}^2}{|v|_{\tilde{\mathcal{S}}_i^{(NN)}}^2} \right)^{-1}.$$

Proof. We want to split u into a sum of local contributions, while being able to uniformly control the $\hat{\mathcal{S}}_i$ -norm of these contributions u_i with the global \mathcal{S} -norm of u to apply Lemma 1.

For any u and $i \geq 1$, we decompose $D_i \mathcal{R}_{\Gamma_i} u = u_i^0 + u_i^\perp$ where $u_i^0 \in V_i^0$ and $u_i^\perp \in \tilde{V}_i^\perp$.

We then define $u_0 = (V_0^T \tilde{\mathcal{S}} V_0)^\dagger V_0^T \mathcal{S} u$ such that $V_0 u_0 = \tilde{\mathcal{P}}_0 u$. We can use the facts that $\sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T D_i \mathcal{R}_{\Gamma_i} = \mathcal{I}$ and $\sum_{i=0}^N \mathcal{R}_{\Gamma_i}^T u_i^0 \in V_0 \subset \ker(\mathcal{I} - \tilde{\mathcal{P}}_0)$ to obtain the decomposition

$$\begin{aligned} u &= \tilde{\mathcal{P}}_0 u + (\mathcal{I} - \tilde{\mathcal{P}}_0) u = V_0 u_0 + (\mathcal{I} - \tilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T D_i \mathcal{R}_{\Gamma_i} u \\ &= V_0 u_0 + (\mathcal{I} - \tilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T (u_i^0 + u_i^\perp) = V_0 u_0 + (\mathcal{I} - \tilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T u_i^\perp \\ &= \sum_{i=0}^N \mathcal{I}_i u_i \quad \text{where} \quad \mathcal{I}_0 = V_0, \quad \mathcal{I}_i = (\mathcal{I} - \tilde{\mathcal{P}}_0) \mathcal{R}_{\Gamma_i} \quad \text{and} \quad u_i = u_i^\perp. \end{aligned}$$

Since $\tilde{\mathcal{P}}_0$ is a $\tilde{\mathcal{S}}$ -orthogonal projection, we have:

$$|u_0|_{\tilde{\mathcal{S}}_0}^2 = |u_0|_{V_0^T \tilde{\mathcal{S}} V_0}^2 = |V_0 u_0|_{\tilde{\mathcal{S}}}^2 = |\tilde{\mathcal{P}}_0 u|_{\tilde{\mathcal{S}}}^2 \leq |u|_{\tilde{\mathcal{S}}}^2 \quad (2.7)$$

Let

$$C = \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\tilde{\mathcal{S}}_i}^2}{|v|_{\tilde{\mathcal{S}}_i^{(NN)}}^2} = \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\tilde{\mathcal{S}}_i}^2}{|v|_{D_i^{-1} \tilde{\mathcal{S}}_i D_i^{-1}}^2}.$$

We can then use equations (2.6), (2.1) and (2.7):

$$\begin{aligned} |u_i^\perp|_{\tilde{\mathcal{S}}_i}^2 &\leq C |u_i^\perp|_{D_i^{-1} \tilde{\mathcal{S}}_i D_i^{-1}}^2 \leq C |u_i^\perp + u_i^0|_{D_i^{-1} \tilde{\mathcal{S}}_i D_i^{-1}}^2 = C |\mathcal{R}_{\Gamma_i} u|_{\tilde{\mathcal{S}}_i}^2, \\ \sum_{i=1}^N |u_i^\perp|_{\tilde{\mathcal{S}}_i}^2 &\leq C \sum_{i=1}^N |\mathcal{R}_{\Gamma_i} u|_{\tilde{\mathcal{S}}_i}^2 = C |u|_{\sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \tilde{\mathcal{S}}_i \mathcal{R}_{\Gamma_i}}^2 = C |u|_{\tilde{\mathcal{S}}}^2, \\ |u_0|_{\tilde{\mathcal{S}}_0}^2 + \sum_{i=1}^N |u_i^\perp|_{\tilde{\mathcal{S}}_i}^2 &\leq (1 + C) |u|_{\tilde{\mathcal{S}}}^2 \leq \omega_+ (1 + C) |u|_{\tilde{\mathcal{S}}}^2, \end{aligned} \quad (2.8)$$

and the local norms are controlled by the global norm. Then, applying Lemma 1, we get

$$\lambda_{\min}(\tilde{\mathcal{M}}_{AS,D}\mathcal{S}) \geq \frac{1}{\omega_+} \left(1 + \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\tilde{\mathcal{S}}_i}^2}{|v|_{\tilde{\mathcal{S}}_i^{(NN)}}^2} \right)^{-1} \square$$

Now that we proved a lower bound for the spectrum of $\tilde{\mathcal{M}}_{NN,D}\mathcal{S}$, we will prove an upper bound in Lemma 5. We first recall a classic upper bound for aS preconditioners in Lemma 4 since it explains the origin of the Rayleigh quotient in the bound for the general case.

Lemma 4. Upper bound for the approximate additive Schwarz preconditioner

$$\text{Let } \tilde{\mathcal{M}}_{aS,D} = V_0 \tilde{\mathcal{S}}_0^\dagger V_0^T + (\mathcal{I} - \tilde{\mathcal{P}}_0) \left(\sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \tilde{\mathcal{S}}_i^{(aS)-1} \mathcal{R}_{\Gamma_i} \right) (\mathcal{I} - \tilde{\mathcal{P}}_0)^T.$$

Then,

$$\lambda_{\max}(\tilde{\mathcal{M}}_{aS,D}\mathcal{S}) \leq \frac{1}{\omega_-} \max_{1 \leq i \leq N} (N_i + 1).$$

Proof. This lemma is a particular case of Lemma 5 which is proven below. \square

Lemma 5. Upper bound for the approximate Abstract Schwarz preconditioner

$$\lambda_{\max}(\tilde{\mathcal{M}}_{AS,D}\mathcal{S}) \leq \frac{1}{\omega_-} \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \hat{V}_i^\perp} \frac{|v|_{\hat{\mathcal{S}}_i^{(aS)}}^2}{|v|_{\hat{\mathcal{S}}_i}^2} \right).$$

Proof. First, let us remark that

$$\tilde{\mathcal{M}}_{AS,D}\tilde{\mathcal{S}}u = V_0\hat{\mathcal{S}}_0^\dagger V_0^T\tilde{\mathcal{S}}u + (\mathcal{I} - \tilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \hat{\mathcal{S}}_i^\dagger \mathcal{R}_{\Gamma_i} (\mathcal{I} - \tilde{\mathcal{P}}_0)^T \tilde{\mathcal{S}}u = u_0 + (\mathcal{I} - \tilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T u_i$$

where $u_0 = \tilde{\mathcal{P}}_0 u$ and u_i is the orthogonal projection of $\hat{\mathcal{S}}_i^\dagger \mathcal{R}_{\Gamma_i} (\mathcal{I} - \tilde{\mathcal{P}}_0)^T \tilde{\mathcal{S}}u$ onto $\text{range}(\hat{\mathcal{S}}_i)$ along $\ker(\hat{\mathcal{S}}_i) \subset V_i^0 \subset \ker[(\mathcal{I} - \tilde{\mathcal{P}}_0)\mathcal{R}_{\Gamma_i}]$.

As a consequence, $u_i \in \hat{V}_i^\perp$:

$$u_i^T \hat{\mathcal{S}}_i V_i^0 = u^T \tilde{\mathcal{S}} (\mathcal{I} - \tilde{\mathcal{P}}_0) \mathcal{R}_{\Gamma_i}^T \hat{\mathcal{S}}_i^\dagger \hat{\mathcal{S}}_i V_i^0 = u^T \tilde{\mathcal{S}} (\mathcal{I} - \tilde{\mathcal{P}}_0) \mathcal{R}_{\Gamma_i}^T V_i^0 = 0.$$

Then,

$$\begin{aligned} |\tilde{\mathcal{M}}_{AS,D}\tilde{\mathcal{S}}u|_{\hat{\mathcal{S}}}^2 &= |u_0|_{\hat{\mathcal{S}}}^2 + |(\mathcal{I} - \tilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T u_i|_{\hat{\mathcal{S}}}^2 \leq |u_0|_{\hat{\mathcal{S}}}^2 + \left| \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T u_i \right|_{\hat{\mathcal{S}}}^2 \\ &\leq |u_0|_{\hat{\mathcal{S}}}^2 + \sum_{i=1}^N (N_i + 1) |\mathcal{R}_{\Gamma_i}^T u_i|_{\hat{\mathcal{S}}}^2 = |u_0|_{\hat{\mathcal{S}}}^2 + \sum_{i=1}^N (N_i + 1) |u_i|_{\mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T}^2 \end{aligned}$$

where we used the fact that

$$\begin{aligned} 0 &\leq \sum_{\mathcal{R}_{\Gamma_i}^T \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_j} \neq 0} |\mathcal{R}_{\Gamma_i}^T u_i - \mathcal{R}_{\Gamma_i}^T u_j|_{\hat{\mathcal{S}}}^2 = 2 \left(\sum_{\mathcal{R}_{\Gamma_i}^T \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_j} \neq 0} |\mathcal{R}_{\Gamma_i}^T u_i|_{\hat{\mathcal{S}}}^2 - \sum_{\mathcal{R}_{\Gamma_i}^T \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_j} \neq 0} u_i^T \mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_j}^T u_j \right) \\ &\leq 2 \left(\sum_{i=1}^N (N_i + 1) |\mathcal{R}_{\Gamma_i}^T u_i|_{\hat{\mathcal{S}}}^2 - \left| \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T u_i \right|_{\hat{\mathcal{S}}}^2 \right). \end{aligned} \quad (2.9)$$

Let us define

$$C = \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \hat{V}_i^\perp} \frac{|v|_{\hat{\mathcal{S}}_i^{(aS)}}^2}{|v|_{\hat{\mathcal{S}}_i}^2} \right) = \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \hat{V}_i^\perp} \frac{|v|_{\mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T}^2}{|v|_{\hat{\mathcal{S}}_i}^2} \right).$$

We can now write

$$\begin{aligned} |\tilde{\mathcal{M}}_{AS,D}\tilde{\mathcal{S}}u|_{\hat{\mathcal{S}}}^2 &\leq C |u_0|_{\hat{\mathcal{S}}}^2 + C \sum_{i=1}^N |u_i|_{\hat{\mathcal{S}}_i}^2 = C u^T \tilde{\mathcal{P}}_0^T \tilde{\mathcal{S}} u_0 + C \sum_{i=1}^N u^T \tilde{\mathcal{S}} (\mathcal{I} - \tilde{\mathcal{P}}_0) \mathcal{R}_{\Gamma_i}^T \hat{\mathcal{S}}_i^\dagger \hat{\mathcal{S}}_i u_i \\ &= C u^T \tilde{\mathcal{S}} \tilde{\mathcal{M}}_{AS,D} \tilde{\mathcal{S}} u \leq C |u|_{\hat{\mathcal{S}}} |\tilde{\mathcal{M}}_{AS,D} \tilde{\mathcal{S}} u|_{\hat{\mathcal{S}}} \\ |\tilde{\mathcal{M}}_{AS,D} \tilde{\mathcal{S}} u|_{\hat{\mathcal{S}}} &\leq C |u|_{\hat{\mathcal{S}}}, \end{aligned}$$

and use the same strategy as in Le Tallec and Vidrascu [1998] to obtain our result:

$$\begin{aligned} \lambda_{\max}(\tilde{\mathcal{M}}_{AS,D}\mathcal{S}) &= \max_{v \in V} \frac{|v|_{\hat{\mathcal{S}}}^2}{|v|_{\tilde{\mathcal{M}}_{AS,D}^{-1}}^2} \leq \max_{v \in V} \frac{1}{\omega_-} \frac{|v|_{\hat{\mathcal{S}}}^2}{|v|_{\tilde{\mathcal{M}}_{AS,D}^{-1}}^2} \leq \max_{v \in V} \frac{1}{\omega_-} \frac{|\tilde{\mathcal{M}}_{AS,D} \tilde{\mathcal{S}} v|_{\hat{\mathcal{S}}}^2}{|v|_{\hat{\mathcal{S}}}^2} \leq \frac{C}{\omega_-}, \\ \lambda_{\max}(\tilde{\mathcal{M}}_{AS,D}\mathcal{S}) &\leq \frac{1}{\omega_-} \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \hat{V}_i^\perp} \frac{|v|_{\hat{\mathcal{S}}_i^{(aS)}}^2}{|v|_{\hat{\mathcal{S}}_i}^2} \right). \quad \square \end{aligned}$$

3 Building the coarse space via generalized eigenproblems

The bound in Theorem 1 is traditionally estimated through functional analysis after a coarse space has been chosen. A more modern approach is to build the coarse space V_0 by solving a generalized eigenproblem in each subdomain in order to control the Rayleigh quotient as proposed by [Spillane and Rixen, 2012, Spillane et al., 2013] for aS and NN, respectively. This approach has also been successfully applied to other AS variants such as the SORAS method Haferssas et al. [2015], in which case two eigenproblems are needed.

The connection between the GenEO method and Theorem 1 comes from the following lemma:

Lemma 6. Bound on the Rayleigh quotient

Let \mathcal{A} be a SPSP matrix, and \mathcal{B} a SPD matrix.

If $V_\eta = \text{span}(\{p, \mathcal{A}p = \lambda \mathcal{B}p, \lambda \leq \eta\})$ and $V_\eta^{\perp \mathcal{A}} = \{u \in \text{range}(\mathcal{A}), \forall v \in V_\eta, u^T \mathcal{A}v = 0\}$,

$$\text{then } \sup_{u \in V_\eta^{\perp \mathcal{A}}} \frac{|u|_{\mathcal{B}}^2}{|u|_{\mathcal{A}}^2} \leq \frac{1}{\eta}.$$

Proof. Since \mathcal{B} is SPD, the generalized eigenproblem $\mathcal{A}p = \lambda \mathcal{B}p$ has solutions (λ_k, p_k) with $p_k^T \mathcal{B}p_l = \delta_{kl}$ and $p_k^T \mathcal{A}p_l = \lambda_k \delta_{kl}$.

Now, let $u \in V_\eta^{\perp \mathcal{A}}$. We can project u on the basis $(p_k)_k$: $u = \sum_k \alpha_k p_k$.

If k is such that $\lambda_k \leq \eta$, then $p_k \in V_\eta$ and $0 = u^T \mathcal{A}p_k = \lambda_k \alpha_k$. As a consequence, $\alpha_k = 0$ because if $\lambda_k = 0$, $p_k \in \ker(\mathcal{A}) = (\text{range}(\mathcal{A}))^\perp \perp u$ and $\alpha_k = u^T p_k = 0$. This leads to

$$\frac{|u|_{\mathcal{B}}^2}{|u|_{\mathcal{A}}^2} = \frac{\sum_{\lambda_k > \eta} \alpha_k^2}{\sum_{\lambda_k > \eta} \lambda_k \alpha_k^2} \leq \frac{1}{\eta}. \quad \square$$

Following the GenEO methodology, we propose to build the coarse space V_0 by solving two generalized eigenproblems to control the condition number of approximate AS preconditioners through two parameters $\alpha > 0$ and $\beta \geq 1$.

Theorem 2. If $\hat{\mathcal{S}}_i$ is SPD and the coarse space is defined as $V_0 = \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T V_i^0$ with

$$\begin{aligned} V_i^0 = \text{span}(\{ & p_k^i, \quad \tilde{\mathcal{S}}_i^{(NN)} p_k^i = \lambda_k^i \hat{\mathcal{S}}_i p_k^i, \quad \lambda_k^i \leq \alpha^{-1} \} \\ & \cup \{ p_k^i, \quad \hat{\mathcal{S}}_i p_k^i = \lambda_k^i \tilde{\mathcal{S}}_i^{(aS)} p_k^i, \quad \lambda_k^i \leq (N_i + 1)\beta^{-1} \}) \end{aligned}$$

then, we can bound the condition number

$$\kappa(\tilde{\mathcal{M}}_{AS,D}\mathcal{S}) \leq \frac{\omega_+}{\omega_-} (1 + \alpha) \beta.$$

Proof. Using Lemma 6 and the definition of \tilde{V}_i^{\perp} and \hat{V}_i^{\perp} , we can bound the Rayleigh quotients

$$\sup_{v \in \tilde{V}_i^{\perp}} \frac{|v|_{\hat{\mathcal{S}}_i}^2}{|v|_{\tilde{\mathcal{S}}_i^{(NN)}}^2} \leq \alpha, \quad \sup_{v \in \hat{V}_i^{\perp}} \frac{|v|_{\tilde{\mathcal{S}}_i^{(aS)}}^2}{|v|_{\hat{\mathcal{S}}_i}^2} \leq \frac{\beta}{N_i + 1}.$$

Replacing these bounds in Theorem 1 gives the result. □

Corollary 2. In the NN or aS cases, for any $\alpha \geq 1$, we can define

$$V_i^0 = \text{span}(\{ p_k^i, \quad \tilde{\mathcal{S}}_i^{(NN)} p_k^i = \lambda_k^i \tilde{\mathcal{S}}_i^{(aS)} p_k^i, \quad \lambda_k^i \leq \alpha^{-1} \}).$$

Then, Corollary 1 and Lemma 6 give

$$\kappa(\tilde{\mathcal{M}}_{aS,D}\mathcal{S}) \leq \frac{\omega_+}{\omega_-} (1 + \alpha) N_c, \quad \kappa(\tilde{\mathcal{M}}_{NN,D}\mathcal{S}) \leq \frac{\omega_+}{\omega_-} \alpha N_c.$$

4 Additive Coarse Correction

The preconditioner $\tilde{\mathcal{M}}_{AS,D}$ separates the part of the solution that is in V_0 (on which a direct coarse solve is performed through $\tilde{\mathcal{S}}_0^\dagger$), from its $\tilde{\mathcal{S}}$ -orthogonal part (on which the local preconditioner $\mathcal{M}_1 = \sum_{i=1}^N \mathcal{R}_i^T \tilde{\mathcal{S}}_i^\dagger \mathcal{R}_i$ is used to accelerate convergence). Eigenvalues λ corresponding to vectors in the coarse space V_0 are shifted to 1 by the coarse solve, and to 0 by the projection step $(\mathcal{I} - \tilde{\mathcal{P}}_0)$, so the overall effect of the deflated preconditioner is to shift them to 1 exactly. However, this step is quite expensive since it involves an additional matrix-vector product. If we skip this projection step, we get an approximate additive two-level preconditioner $\tilde{\mathcal{M}}_{AS,2}$ similar to $\mathcal{M}_{AS,2}$ presented in Equation (1.5). Without the projection step eigenvalues are shifted to $1 + \lambda$. As a result, this coarse correction applied on big eigenvalues only makes them bigger, thus hampering convergence. This additive coarse correction can only be effective to tackle the lower part of the spectrum since small eigenvalues $\lambda \ll 1$ are shifted to $1 + \lambda \approx 1$.

The one-level aS method has already an upper bound on the spectrum (see Lemma 4), and only the lower bound needs to be recovered, making it an ideal candidate for an additive coarse correction. In this section, we show that in the approximate aS case, when $\hat{\mathcal{S}}_i = \tilde{\mathcal{S}}_i^{(aS)} = \mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T$, the projection step can be removed without losing robustness. Namely, we still have a bound for the condition number of the additive two-level aS method independent of the number of subdomains.

Theorem 3. Condition number of the 2-level approximate aS preconditioners

Let $\mathcal{M}_{aS,2} = V_0 \tilde{\mathcal{S}}_0^\dagger V_0^T + \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \tilde{\mathcal{S}}_i^{(aS)-1} \mathcal{R}_{\Gamma_i}$ and $N_c = \max_{1 \leq i \leq N} (N_i + 1)$.

Then, we can bound the condition number

$$\kappa(\mathcal{M}_{aS,2}\mathcal{S}) \leq \frac{\omega_+}{\omega_-} \left[N_c + 1 + (N_c + 2) \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\tilde{\mathcal{S}}_i^{(aS)}}^2}{|v|_{\tilde{\mathcal{S}}_i^{(NN)}}^2} \right] (N_c + 1).$$

For any $\alpha > 0$, if we choose

$$V_i^0 = \text{span}(\{p_k^i, \tilde{\mathcal{S}}_i^{(NN)} p_k^i = \lambda_k^i \tilde{\mathcal{S}}_i^{(aS)} p_k^i, \lambda_k^i \leq \alpha^{-1}\}),$$

we have

$$\kappa(\mathcal{M}_{aS,2}\mathcal{S}) \leq \frac{\omega_+}{\omega_-} [N_c + 1 + \alpha(N_c + 2)] (N_c + 1).$$

Proof. If we apply Lemma 4 without a coarse space and consider V_0 as another subdomain in the decomposition, we get

$$\lambda_{\max}(\mathcal{M}_{aS,2}\mathcal{S}) \leq \frac{1}{\omega_-} (N_c + 1).$$

The lower bound is a consequence of Lemma 1. We define $u_i^0 \in V_i^0$ and $u_i^\perp \in \tilde{V}_i^{\perp}$ such that $D_i \mathcal{R}_{\Gamma_i} u = u_i^0 + u_i^\perp$ as in the proof of Lemma 3. We now introduce u_0 such that $V_0 u_0 = \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T u_i^0$, and $u = V_0 u_0 + \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T u_i^\perp$.

We get from Equation (2.8) that

$$\sum_{i=1}^N |u_i^\perp|_{\tilde{\mathcal{S}}_i^{(aS)}}^2 = \sum_{i=1}^N |u_i^\perp|_{\tilde{\mathcal{S}}_i}^2 \leq C |u|_{\tilde{\mathcal{S}}}^2 \quad \text{with} \quad C = \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\tilde{\mathcal{S}}_i^{(aS)}}^2}{|v|_{\tilde{\mathcal{S}}_i^{(NN)}}^2} = \frac{|v|_{\mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T}^2}{|v|_{D_i^{-1} \tilde{\mathcal{S}}_i D_i^{-1}}^2}.$$

Then, we can use the same method as in Equation (2.9):

$$\begin{aligned} |u_0|_{\tilde{\mathcal{S}}}^2 &= |u - \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T u_i^\perp|_{\tilde{\mathcal{S}}}^2 \leq (N_c + 1) \left(|u|_{\tilde{\mathcal{S}}}^2 + \sum_{i=1}^N |\mathcal{R}_{\Gamma_i}^T u_i^\perp|_{\tilde{\mathcal{S}}}^2 \right) \\ &= (N_c + 1) \left(|u|_{\tilde{\mathcal{S}}}^2 + \sum_{i=1}^N |u_i^\perp|_{\mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T}^2 \right) \leq (N_c + 1)(1 + C) |u|_{\tilde{\mathcal{S}}}^2 \\ |u_0|_{\tilde{\mathcal{S}}}^2 + \sum_{i=1}^N |u_i^\perp|_{\mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T}^2 &\leq [N_c + 1 + (N_c + 2)C] |u|_{\tilde{\mathcal{S}}}^2 \leq \omega_+ [N_c + 1 + (N_c + 2)C] |u|_{\tilde{\mathcal{S}}}^2. \end{aligned}$$

We then use Lemma 1 with $\mathcal{I}_0 = V_0$, $\mathcal{I}_i = \mathcal{R}_{\Gamma_i}^T$ and $\mathcal{B}_i = \mathcal{I}_i \tilde{\mathcal{S}} \mathcal{I}_i^T$ to get the bound

$$\lambda_{\min}(\mathcal{M}_{a,S,2}\mathcal{S}) \geq \frac{1}{\omega_+} \left[N_c + 1 + (N_c + 2) \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\mathcal{R}_{\Gamma_i} \tilde{\mathcal{S}} \mathcal{R}_{\Gamma_i}^T}^2}{|v|_{\tilde{\mathcal{S}}^{(NN)}}^2} \right]^{-1}.$$

We can then conclude with Lemma 6. \square

5 Numerical experiments

5.1 Experimental setup

The methods introduced in sections 2, 3 and 4 are tested on a problem similar to what is presented in Spillane et al. [2013]. We use the Finite Element Method to solve a heterogeneous diffusion equation $\nabla(K\nabla u) = 1$ in a 3D stratified medium. The domain $[0, N] \times [0, 6] \times [0, 1]$ is discretized on a regular mesh of $(5N + 1) \times 31 \times 6$ nodes. The domain is divided into N identical subdomains along the first axis. Along the second axis, it is divided into 10 layers (of $5N \times 3 \times 5$ elements each) of alternating conductivity $K_1 = 1$ and $K_2 = K$. A Dirichlet boundary condition is applied on the left of the domain ($x = 0$), a Neumann condition on every other boundary.

Although for a matter of conciseness we have chosen to express our algorithm on a matrix \mathcal{S} , all the equations are valid outside the scope of substructuring methods, and the matrix noted \mathcal{S} in Equation (2.1) can either be:

- the finite-element discretisation of the global problem, in which case the preconditioner is said to be applied on the original matrix \mathcal{A} and the Abstract Schwarz method is noted (AS/\mathcal{A}) ;
- the substructuring matrix obtained by eliminating the interior variables from \mathcal{A} , in which case the preconditioner is said to be applied on the Schur matrix (AS/S) .

Using the decompositions $\mathcal{A} = \sum_{i=1}^N \mathcal{A}_i$ or $\mathcal{S} = \sum_{i=1}^N \mathcal{S}_i$, the aS and NN preconditioners introduced in Equation (1.3) can be built on \mathcal{A} or \mathcal{S} . We also test our method on an additional AS preconditioner \mathcal{M}_{Ro} built by defining $\tilde{\mathcal{S}}_i = \mathcal{S}_i + \mathcal{I}$ or $\tilde{\mathcal{A}}_i = \mathcal{A}_i + \mathcal{I}$ in Equation (2.4), where \mathcal{I} is the identity matrix of the same size as $\tilde{\mathcal{S}}_i$ or $\tilde{\mathcal{A}}_i$, respectively.

5.2 Controlling the condition number using deflation

We proved in Section 3 that it is possible to control the condition number $\kappa(\tilde{\mathcal{M}}_{AS,D}\mathcal{S})$ of approximate AS methods through some parameters α and β . Here, we do not use any approximation and $\tilde{\mathcal{S}}_i = \mathcal{S}_i$, leading to $\omega_- = \omega_+ = 1$. In order to compare the three methods, we first choose a bound χ and then we choose α and β such that $\kappa \leq \chi$:

- for aS (resp. NN), Corollary 2 states that $\kappa \leq (1 + \alpha)N_c$ (resp. $\kappa \leq \alpha N_c$). We choose $\alpha = \chi/N_c - 1$ (resp. $\alpha = \chi/N_c$).
- for Ro (or any other AS preconditioner), Theorem 2 states that $\kappa \leq (1 + \alpha)\beta$ and we choose $\alpha = \sqrt{1/4 + \chi} - 1/2$ and $\beta = \sqrt{1/4 + \chi} + 1/2$.

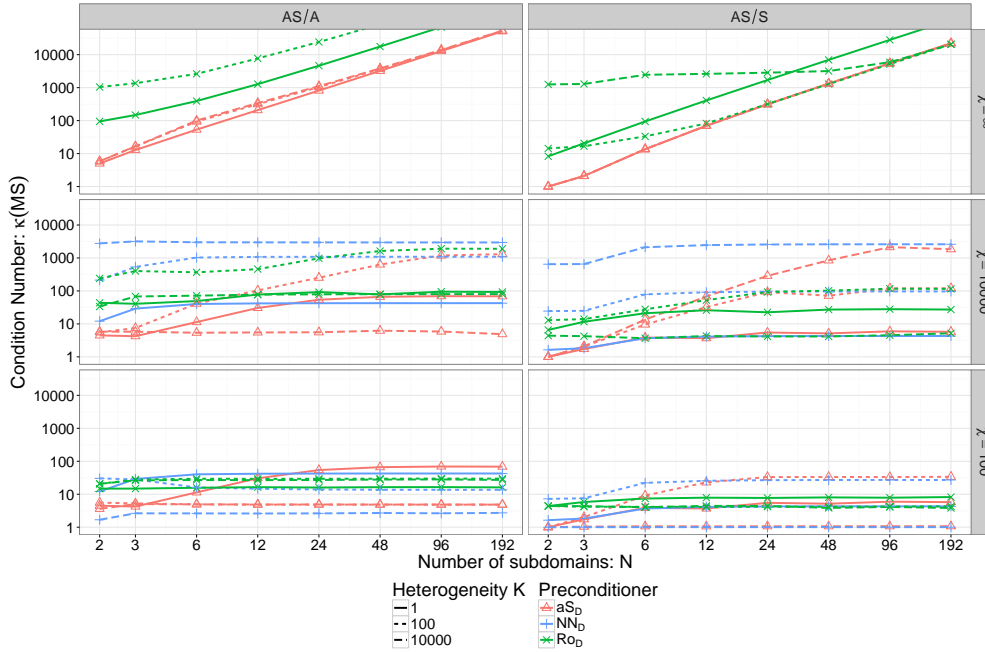


Figure 1: Controlling the condition number using deflation. Whatever the chosen target χ , we ensure that the condition number of the iterative problem $\kappa(\mathcal{MS})$ remains below χ . Each preconditioner can be applied either on the original matrix \mathcal{A} (AS/A), left, or in a substructuring context on \mathcal{S} (AS/S), right.

When we do not impose an upper bound ($\chi = \infty$), no coarse space is used. In that case, NN is singular and only results for aS and Ro are presented. We observe (Figure 1) that the condition number κ grows quadratically with the number of subdomains N and that the number of iterations to reach convergence (Figure 2) is proportional to the number of subdomains (beware of the log scales). This lack of scalability is the main motivation for using a two-level method.

We also note that without a coarse space, our aS preconditioner outperforms the Ro preconditioner, especially when the heterogeneity K is high: the aS preconditioner performs a more appropriate local solve than the very basic Ro preconditioner. As expected, the condition number is also lower when working on the Schur matrix \mathcal{S} instead of \mathcal{A} , since all the interior unknowns are solved using a direct method and do not appear anymore in the iterative solve.

When we impose an upper bound on the condition number $\chi = 10000$ or $\chi = 100$, the condition number κ effectively drops below this bound, independently of the number of subdomains N , the local preconditioner aS, NN or Ro, the heterogeneity K and the choice of solving \mathcal{A} or \mathcal{S} .

However, this control on the condition number comes at the expense of having to use a direct solve on a coarse space V_0 whose dimension can be quite high. Each subdomain computes a local

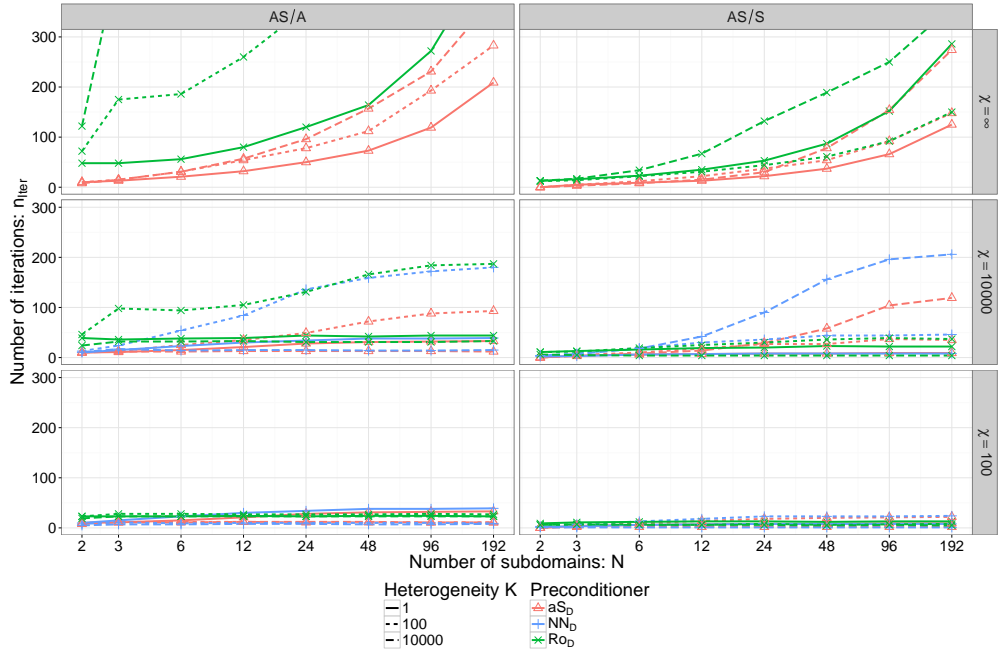


Figure 2: The number of iterations can be controlled independently of the number of subdomains N and of the heterogeneity K in the matrix.

coarse space V_i^0 of dimension $n_v^{(i)}$ (Figure 3) and the size of the global coarse space therefore grows linearly with the number of subdomains.

Since without deflation ($\chi = \infty$) the Ro preconditioner applied on the original matrix \mathcal{A} does not perform very well in the heterogeneous case, the size of the coarse space necessary to obtain a condition number below the target χ is very high (up to 87 vectors per subdomain). However, using a better local preconditioner such as aS or NN can greatly reduce the size of the coarse space, as well as working on the Schur matrix \mathcal{S} instead of \mathcal{A} .

5.3 Working with a fixed coarse space size

We showed in the previous section that we can effectively control the condition number κ of the method by building the coarse space using two parameters α and β as presented in Theorem 2. However, this can lead to an impractically large coarse space and we now consider the context where the size n_v of the local subspace in each subdomain is chosen *a priori*. Instead of choosing the coarse space by comparing the eigenvalues to a threshold, we thus keep the eigenvectors associated with the n_v smallest eigenvalues. Once the coarse space is computed, we know what threshold would have led us to keep the same number of vectors and we can get, *a posteriori*, a bound on the condition number of the method: if λ_{n_v+1} is the lowest eigenvalue corresponding to a vector not in the coarse space, Theorem 2 ensures that $\kappa(\mathcal{M}_{Ro,D} \mathcal{S}) \leq N_c(1 + 1/\lambda_{n_v+1})/\lambda_{n_v+1}$. As in Section 5.2, this bound can be improved for NN and aS preconditioners using Corollary 2 and Theorem 3:

- $\kappa(\mathcal{M}_{NN,D} \mathcal{S}) \leq N_c/\lambda_{n_v+1}$;
- $\kappa(\mathcal{M}_{aS,D} \mathcal{S}) \leq N_c(1 + 1/\lambda_{n_v+1})$;

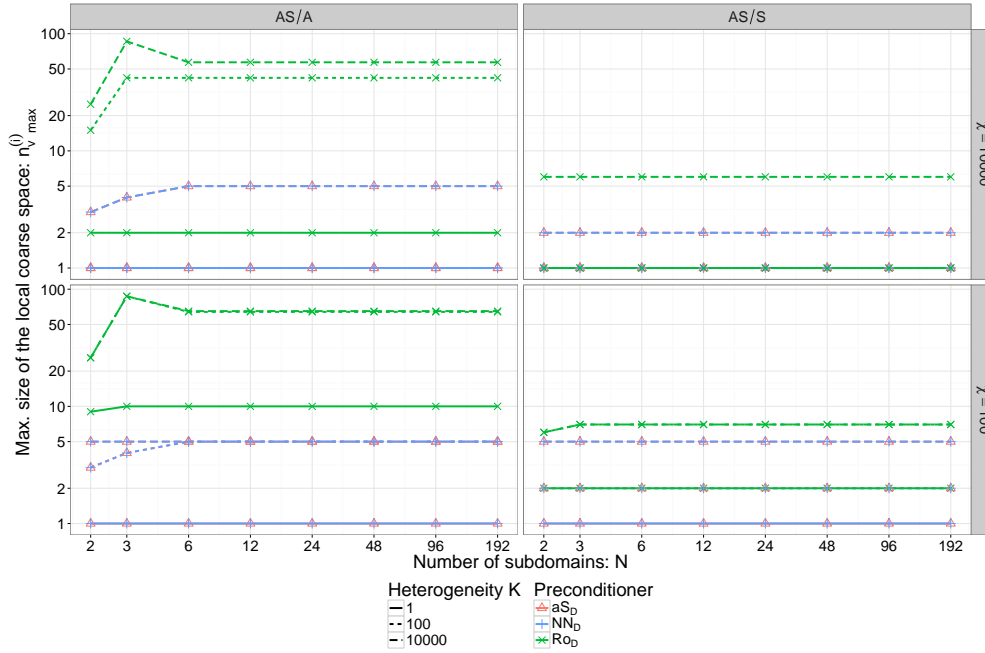


Figure 3: In most cases, just a few vectors per subdomain are enough. However, the robustness of the method can imply a large local coarse space V_i^0 in some cases.

- $\kappa(\mathcal{M}_{aS,2} \mathcal{S}) \leq (N_c + 1) [N_c + 1 + (N_c + 2)/\lambda_{n_v+1}]$.

Despite NN_D has the lowest bound of these three preconditioners and aS_2 has the largest bound, Figure 4 shows that in practice aS_2 and aS_D have similar behaviors and both perform much better than NN_D when n_v is small.

The factorization of the interiors to compute the Schur matrix \mathcal{S} introduces some new connectivity between the domains: through the adjacency graph of \mathcal{S} , a domain is not only connected to its original neighbors but to their respective neighbors too. As a result, in our example, $N_c = 3$ when working on \mathcal{A} , but $N_c = 5$ when the Schur matrix \mathcal{S} is used. Therefore, the theoretical bounds are higher with \mathcal{S} than with \mathcal{A} . However, in practice, the factorization of the interiors significantly improves the condition number performed in AS/S methods.

Although the computation of $\hat{\mathcal{S}}_i$ is more expensive in the aS case compared to NN since some communication is needed, our results show that this information exchanged between neighboring subdomains is useful since it greatly improves the performance of the preconditioner. Moreover, the ability to skip the projection step in aS_2 saves one matrix-vector product per iteration, yielding a very robust and efficient two-level preconditioner.

5.4 Empirical study on the impact of sparsification

In order to limit the memory cost of using the Schur, it is possible to approximate the dense matrix \mathcal{S}_i by a sparse matrix $\tilde{\mathcal{S}}_i$, by dropping some entries in the matrix. This process is called *sparsification*. In a very heterogeneous medium ($K \gg 1$), some entries in \mathcal{S} corresponding to couplings between unknown separated by a low-conductivity layer, are negligible. We use the symmetry-preserving strategy of dropping s_{ij} if $|s_{ij}| \leq \epsilon(s_{ii} + s_{jj})$, where ϵ is a parameter that

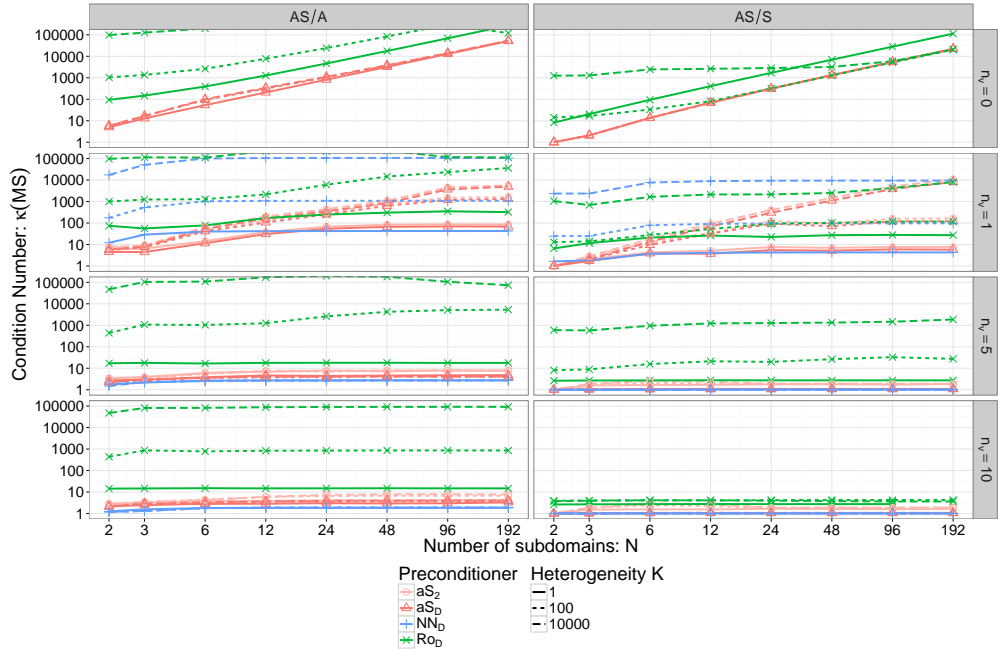


Figure 4: Fixing the size n_v of the local coarse space V_i^0 instead of thresholds α and β , we are still able to significantly reduce the condition number of the methods. The best convergence results in our study are obtained using the aS_D/S method.

controls the sparsity.

The benefits of sparsification are evaluated by assessing the proportion of non-zero elements in the Cholesky factorization $\hat{S}_i = LL^T$ of the local preconditioner. In Figure 5, we evaluate the impact of sparsification on the robustness of the method. It appears that, up to a certain level, we are still able to find a good coarse space despite having significantly reduced the memory footprint of the preconditioner. For instance, with a sparsity parameter of $\epsilon = 0.001$, although 88.8% of the entries in the factorization of the preconditioner are dropped, our coarse space with $n_v = 5$ vectors per subdomain still significantly improves the convergence.

6 Conclusion

In this paper, we have proposed a new class of AS preconditioners, so called approximate AS preconditioners. These preconditioners are fully algebraic in the sense that they do not require the element matrices or any other information apart from the SPSD subdomain matrices which naturally arise from the local assembly of the subdomain problem. This class is wide as it consists of all AS preconditioners, provided that their coarse space results of the assembly of local components that contain the kernel of some local operators (Definition 1). In particular, it generalizes the class of approximate NN preconditioners introduced in Le Tallec and Vidrascu [1998] (named *generalized NN* in the original paper). We exhibited a bound on the condition number of all approximate deflated AS preconditioners (Theorem 1). This bound depends on a generalized Rayleigh quotient and generalizes the result from Le Tallec and Vidrascu [1998] beyond the class of approximate NN methods. Applying a GenEO procedure on two well chosen

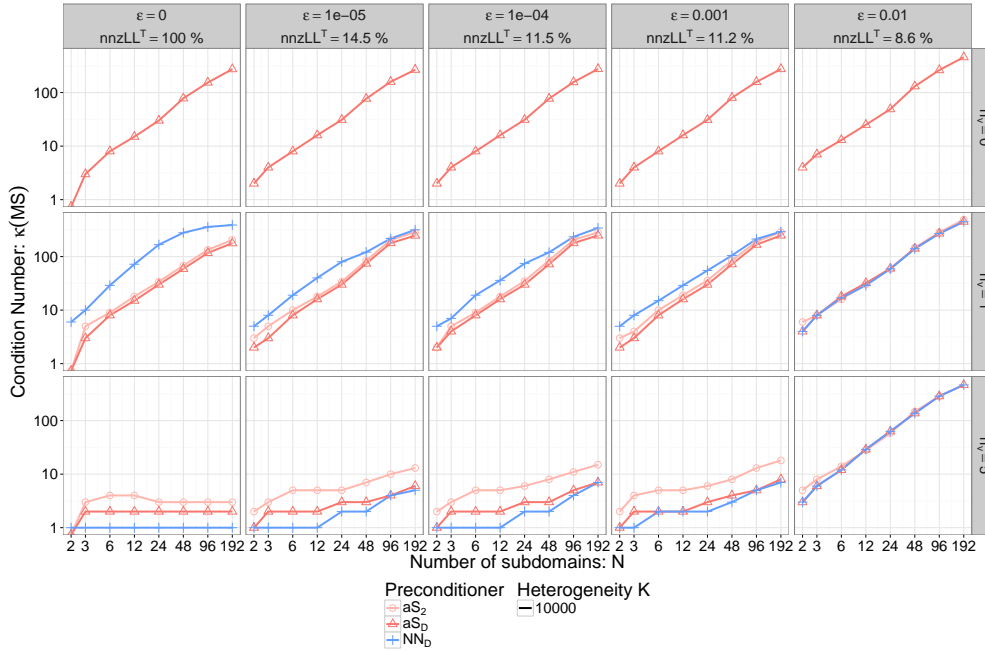


Figure 5: Up to a certain level, the sparsification does not break the robustness of the method: using a big enough coarse space ($n_v = 5$), it is possible to discard 88.8% of the entries in the factorization of the preconditioner without losing convergence.

generalized eigenproblems, we proposed a procedure to explicitly compute the coarse space while bounding this Rayleigh quotient leading to a bound on the condition number (hence on the number of iterations of PCG) independent of the number of subdomains. We also showed that a similar bound can be obtained when the coarse space is applied additively for the subclass of newly introduced approximate aS methods.

The results presented in this paper can be readily derived to retrieve the bounds previously obtained via generalized eigenproblems in the particular cases of aS (applied on the original problem, aS/\mathcal{A}) Spillane et al. [2013], NN Spillane and Rixen [2012] and optimized Robin (SORAS) Haferssas et al. [2015]. It also generalizes these results when used with approximate local solvers. Furthermore, they allowed us to define a coarse space for the aS method applied on the Schur complement (aS/S) Carvalho et al. [2001], leading to a new powerful and versatile substructuring method, for which the coarse space can be applied either with deflation or additively, and with the freedom of relying on an approximate local Schur complement. Numerical experiments illustrated these statements.

This work shall motivate the design of a high-performance parallel library for the proposed approximate abstract Schwarz method, following the design of high-performance coarse grid correction introduced in Jolivet et al. [2013]. More particularly, we plan to assess the special case of aS/S in the MaPHYs framework Giraud et al. [2008], Agullo et al. [2011]. Another challenge opened by the present study is to determine an explicit procedure to perform the approximation while achieving a given fixed bound on the condition number. Finally, it might also be interesting to study the effects of the method on the spectrum and on the empirical convergence of non symmetric test cases.

7 Acknowledgments

The authors would like to thank Nicole Spillane for proofreading an early draft of this manuscript. Experiments presented in this paper were carried out using the PLAFRIM experimental testbed, being developed under the Inria PlaFRIM development action with support from Bordeaux INP, LABRI and IMB and other entities: Conseil Régional d'Aquitaine, Université de Bordeaux and CNRS (and ANR in accordance to the programme d'investissements d'Avenir). Finally, this work was partially supported by the French research agency ANR in the framework of the DEDALES project (ANR-14-CE23-0005), in particular the PhD thesis of the third author (in the alphabetical order) was funded by this project.

References

- Emmanuel Agullo, Luc Giraud, Abdou Guermouche, and Jean Roman. Parallel hierarchical hybrid linear solvers for emerging computing platforms. *Comptes Rendus Mécanique*, 339(2–3):96–103, February 2011. ISSN 1631-0721. doi: 10.1016/j.crme.2010.11.005. URL <http://www.sciencedirect.com/science/article/pii/S1631072110002068>.
- L. M. Carvalho, L. Giraud, and G. Meurant. Local preconditioners for two-level non-overlapping domain decomposition methods. *Numerical Linear Algebra with Applications*, 8(4):207–227, 2001.
- Victorita Dolean, Pierre Jolivet, and Frédéric Nataf. *An introduction to domain decomposition methods: algorithms, theory, and parallel implementation*. Number 144 in Other titles in applied mathematics. Society for Industrial and Applied Mathematics, Philadelphia, 2015. ISBN 978-1-61197-405-8. URL <https://hal.archives-ouvertes.fr/ce1-01100932/document>.
- Luc Giraud, Azzam Haidar, and Layne T. Watson. Parallel scalability study of hybrid preconditioners in three dimensions. *Parallel Computing*, 34(6):363–379, 2008. URL <http://www.sciencedirect.com/science/article/pii/S0167819108000197>.
- Ryadh Haferssas, Pierre Jolivet, and Frédéric Nataf. A robust coarse space for optimized schwarz methods: SORAS-GenEO-2. *Comptes Rendus Mathématique*, 353(10):959–963, October 2015. ISSN 1631-073X. doi: 10.1016/j.crma.2015.07.014. URL <http://www.sciencedirect.com/science/article/pii/S1631073X15002162>.
- Pierre Jolivet, Frédéric Hecht, Frédéric Nataf, and Christophe Prud'homme. Scalable domain decomposition preconditioners for heterogeneous elliptic problems. In *Proceedings of the International Conference on High Performance Computing, Networking, Storage and Analysis, SC '13*, pages 80:1–80:11, New York, NY, USA, 2013. ACM. ISBN 978-1-4503-2378-9. doi: 10.1145/2503210.2503212. URL <http://doi.acm.org/10.1145/2503210.2503212>.
- Patrick Le Tallec and Marina Vidrascu. Generalized Neumann-Neumann preconditioners for iterative substructuring. 1998.
- Jan Mandel. Balancing domain decomposition. *Communications in Numerical Methods in Engineering*, 9(3):233–241, March 1993. ISSN 1099-0887. doi: 10.1002/cnm.1640090307. URL <http://onlinelibrary.wiley.com/doi/10.1002/cnm.1640090307/abstract>.
- T. Mathew. *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*. Springer Lecture Notes in Computational Science and Engineering. Springer, 2008.

N. Spillane, V. Dolean, P. Hauret, F. Nataf, C. Pechstein, and R. Scheichl. Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems in the overlaps. *Numerische Mathematik*, 126(4):741–770, August 2013. ISSN 0029-599X, 0945-3245. doi: 10.1007/s00211-013-0576-y. URL <http://link.springer.com/article/10.1007/s00211-013-0576-y>.

Nicole Spillane and Daniel J. Rixen. Automatic spectral coarse spaces for robust FETI and BDD algorithms. 2012. URL <http://hal.upmc.fr/hal-00756994>.

Andrea Toselli and Olof Widlund. *Domain decomposition methods: algorithms and theory*, volume 3. Springer, 2005. URL <http://link.springer.com/content/pdf/10.1007/b137868.pdf>.



**RESEARCH CENTRE
BORDEAUX – SUD-OUEST**

200 avenue de la Vielle Tour
33405 Talence Cedex

Publisher
Inria
Domaine de Voluceau - Rocquencourt
BP 105 - 78153 Le Chesnay Cedex
inria.fr

ISSN 0249-6399