

Robust preconditioners via generalized eigenproblems for hybrid sparse linear solvers

Emmanuel Agullo, Luc Giraud, Louis Poirel

► **To cite this version:**

Emmanuel Agullo, Luc Giraud, Louis Poirel. Robust preconditioners via generalized eigenproblems for hybrid sparse linear solvers. *SIAM Journal on Matrix Analysis and Applications*, Society for Industrial and Applied Mathematics, 2019, 40 (2), pp.417-439. 10.1137/17M1153765 . hal-02074474

HAL Id: hal-02074474

<https://hal.inria.fr/hal-02074474>

Submitted on 20 Mar 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.

ROBUST PRECONDITIONERS VIA GENERALIZED EIGENPROBLEMS FOR HYBRID SPARSE LINEAR SOLVERS

E. AGULLO, L. GIRAUD, L. POIREL

Abstract. The solution of large sparse linear systems is one of the most time consuming kernels in many numerical simulations. The domain decomposition community has developed many efficient and robust methods in the last decades. While many of these solvers fall into the abstract Schwarz (aS) framework, their robustness has originally 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 or improve 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. Numerical experiments illustrate these statements.

Key words. preconditioning, SPD linear systems, robust, scalable, coarse space, generalized eigenvalue, parallel hybrid (direct/iterative) solver

AMS subject classifications. 15A06, 65F08, 65F10, 15A12, 65N55

1. Introduction. Many scientific or engineering applications require at some point the solution of large sparse linear systems in parallel. Once the specific problem has been discretized, the resulting matrix equation can be solved using either an external general purpose linear solver, or a more specific solver tailored to the particular problem. With the first approach, referred to as the algebraic approach, the user can benefit with little integration effort from the developments and optimizations of black-box libraries which perform very well on modern architectures [4, 19].

On the other hand, the second approach often allows additional optimizations that further exploit additional characteristics of the underlying problem and requires a tighter integration of the solver within the application code. A widely used class of methods that fall in this latter category are domain decomposition methods (DDM) [10, 20, 28, 30, 35], which are inherently parallel and provide robust and scalable solvers for a wide range of physical problems.

In this article, we aim at combining the advantages of both these approaches. For that, while remaining as algebraic as possible, we identify some key information to be provided to the solver alongside the matrix. For symmetric positive definite (SPD) problems we show that providing the matrix in a distributed fashion, as a sum of symmetric positive semi-definite (SPSD) matrices, is enough to build a robust and scalable hybrid solver. This is a common situation when applying a finite element method over a partitioned mesh, but the methods presented in this article are not limited to this particular case: for instance, more complex discretizations such as the hybridizable discontinuous Galerkin method [7] can be used instead.

The linear system to be solved is

$$(1.1) \quad \mathcal{K}u = f,$$

46 where \mathcal{K} is a $n \times n$ sparse SPD matrix that does not need to be known explicitly. Instead, the
 47 parallel application provides \mathcal{K} to the solver as a sum $\mathcal{K} = \sum_{i=1}^N \mathcal{K}_i^{(g)}$ of N SPSD matrices $\mathcal{K}_i^{(g)}$.
 48 Even though $\mathcal{K}_i^{(g)}$ is of size $n \times n$, in practical applications it has only n_i non-zero rows (and columns),
 49 meaning that this matrix represents the interaction of only a subset of the unknowns from the global
 50 problem. We define the *global domain* $\Omega = \{1, \dots, n\}$ as the set of row (or column) indices in \mathcal{K} ,
 51 and the *subdomain* $\Omega_i = \{\omega_1^{(i)}, \omega_2^{(i)}, \dots, \omega_{n_i}^{(i)}\}$ as the set of indices of the non-zero rows and columns
 52 in $\mathcal{K}_i^{(g)}$ (Ω_i is the set of vertices in the adjacency graph of $\mathcal{K}_i^{(g)}$). We introduce the $n_i \times n$ canonical
 53 restriction matrix \mathcal{R}_{Ω_i} from Ω to Ω_i , such that for any vector $u = (u_1, \dots, u_n) \in \mathbb{R}^n$, $\mathcal{R}_{\Omega_i} u$ is
 54 the vector $(u_{\omega_1^{(i)}}, \dots, u_{\omega_{n_i}^{(i)}}) \in \mathbb{R}^{n_i}$. Then, we define the $n_i \times n_i$ SPSD matrix $\mathcal{K}_i = \mathcal{R}_{\Omega_i} \mathcal{K}_i^{(g)} \mathcal{R}_{\Omega_i}^T$,
 55 referred to as the *local matrix* of subdomain Ω_i , leading to

$$56 \quad (1.2) \quad \mathcal{K} = \sum_{i=1}^N \mathcal{R}_{\Omega_i}^T \mathcal{K}_i \mathcal{R}_{\Omega_i}.$$

58 The unknowns in any subdomain Ω_i can be partitioned into an interior $\mathcal{I}_i = \{\omega \in \Omega_i \text{ s.t. } \forall j \neq$
 59 $i, \omega \notin \Omega_j\}$ and an interface $\Gamma_i = \{\omega \in \Omega_i \text{ s.t. } \exists j \neq i \omega \in \Omega_j\} = \Omega_i \setminus \mathcal{I}_i$. If an unknown
 60 $\omega \in \Omega_i$ appears in at least one other subdomain, then $\omega \in \Gamma_i$, otherwise $\omega \in \mathcal{I}_i$. This yields a
 61 partition of the global domain $\Omega = \{1, \dots, n\} = \mathcal{I}_1 \cup \dots \cup \mathcal{I}_N \cup \Gamma$ where $\Gamma = \Gamma_1 \cup \dots \cup \Gamma_N$ is the
 62 global interface.

63 Then, eliminating in parallel the interior unknowns following for instance [Section 2, 25] the
 64 original system (1.1) reduces to a Schur problem defined on the interface Γ

$$65 \quad (1.3) \quad \mathcal{S} u_\Gamma = \tilde{f}_\Gamma, \quad \mathcal{S} = \sum_{i=1}^N \mathcal{R}_{\Gamma_i}^T \mathcal{S}_i \mathcal{R}_{\Gamma_i},$$

67 where the global Schur matrix \mathcal{S} is SPD and the local Schur matrices \mathcal{S}_i are SPSD. Using the
 68 classical index notation for referring to sub-blocks of matrices and vectors, we have $\mathcal{S} = \mathcal{K}_{\Gamma\Gamma} -$
 69 $\sum_{i=1}^N \mathcal{K}_{\Gamma\mathcal{I}_i} \mathcal{K}_{\mathcal{I}_i\mathcal{I}_i}^{-1} \mathcal{K}_{\mathcal{I}_i\Gamma}$, $\tilde{f}_\Gamma = f_\Gamma - \sum_{i=1}^N \mathcal{K}_{\Gamma\mathcal{I}_i} \mathcal{K}_{\mathcal{I}_i\mathcal{I}_i}^{-1} f_{\mathcal{I}_i}$ and $\mathcal{S}_i = \mathcal{K}_{\Gamma_i\Gamma_i} - \mathcal{K}_{\Gamma_i\mathcal{I}_i} \mathcal{K}_{\mathcal{I}_i\mathcal{I}_i}^{-1} \mathcal{K}_{\mathcal{I}_i\Gamma_i}$. From the
 70 interface solution u_Γ , the solution in \mathcal{I}_i can be computed as $u_{\mathcal{I}_i} = \mathcal{K}_{\mathcal{I}_i\mathcal{I}_i}^{-1} (f_{\mathcal{I}_i} - \mathcal{K}_{\mathcal{I}_i\mathcal{I}_i}^{-1} \mathcal{K}_{\mathcal{I}_i\Gamma} u_\Gamma)$.

71 Algebraically, the problems (1.1) and (1.3) are very similar; their only difference is that even
 72 when \mathcal{K}_i is sparse, \mathcal{S}_i is in general a dense matrix (as soon as \mathcal{K}_i is irreducible). Although eliminating
 73 the interior unknowns is often associated with specific DDM such as BDD [8, 26] or FETI [13], it is in
 74 fact an optional step in the solution of Problem (1.1) and most domain decomposition methods can
 75 be applied either directly on \mathcal{K} or, after eliminating the interior unknowns, on \mathcal{S} . This elimination
 76 step may take time and consume memory, but it allows us to reduce the size and the condition
 77 number of the linear system (\mathcal{S}) to be solved [5, 27], making it a useful optional preprocessing.
 78 Since the theory presented in sections 2 and 3 can be applied to solve either the original problem
 79 in (1.1) or the reduced Schur problem in (1.3), we write them in a general form as

$$80 \quad (1.4) \quad \mathcal{A} x = b, \quad \mathcal{A} = \sum_{i=1}^N \mathcal{R}_i^T \mathcal{A}_i \mathcal{R}_i,$$

82 where the global SPD matrix \mathcal{A} , the local SPSD matrices \mathcal{A}_i , and the restriction matrices \mathcal{R}_i can
 83 represent \mathcal{K} , \mathcal{K}_i and \mathcal{R}_{Ω_i} or \mathcal{S} , \mathcal{S}_i and \mathcal{R}_{Γ_i} when solving (1.1) or (1.3), respectively. When needed,
 84 a specific method M will be noted M/\mathcal{K} or M/\mathcal{S} to specify on which problem this method is

85 applied. In both cases, \mathcal{A} is SPD, assuming that the \mathcal{A}_i are assigned to different computing units,
 86 Problem (1.4) can be solved in parallel using the preconditioned conjugate gradient method (PCG).

87 A good preconditioner \mathcal{M} for (1.4) should have the two following properties: (1) \mathcal{M} is SPD
 88 and *close* to \mathcal{A}^{-1} , in the sense that the condition number $\kappa(\mathcal{M}\mathcal{A})$ should be as small as possible;
 89 (2) it is easy to compute $\mathcal{M}u$ for any vector u (at least much easier than $\mathcal{A}^{-1}u$). DDM are often
 90 used to build such preconditioners of the form

$$91 \quad (1.5) \quad \mathcal{M}_{aS} = \sum_{i=1}^N \mathcal{R}_i^T \widehat{\mathcal{A}}_i^\dagger \mathcal{R}_i$$

92 where $\widehat{\mathcal{A}}_i$ is a local problem associated with \mathcal{A} on subdomain i , and \dagger represents a pseudo-inverse.

93 These preconditioners have been studied for a long time using the abstract Schwarz (aS) theory
 94 (see, e.g., [10, 35] for recent overviews). Two particular cases of preconditioners that fit this
 95 description are the Neumann-Neumann (NN) preconditioner [26], with $\widehat{\mathcal{A}}_i = D_i^{-1} \mathcal{A}_i D_i^{-1}$, and the
 96 Additive Schwarz (AS) preconditioner, with $\widehat{\mathcal{A}}_i = \mathcal{R}_i \mathcal{A} \mathcal{R}_i^T$

$$98 \quad (1.6) \quad \mathcal{M}_{NN} = \sum_{i=1}^N \mathcal{R}_i^T D_i \mathcal{A}_i^\dagger D_i \mathcal{R}_i, \quad \mathcal{M}_{AS} = \sum_{i=1}^N \mathcal{R}_i^T (\mathcal{R}_i \mathcal{A} \mathcal{R}_i^T)^{-1} \mathcal{R}_i,$$

99 where $(D_i)_{i=1}^N$ is a partition of unity such that $\sum_{i=1}^N \mathcal{R}_i^T D_i \mathcal{R}_i = I_n$ and I_n is the $n \times n$ identity
 100 matrix. These two preconditioners are of particular importance, but any other SPSD matrix can
 101 be used as the local preconditioner $\widehat{\mathcal{A}}_i$ in (1.5).
 102

103 Unless $\widehat{\mathcal{A}}_i$ perfectly mimics the global action of \mathcal{A} in subdomain Ω_i , $\kappa(\mathcal{M}_{aS}\mathcal{A})$ may significantly
 104 increase with the number N of subdomains, leading to a non scalable numerical method.

105 Furthermore, if $\widehat{\mathcal{A}}_i$ is singular, the pseudo-inverse is only defined up to an element in its null-
 106 space $\ker(\widehat{\mathcal{A}}_i)$. To solve these two problems, a coarse space V_0 such that $\mathcal{R}_i^T \ker(\widehat{\mathcal{A}}_i) \subset V_0$
 107 can be introduced, leading to the deflated aS preconditioner

$$108 \quad (1.7) \quad \mathcal{M}_{aS,D} = V_0(V_0^T \mathcal{A} V_0)^\dagger V_0^T + (I_n - \mathcal{P}_0) \left(\sum_{i=1}^N \mathcal{R}_i^T \widehat{\mathcal{A}}_i^\dagger \mathcal{R}_i \right) (I_n - \mathcal{P}_0)^T$$

109 where $\mathcal{P}_0 = V_0(V_0^T \mathcal{A} V_0)^\dagger V_0^T \mathcal{A}$ is the \mathcal{A} -orthogonal projection onto V_0 . A simpler additive two-
 110 level preconditioner can also be obtained by just adding the coarse component to the one-level
 111 preconditioner
 112

$$113 \quad (1.8) \quad \mathcal{M}_{aS,2} = V_0(V_0^T \mathcal{A} V_0)^\dagger V_0^T + \sum_{i=1}^N \mathcal{R}_i^T \widehat{\mathcal{A}}_i^\dagger \mathcal{R}_i.$$

114 While previous works had proposed bounds on the condition number $\kappa(\mathcal{M}\mathcal{A})$ on particular
 115 numerical cases, often relying on analytical assumptions, Le Tallec and Vidrascu [25] derived an
 116 algebraic bound for a new class of preconditioners, relying on the generalized Rayleigh quotient of
 117 two local matrices. These preconditioners are called *generalized NN* in the original article; however,
 118 because the generalization consists of handling an approximate matrix, we will instead refer to them
 119 as *approximate NN* preconditioners in the present article. The approximation is not related to the
 120 use of inexact solvers to compute the preconditioner, but to the use of an approximation matrix
 121

122 $\tilde{\mathcal{A}}$ instead of \mathcal{A} in the construction of the preconditioner. The approximate NN preconditioner is
123 in fact an exact algebraic NN preconditioner for $\tilde{\mathcal{A}}$. Then, this approximate preconditioner is used
124 to accelerate the convergence of PCG applied on the exact matrix \mathcal{A} , guaranteeing a convergence
125 towards the actual solution of Equation (1.4).

126 This class of approximate NN preconditioners generalizes classical NN but does not cover
127 the whole aS class of preconditioners. Note, for instance, that AS cannot be expressed as a NN
128 preconditioner. The first contribution (Section 2) of this article is to extend the result from [25]
129 by using a generic local preconditioner and cover a broader range of aS methods, which we name
130 *approximate deflated aS* methods and consist of all deflated aS methods whose coarse grid consists of
131 the assembly of local components that contain the kernel of some local operators (that are formally
132 introduced below, in Definition 1). Interestingly, the bound we exhibit (Theorem 2) highlights the
133 key position of NN and AS among other local preconditioners in the Schwarz framework: they
134 provide two bounds on the spectrum of the preconditioned operator, and the convergence of any
135 aS local preconditioner can be evaluated by comparing it to these two well-known methods.

136 This bound depends on generalized Rayleigh quotients which are traditionally estimated using
137 functional analysis. Alternatively, we propose to control these Rayleigh quotients algebraically by
138 building the coarse space using eigenvectors of well chosen generalized eigenproblems (Theorem 10).
139 For that, we follow the Generalized Eigenvalue in the Overlap (GenEO) procedure [33]. This
140 second contribution (Section 3) results in an explicit procedure for building a robust coarse space
141 of any approximate deflated aS method leading to a bound on the condition number (hence on
142 the number of iterations of PCG) independent of the number of subdomains. This result can be
143 readily applied to retrieve or improve the bounds previously obtained via generalized eigenproblems
144 in the particular cases of AS/ \mathcal{K} [33], NN/ \mathcal{S} [34] and optimized Robin (SORAS/ \mathcal{K}) [18]. It also
145 generalizes these results to the approximate case. The idea of building a coarse space by solving
146 local eigenproblems in each subdomain was introduced in [15, 29]; it was successfully applied for
147 other DDM such as FETI-DP [12] or BDDC [9] in [22, 23, 24].

148 The third contribution (Section 4) of this paper is that the application of the considered coarse
149 grid correction in an additive fashion is robust in the approximate AS case (although it is not robust
150 for aS methods in general). The bound we obtain (Theorem 12) can be applied for retrieving the
151 bound obtained in [33], when the coarse correction is applied additively to the AS method on the
152 original matrix (AS/ \mathcal{K}). When working on the Schur matrix (AS/ \mathcal{S}) [6], the bound is still valid
153 and leads, as commented in [15], to a smaller coarse space compared to AS/ \mathcal{K} .

154 Numerical experiments illustrate our discussion in Section 5. A high performance implementa-
155 tion of the coarse grid correction of one particular, consistently robust method (AS/ \mathcal{S}) has further-
156 more been implemented in the high-performance MaPHyS¹ hybrid (direct/iterative) sparse linear
157 solver [2, 3] to eventually assess its performance on a modern parallel computer (Section 5.5) and
158 make this scalable method available to the scientific community.

159 The paper is organized as follows. Section 2 introduces a new class of approximate (deflated)
160 aS preconditioners and provides a bound on their condition number, which depends on generalized
161 Rayleigh quotients. Applying the GenEO procedure on two well chosen generalized eigenproblems,
162 Section 3 proposes a procedure to explicitly compute the coarse space while bounding these Rayleigh
163 quotients leading to a bound on the condition number (hence on the number of iterations of PCG)
164 independent of the number of subdomains. Section 4 shows that a similar result (and procedure)
165 can be obtained when the coarse grid correction is additively applied, in the case of approximate

¹See <https://gitlab.inria.fr/solverstack/maphys/>

166 AS problems. Numerical experiments illustrate our discussion in Section 5 before concluding in
 167 Section 6.

168 **2. Approximate abstract Schwarz preconditioners.** In this section, we first define a class
 169 of approximate aS preconditioners, which combine a local preconditioner $\widehat{\mathcal{A}}_i$, an approximate matrix
 170 $\widetilde{\mathcal{A}}$ and a coarse space V_0 in Section 2.1. We then provide a bound on the condition number of this
 171 class of methods in Section 2.2, whose proof is provided in Section 2.3.

172 **2.1. Context.**

173 DEFINITION 1 (Approximate abstract Schwarz preconditioner $\widetilde{\mathcal{M}}_{aS,D}$).

174 *In order to build such a preconditioner for Problem (1.4), we need the three following ingredients:*

- 175 1. a set of symmetric positive semi-definite (SPSD) local preconditioners $\widehat{\mathcal{A}}_i$,
 176 2. an approximation $\widetilde{\mathcal{A}}$ of \mathcal{A} such that

177 (2.1)
$$\exists (\widetilde{\mathcal{A}}_i)_{i=1}^N, \quad \widetilde{\mathcal{A}} = \sum_{i=1}^N \mathcal{R}_i^T \widetilde{\mathcal{A}}_i \mathcal{R}_i \text{ and } \widetilde{\mathcal{A}}_i \text{ is SPSPD,}$$

178 (2.2)
$$\exists \omega_-, \omega_+ > 0, \quad \forall v \in V \quad \omega_- v^T \mathcal{A} v \leq v^T \widetilde{\mathcal{A}} v \leq \omega_+ v^T \mathcal{A} v,$$

- 180 3. and a coarse space V_0 such that

181 (2.3)
$$\exists (V_i^0)_{i=1}^N, \quad V_0 = \sum_{i=1}^N \mathcal{R}_i^T V_i^0 \quad \text{with} \quad \ker(\widehat{\mathcal{A}}_i) + \ker(\widetilde{\mathcal{A}}_i^{(NN)}) \subset V_i^0,$$

182 where $\widetilde{\mathcal{A}}_i^{(NN)} = D_i^{-1} \widetilde{\mathcal{A}}_i D_i^{-1}$.

184 We can then define a coarse matrix $\widetilde{\mathcal{A}}_0 = V_0^T \widetilde{\mathcal{A}} V_0$, a coarse projection $\widetilde{\mathcal{P}}_0 = V_0 \widetilde{\mathcal{A}}_0^\dagger V_0^T \widetilde{\mathcal{A}}$ and the
 185 approximate aS preconditioner is then defined as

186 (2.4)
$$\widetilde{\mathcal{M}}_{aS,D} = V_0 \widetilde{\mathcal{A}}_0^\dagger V_0^T + (I_n - \widetilde{\mathcal{P}}_0) \left(\sum_{i=1}^N \mathcal{R}_i^T \widehat{\mathcal{A}}_i^\dagger \mathcal{R}_i \right) (I_n - \widetilde{\mathcal{P}}_0)^T.$$

188 Note that the matrix $\widetilde{\mathcal{A}}_i^{(NN)}$ introduced in (2.3) is the local matrix in the approximate NN pre-
 189 conditioner $\widetilde{\mathcal{M}}_{NN,D}$ with the algebraic decomposition from (2.1). The matrices D_i can be any
 190 partition of unity as in (1.6). $\widetilde{\mathcal{A}}_i^{(NN)}$ is a scaled version of the local matrix $\widetilde{\mathcal{A}}_i$ in the approximation
 191 $\widetilde{\mathcal{A}}$ of \mathcal{A} .

192 When no approximation is used, after a suitable initialization, $\widetilde{\mathcal{M}}_{aS,D}$ can be replaced by
 193 $(I_n - \widetilde{\mathcal{P}}_0) \left(\sum_{i=1}^N \mathcal{R}_i^T \widehat{\mathcal{A}}_i^\dagger \mathcal{R}_i \right)$ in the PCG iterations, as noted in [26].

194 **2.2. Convergence result for $\widetilde{\mathcal{M}}_{aS,D}$.** In each subdomain, we note $N_i = \#\{j \neq i, \mathcal{R}_i \widetilde{\mathcal{A}} \mathcal{R}_j^T \neq$
 195 $0\}$ the number of neighbors through the connectivity graph of $\widetilde{\mathcal{A}}$. We also define two local subspaces
 196 \widehat{V}_i^\perp and \widetilde{V}_i^\perp as the orthogonal spaces of V_i^0 for the inner products inferred by $\widehat{\mathcal{A}}_i$ in $\text{range}(\widehat{\mathcal{A}}_i)$ and
 197 $\widetilde{\mathcal{A}}_i^{(NN)}$ in $\text{range}(\widetilde{\mathcal{A}}_i^{(NN)})$ respectively. Then,

198 (2.5)
$$\text{range}(\mathcal{R}_i) = \widehat{V}_i^\perp \oplus V_i^0 = \widetilde{V}_i^\perp \oplus V_i^0,$$

199 (2.6)
$$\forall u \in V_i^0, \forall v \in \widehat{V}_i^\perp, \forall w \in \widetilde{V}_i^\perp \quad u^T \widehat{\mathcal{A}}_i v = u^T \widetilde{\mathcal{A}}_i^{(NN)} w = 0.$$

201 Finally, for any SPSP 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
 202 SPD, we note it $\|u\|_{\mathcal{B}}$.

203 THEOREM 2 (Convergence result for approximate aS).

204 The condition number of the preconditioned matrix $\widetilde{\mathcal{M}}_{aS,D}\mathcal{A}$ is bounded by

$$205 \quad \kappa(\widetilde{\mathcal{M}}_{aS,D}\mathcal{A}) \leq \frac{\omega_+}{\omega_-} \left(1 + \max_{1 \leq i \leq N} \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widehat{\mathcal{A}}_i}^2}{|v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2} \right) \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2}{|v|_{\widehat{\mathcal{A}}_i}^2} \right),$$

206 where $\widetilde{\mathcal{A}}_i^{(NN)} = D_i^{-1} \widehat{\mathcal{A}}_i D_i^{-1}$ and $\widetilde{\mathcal{A}}_i^{(AS)} = \mathcal{R}_i \widetilde{\mathcal{A}} \mathcal{R}_i^T$.

207 We see three factors in this bound:

- 208 • The first one, with ω_+ and ω_- , controls the quality of the approximation $\widetilde{\mathcal{A}}$. If no approx-
 209 imation is used, then $\widetilde{\mathcal{A}} = \mathcal{A}$ and $\omega_- = \omega_+ = 1$.
- 210 • The second one is a generalized Rayleigh quotient between the local preconditioner $\widehat{\mathcal{A}}_i$ and
 211 the approximate NN preconditioner $\widetilde{\mathcal{A}}_i^{(NN)} = D_i^{-1} \widehat{\mathcal{A}}_i D_i^{-1}$ defined in [25].
- 212 • The last one is a generalized Rayleigh quotient between the local preconditioner $\widehat{\mathcal{A}}_i$ and an
 213 approximate AS preconditioner $\widetilde{\mathcal{A}}_i^{(AS)} = \mathcal{R}_i \widetilde{\mathcal{A}} \mathcal{R}_i^T$.

214 As for $\widetilde{\mathcal{A}}_i^{(NN)}$ above with NN, $\widetilde{\mathcal{A}}_i^{(AS)} = \mathcal{R}_i \widetilde{\mathcal{A}} \mathcal{R}_i^T$ is an algebraic generalization of the local matrix
 215 in the AS preconditioner in Equation (1.6), built upon the approximation $\widetilde{\mathcal{A}}$ instead of \mathcal{A} .

216 *Proof.* The proof of Theorem 2 is a direct consequence of lemmas 6 and 8 in Section 2.3, using
 217 the definition of

$$218 \quad \kappa(\widetilde{\mathcal{M}}_{aS,D}\mathcal{A}) = \frac{\lambda_{\max}(\widetilde{\mathcal{M}}_{aS,D}\mathcal{A})}{\lambda_{\min}(\widetilde{\mathcal{M}}_{aS,D}\mathcal{A})}. \quad \square$$

219 COROLLARY 3 (Convergence results for approximate AS and approximate NN).

220 We define the approximate AS and NN preconditioners $\widetilde{\mathcal{M}}_{AS,D}$ and $\widetilde{\mathcal{M}}_{NN,D}$ by replacing $\widehat{\mathcal{A}}_i$
 221 with $\widetilde{\mathcal{A}}_i^{(AS)}$ or $\widetilde{\mathcal{A}}_i^{(NN)}$ respectively in Equation (2.4). We also define $N_c = \max_{1 \leq i \leq N} (N_i + 1)$.

222 Then, the condition numbers of $\widetilde{\mathcal{M}}_{NN,D}\mathcal{A}$ and $\widetilde{\mathcal{M}}_{AS,D}\mathcal{A}$ are bounded by

$$223 \quad \kappa(\widetilde{\mathcal{M}}_{AS,D}\mathcal{A}) \leq \frac{\omega_+}{\omega_-} \left(1 + \max_{1 \leq i \leq N} \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2}{|v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2} \right) N_c,$$

$$224 \quad \kappa(\widetilde{\mathcal{M}}_{NN,D}\mathcal{A}) \leq \frac{\omega_+}{\omega_-} \max \left(1, \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2}{|v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2} \right) N_c.$$

225 *Proof.* The proof of Corollary 3 is a consequence of lemmas 6 and 7 for AS, and lemmas 5 and 8
 226 for NN. □

227 Note that the bound for $\widetilde{\mathcal{M}}_{NN,D}$ in Corollary 3 is the same as in [Theorem 1, 25]. This bound
 228 is tighter than the bound obtained by setting $\widehat{\mathcal{A}}_i = \widetilde{\mathcal{A}}_i^{(NN)}$ in Theorem 2; this comes from the fact
 229 that the bound in Lemma 5 is also tighter than its generalization in Lemma 6.

230 The similarity of the bounds for AS and NN in Corollary 3 shows that the convergence of
 231 these two methods are governed by the same quantity $\sup_{v \in \widetilde{V}_i^\perp} |v|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2 / |v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2$. As a result, with

235 the same coarse space, we expect the AS/S method [6] to show the same convergence behavior
 236 as the BDD method (NN/S) [26] or its dual counterpart FETI [13]. Although AS require more
 237 communication than NN (each subdomain i has to send the matrix block $\mathcal{R}_j \mathcal{R}_i^T \tilde{\mathcal{A}}_i \mathcal{R}_i \mathcal{R}_j^T$ to each
 238 neighbor j) to setup the preconditioner, one advantage of using AS over NN is that the local
 239 preconditioner $\tilde{\mathcal{A}}_i^{(NN)}$ is often singular in some domains while $\tilde{\mathcal{A}}_i^{(AS)}$ remains SPD, and $\tilde{\mathcal{A}}_i^{(AS)-1} u_i$
 240 is easier and faster to compute than $\tilde{\mathcal{A}}_i^{(NN)\dagger} u_i$.

241 **2.3. Proof of Theorem 2.** To estimate the condition number of $\widetilde{\mathcal{M}}_{a,S,D}\mathcal{A}$, we need to bound
 242 the spectrum of this operator from above and below. The lower bound is a consequence of the
 243 Stable Decomposition Lemma as stated in [35].

244 LEMMA 4 (Stable decomposition lemma).

245 *If there exists a constant C_0 , local matrices \mathcal{B}_i and extension operators \mathcal{I}_i , such that $\ker(\mathcal{B}_i) \subset$
 246 $\ker(\mathcal{I}_i)$ and every $u \in V$ admits a decomposition*

$$247 \quad u = \sum_{i=0}^N \mathcal{I}_i u_i, \quad \{u_i \in V_i, 0 \leq i \leq N\} \quad \text{that satisfies} \quad \sum_{i=0}^N |u_i|_{\mathcal{B}_i}^2 \leq C_0^2 \|u\|_{\mathcal{A}}^2.$$

249 *Then*

$$250 \quad \lambda_{\min}(\mathcal{M}\mathcal{A}) \geq C_0^{-2}, \quad \text{where} \quad \mathcal{M} = \sum_{i=0}^N \mathcal{I}_i \mathcal{B}_i^\dagger \mathcal{I}_i^T.$$

252 *Proof.* see, e.g., Lemma 2.5 in [35]. □

253 Then, although it is not directly used in the proof of Theorem 2, we first expose in Lemma 5
 254 a lower bound for the spectrum of NN ($\widehat{\mathcal{A}}_i = \tilde{\mathcal{A}}_i^{(NN)}$) as it provides a good insight on the reason
 255 behind the Rayleigh quotients in the bound presented in Lemma 6 for the general case.

256 LEMMA 5 (Lower bound for the approximate Neumann-Neumann preconditioner).

$$257 \quad \text{Let } \widetilde{\mathcal{M}}_{NN,D} = V_0 \tilde{\mathcal{A}}_0^\dagger V_0^T + (I_n - \tilde{\mathcal{P}}_0) \left(\sum_{i=1}^N \mathcal{R}_i^T \tilde{\mathcal{A}}_i^{(NN)\dagger} \mathcal{R}_i \right) (I_n - \tilde{\mathcal{P}}_0)^T.$$

258 *Then,*

$$259 \quad \lambda_{\min}(\widetilde{\mathcal{M}}_{NN,D}\mathcal{A}) \geq \frac{1}{\omega_+}.$$

261 *Proof.* This is a consequence of Lemma 4 (see Theorem 1 in [25]). □

262 If, instead of $\tilde{\mathcal{A}}_i^{(NN)}$, another local preconditioner $\widehat{\mathcal{A}}_i$ is used, there is no change on the bound
 263 if we restrict the operators to the coarse space V_0 since the application of the local preconditioner
 264 is preceded and followed by projections $(I_n - \tilde{\mathcal{P}}_0)$ and $(I_n - \tilde{\mathcal{P}}_0)^T$. However, in the orthogonal of
 265 the coarse space, the bound has to change and reflect the difference between $\tilde{\mathcal{A}}_i^{(NN)}$ and $\widehat{\mathcal{A}}_i$. As is
 266 proved in Lemma 6, the lower bound on the spectrum of $\widetilde{\mathcal{M}}_{a,S,D}\mathcal{A}$ can be deduced from the bound
 267 for $\widetilde{\mathcal{M}}_{NN,D}\mathcal{A}$ in Lemma 5 by adding a correction related to the generalized Rayleigh quotient
 268 between $\tilde{\mathcal{A}}_i^{(NN)}$ and $\widehat{\mathcal{A}}_i$ in the orthogonal of the coarse space.

269 LEMMA 6 (Lower bound for the approximate abstract Schwarz preconditioner).

$$270 \quad \lambda_{\min}(\widetilde{\mathcal{M}}_{aS,D}\mathcal{A}) \geq \frac{1}{\omega_+} \left(1 + \max_{1 \leq i \leq N} \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i}^2}{|v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2} \right)^{-1}.$$

271
 272 *Proof.* We want to split u into a sum of local contributions, while being able to uniformly
 273 control the $\widetilde{\mathcal{A}}_i$ -norm of these contributions u_i with the global \mathcal{A} -norm of u to apply Lemma 4. For
 274 any u and $i \geq 1$, we decompose $D_i \mathcal{R}_i u = u_i^0 + u_i^\perp$ where $u_i^0 \in V_i^0$ and $u_i^\perp \in \widetilde{V}_i^\perp$. We then define
 275 $u_0 = (V_0^T \widetilde{\mathcal{A}} V_0)^\dagger V_0^T \mathcal{A} u$ such that $V_0 u_0 = \widetilde{\mathcal{P}}_0 u$. We can use the facts that $\sum_{i=1}^N \mathcal{R}_i^T D_i \mathcal{R}_i = I_n$ and
 276 $\sum_{i=0}^N \mathcal{R}_i^T u_i^0 \in V_0 \subset \ker(I_n - \widetilde{\mathcal{P}}_0)$ to obtain the decomposition

$$277 \quad \begin{aligned} u &= \widetilde{\mathcal{P}}_0 u + (I_n - \widetilde{\mathcal{P}}_0) u = V_0 u_0 + (I_n - \widetilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_i^T D_i \mathcal{R}_i u \\ 278 \quad &= V_0 u_0 + (I_n - \widetilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_i^T (u_i^0 + u_i^\perp) = V_0 u_0 + (I_n - \widetilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_i^T u_i^\perp \\ 279 \quad &= \sum_{i=0}^N \mathcal{I}_i u_i \quad \text{where } \mathcal{I}_0 = V_0, \quad \mathcal{I}_i = (I_n - \widetilde{\mathcal{P}}_0) \mathcal{R}_i \quad \text{and} \quad u_i = u_i^\perp. \end{aligned}$$

281 Since $\widetilde{\mathcal{P}}_0$ is a $\widetilde{\mathcal{A}}$ -orthogonal projection, it holds that:

$$282 \quad (2.7) \quad |u_0|_{\widetilde{\mathcal{A}}_0}^2 = |u_0|_{V_0^T \widetilde{\mathcal{A}} V_0}^2 = |V_0 u_0|_{\widetilde{\mathcal{A}}}^2 = |\widetilde{\mathcal{P}}_0 u|_{\widetilde{\mathcal{A}}}^2 \leq |u|_{\widetilde{\mathcal{A}}}^2$$

284 Let

$$285 \quad C = \max_{1 \leq i \leq N} \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i}^2}{|v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2} = \max_{1 \leq i \leq N} \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i}^2}{|v|_{D_i^{-1} \widetilde{\mathcal{A}}_i D_i^{-1}}^2}.$$

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

$$287 \quad \begin{aligned} |u_i^\perp|_{\widetilde{\mathcal{A}}_i}^2 &\leq C |u_i^\perp|_{D_i^{-1} \widetilde{\mathcal{A}}_i D_i^{-1}}^2 \leq C |u_i^\perp + u_i^0|_{D_i^{-1} \widetilde{\mathcal{A}}_i D_i^{-1}}^2 = C |\mathcal{R}_i u|_{\widetilde{\mathcal{A}}_i}^2, \\ 288 \quad (2.8) \quad \sum_{i=1}^N |u_i^\perp|_{\widetilde{\mathcal{A}}_i}^2 &\leq C \sum_{i=1}^N |\mathcal{R}_i u|_{\widetilde{\mathcal{A}}_i}^2 = C |u|_{\sum_{i=1}^N \mathcal{R}_i^T \widetilde{\mathcal{A}}_i \mathcal{R}_i}^2 = C |u|_{\widetilde{\mathcal{A}}}^2, \\ 289 \quad |u_0|_{\widetilde{\mathcal{A}}_0}^2 + \sum_{i=1}^N |u_i^\perp|_{\widetilde{\mathcal{A}}_i}^2 &\leq (1 + C) |u|_{\widetilde{\mathcal{A}}}^2 \leq \omega_+ (1 + C) |u|_{\widetilde{\mathcal{A}}}^2, \end{aligned}$$

291 and the local norms are controlled by the global norm. Then, applying Lemma 4, we get

$$292 \quad \lambda_{\min}(\widetilde{\mathcal{M}}_{aS,D}\mathcal{A}) \geq \frac{1}{\omega_+} \left(1 + \max_{1 \leq i \leq N} \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i}^2}{|v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2} \right)^{-1}. \quad \square$$

294 Now that we proved a lower bound for the spectrum of $\widetilde{\mathcal{M}}_{NN,D}\mathcal{A}$, we will prove an upper
 295 bound in Lemma 8. We first recall a classic upper bound for AS preconditioners in Lemma 7 since
 296 it explains the origin of the Rayleigh quotient in the bound for the general case.

297 LEMMA 7 (Upper bound for the approximate Additive Schwarz preconditioner).

298 Let $\widetilde{\mathcal{M}}_{AS,D} = V_0 \widetilde{\mathcal{A}}_0^\dagger V_0^T + (I_n - \widetilde{\mathcal{P}}_0) \left(\sum_{i=1}^N \mathcal{R}_i^T \widetilde{\mathcal{A}}_i^{(AS)-1} \mathcal{R}_i \right) (I_n - \widetilde{\mathcal{P}}_0)^T$.

299 Then,

$$300 \lambda_{\max}(\widetilde{\mathcal{M}}_{AS,D}\mathcal{A}) \leq \frac{1}{\omega_-} \max_{1 \leq i \leq N} (N_i + 1).$$

301 *Proof.* This lemma is a particular case of Lemma 8 which is proven below. \square

302 LEMMA 8 (Upper bound for the approximate abstract Schwarz preconditioner).

$$303 \lambda_{\max}(\widetilde{\mathcal{M}}_{aS,D}\mathcal{A}) \leq \frac{1}{\omega_-} \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \widehat{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2}{|v|_{\widetilde{\mathcal{A}}_i}^2} \right).$$

305 *Proof.* First, let us remark that

$$306 \widetilde{\mathcal{M}}_{aS,D}\widetilde{\mathcal{A}}u = V_0 \widetilde{\mathcal{A}}_0^\dagger V_0^T \widetilde{\mathcal{A}}u + (I_n - \widetilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_i^T \widetilde{\mathcal{A}}_i^\dagger \mathcal{R}_i (I_n - \widetilde{\mathcal{P}}_0)^T \widetilde{\mathcal{A}}u = u_0 + (I_n - \widetilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_i^T u_i$$

308 where $u_0 = \widetilde{\mathcal{P}}_0 u$ and u_i is the orthogonal projection of $\widetilde{\mathcal{A}}_i^\dagger \mathcal{R}_i (I_n - \widetilde{\mathcal{P}}_0)^T \widetilde{\mathcal{A}}u$ onto $\text{range}(\widehat{\mathcal{A}}_i)$ along $\ker(\widehat{\mathcal{A}}_i) \subset V_i^0 \subset \ker[(I_n - \widetilde{\mathcal{P}}_0)\mathcal{R}_i^T]$.

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

$$311 \frac{1}{2} u_i^T \widehat{\mathcal{A}}_i V_i^0 = u^T \widetilde{\mathcal{A}} (I_n - \widetilde{\mathcal{P}}_0) \mathcal{R}_i^T \widehat{\mathcal{A}}_i^\dagger \widehat{\mathcal{A}}_i V_i^0 = u^T \widetilde{\mathcal{A}} (I_n - \widetilde{\mathcal{P}}_0) \mathcal{R}_i^T V_i^0 = 0.$$

313 Then,

$$314 |\widetilde{\mathcal{M}}_{aS,D}\widetilde{\mathcal{A}}u|_{\widetilde{\mathcal{A}}}^2 = |u_0|_{\widetilde{\mathcal{A}}}^2 + |(I_n - \widetilde{\mathcal{P}}_0) \sum_{i=1}^N \mathcal{R}_i^T u_i|_{\widetilde{\mathcal{A}}}^2 \leq |u_0|_{\widetilde{\mathcal{A}}}^2 + \left| \sum_{i=1}^N \mathcal{R}_i^T u_i \right|_{\widetilde{\mathcal{A}}}^2$$

$$315 \leq |u_0|_{\widetilde{\mathcal{A}}}^2 + \sum_{i=1}^N (N_i + 1) |\mathcal{R}_i^T u_i|_{\widetilde{\mathcal{A}}}^2 = |u_0|_{\widetilde{\mathcal{A}}}^2 + \sum_{i=1}^N (N_i + 1) |u_i|_{\mathcal{R}_i \widetilde{\mathcal{A}} \mathcal{R}_i^T}^2$$

317 where we used the fact that

$$318 0 \leq \sum_{\substack{1 \leq i, j \leq N \\ \mathcal{R}_i^T \widetilde{\mathcal{A}} \mathcal{R}_j \neq 0}} |\mathcal{R}_i^T u_i - \mathcal{R}_j^T u_j|_{\widetilde{\mathcal{A}}}^2 = 2 \left(\sum_{\substack{1 \leq i, j \leq N \\ \mathcal{R}_i^T \widetilde{\mathcal{A}} \mathcal{R}_j \neq 0}} |\mathcal{R}_i^T u_i|_{\widetilde{\mathcal{A}}}^2 - \sum_{\substack{1 \leq i, j \leq N \\ \mathcal{R}_i^T \widetilde{\mathcal{A}} \mathcal{R}_j \neq 0}} u_i^T \mathcal{R}_i \widetilde{\mathcal{A}} \mathcal{R}_j^T u_j \right)$$

$$319 (2.9) \leq 2 \left(\sum_{i=1}^N (N_i + 1) |\mathcal{R}_i^T u_i|_{\widetilde{\mathcal{A}}}^2 - \left| \sum_{i=1}^N \mathcal{R}_i^T u_i \right|_{\widetilde{\mathcal{A}}}^2 \right).$$

321 Let us define

$$322 C = \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \widehat{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2}{|v|_{\widetilde{\mathcal{A}}_i}^2} \right) = \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \widehat{V}_i^\perp} \frac{|v|_{\mathcal{R}_i \widetilde{\mathcal{A}} \mathcal{R}_i^T}^2}{|v|_{\widetilde{\mathcal{A}}_i}^2} \right).$$

323 We can now write

$$\begin{aligned}
324 \quad |\widetilde{\mathcal{M}}_{a,S,D}\widetilde{\mathcal{A}}u|_{\widetilde{\mathcal{A}}}^2 &\leq C|u_0|_{\widetilde{\mathcal{A}}}^2 + C \sum_{i=1}^N |u_i|_{\widehat{\mathcal{A}}_i}^2 = Cu^T \widetilde{\mathcal{P}}_0^T \widetilde{\mathcal{A}}u_0 + C \sum_{i=1}^N u^T \widetilde{\mathcal{A}}(I_n - \widetilde{\mathcal{P}}_0)\mathcal{R}_i^T \widehat{\mathcal{A}}_i^\dagger \widehat{\mathcal{A}}_i u_i \\
325 \quad &= Cu^T \widetilde{\mathcal{A}}\widetilde{\mathcal{M}}_{a,S,D}\widetilde{\mathcal{A}}u \leq C|u|_{\widetilde{\mathcal{A}}}| \widetilde{\mathcal{M}}_{a,S,D}\widetilde{\mathcal{A}}u|_{\widetilde{\mathcal{A}}} \\
326 \quad |\widetilde{\mathcal{M}}_{a,S,D}\widetilde{\mathcal{A}}u|_{\widetilde{\mathcal{A}}} &\leq C|u|_{\widetilde{\mathcal{A}}},
\end{aligned}$$

328 and use the same strategy as in [25] to obtain our result:

$$\begin{aligned}
329 \quad \lambda_{\max}(\widetilde{\mathcal{M}}_{a,S,D}\mathcal{A}) &= \max_{v \in V} \frac{|v|_{\widetilde{\mathcal{A}}}^2}{|v|_{\widetilde{\mathcal{M}}_{a,S,D}^{-1}}^2} \leq \max_{v \in V} \frac{1}{\omega_-} \frac{|v|_{\widetilde{\mathcal{A}}}^2}{|v|_{\widetilde{\mathcal{M}}_{a,S,D}^{-1}}^2} \leq \max_{v \in V} \frac{1}{\omega_-} \frac{|\widetilde{\mathcal{M}}_{a,S,D}\widetilde{\mathcal{A}}v|_{\widetilde{\mathcal{A}}}^2}{|v|_{\widetilde{\mathcal{A}}}^2} \leq \frac{C}{\omega_-}, \\
330 \quad \lambda_{\max}(\widetilde{\mathcal{M}}_{a,S,D}\mathcal{A}) &\leq \frac{1}{\omega_-} \max \left(1, \max_{1 \leq i \leq N} (N_i + 1) \sup_{v \in \widehat{V}_i^\perp} \frac{|v|_{\widehat{\mathcal{A}}_i^{(AS)}}^2}{|v|_{\widehat{\mathcal{A}}_i}^2} \right). \quad \square \\
331
\end{aligned}$$

332 **3. Building the coarse space via generalized eigenproblems.** The bound in Theorem 2
333 has originally been estimated through functional analysis after a coarse space has been chosen. A
334 more algebraic approach is to build the coarse space V_0 by solving a generalized eigenproblem in
335 each subdomain in order to control the Rayleigh quotient as proposed by [33, 34] for AS/ \mathcal{K} and
336 NN/ \mathcal{S} , respectively. This approach has also been successfully applied to other aS variants such
337 as the SORAS method [18], in which case two eigenproblems are needed. The case where the
338 correction is applied additively as in [11, 15, 33] for AS is treated in Section 4.

339 The connection between the GenEO method and Theorem 2 comes from the following lemma:

340 LEMMA 9 (Bound on the Rayleigh quotient).

341 Let \mathcal{B} be a SPSP matrix, \mathcal{C} a SPD matrix and $\eta > 0$ be a parameter.

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

$$343 \quad \text{then } \sup_{u \in V_\eta^{\perp\mathcal{B}}} \frac{|u|_{\mathcal{C}}^2}{|u|_{\mathcal{B}}^2} \leq \frac{1}{\eta}.$$

344 *Proof.* Since \mathcal{C} is SPD, the generalized eigenproblem $\mathcal{B}p = \lambda\mathcal{C}p$ has solutions (λ_k, p_k) with
345 $p_k^T \mathcal{C}p_l = \delta_{kl}$ and $p_k^T \mathcal{B}p_l = \lambda_k \delta_{kl}$.

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

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

$$349 \quad \frac{|u|_{\mathcal{C}}^2}{|u|_{\mathcal{B}}^2} = \frac{\sum_{\lambda_k > \eta} \alpha_k^2}{\sum_{\lambda_k > \eta} \lambda_k \alpha_k^2} \leq \frac{1}{\eta}. \quad \square \\
350$$

351 Following the GenEO methodology, we propose to build the coarse space V_0 by solving two gen-
352 eralized eigenproblems to control the condition number of approximate aS preconditioners through
353 two parameters $\alpha > 0$ and $\beta \geq 1$.

354 THEOREM 10 (Condition number of aS preconditioners). *If $\widehat{\mathcal{A}}_i$ is SPD and the coarse space*
 355 *is defined as $V_0 = \sum_{i=1}^N \mathcal{R}_i^T V_i^0$ with*

$$356 \quad V_i^0 = \text{span}(\{p_k^i, \widetilde{\mathcal{A}}_i^{(NN)} p_k^i = \lambda_k^i \widehat{\mathcal{A}}_i p_k^i, \lambda_k^i \leq \alpha^{-1}\}$$

$$357 \quad \cup \{p_k^i, \widehat{\mathcal{A}}_i p_k^i = \lambda_k^i \widetilde{\mathcal{A}}_i^{(AS)} p_k^i, \lambda_k^i \leq (N_i + 1)\beta^{-1}\})$$

359 *then, we can bound the condition number*

$$360 \quad \kappa(\widetilde{\mathcal{M}}_{aS,D}\mathcal{A}) \leq \frac{\omega_+}{\omega_-} (1 + \alpha) \beta.$$

361 *Proof.* Using Lemma 9 and the definition of \widetilde{V}_i^\perp and \widehat{V}_i^\perp in 2.2, we can bound the Rayleigh
 362 quotients

$$363 \quad \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widehat{\mathcal{A}}_i}^2}{|v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2} \leq \alpha, \quad \sup_{v \in \widehat{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2}{|v|_{\widehat{\mathcal{A}}_i}^2} \leq \frac{\beta}{N_i + 1}.$$

365 Replacing these bounds in Theorem 2 gives the result. □

366 COROLLARY 11. *In the NN or AS cases, for any $\alpha \geq 1$, we can define*

$$367 \quad V_i^0 = \text{span}(\{p_k^i, \widetilde{\mathcal{A}}_i^{(NN)} p_k^i = \lambda_k^i \widetilde{\mathcal{A}}_i^{(AS)} p_k^i, \lambda_k^i \leq \alpha^{-1}\}).$$

369 *Then, Corollary 3 and Lemma 9 give*

$$370 \quad \kappa(\widetilde{\mathcal{M}}_{AS,D}\mathcal{A}) \leq \frac{\omega_+}{\omega_-} (1 + \alpha) N_c, \quad \kappa(\widetilde{\mathcal{M}}_{NN,D}\mathcal{A}) \leq \frac{\omega_+}{\omega_-} \alpha N_c.$$

372 If $\alpha^{-1} = \min_{\lambda_k^i \neq 0} (\lambda_k^i)$, then $V_i^0 = \ker(\widetilde{\mathcal{A}}_i^{(NN)}) = D_i \ker(\widetilde{\mathcal{A}}_i)$ and the resulting coarse space for NN
 373 is exactly the same as in the BDD algorithm.

374 With small variations in the generalized eigenproblems considered, Theorem 10 and Corollary 11
 375 retrieve or improve previous GenEO results and generalize them to the approximate case: AS/ \mathcal{K}
 376 [32, 33], NN/ \mathcal{S} [34] and SORAS [18].

377 4. Additive Coarse Correction.

378 **4.1. Context.** The preconditioner $\widetilde{\mathcal{M}}_{aS,D}$ separates the part of the solution that is in V_0
 379 (on which a direct coarse solve is performed through $\widetilde{\mathcal{A}}_0^\dagger$), from its $\widetilde{\mathcal{A}}$ -orthogonal part (on which
 380 the local preconditioner $\mathcal{M}_{aS} = \sum_{i=1}^N \mathcal{R}_i^T \widehat{\mathcal{A}}_i^\dagger \mathcal{R}_i$ is used to accelerate convergence). Eigenvalues
 381 or Rayleigh quotients λ corresponding to vectors in the coarse space V_0 are shifted to 1 by the
 382 coarse solve, and to 0 by the projection steps $(I_n - \widetilde{\mathcal{P}}_0)$ and $(I_n - \widetilde{\mathcal{P}}_0)^T$, so the overall effect of the
 383 deflated preconditioner is to shift them to 1 exactly. If we skip these projection steps, we obtain an
 384 approximate additive two-level preconditioner $\widetilde{\mathcal{M}}_{aS,2}$ similar to $\mathcal{M}_{aS,2}$ presented in Equation (1.8).
 385 In this case, without the projection steps eigenvalues are shifted to $1 + \lambda$. As a result, this coarse
 386 correction applied on big eigenvalues only makes them bigger, thus hampering convergence. This
 387 additive coarse correction can only be effective to tackle the lower part of the spectrum since small
 388 eigenvalues $\lambda \ll 1$ are shifted to $1 + \lambda \approx 1$.

389 The one-level AS method has already an upper bound on the spectrum (see Lemma 7), and
 390 only the lower bound needs to be recovered, making it an ideal candidate for an additive coarse
 391 correction. In this section, we show that in the approximate AS case, when $\widehat{\mathcal{A}}_i = \widetilde{\mathcal{A}}_i^{(AS)} = \mathcal{R}_i \widetilde{\mathcal{A}} \mathcal{R}_i^T$,
 392 the projection steps can be removed without losing robustness. Namely, we still have a bound for the
 393 condition number of the additive two-level AS method independent of the number of subdomains.

394 **THEOREM 12** (Condition number of the 2-level approximate AS preconditioner).

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

396 Then, we can bound the condition number

$$397 \quad \kappa(\mathcal{M}_{AS,2}\mathcal{A}) \leq \frac{\omega_+}{\omega_-} \left[N_c + 1 + (N_c + 2) \max_{1 \leq i \leq N} \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2}{|v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2} \right] (N_c + 1).$$

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

$$400 \quad V_i^0 = \text{span}(\{p_k^i, \widetilde{\mathcal{A}}_i^{(NN)} p_k^i = \lambda_k^i \widetilde{\mathcal{A}}_i^{(AS)} p_k^i, \lambda_k^i \leq \alpha^{-1}\}),$$

402 it holds that

$$403 \quad \kappa(\mathcal{M}_{AS,2}\mathcal{A}) \leq \frac{\omega_+}{\omega_-} [N_c + 1 + \alpha(N_c + 2)] (N_c + 1).$$

405 Theorem 12 generalizes [Theorem 4.40, 33] to the approximate case, while improving the
 406 bound.

407 A spectral coarse space composed of eigenvectors of a generalized eigenproblem was earlier
 408 proposed in [11, 15]. In those studies, the authors also discuss the analytical and numerical interest
 409 of using AS,2/S instead of the more traditional AS,2/K to reduce the size of the coarse space. In
 410 comparison, our method is more algebraic in the sense that it does not need a stable interpolation
 411 operator, nor the mass matrix.

412 *Proof.* If we apply Lemma 7 without a coarse space and consider V_0 as another subdomain in
 413 the decomposition, we get

$$414 \quad \lambda_{\max}(\mathcal{M}_{AS,2}\mathcal{A}) \leq \frac{1}{\omega_-} (N_c + 1).$$

415 The lower bound is a consequence of Lemma 4. We define $u_i^0 \in V_i^0$ and $u_i^\perp \in \widetilde{V}_i^\perp$ such that
 416 $D_i \mathcal{R}_i u = u_i^0 + u_i^\perp$ as in the proof of Lemma 6. We now introduce u_0 such that $V_0 u_0 = \sum_{i=1}^N \mathcal{R}_i^T u_i^0$,
 417 and $u = V_0 u_0 + \sum_{i=1}^N \mathcal{R}_i^T u_i^\perp$.

418 We get from Equation (2.8) that

$$419 \quad \sum_{i=1}^N |u_i^\perp|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2 = \sum_{i=1}^N |u_i^\perp|_{\widetilde{\mathcal{A}}_i}^2 \leq C |u|_{\widetilde{\mathcal{A}}}^2 \quad \text{with} \quad C = \max_{1 \leq i \leq N} \sup_{v \in \widetilde{V}_i^\perp} \frac{|v|_{\widetilde{\mathcal{A}}_i^{(AS)}}^2}{|v|_{\widetilde{\mathcal{A}}_i^{(NN)}}^2} = \frac{|v|_{\mathcal{R}_i \widetilde{\mathcal{A}} \mathcal{R}_i^T}^2}{|v|_{D_i^{-1} \widetilde{\mathcal{A}} D_i^{-1}}^2}.$$

420

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

$$\begin{aligned}
422 \quad |u_0|_{\tilde{\mathcal{A}}}^2 &= |u - \sum_{i=1}^N \mathcal{R}_i^T u_i^\perp|_{\tilde{\mathcal{A}}}^2 \leq (N_c + 1) \left(|u|_{\tilde{\mathcal{A}}}^2 + \sum_{i=1}^N |\mathcal{R}_i^T u_i^\perp|_{\tilde{\mathcal{A}}}^2 \right) \\
423 \quad &= (N_c + 1) \left(|u|_{\tilde{\mathcal{A}}}^2 + \sum_{i=1}^N |u_i^\perp|_{\mathcal{R}_i \tilde{\mathcal{A}} \mathcal{R}_i^T}^2 \right) \leq (N_c + 1)(1 + C) |u|_{\tilde{\mathcal{A}}}^2 \\
424 \quad |u_0|_{\tilde{\mathcal{A}}}^2 + \sum_{i=1}^N |u_i^\perp|_{\mathcal{R}_i \tilde{\mathcal{A}} \mathcal{R}_i^T}^2 &\leq [N_c + 1 + (N_c + 2)C] |u|_{\tilde{\mathcal{A}}}^2 \leq \omega_+ [N_c + 1 + (N_c + 2)C] |u|_{\tilde{\mathcal{A}}}^2. \\
425
\end{aligned}$$

426 We then use Lemma 4 with $\mathcal{I}_0 = V_0$, $\mathcal{I}_i = \mathcal{R}_i^T$ and $\mathcal{B}_i = \mathcal{I}_i^T \tilde{\mathcal{A}} \mathcal{I}_i$ to get the bound

$$427 \quad \lambda_{\min}(\mathcal{M}_{AS,2\mathcal{A}}) \geq \frac{1}{\omega_+} \left[N_c + 1 + (N_c + 2) \max_{1 \leq i \leq N} \sup_{v \in \tilde{V}_i^\perp} \frac{|v|_{\mathcal{R}_i \tilde{\mathcal{A}} \mathcal{R}_i^T}^2}{|v|_{\tilde{\mathcal{A}}^{(NN)}}^2} \right]^{-1}.$$

429 We can then conclude with Lemma 9. □

430 5. Numerical experiments.

431 **5.1. Experimental setup.** The methods introduced in sections 2, 3 and 4 are tested on a
432 problem similar to what is presented in [33]. We use the Finite Element Method (FEM) with Q1
433 elements to solve a heterogeneous diffusion equation $\nabla \cdot (k \nabla u) = 1$ in a 3D stratified medium. The
434 domain $[0, N] \times [0, 6] \times [0, 1]$ is discretized on a regular mesh of $(5N + 1) \times 31 \times 6$ nodes. The
435 domain is divided into N identical subdomains along the first axis. Along the second axis, it is
436 divided into 10 layers (of $5N \times 3 \times 5$ elements each) of alternating conductivity $k = 1$ and $k = K$
437 (K is a heterogeneity parameter). A Dirichlet boundary condition is applied on the left of the
438 domain ($x = 0$), a Neumann condition on every other boundary. Using a FEM discretization on
439 each subdomain gives rise naturally to a set of local SPSD matrices and a global matrix that is
440 SPD. The geometry and 1D partitioning of this test case are chosen so as to emphasize the effects of
441 using a coarse grid correction: indeed, without a coarse correction, the number of iterations grows
442 as $O(N^{1/d})$ where d is the dimension of the partitioning. Using a 3D partitioning of the global
443 domain, one would need more than 7M subdomains (192^3) to illustrate the same effect as in the
444 experiments presented here with a 1D partitioning and 192 subdomains. The layered structure of
445 the domain is introduced to deteriorate the condition number of the local subproblems. Since all
446 subdomains (except the first and last ones) are identical, the bound on the condition number of the
447 method in Theorem 2 is independent of N if at least the kernels of $\tilde{\mathcal{A}}_i^{(NN)}$ and $\hat{\mathcal{A}}_i$ are included in V_i^0 ;
448 a coarse space that only includes these kernels (as in BDD for instance) thus yields a method that
449 can be considered as robust in this regard, while being considerably simpler to compute than the
450 coarse space proposed in this article. However, the condition number still depends on the inverse
451 of the smallest eigenvalues not included in the coarse space, which can be quite close to 0 if the
452 local problems are ill-conditioned (*i.e.*, if K is big). As a result, the condition number, although
453 independent of N , can still be too large for the iterative solver to converge in a reasonable number
454 of iterations. Building the coarse space by solving the generalized eigenproblems as proposed in
455 Section 3 yields a more robust method in the sense that the condition number of the method can
456 be controlled independently of both N , K , and the particular choice of a local preconditioner. We
457 consider three aS methods: the AS and NN preconditioners introduced in Equation (1.6) and a

458 Shifted (Sh) preconditioner whose local matrix is obtained by shifting the diagonal of $\tilde{\mathcal{A}}_i$ by 1 to
 459 remove its potential singularity: $\tilde{\mathcal{M}}_{Sh} = \sum_{i=1}^N \mathcal{R}_i^T (\tilde{\mathcal{A}}_i + I_{n_i})^\dagger \mathcal{R}_i$ where I_{n_i} is the identity matrix
 460 of same size as \mathcal{A}_i . If built on the Schur matrix, $\tilde{\mathcal{M}}_{Sh}$ is a (non-optimized) Robin preconditioner.
 461 The optimization of the Robin condition as proposed in [16] is not considered here as it is out of
 462 the scope of this paper. It is introduced as an example of a more generic aS preconditioner than
 463 AS and NN; as such, two generalized eigenproblems need to be solved to compute the coarse space
 464 for Sh as opposed to only one for AS and NN. Each of these method is assessed with $\mathcal{A} = \mathcal{K}$ or
 465 $\mathcal{A} = \mathcal{S}$. Equation (1.4) can therefore either result from:

- 466 • the FEM discretization (1.1) of the global problem, in which case the preconditioner is said
 467 to be applied on the original matrix \mathcal{K} and the abstract Schwarz method is noted aS/ \mathcal{K} ;
- 468 • or the structuring system (1.3) obtained by eliminating the interior variables from Equa-
 469 tion (1.1), in which case the preconditioner is said to be applied on the Schur matrix \mathcal{S} and
 470 the method is noted aS/ \mathcal{S} .

471 We study the numerical behaviour of these methods under the constraint of a bounded condition
 472 number or an imposed coarse space size in sections 5.2 and 5.3, respectively. We then study
 473 the approximate case with an empirical approach in Section 5.4, using a so-called *sparsification*
 474 technique. Our numerical results overall confirm [11, 15] regarding the numerical interest of using
 475 AS,2/ \mathcal{S} instead of the more traditional AS,2/ \mathcal{K} method to reduce the size of the coarse space.
 476 Section 5.5 eventually illustrates the parallel behavior of that promising variant.

477 The partition of unity D_i is computed using the diagonal values of \mathcal{A}_i . The condition numbers
 478 of the preconditioned matrices are estimated using the eigenvalues of the tridiagonal Lanczos ma-
 479 trix computed during the PCG iterations (see, e.g., [14]). The stopping criterion is based on the
 480 normwise backward error $\|b - \mathcal{A}x_k\|/\|b\| \leq 10^{-6}$.

481 **5.2. Imposing an *a priori* bound on the condition number.** We proved in Section 3 that
 482 it is possible to control the condition number $\kappa(\tilde{\mathcal{M}}_{aS,D}\mathcal{A})$ of aS methods through some parameters
 483 α and β . For now, we do not use any approximation (whose effects are the object of Section 5.4),
 484 hence $\tilde{\mathcal{A}}_i = \mathcal{A}_i$ and $\omega_- = \omega_+ = 1$. In order to compare the three methods, we first choose a bound
 485 χ and then we choose α and β such that $\kappa \leq \chi$:

- 486 • for AS (resp. NN), Corollary 11 states that $\kappa \leq (1 + \alpha)N_c$ (resp. $\kappa \leq \alpha N_c$). We choose
 487 $\alpha = \chi/N_c - 1$ (resp. $\alpha = \chi/N_c$).
- 488 • for Sh (or any other aS preconditioner), Theorem 10 states that $\kappa \leq (1 + \alpha)\beta$ and we choose
 489 $\alpha = \sqrt{1/4 + \chi} - 1/2$ and $\beta = \sqrt{1/4 + \chi} + 1/2$.

490 When we do not impose an upper bound ($\chi = \infty$), no coarse space is used and results are
 491 presented only for AS and Sh. We observe (Figure 1) that the condition number κ grows quadrat-
 492 ically with the number of subdomains N and that the number of iterations to reach convergence
 493 (Figure 2) is proportional to the number of subdomains (note the log scale for the x -axis). This
 494 lack of scalability is the main motivation for using a two-level method. We also note that, with-
 495 out a coarse space, our AS preconditioner outperforms the Sh preconditioner, especially when the
 496 heterogeneity K is high: the AS preconditioner performs a more appropriate local solve than the
 497 very basic Sh preconditioner. As expected, the condition number is also lower when working on the
 498 Schur matrix \mathcal{S} instead of \mathcal{K} , since all the interior unknowns are solved using a direct method and
 499 do not appear anymore in the iterative process.

500 When we impose an upper bound on the condition number ($\chi = 10,000$ or $\chi = 100$), we observe
 501 that the condition number κ does indeed drop below the prescribed bound χ , independently of the
 502 number of subdomains N , the local preconditioner AS, NN or Sh, the heterogeneity K and the

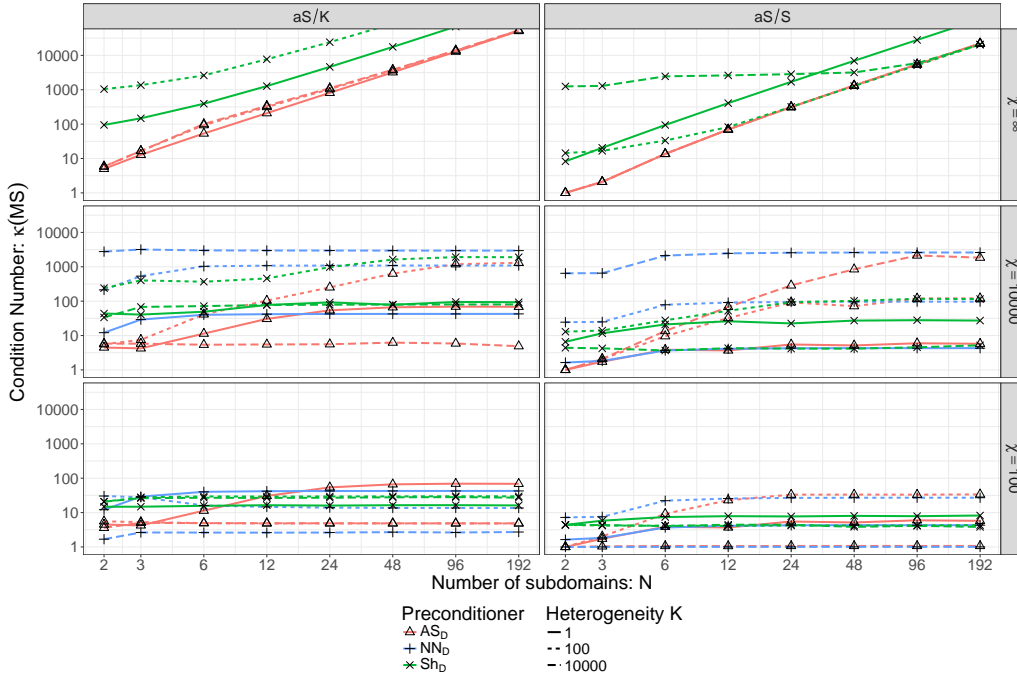


FIG. 1. Imposing an *a priori* bound χ on the condition number using deflation. Whatever the chosen target χ , we ensure that the condition number of the iterative problem $\kappa(\mathcal{M}\mathcal{A})$ remains below χ . Each preconditioner (AS, NN, Sh) can be applied either on the original matrix \mathcal{K} (aS/ \mathcal{K}), left, or in a substructuring context on \mathcal{S} (aS/ \mathcal{S}), right.

503 choice of operating on \mathcal{K} or \mathcal{S} . However, this *a priori* control on the condition number comes at
 504 the expense of having to use a direct solve on a coarse space V_0 whose dimension can be quite
 505 large. Each subdomain computes a local coarse space V_i^0 of dimension $n_v^{(i)}$ (Figure 3) and the size
 506 of the global coarse space therefore grows linearly with the number of subdomains. Since without
 507 deflation ($\chi = \infty$) the Sh preconditioner applied to the original matrix \mathcal{K} does not perform very
 508 well in the heterogeneous case, the size of the coarse space necessary to obtain a condition number
 509 below the target χ is very large (up to 87 vectors per subdomain). However, using a better local
 510 preconditioner such as AS or NN can greatly reduce the size of the coarse space, as well as working
 511 on the Schur matrix \mathcal{S} instead of \mathcal{K} .

512 **5.3. Imposing an *a priori* coarse space size.** We showed in the previous section that we
 513 can effectively control the condition number κ of the method by building the coarse space using
 514 two parameters α and β as presented in Theorem 10. However, this can lead to an impractically
 515 large coarse space and we now consider the context where the size n_v of the local subspace in each
 516 subdomain is chosen *a priori*. Instead of choosing the coarse space by comparing the eigenvalues
 517 to a threshold, we thus keep the eigenvectors associated with the n_v smallest eigenvalues. Once the
 518 coarse space is computed, we know what threshold would have led us to keep the same number of
 519 vectors and we can get, *a posteriori*, a bound on the condition number of the method: if λ_{n_v+1} is
 520 the lowest eigenvalue corresponding to a vector not in the coarse space, Theorem 10 ensures that

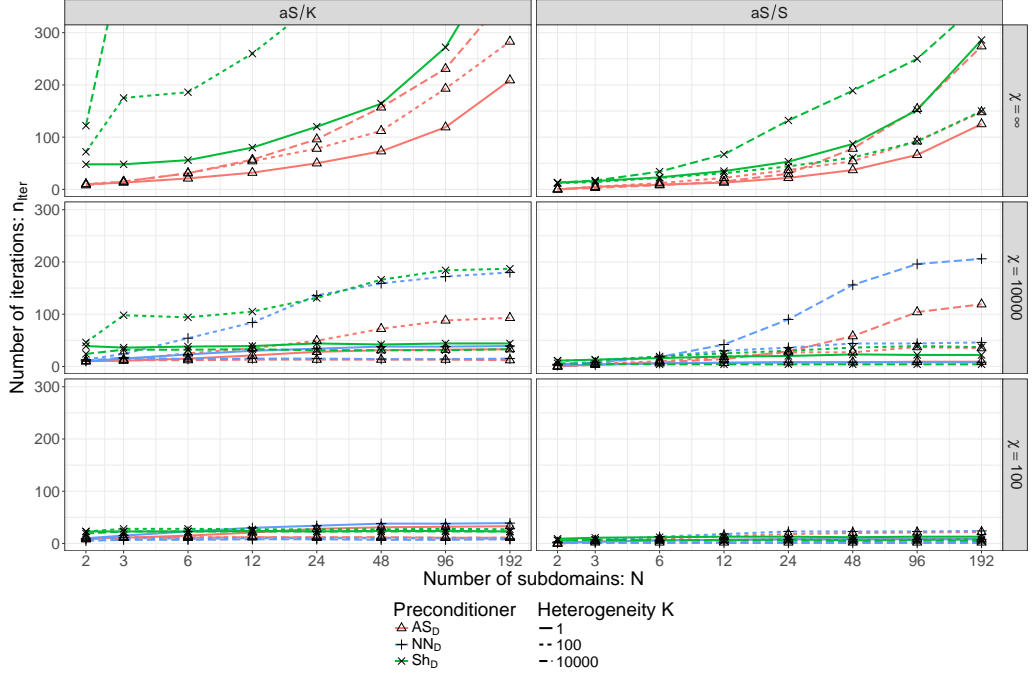


FIG. 2. Number of iterations when imposing an a priori bound χ on the condition number.

521 $\kappa(\mathcal{M}_{\text{Sh},D} \mathcal{A}) \leq N_c(1 + 1/\lambda_{n_v+1})/\lambda_{n_v+1}$. As in Section 5.2, this bound can be improved for NN and
 522 AS preconditioners using Corollary 11 and Theorem 12:

- 523
- $\kappa(\mathcal{M}_{\text{NN},D} \mathcal{A}) \leq N_c/\lambda_{n_v+1}$;
 - $\kappa(\mathcal{M}_{\text{AS},D} \mathcal{A}) \leq N_c(1 + 1/\lambda_{n_v+1})$;
 - $\kappa(\mathcal{M}_{\text{AS},2} \mathcal{A}) \leq (N_c + 1)[N_c + 1 + (N_c + 2)/\lambda_{n_v+1}]$.

526 The Schur matrix \mathcal{S} is smaller and better conditioned [5, 27] than the original matrix \mathcal{K} .
 527 Furthermore, in a 2-level domain decomposition framework, eliminating the interior unknowns
 528 significantly improves the convergence by reducing the size of the coarse space needed to take into
 529 account the physical heterogeneity in the domain [15]. In accordance with these theoretical results,
 530 Figure 4 highlights the benefits of operating on \mathcal{S} (Figure 4, right) instead of \mathcal{K} (left): the condition
 531 number is consistently smaller when applying any aS method on \mathcal{S} instead of \mathcal{K} . Without a coarse
 532 space ($n_v = 0$, top), the results are consistent with Figure 1, top ($\chi = \infty$): the condition number
 533 κ increases with the number of subdomains N . Choosing $n_v = 1$, our coarse space reduces to a
 534 classical partition-of-unity coarse space [31] and is sufficient in the homogeneous case ($K = 1$, plain
 535 lines); we notice that NN,D/ \mathcal{S} then reduces to classical BDD where the condition number does
 536 not depend on N but remains fairly large for large values of K . However, in the heterogeneous
 537 cases ($K = 100$ or $10,000$, dashed lines), this simpler coarse space is not enough to get a scalable
 538 method: one eigenvector per high-conductivity inclusion is needed in the coarse space to build a
 539 robust method [15]. In our case, with 5 high-conductivity layers passing through all the subdomains,
 540 $n_v = 5$ eigenvectors are enough to bound the condition number for AS/ \mathcal{S} and NN/ \mathcal{S} . Using the
 541 Sh/ \mathcal{S} method, since two eigenproblems are solved in each subdomain, 10 vectors are needed to get

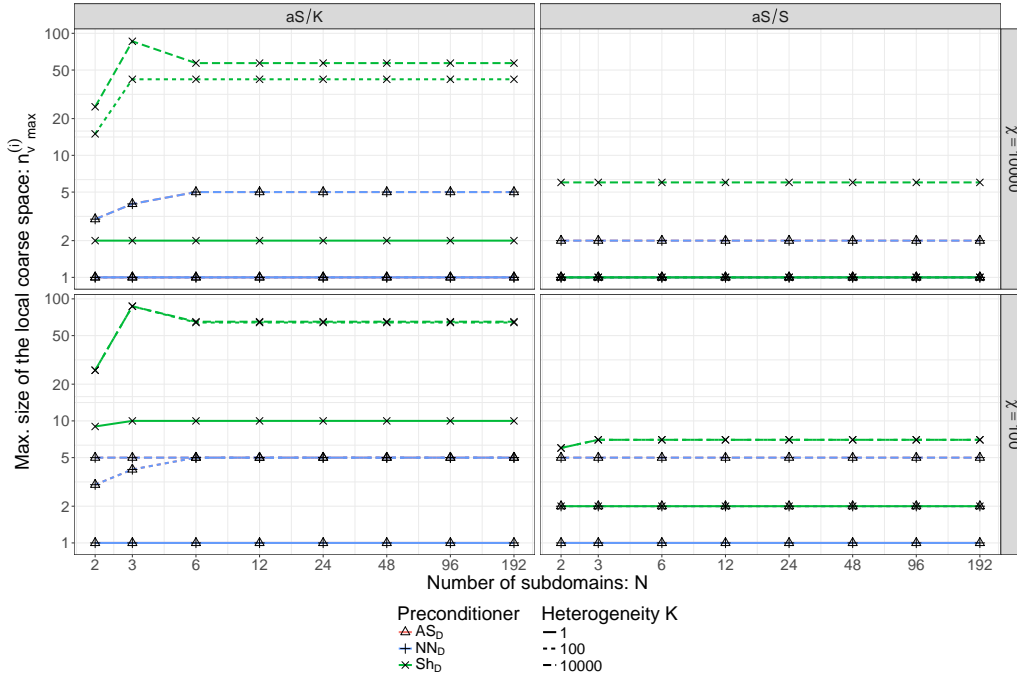


FIG. 3. Maximum size of the local coarse space when imposing an a priori bound χ on the condition number. Note that AS and NN overlap with each other. In most cases, only few vectors per subdomain are enough but the least robust methods can induce a relatively large local coarse space V_i^0 in some cases.

542 a good convergence (bottom right).

543 With a large enough coarse space, the three methods NN,D/S, AS,2/S and AS,D/S perform
 544 quite similarly, with a slight advantage for NN. However, when the coarse space is too small ($n_v = 1$
 545 and $K = 10,000$ for instance), AS,2/S and AS,D/S have a significantly smaller condition number
 546 than NN,D/S, and they appear more robust. As a consequence, we will choose for our proposed
 547 high performance implementation to focus on the AS,2/S method (Section 5.5).

548 **5.4. Approximate case: Empirical study of the impact of sparsification.** The con-
 549 vergence results for approximate aS methods in sections 2, 3 and 4 apply for both aS/K and aS/S
 550 cases. However, for a matter of conciseness, we now only focus on the latter context for illustrating
 551 the impact of approximation, as the above experiments showed the numerical benefits of operating
 552 on the Schur complement. For that, we approximate the dense matrix \mathcal{S}_i with a sparse matrix $\tilde{\mathcal{S}}_i$, by
 553 dropping some entries in the matrix. This process is called *sparsification*. In a very heterogeneous
 554 medium ($K \gg 1$), some entries in \mathcal{S} corresponding to couplings between unknown separated by a
 555 low-conductivity layer, are negligible. We use the symmetry-preserving strategy of dropping s_{ij} if
 556 $|s_{ij}| \leq \epsilon(s_{ii} + s_{jj})$, where ϵ is a parameter that controls the sparsity (see, e.g., [6]).

557 The benefits of sparsification are evaluated by assessing the proportion $nnz(LL^T)$ of non-zero
 558 elements in the Cholesky factorization $\hat{\mathcal{S}}_i = LL^T$ of the local preconditioner. In Figure 5, we
 559 evaluate the impact of sparsification on the robustness of the method. It appears that, up to a
 560 certain level, we are still able to find a robust coarse space despite having significantly reduced

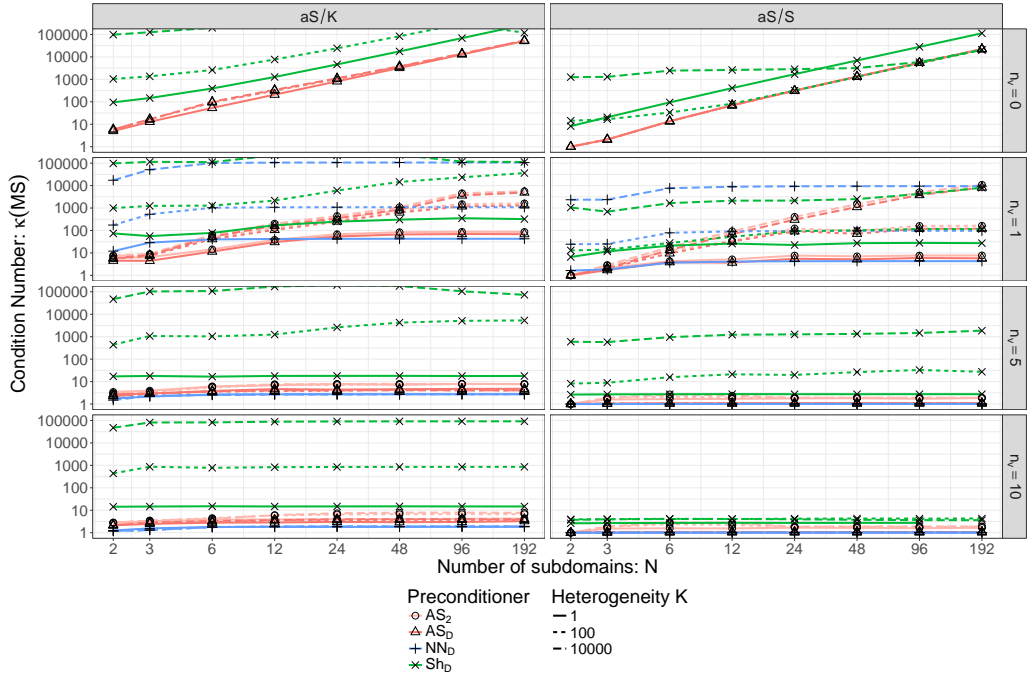


FIG. 4. Condition number when imposing an a priori size n_v for the local coarse space V_i^0 . We are still able to significantly reduce the condition number of the methods. The best convergence results are obtained with the AS,D/S method.

561 the memory footprint of the preconditioner. For instance, with a sparsity parameter of $\epsilon = 0.001$,
 562 although 88.8% of the entries in the factorization of the preconditioner are dropped, our coarse
 563 space with $n_v = 5$ vectors per subdomain still significantly improves the convergence.

564 These results are very promising as they show we can efficiently apply an approximate scheme
 565 to reduce the complexity of two-level aS methods. However, the considered sparsification technique
 566 is delicate for ensuring an *a priori* condition number. Approximation through hierarchical matri-
 567 ces [17] might better fit this objective, for bounding ω_- and ω_+ and ensure theorems 10 and 12
 568 apply. This is left for future work (see [1] for preliminary investigations in this direction) and we do
 569 not consider approximation techniques in the high performance implementation we propose below.

570 **5.5. Performance of AS,2/S on a modern parallel computer.** The excellent numerical
 571 properties exhibited above by the AS,2/S method motivated the design of an high-performance
 572 code of that variant. For that, we relied on the MaPHyS package and we added a coarse grid
 573 correction to the baseline, one-level AS/S variant [3] for the purpose of the present study. MaPHyS
 574 is a parallel hybrid (direct/iterative) sparse linear solver. Its *Setup* step relies on third-party sparse
 575 direct solvers for efficiently performing the elimination of the interior variables and computing the
 576 local Schur complement \mathcal{S}_i . Subdomains are processed concurrently, each subdomain being associ-
 577 ated with a process. The computation of the one-level preconditioner (still within the *Setup* step)
 578 is then performed with neighbor-to-neighbor communications. The *Solve* step consists of classical
 579 preconditioned conjugate gradient iterations. In particular, global synchronizations are only re-

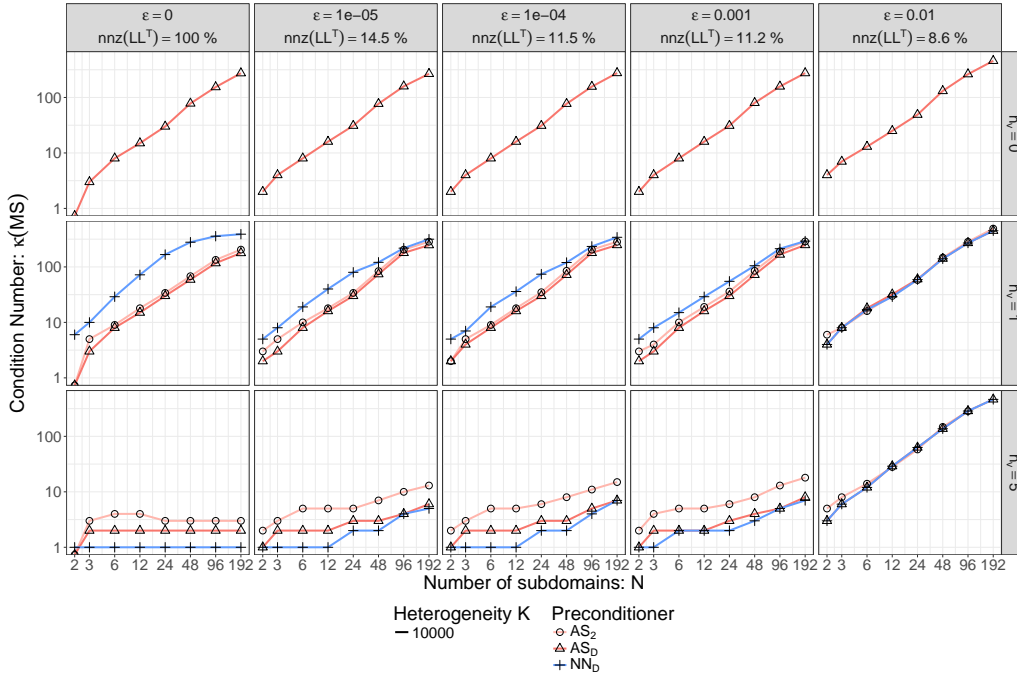


FIG. 5. Up to a certain level, the sparsification does not break the robustness of the method: using a big enough coarse space ($n_v = 5$), it is possible to discard 88.8% of the entries in the factorization of the preconditioner without losing convergence.

580 required for computing dot products while the matrix-vector product can be performed concurrently
 581 on each subdomain and the application of the (one-level) preconditioner only requires neighbor-to-
 582 neighbor communications. We extended MaPHYs to ensure a coarse grid correction as follows. In
 583 the *Setup* step, the generalized eigenproblems are processed concurrently on each subdomain; the
 584 matrix associated with the resulting coarse space is then assembled and factorized using a third-
 585 party parallel sparse direct solver. In the *Solve* step, a coarse solve is added in the application of the
 586 preconditioner at each iteration. Due to the nature of the coarse space, these operations add global
 587 communications and synchronizations in the algorithm and particular care must be taken in their
 588 implementation in order to achieve good scalability and parallel efficiency. Several parallelization
 589 strategies for the coarse correction are currently investigated and will be discussed in a future work.
 590 In the current experiment, the coarse matrix \mathcal{A}_0 is assembled and factorized redundantly on disjoint
 591 sub-communicators (obtained by splitting the global one) in order to reduce the number of global
 592 communications during the solve step.

593 We now present a weak scalability study conducted on test cases similar to the ones introduced
 594 in Section 5.1, but with larger subdomains. Each subdomain is indeed a cube discretized on
 595 a $31 \times 31 \times 31$ mesh with 29,791 unknowns. There are now 6 alternating conductivity layers
 596 ($K = 10,000$), and we consider a scenario with an imposed coarse space size (as in Section 5.3)
 597 using 3 vectors per subdomain. No approximation is performed. The same stopping criterion as
 598 above is used. The experiments have been conducted on the Occigen machine at CINES. Each node
 599 is composed of two Haswell (E5-2690V3) 12-core processors running at 2.6 GHz. A subdomain is

600 associated with a process, binded on a CPU core. MaPHyS was compiled with Intel 17.0 and Intel
 601 MPI 2017.0.098. All dense operations are performed with the Intel Math Kernel Library (MKL)
 602 2017 (including the Lapack `dsygvx` routine for solving the eigenproblems, that allows one to only
 603 compute a targeted subset of eigenpairs). Sparse factorizations are performed with the MUMPS
 604 5.0.2 sparse direct solver [4] together with the ParMetis 4.0.3 partitioner [21].

605 Table 1 compares the behavior of our extension of MaPHyS relying on the proposed coarse grid
 606 correction described above (AS,2/S) with the baseline, one-level version of MaPHyS [3] (AS/S).
 607 The number of subdomains N , which is equal to the number of MPI processes and CPU cores used
 608 for the respective computation, the total number of unknowns $n = (30N + 1) \times 31 \times 31$ and the size
 609 of the coarse space n_0 are provided in the table along with the maximum (among all subdomains)
 610 time in seconds needed to perform the *Setup* step, the *Solve* step or both steps (*Total*) and the
 611 number of PCG iterations performed during the *Solve* step, for both the AS/S method (left) and
 612 the AS,2/S method (right). The *Setup* step includes the time spent in the factorization of the
 613 local matrices and the computation of the local Schur complement matrix using a sequential sparse
 614 direct solver, the assembly and factorization of the local Schur complement, the solution of the
 615 generalized eigenproblems, the construction and the factorization of the coarse matrix. The *Solve*
 616 step corresponds to the PCG iterations and the final computation of the interior unknowns. We
 617 observe that the addition of the coarse correction increases the *Setup* time and the individual cost
 618 of each iteration (up to a factor 2), mainly due to the induced global communications. On the
 619 other hand, the number of iterations of AS,2/S remains stable, leading to a drastically overall
 620 reduced *Solve* time compared to the baseline AS/S method (up to a factor 37 when the 44,283,841
 621 unknowns are distributed among 1,536 subdomains). As a consequence, in a scenario consisting of
 622 solving a linear system with a single right-hand side, the coarse grid usage reduces the total time
 623 to solution (*Setup* + *Solve*) when the number of subdomains (and CPU cores) is equal to or higher
 624 than 384. In another common application scenario where multiple (say, p), successive, right-hand
 625 sides must be solved, the total time to solution (*Setup* + p *Solve*) may then essentially be governed
 626 by the *Solve* step if p is large. In that latter case, the benefits of the coarse grid may then thus be
 627 tremendous on large scale computers.

TABLE 1

A weak scalability study was performed using the MaPHyS parallel solver. The Setup, Solve and Total times are the max among all subdomains, in seconds (s). Each subdomain is associated with one MPI process binded on one CPU core. N is the number of subdomains, n is the size of \mathcal{K} and n_0 is the size of the coarse space. Without coarse correction, the Setup time remains stable, whereas the Solve time grows linearly with the number of domains. The coarse correction adds to the Setup time but keeps the number of iterations constant, thus improving the scalability. Without coarse correction, no convergence was achieved on 3,072 domains.

N	n	n_0	AS/S				AS,2/S			
			<i>Setup</i>	<i>Solve</i>	<i>Total</i>	# iter	<i>Setup</i>	<i>Solve</i>	<i>Total</i>	# iter
24	692k	72	3,64	0,47	4,12	33	6,13	0,30	6,44	15
48	1.4M	144	3,67	0,87	4,54	62	6,52	0,30	6,83	15
96	2.8M	288	3,79	1,62	5,41	119	6,52	0,31	6,84	15
192	5.6M	576	3,75	3,17	6,92	233	6,59	0,33	6,92	15
384	11.1M	1.1k	3,87	5,02	8,90	371	6,61	0,32	6,93	14
768	22.1M	2.3k	3,78	8,30	12,1	609	6,61	0,33	6,95	14
1536	44.3M	4.6k	4,13	15,1	19,2	1,077	6,96	0,40	7,38	14
3072	88.6M	9.2k	-	-	-	-	7,24	0,42	7,70	14

628 **6. Conclusion.** In this paper, we have proposed a new class of aS preconditioners, so-called
629 approximate aS preconditioners. These preconditioners are fully algebraic in the sense that they
630 do not require any other information apart from SPSD subdomain matrices. This class is wide as
631 it consists of all aS preconditioners, provided that their coarse space results from the assembly of
632 local components that contain the kernel of some local operators (Definition 1). In particular, it
633 generalizes the class of approximate NN preconditioners introduced in [25] (named *generalized NN*
634 in the original paper). We exhibited a bound on the condition number of all approximate deflated aS
635 preconditioners (Theorem 2). This bound depends on generalized Rayleigh quotients and generalizes
636 the result from [25] beyond the class of approximate NN methods. Applying a GenEO procedure
637 on two well chosen generalized eigenproblems, we proposed to explicitly compute the coarse space
638 while bounding these Rayleigh quotients leading to a bound on the condition number (hence on the
639 number of iterations of PCG) independent of the number of subdomains. We also showed that a
640 similar bound can be obtained when the coarse space is applied additively for the subclass of newly
641 introduced approximate AS methods.

642 The results presented in this paper can be readily derived to retrieve the bounds previously
643 obtained via generalized eigenproblems in the particular cases of AS/ \mathcal{K} [11, 33], NN/ \mathcal{S} [34] and
644 optimized Robin (SORAS) [18]. It also generalizes these results when used with approximate local
645 solvers. Furthermore, they allowed us to define a coarse space for the AS method applied on the
646 Schur complement (AS/ \mathcal{S}) [6], leading to an extremely robust substructuring method, for which
647 the coarse space can be applied either with deflation or additively, and with the freedom of relying
648 on an approximate local Schur complement. Numerical experiments illustrated these statements.
649 In particular, they motivated an high-performance design of a coarse grid correction for AS/ \mathcal{S} . We
650 implemented it within the MaPHyS package. Parallel experiments showed the significant benefits
651 that the resulting AS,2/ \mathcal{S} solver could bring.

652 A challenge opened by the present study is to determine an explicit procedure to perform the
653 approximation while achieving a given *a priori* bound on the condition number. We also plan to
654 study the effects of the method on the spectrum and on the empirical convergence of non symmetric
655 test cases.

656 **7. Acknowledgments.** The authors would like to thank Nicole Spillane for proofreading
657 an early draft of this manuscript. Experiments presented in sections 5.2 to 5.4 were carried
658 out using the PLAFRIM experimental testbed, being developed under the Inria PlaFRIM de-
659 velopment action with support from Bordeaux INP, LABRI and IMB and other entities: Conseil
660 Régional d’Aquitaine, Université de Bordeaux and CNRS (and ANR in accordance to the pro-
661 gramme d’investissements d’Avenir). Experiments presented in Section 5.5 were performed on the
662 GENCI Occigen cluster at CINES by Matthieu Kuhn and Gilles Marait. We would also like to
663 thank the anonymous referees whose constructive comments enabled us to significantly improve
664 the manuscript. Finally, this work was partially supported by the French research agency ANR in
665 the framework of the DEDALES project (ANR-14-CE23-0005), in particular the PhD thesis of the
666 third author (in the alphabetical order) was funded by this project.

667 **References.**

- 668 [1] E. AGULLO, E. DARVE, L. GIRAUD, AND Y. HARNESS, *Low-Rank Factorizations in Data*
669 *Sparse Hierarchical Algorithms for Preconditioning Symmetric Positive Definite Matrices*,
670 *SIAM Journal on Matrix Analysis and Applications*, 39 (2018), pp. 1701–1725, <https://hal.inria.fr/hal-01940053>.
671 [2] E. AGULLO, L. GIRAUD, A. GUERMOUCHE, AND J. ROMAN, *Parallel hierarchical hy-*
672

- 673 *brid linear solvers for emerging computing platforms*, *Comptes Rendus Mécanique*, 339
674 (2011), pp. 96–103, <https://doi.org/10.1016/j.crme.2010.11.005>, <http://www.sciencedirect.com/science/article/pii/S1631072110002068> (accessed 2018-03-08).
- 675
- 676 [3] E. AGULLO, L. GIRAUD, S. NAKOV, AND J. ROMAN, *Hierarchical hybrid sparse linear solver*
677 *for multicore platforms*, report, INRIA Bordeaux, Oct. 2016, [https://hal.inria.fr/hal-01379227/](https://hal.inria.fr/hal-01379227/document)
678 [document](https://hal.inria.fr/hal-01379227/document) (accessed 2018-03-26).
- 679 [4] P. R. AMESTOY, I. S. DUFF, J.-Y. L'EXCELLENT, AND J. KOSTER, *A fully asynchronous*
680 *multifrontal solver using distributed dynamic scheduling*, *SIAM Journal on Matrix Analysis*
681 *and Applications*, 23 (2001), pp. 15–41.
- 682 [5] S. C. BRENNER, *The condition number of the Schur complement in domain decomposition*,
683 *Numerische Mathematik*, 83 (1999), pp. 187–203.
- 684 [6] L. M. CARVALHO, L. GIRAUD, AND G. MEURANT, *Local preconditioners for two-level non-*
685 *overlapping domain decomposition methods*, *Numerical linear algebra with applications*, 8
686 (2001), pp. 207–227.
- 687 [7] B. COCKBURN, J. GOPALAKRISHNAN, AND R. LAZAROV, *Unified hybridization of discontin-*
688 *uous Galerkin, mixed, and continuous Galerkin methods for second order elliptic problems*,
689 *SIAM Journal on Numerical Analysis*, 47 (2009), pp. 1319–1365.
- 690 [8] Y.-H. DE ROECK AND P. LE TALLEC, *Analysis and test of a local domain decomposition*
691 *preconditioner*, in *Fourth International Symposium on Domain Decomposition Methods for*
692 *Partial Differential Equations*, vol. 4, 1991.
- 693 [9] C. R. DOHRMANN, *A preconditioner for substructuring based on constrained energy minimiza-*
694 *tion*, *SIAM Journal on Scientific Computing*, 25 (2003), pp. 246–258.
- 695 [10] V. DOLEAN, P. JOLIVET, AND F. NATAF, *An Introduction to Domain Decomposition Methods:*
696 *Algorithms, Theory, and Parallel Implementation*, vol. 144, SIAM, 2015.
- 697 [11] Y. EFENDIEV, J. GALVIS, R. LAZAROV, AND J. WILLEMS, *Robust domain decomposition*
698 *preconditioners for abstract symmetric positive definite bilinear forms*, *ESAIM: Mathematical*
699 *Modelling and Numerical Analysis*, 46 (2012), pp. 1175–1199, [http://www.esaim-m2an.org/](http://www.esaim-m2an.org/articles/m2an/abs/2012/05/m2an110073/m2an110073.html)
700 [articles/m2an/abs/2012/05/m2an110073/m2an110073.html](http://www.esaim-m2an.org/articles/m2an/abs/2012/05/m2an110073/m2an110073.html) (accessed 2017-06-08).
- 701 [12] C. FARHAT, M. LESOINNE, P. LETALLEC, K. PIERSON, AND D. RIXEN, *FETI-DP: A*
702 *dual–primal unified FETI method—part I: A faster alternative to the two-level FETI method*,
703 *International journal for numerical methods in engineering*, 50 (2001), pp. 1523–1544.
- 704 [13] C. FARHAT AND F.-X. ROUX, *A method of finite element tearing and interconnecting and its*
705 *parallel solution algorithm*, *International Journal for Numerical Methods in Engineering*, 32
706 (1991), pp. 1205–1227.
- 707 [14] V. FRAYSSÉ AND L. GIRAUD, *A set of conjugate gradient routines for real and complex arith-*
708 *metics*, CERFACS Technical Report TR/PA/00/47, (2000).
- 709 [15] J. GALVIS AND Y. EFENDIEV, *Domain Decomposition Preconditioners for Multiscale Flows in*
710 *High-Contrast Media*, *Multiscale Modeling & Simulation*, 8 (2010), pp. 1461–1483, [https://doi.](https://doi.org/10.1137/090751190)
711 [org/10.1137/090751190](https://doi.org/10.1137/090751190), <http://epubs.siam.org/doi/abs/10.1137/090751190> (accessed 2016-
712 09-15).
- 713 [16] M. J. GANDER, *Optimized schwarz methods*, *SIAM Journal on Numerical Analysis*, 44 (2006),
714 pp. 699–731.
- 715 [17] L. GRASEDYCK AND W. HACKBUSCH, *Construction and Arithmetics of H-Matrices*, *Comput-*
716 *ing*, 70 (2003), pp. 295–334, <https://doi.org/10.1007/s00607-003-0019-1>, [https://link.springer.](https://link.springer.com/article/10.1007/s00607-003-0019-1)
717 [com/article/10.1007/s00607-003-0019-1](https://link.springer.com/article/10.1007/s00607-003-0019-1) (accessed 2018-03-26).
- 718 [18] R. HAFERSAS, P. JOLIVET, AND F. NATAF, *An Additive Schwarz Method Type Theory for*

- 719 *Lions's Algorithm and a Symmetrized Optimized Restricted Additive Schwarz Method*, SIAM
720 Journal on Scientific Computing, 39 (2017), pp. A1345–A1365.
- 721 [19] P. HÉNON, P. RAMET, AND J. ROMAN, *PASTIX: A high-performance parallel direct solver*
722 *for sparse symmetric positive definite systems*, Parallel Computing, 28 (2002), pp. 301–321.
- 723 [20] P. JOLIVET, F. HECHT, F. NATAF, AND C. PRUD'HOMME, *Scalable Domain Decomposition*
724 *Preconditioners for Heterogeneous Elliptic Problems*, in Proceedings of the International
725 Conference on High Performance Computing, Networking, Storage and Analysis, SC '13,
726 New York, NY, USA, 2013, ACM, pp. 80:1–80:11, <https://doi.org/10.1145/2503210.2503212>,
727 <http://doi.acm.org/10.1145/2503210.2503212> (accessed 2018-03-26).
- 728 [21] G. KARYPIS AND V. KUMAR, *MeTis: Unstructured Graph Partitioning and Sparse Matrix*
729 *Ordering System, Version 4.0*, 2009, <http://www.cs.umn.edu/~metis>.
- 730 [22] A. KLAWONN, M. KUHN, AND O. RHEINBACH, *Adaptive coarse spaces for FETI-DP in three*
731 *dimensions*, SIAM Journal on Scientific Computing, 38 (2016), pp. A2880–A2911.
- 732 [23] A. KLAWONN, M. KÜHN, AND O. RHEINBACH, *Adaptive FETI-DP and BDDC methods with*
733 *a generalized transformation of basis for heterogeneous problems*, Electronic Transactions on
734 Numerical Analysis, 49 (2018), pp. 1–27.
- 735 [24] A. KLAWONN, P. RADTKE, AND O. RHEINBACH, *A comparison of adaptive coarse spaces for*
736 *iterative substructuring in two dimensions*, Electron. Trans. Numer. Anal, 45 (2016), pp. 75–
737 106.
- 738 [25] P. LE TALLEC AND M. VIDRASCU, *Generalized Neumann-Neumann preconditioners for iter-*
739 *ative substructuring*, in Domain Decomposition Methods in Sciences and Engineering, 1998.
- 740 [26] J. MANDEL, *Balancing domain decomposition*, International Journal for Numerical Methods
741 in Biomedical Engineering, 9 (1993), pp. 233–241.
- 742 [27] L. MANSFIELD, *On the Conjugate Gradient Solution of the Schur Complement System Ob-*
743 *tained from Domain Decomposition*, SIAM Journal on Numerical Analysis, 27 (1990), pp. 1612–
744 1620, <https://doi.org/10.1137/0727094>, <http://epubs.siam.org/doi/abs/10.1137/0727094> (ac-
745 cessed 2015-07-16).
- 746 [28] T. P. A. MATHEW, *Domain Decomposition Methods for the Numerical Solution of Partial*
747 *Differential Equations*, vol. 61, Springer Science & Business Media, 2008.
- 748 [29] F. NATAF, H. XIANG, V. DOLEAN, AND N. SPILLANE, *A Coarse Space Construction Based on*
749 *Local Dirichlet-to-Neumann Maps*, SIAM Journal on Scientific Computing, 33 (2011), pp. 1623–
750 1642, <https://doi.org/10.1137/100796376>, <http://epubs.siam.org/doi/abs/10.1137/100796376>
751 (accessed 2014-11-18).
- 752 [30] A. QUARTERONI AND A. VALLI, *Domain Decomposition Methods for Partial Differential Equa-*
753 *tions*, Oxford University Press, 1999.
- 754 [31] M. SARKIS, *Partition of unity coarse spaces: Enhanced versions, discontinuous coefficients and*
755 *applications to elasticity*, Domain decomposition methods in science and engineering, (2003),
756 pp. 149–158.
- 757 [32] N. SPILLANE, *Méthodes de Décomposition de Domaine Robustes Pour Les Problèmes*
758 *Symétriques Définis Positifs*, PhD Thesis, Paris 6, 2014.
- 759 [33] N. SPILLANE, V. DOLEAN, P. HAURET, F. NATAF, C. PECHSTEIN, AND R. SCHE-
760 ICHL, *Abstract robust coarse spaces for systems of PDEs via generalized eigenproblems*
761 *in the overlaps*, Numerische Mathematik, 126 (2014), pp. 741–770, [https://doi.org/](https://doi.org/10.1007/s00211-013-0576-y)
762 [10.1007/s00211-013-0576-y](https://doi.org/10.1007/s00211-013-0576-y), <https://link.springer.com/article/10.1007/s00211-013-0576-y> (ac-
763 cessed 2018-03-26).
- 764 [34] N. SPILLANE AND D. J. RIXEN, *Automatic spectral coarse spaces for robust finite element tear-*

765 *ing and interconnecting and balanced domain decomposition algorithms*, International Journal
766 for Numerical Methods in Engineering, 95 (2013), pp. 953–990, [https://doi.org/10.1002/nme.](https://doi.org/10.1002/nme.4534)
767 [4534](https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.4534), <https://onlinelibrary.wiley.com/doi/abs/10.1002/nme.4534>.
768 [35] A. TOSELLI AND O. WIDLUND, *Domain Decomposition Methods-Algorithms and Theory*,
769 vol. 34, Springer Science & Business Media, 2006.